Matrix Padé Approximation: Recursive Computations *

Xu Guo-liang †
Computing Center, Academia Sinica, Beijing, China
A. Bultheel
Department of Computer Science, K.U.Leuven, Belgium

Abstract

In this paper we consider computational aspects of the matrix Padé approximants whose definitions and properties were considered in an accompanying paper. The notations and definitions are briefly reviewed and the solution of the linear systems of Padé equations is studied. In certain cases the computations can be done recursively.

1 Introduction

In [1] the authors have studied the general matrix Padé approximation problem. It turned out that it was not only necessary to consider left and right approximants, a duality imposed by the non commutativity of the matrix multiplication, but also type I and type II approximants, depending on the normalization of the denominator. In this paper, we shall briefly review the definitions and notations from [1] and then consider some possible methods to compute the approximants.

We shall consider matrix Padé approximants for $f(z)$ with

$$f(z) = \sum_{k=0}^{\infty} c_k z^k, \qquad c_k \in \mathbf{C}^{p\times m}, \tag{1.1}$$

where $\mathbf{C}^{p\times m}$ consists of all $p \times m$ matrices with their elements in the complex plane $\mathbf{C}$.

Furthermore set

$$\mathbf{H}_k := \{p(z) : p(z) = \sum_{i=0}^{k} a_i z^i, \quad a_i \in \mathbf{C}\} \text{ and } \mathbf{E}_k := \{e(z) : e(z) = \sum_{i=k+1}^{\infty} a_i z^i, \quad a_i \in \mathbf{C}\}.$$

The nonnegative integers are denoted by $\mathbf{Z}_+$. For $V = (v_{ij}) \in \mathbf{Z}_+^{p\times m}$, set

$$\mathbf{H}_V^{p\times m} = \{P(z) = (p_{i,j}(z))_{i,j=1}^{p,m} : p_{ij}(z) \in \mathbf{H}_{v_{ij}}\},$$

$$\mathbf{E}_V^{p\times m} = \{R(z) = (r_{i,j}(z))_{i,j=1}^{p,m} : r_{ij}(z) \in \mathbf{E}_{v_{ij}}\}.$$

We look for rational matrix approximants $N(z)M(z)^{-1}$ of $f(z)$ with $N(z) \in \mathbf{H}_V^{p\times m}$, $M(z) \in \mathbf{H}_U^{m\times m}$, and such that $f(z)M(z) - N(z) \in \mathbf{E}_E^{p\times m}$. So normally $V, E \in \mathbf{Z}_+^{p\times m}$ while $U \in \mathbf{Z}_+^{m\times m}$. We shall however use degrees $V \in \mathbf{Z}_+^{p\times 1}$ or $V \in \mathbf{Z}_+^{1\times m}$ or even $V \in \mathbf{Z}_+$. In this situation, we suppose that we denote row degrees, column degrees, or a constant overall degree : V is then thought to be expanded to a full $p \times m$ matrix with constant row elements or constant column elements or constants all over. Similar considerations are valid for U and E.

Since the treatment of left and right approximants is completely parallel, we shall only treat the right approximants and drop the prefix "right" from the terminology.

We are now ready to give the definitions of type I and type II matrix Padé approximants (MPAs). For a type I MPA, the normalization is a condition on the constant term of the denominator.

*The work was done while the first author worked at the Department of Computer Science, K.U.Leuven

†Supported by the National Science Foundation of China for youth

Definition 1.1 (Type I MPA) *Let $f(z) \in \mathbf{C}^{p\times m}[[z]]$ be a (formal) power series with coefficients in $\mathbf{C}^{p\times m}$ and let $V = (v_{ij}) \in \mathbf{Z}_+^{p\times m}$, $U = (u_{ij}) \in \mathbf{Z}_+^{m\times m}$ and $W = (w_{ij}) \in \mathbf{Z}_+^{p\times m}$ such that*

$$\sum_{k=1}^{p} w_{kj} = \sum_{k=1}^{p} v_{kj} + \sum_{k=1}^{m} u_{kj}, \quad j = 1, 2, \ldots, m, \tag{1.2}$$

and $W - V \geq 0$. Then the Padé approximation problem of the first type is denoted by $(V, U, W)_I^f$ and consists in finding polynomials $N(z) \in \mathbf{H}_V^{p\times m}$ and $M(z) \in \mathbf{H}_U^{m\times m}$ such that

1. $f(z)M(z) - N(z) \in \mathbf{E}_W^{p\times m}$.
2. $M(0) = I$.

The collection of all rational functions $N(z)M(z)^{-1}$ formed by the solutions of $(V, U, W)_I^f$ is denoted by $[V, U, W]_I^f$.

For a type II MPA, the normalization condition is on the leading coefficient.

Definition 1.2 (Type II MPA) *Let $f(z) \in \mathbf{C}^{p\times m}[[z]]$ be a (formal) power series with coefficients in $\mathbf{C}^{p\times m}$ and let $V = (v_{ij}) \in \mathbf{Z}_+^{p\times m}$, $U = (u_j) \in \mathbf{Z}_+^{m\times 1}$ and $W = (w_{ij}) \in \mathbf{Z}_+^{p\times m}$ such that*

$$\sum_{k=1}^{p} w_{kj} = \sum_{k=1}^{p} v_{kj} + \sum_{\substack{u_i > u_j \\ 1 \leq i \leq m}} u_j + \sum_{\substack{u_i \leq u_j \\ 1 \leq i \leq m}} u_i + \sum_{\substack{u_i > u_j \\ 1 \leq i < j}} 1, \quad j = 1, 2, \ldots, m, \tag{1.3}$$

and $W - V \geq 0$. Then the Padé approximation problem of the second type is denoted by $(V, U, W)_{II}^f$ and consists in finding polynomials $N(z) \in \mathbf{H}_V^{p\times m}$ and $M(z) \in \mathbf{H}_U^{m\times m}$ such that

1. $f(z)M(z) - N(z) \in \mathbf{E}_W^{p\times m}$.
2. *$M(z)$ is $C1$ canonical .*

The solution set of $(V, U, W)_{II}^f$ is denoted by $[V, U, W]_{II}^f$.

A $C1$ canonical matrix $M(z) \in \mathbf{H}_V^{m\times m}$ is the following

Definition 1.3 ($C1$ canonical) *We say that $M(z) \in \mathbf{H}_V^{m\times m}$ is $C1$ canonical iff*
(i) There is a vector $U \in \mathbf{Z}_+^{m\times 1}$ such that $M(z) \in \mathbf{H}_U^{m\times m} \cap \mathbf{H}_{U^T}^{m\times m}$.
(ii) The leading row coefficient matrix M^{hr} of $M(z)$ is the unit matrix.
(iii) The leading column coefficient matrix M^{hc} of $M(z)$ is unit upper triangular.

It was shown in [1] that the denominator coefficients are solutions of linear systems which we shall presently describe. For

$$g(z) = \sum_{k=0}^{\infty} a_k z^k, \quad a_k \in \mathbf{C}^{s\times t}, \qquad a_k = 0, \quad k < 0,$$

let

$$T_{mn}^l(g) = \begin{bmatrix} a_l & a_{l-1} & \cdots & a_{l-n+1} \\ a_{l+1} & a_l & \cdots & a_{l-n+2} \\ \cdots & \cdots & \cdots & \cdots \\ a_{l+m-1} & a_{l+m-2} & \cdots & a_{l+m-n} \end{bmatrix} \in \mathbf{C}^{ms\times nt}$$

be a block Toeplitz matrix. Let $f = (f_{ij}) \in \mathbf{C}^{p\times m}[[z]]$ be given. For the problem $(V, U, W)_I^f$ we introduce the matrices $E = (e_{ij}) = W - V$ and

$$H_I^j(V, U, W) = \begin{bmatrix} T_{e_{1j},u_{1j}}^{v_{1j}}(f_{11}) & \cdots & T_{e_{1j},u_{mj}}^{v_{1j}}(f_{1m}) \\ \cdots & \cdots & \cdots \\ T_{e_{pj},u_{1j}}^{v_{pj}}(f_{p1}) & \cdots & T_{e_{pj},u_{mj}}^{v_{pj}}(f_{pm}) \end{bmatrix} \text{ and } B_I^j(V, U, W) = \begin{bmatrix} T_{e_{1j},1}^{v_{1j}+1}(f_{1j}) \\ \cdots \\ T_{e_{pj},1}^{v_{pj}+1}(f_{pj}) \end{bmatrix}.$$

Using these notations, the problem of determining the j-th column of the denominator $M(z)$ in $(V,U,W)_I^f$ can be expressed by the following equations.

$$H_I^j(V,U,W)\ X = -B_I^j(V,U,W). \tag{1.4}$$

If the j-th column $M_j(z)$ of $M(z)$ is determined, the j-th column $N_j(z)$ of the numerator $N(z)$ can be easily found from $N_j(z) = (fM_j)^{(V_j)}(z)$, where V_j is the degree vector of $N_j(z)$ and $(fM_j)^{(V_j)}$ is $f(z)M_j(z) \bmod z^{V_j}$.

For type II problem $(V,U,W)_{II}^f$ with $V, W \in \mathbf{Z}_+^{p\times m}$ and $U \in \mathbf{Z}_+^{m\times 1}$ we associate with the sequence $(u_i)_1^m$, the matrices

$$C(U) = (c_{ij})_{i,j=1}^m, \quad T(U) = (t_{ij})_{i,j=1}^m \quad \text{and} \quad S(U) = (s_{ij})_{i,j=1}^m, \tag{1.5}$$

where

$$c_{ij} = \begin{cases} 1, & \text{for } i \geq j, \text{ or } u_i \leq u_j \text{ and } i < j \\ 0, & \text{otherwise ,} \end{cases} \tag{1.6}$$

$$t_{ij} = \min\{u_i, u_j\} - c_{ij}, \quad \text{and } s_{ij} = t_{ij} + 1. \tag{1.7}$$

Using these notations, the problem for determining the j-th column of the denominator $M(z)$ in the problem $(V,U,W)_{II}^f$ can be expressed by the following set of equations.

$$H_{II}^j(V,U,W)\ X = -B_{II}^j(V,U,W), \tag{1.8}$$

where

$$H_{II}^j(V,U,W) = \begin{bmatrix} T_{e_{1j},s_{1j}}^{v_{1j}+1}(f_{11}) & \cdots & T_{e_{1j},s_{mj}}^{v_{1j}+1}(f_{1m}) \\ \cdots & \cdots & \cdots \\ T_{e_{pj},s_{1j}}^{v_{pj}+1}(f_{p1}) & \cdots & T_{e_{pj},s_{mj}}^{v_{pj}+1}(f_{pm}) \end{bmatrix}, \quad B_{II}^j(V,U,W) = \begin{bmatrix} T_{e_{1j},1}^{v_{1j}-u_j+1}(f_{1j}) \\ \cdots \\ T_{e_{pj},1}^{v_{pj}-u_j+1}(f_{pj}) \end{bmatrix},$$

and $E = (e_{ij}) = W - V$.

All the details about the definitions and these systems can be found in [1]. We shall now concentrate on the computations.

2 Recursive Computation

We assume now that $V, W \in \mathbf{Z}_+^{p\times 1}$ and $U \in \mathbf{Z}_+^{m\times 1}$, so that we don't have to mention this condition every time. In this case $H_I^j(V,U,W)$ does not depend on j. So we denote it as $H_I(V,U,W)$. On the other hand, we mainly consider the computation of type I right MPAs. The computation of the second type will be shown to be equivalent to the computation of the first type under some conditions. So the right subscript I or II in the notations will be deleted if there is no confusion. Consider the set

$$[V,U,W] = \{(N,M) : NM^{-1} \in [V,U,W]^f\}.$$

The problem we want to solve here is to compute $[V,U,W]$ from two "previous" ones. We shall need some normality condition for f which is different from the one we have defined earlier (see [1]). We call the new concept I-normality. It is defined as follows :

Definition 2.1 (I-normal) *If for any V, U and W which satisfy condition (1.2), the matrix $H(V,U,W)$ is nonsingular. Then we say that f is I-normal.*

For $(N, M) \in [V, U, W]$, we introduce the following notations for the elements of the numerator, denominator and residual

$$N_{ij}(z) = \sum_{k=0}^{v_i} n_{ij}^{(k)} z^k, \quad M_{ij}(z) = \sum_{k=0}^{u_i} m_{ij}^{(k)} z^k \text{ and } (fM - N)_{ij}(z) = \sum_{k=0}^{\infty} e_{ij}^{(k)} z^k.$$

We introduce the following matrices:

$$N(V, U, W, V') = \begin{bmatrix} N_1^{(v_1'+1)}(V, U, W) \\ \vdots \\ N_1^{(v_1)}(V, U, W) \\ \vdots \\ N_p^{(v_p'+1)}(V, U, W) \\ \vdots \\ N_p^{(v_p)}(V, U, W) \end{bmatrix}, \quad M(V, U, W) = \begin{bmatrix} M_1^{(0)}(V, U, W) \\ \vdots \\ M_1^{(u_1)}(V, U, W) \\ \vdots \\ M_m^{(0)}(V, U, W) \\ \vdots \\ M_m^{(u_m)}(V, U, W) \end{bmatrix},$$

and

$$E(V, U, W, W') = \begin{bmatrix} E_1^{(w_1+1)}(V, U, W) \\ \vdots \\ E_1^{(w_1')}(V, U, W) \\ \vdots \\ E_p^{(w_p+1)}(V, U, W) \\ \vdots \\ E_p^{(w_p')}(V, U, W) \end{bmatrix},$$

where

$$N_i^{(k)}(V, U, W) = [n_{i1}^{(k)} \ n_{i2}^{(k)} \ \cdots \ n_{im}^{(k)}], \quad M_i^{(k)}(V, U, W) = [m_{i1}^{(k)} \ m_{i2}^{(k)} \ \cdots \ m_{im}^{(k)}], \text{ and}$$

$$E_i^{(k)}(V, U, W) = [e_{i1}^{(k)} \ e_{i2}^{(k)} \ \cdots \ e_{im}^{(k)}].$$

The degrees $V' + 1$ and orders W' satisfy

$$V' + 1 \in \mathbf{Z}_+^{p \times 1} \quad \text{and} \quad V' \le V, \ \ W' \ge W.$$

It can be seen that $N(V, U, W, -1)$ is the coefficient matrix of N, $M(V, U, W)$ is the coefficient matrix of M, and $E(V, U, W, \infty)$ is the coefficient matrix of $R = fM - N$. Hence from the definition of MPA, we have

$$H(V' + 1, U + 1, V + 1) M(V, U, W) = N(V, U, W, V'), \tag{2.9}$$

$$H(V + 1, U + 1, W + 1) M(V, U, W) = 0, \tag{2.10}$$

$$H(W + 1, U + 1, W' + 1) M(V, U, W) = E(V, U, W, W'). \tag{2.11}$$

Lemma 2.1 *Let f be I–normal. Then for any V, U, W satisfying (1.2), we have*

(i) The matrix

$$\begin{bmatrix} N(V, U, W, V') \\ E(V, U, W, W') \end{bmatrix} \text{ is nonsingular,} \tag{2.12}$$

where

$$-1 \le V' \le V, \ \ W' \ge W \text{ and } |V - V'| + |W' - W| = m. \tag{2.13}$$

(ii) The leading coefficient matrix of the denominator M

$$M^{hc}(V,U,W) = \begin{bmatrix} M_1^{(u_1)}(V,U,W) \\ \vdots \\ M_m^{(u_m)}(V,U,W) \end{bmatrix} \text{ is nonsingular.}$$

(iii) The matrix

$$\begin{bmatrix} N(V,U,W,V') \\ E(V,U,W,W') \end{bmatrix} - \begin{bmatrix} N(V,U+1,W'',V') \\ 0 \end{bmatrix} \text{ is nonsingular,} \tag{2.14}$$

where V', W' satisfy (2.13), $W'' \geq W'$ and $(V,\ U+1,\ W'')$ satisfy (1.2).

Proof. (i) For $N \in \mathbf{Z}_+^{p\times 1}$, define

$$\theta_N^{(j)} = \begin{bmatrix} \theta_{n_j,n_1} & \theta_{n_j,n_2} & \cdots & \theta_{n_j,n_p} \end{bmatrix} \text{ and } I_N^{(j)} = \begin{bmatrix} \theta_{n_j,n_1} & \cdots & \theta_{n_j,n_{j-1}} & I_{n_j} & \theta_{n_j,n_{j+1}} & \cdots & \theta_{n_j,n_p} \end{bmatrix},$$

where $\theta_{q,k} \in \mathbf{C}^{q\times k}$ is a zero matrix. For $N_i \in \mathbf{Z}_+^{p\times 1}$, $i = 0, 1, \ldots, k$, and $N_0 \leq N_1 \leq \cdots \leq N_k$, let

$$L^{(j)}(N_0, N_1, \ldots, N_k) = \begin{bmatrix} I^{(j)}_{N_1-N_0} & \theta^{(j)}_{N_2-N_1} & \cdots & \theta^{(j)}_{N_k-N_{k-1}} \\ \theta^{(j)}_{N_1-N_0} & I^{(j)}_{N_2-N_1} & \cdots & \theta^{(j)}_{N_k-N_{k-1}} \\ \cdots & \cdots & \cdots & \cdots \\ \theta^{(j)}_{N_1-N_0} & \theta^{(j)}_{N_2-N_1} & \cdots & I^{(j)}_{N_k-N_{k-1}} \end{bmatrix},$$

and

$$L(N_0, N_1, \ldots, N_k) = \begin{bmatrix} L^{(1)}(N_0, N_1, \ldots, N_k) \\ L^{(2)}(N_0, N_1, \ldots, N_k) \\ \vdots \\ L^{(p)}(N_0, N_1, \ldots, N_k) \end{bmatrix}.$$

Then

$$H(V'+1, U+1, W'+1) = L(V', V, W, W') \begin{bmatrix} H(V'+1, U+1, V+1) \\ H(V+1, U+1, W+1) \\ H(W+1, U+1, W'+1) \end{bmatrix}. \tag{2.15}$$

Let

$$I(U) = \begin{bmatrix} \theta_U \\ I_U^{(1)} \\ \vdots \\ \theta_U \\ I_U^{(m)} \end{bmatrix} \text{ and } J(U) = \begin{bmatrix} I_U^{(1)} \\ \theta_U \\ \vdots \\ I_U^{(m)} \\ \theta_U \end{bmatrix}$$

where $\theta_U \in \mathbf{C}^{1\times|U|}$ is a zero vector. Then

$$\operatorname{rank} H(V'+1, U+1, W'+1) = \operatorname{rank}\{H(V'+1, U+1, W'+1)\,[I(U)\ M(V,U,W)]\}$$

$$= \operatorname{rank} \begin{bmatrix} H(V', U, V) & H(V'+1, U+1, V+1) M(V,U,W) \\ H(V, U, W) & 0 \\ H(W, U, W') & H(W+1, U+1, W'+1) M(V,U,W) \end{bmatrix}$$

$$= \operatorname{rank} H(V,U,W) + \operatorname{rank} \begin{bmatrix} N(V,U,W,V') \\ E(V,U,W,W') \end{bmatrix}. \tag{2.16}$$

Now (2.12) follows from (2.16) and the I-normality of f.

(ii) Let

$$\bar{M}(V,U,W) = \begin{bmatrix} M_1^{(0)}(V,U,W) \\ \vdots \\ M_1^{(u_1-1)}(V,U,W) \\ \vdots \\ M_m^{(0)}(V,U,W) \\ \vdots \\ M_m^{(u_m-1)}(V,U,W) \end{bmatrix}.$$

Then

$$\begin{aligned}
\operatorname{rank} H(V+1,U,W+1) &= \operatorname{rank} \{ H(V+1,U,W+1) \, [I(U-1) \ \bar{M}(V,U,W)] \} \\
&= \operatorname{rank} [H(V,U-1,W) \quad H(V+1,U,W+1)\bar{M}(V,U,W)] \\
&= \operatorname{rank} \left[H(V,U-1,W) \quad - \left[C^1(V,U,W) \cdots C^m(V,U,W) \right] M^{hc}(V,U,W) \right] \\
&= \operatorname{rank} \left\{ \left[H(V,U-1,W) \ C^1(V,U,W) \cdots C^m(V,U,W) \right] \begin{bmatrix} I_{|U|-m} & 0 \\ 0 & -M^{hc}(V,U,W) \end{bmatrix} \right\} \\
&= \operatorname{rank} \left\{ H(V,U,W) \begin{bmatrix} I_{|U|-m} & 0 \\ 0 & -M^{hc}(V,U,W) \end{bmatrix} \right\} \\
&= \operatorname{rank} \begin{bmatrix} I_{|U|-m} & 0 \\ 0 & -M^{hc}(V,U,W) \end{bmatrix},
\end{aligned}$$

where

$$C^j(V,U,W) = \begin{bmatrix} T_{w_1-v_1,1}^{v_1-u_j+1}(f_{1j}) \\ \vdots \\ T_{w_p-v_p,1}^{v_p-u_j+1}(f_{pj}) \end{bmatrix}.$$

Hence conclusion (ii) holds.

(iii) Let

$$M_1(V,U,W) = \begin{bmatrix} M_1^{(1)}(V,U,W) - M_1^{(1)}(V,U+1,W'') \\ \vdots \\ M_1^{(u_1)}(V,U,W) - M_1^{(u_1)}(V,U+1,W'') \\ M_1^{(u_1+1)}(V,U+1,W'') \\ \vdots \\ M_m^{(1)}(V,U,W) - M_m^{(1)}(V,U+1,W'') \\ \vdots \\ M_m^{(u_m)}(V,U,W) - M_m^{(u_m)}(V,U+1,W'') \\ M_m^{(u_m+1)}(V,U+1,W'') \end{bmatrix}.$$

Noting that $M(0) = I$ and that the leading coefficient matrix $M^{hc}(V, U+1, W'')$ is nonsingular, we have by (2.15) that

$$\begin{aligned}
&\operatorname{rank} H(V', U+1, W') \\
&\quad = \operatorname{rank} \{ H(V', U+1, W') \, [M_1(V, U, W) \;\; J(U)] \} \\
&\quad = \operatorname{rank} \begin{bmatrix} N(V,U,W,V') - N(V,U+1,W'',V') & H(V',U,V)) \\ 0 & H(V,U,W) \\ E(V,U,W,W') & H(W,U,W') \end{bmatrix} \\
&\quad = \operatorname{rank} \left\{ \begin{bmatrix} N(V,U,W,V') \\ E(V,U,W,W') \end{bmatrix} - \begin{bmatrix} N(V,U+1,W'',V') \\ 0 \end{bmatrix} \right\} + \operatorname{rank} H(V,U,W),
\end{aligned}$$

Thus conclusion (iii) holds. □

As a consequence of this we have

Corollary 2.2 *Suppose f is I-normal. Then*

(i) $E(V, U, W, W')$ is nonsingular, when $W' \geq W$ and $|W' - W| = m$.

(ii) $N(V, U, W, V')$ is nonsingular, when $-1 \leq V' \leq V$ and $|V - V'| = m$.

These are two special cases of the conclusion (i) of Lemma 2.1.

Corollary 2.3 *Suppose f is I-normal. If U satisfies $|u_i - u_j| \leq 1 \;\forall\; i, j$, $u_1 \geq u_2 \cdots \geq u_m$, then the two types of MPA problems are equivalent.*

Proof. From conclusion (ii) of Lemma 2.1, it follows that the leading coefficient matrix M^{hc} of $M(z)$ is nonsingular. Multiply (N, M) from the right with M^{hc} and the leading coefficient matrix of M will become the unit matrix. Then the denominator is made $C1$ canonical. On the other hand, $[V, U, W]_{II}$ exist uniquely, since $H_{II}(V, U, W) = H_I(V+1, U, W+1)$. Then the equivalency is proved. □

This is the reason why we don't have to consider the computation of MPAs of type II.

Theorem 2.4 *Suppose $m \geq p$ and f is I-normal. Then, for given V, U, W satisfying (1.2), and $W' \geq W$, $|W' - W| = m$,*

(i) there exists a nonsingular $Q(V, U, W, W') \in \mathbf{C}^{m \times m}$ such that

$$[V+1, U+1, W'+1] = z[V, U, W] Q(V, U, W, W') + [V+1, U, W+1], \tag{2.17}$$

and

$$Q(V, U, W, W') = -E(V, U, W, W')^{-1} E(V+1, U, W+1, W'+1). \tag{2.18}$$

(ii) If $W' \geq W+1$, there exists a rank p matrix $D(V, U, W, W') \in \mathbf{C}^{m \times m}$ such that

$$[V+1, U+1, W'+1] = z[V, U, W] D(V, U, W, W') + [V, U+1, W'], \tag{2.19}$$

and

$$D(V, U, W, W') = -E(V, U, W, W')^{-1} H(W+2, U+2, W'+2) M(V, U+1, W'). \tag{2.20}$$

(iii) If $W' \geq W+1$, there exists a rank p matrix $T(V, U, W, W') \in \mathbf{C}^{m \times m}$ such that

$$[V-1, U+1, W'-1] = [V, U, W] T(V, U, W, W') + [V, U+1, W'](I - T(V, U, W, W')), \tag{2.21}$$

and

$$T(V,U,W,W') =$$

$$-\begin{bmatrix} N(V,U,W,V-1) - N(V,U+1,W',V-1) \\ E(V,U,W,W'-1) \end{bmatrix}^{-1} \begin{bmatrix} N(V,U+1,W',V-1) \\ 0 \end{bmatrix}. \quad (2.22)$$

(iv) If $W' \geq W+1$, there exists a nonsingular $S(V,U,W,W') \in \mathbf{C}^{m\times m}$ such that

$$[V-1,U+1,W'-1] = [V,U,W] + z[V-1,U,W-1]S(V,U,W,W'), \quad (2.23)$$

and

$$S(V,U,W,W') = -\begin{bmatrix} N(V-1,U,W-1,V-2) \\ E(V-1,U,W-1,W'-2) \end{bmatrix}^{-1} \begin{bmatrix} N(V,U,W,V-1) \\ E(V,U,W,W'-1) \end{bmatrix}. \quad (2.24)$$

Proof. (i) Let

$$(N_1,M_1) = [V,U,W], \quad (N_2,M_2) = [V+1,U,W+1], \quad (N_3,M_3) = [V+1,U+1,W'+1].$$

Then by

$$fM_1 - N_1 = O(z^{W+1}), \quad fM_2 - N_2 = O(z^{W+2}) \quad \text{and} \quad fM_3 - N_3 = O(z^{W'+2}),$$

we have that $Q(V,U,W,W')$ satisfies

$$E(V,U,W,W')Q(V,U,W,W') + E(V+1,U,W+1,W'+1) = 0.$$

From Lemma 2.1, it follows that (2.17) and (2.18) hold and $Q(V,U,W,W')$ is nonsingular.

(ii) Let $(N_1,M_1) = [V,U,W]$, $(N_2,M_2) = [V,U+1,W']$ and $(N_3,M_3) = [V+1,U+1,W'+1]$. Then by

$$fM_1 - N_1 = O(z^{W+1}), \quad fM_2 - N_2 = O(z^{W'+1}) \quad \text{and} \quad fM_3 - N_3 = O(z^{W'+2}),$$

we have that $D(V,U,W,W')$ satisfies

$$E(V,U,W,W')D(V,U,W'W') + H(W+2,U+2,W'+2)M(V,U+1,W') = 0.$$

From Lemma 2.1, it follows that (2.19) and (2.20) hold, and

$$\begin{aligned} \operatorname{rank} D(V,U,W,W') &= \operatorname{rank}[H(W+2,U+2,W'+2)M(V,U+1,W')] \\ &= \operatorname{rank}\left\{\begin{bmatrix} H(W+2,U+2,W'+1) \\ H(W'+1,U+2,W'+2) \end{bmatrix} M(V,U+1,W')\right\} \\ &= \operatorname{rank}\begin{bmatrix} 0 \\ H(W+2,U+2,W'+1) \end{bmatrix} = p. \end{aligned}$$

(iii) Let $(N_1,M_1) = [V,U,W]$, $(N_2,M_2) = [V,U+1,W']$ and $(N_3,M_3) = [V-1,U+1,W'-1]$. Then by their definitions

$$fM_1 - N_1 = O(z^{W+1}), \quad fM_2 - N_2 = O(z^{W'+1}) \quad \text{and} \quad fM_3 - N_3 = O(z^{W'}),$$

and thus we have that $T(V,U,W,W')$ satisfies

$$N(V,U,W,V-1)T(V,U,W,W') + N(V,U+1,W',V-1)(I - T(V,U,W,W')) = 0,$$

$$E(V,U,W,W'-1)T(V,U,W,W') = 0.$$

From Lemma 2.1, it follows that (2.21) and (2.22) hold and that

$$\operatorname{rank} T(V,U,W,W') = \operatorname{rank} N(V,U+1,W',V-1) = p.$$

(iv) Let $(N_1,M_1)=[V,U,W]$, $(N_2,M_2)=[V-1,U,W-1]$ and $(N_3,M_3)=[V-1,U+1,W'-1]$. Then by

$$fM_1 - N_1 = O(z^{W+1}), \quad fM_2 - N_2 = O(z^W) \quad \text{and} \quad fM_3 - N_3 = O(z^{W'}),$$

we have that $S(V,U,W,W')$ satisfies

$$N(V,U,W,V-1) + N(V-1,U,W-1,V-2)S(V,U,W,W') = 0,$$

and

$$E(V,U,W,W'-1) + E(V-1,U,W-1,W'-2)S(V,U,W,W') = 0.$$

From Lemma 2.1, it follows that (2.23) and (2.24) hold and that $S(V,U,W,W')$ is nonsingular. The proof of the theorem is now completed. □

Formulas (2.17),(2.19), (2.21) and (2.23) can be written in matrix form. These matrices are the coefficient matrices of numerator and denominator. For example, (2.17) can be written as

$$N(V+1,U+1,W'+1,-1) = \downarrow N(V,U,W,-1)Q(V,U,W,W') + N(V+1,U,W+1,-1),$$

$$M(V+1,U+1,W'+1) = \downarrow M(V,U,W)Q(V,U,W,W') + M(V+1,U,W+1),$$

where

$$\downarrow N(V,U,W,-1) = \begin{bmatrix} \theta_m \\ N_1^{(0)}(V,U,W) \\ \vdots \\ N_1^{(v_1)}(V,U,W) \\ \vdots \\ \theta_m \\ N_p^{(0)}(V,U,W) \\ \vdots \\ N_p^{(v_p)}(V,U,W) \end{bmatrix} \quad \text{and} \quad \downarrow M(V,U,W) = \begin{bmatrix} \theta_m \\ M_1^{(0)}(V,U,W) \\ \vdots \\ M_1^{(u_1)}(V,U,W) \\ \vdots \\ \theta_m \\ M_m^{(0)}(V,U,W) \\ \vdots \\ M_m^{(u_m)}(V,U,W) \end{bmatrix}.$$

For the coefficient matrix of the residual, there is a similar formula. Therefore, formulas (2.17)–(2.24) can be used to compute the MPAs recursively, provided two approximants $[V,U,W]$ and $[V+1,U,W+1]$ or $[V,U,W]$ and $[V,U+1,W']$ are known. If all the numerator degrees are the same, i.e., $U \in \mathbf{Z}_+$, we can start from $[V,0,V] = (f^{(V)}(z),\ I_m)$ and $[V+1,0,V+1] = (f^{(V+1)}(z),\ I_m)$.

From Theorem 2.4 we know that, if $p=m$ all the parameters Q,D,T,S are nonsingular.

Now we want to establish some recursive relations among the parameters Q,D,T,S.

Theorem 2.5 *Suppose $m \geq p$ and f is I–normal. Then, for given V,U,W satisfying (1.2), and W' satisfying*

$$W' \geq W+1, \quad |W'-W| = m,$$

we have for the matrices appearing in the previous theorem, the following relations :

(i)

$$D(V,U,W,W') = Q(V,U,W,W') - S(V+1,U,W+1,W'+1), \tag{2.25}$$

(ii)

$$T(V+1,U,W+1,W'+1) = Q^{-1}(V,U,W,W')D(V,U,W,W'), \tag{2.26}$$

(iii) For $W'' \le W$ *and* $|W+1-W''| = m$,

$$S(V+2,U,W+2,W'+2) = Q(V,U,W,W') - D(V+1,U-1,W'',W+1), \tag{2.27}$$

(iv) For $W'' \le W$ *and* $|W+1-W''| = m$,

$$D(V+1,U-1,W'',W+1) = T(V+1,U-1,W'',W+1)Q(V,U,W,W'). \tag{2.28}$$

Proof. (i) In formula (2.23), changing the parameter (V,U,W,W') to $(V+1,U,W+1,W'+1)$ and subtracting with formula (2.17), we get

$$\begin{aligned}[V+1,U+1,W'+1] - [V,U+1,W'] \quad &= z[V,U,W]\,[Q(V,U,W,W') \\ &\quad -S(V+1,U,W+1,W'+1)]\,.\end{aligned}$$

It follows from (2.19) that the right hand side of the above equality is just $z[V,U,W]D(V,U,W,W')$. Therefore conclusion (i) holds.

(ii) Eliminating $z[V,U,W]$ from (2.17) and (2.19), we get

$$\begin{aligned}[V,U+1,W'] \quad &= [V+1,U,W+1]Q^{-1}(V,U,W,W')D(V,U,W,W') \\ &\quad +[V+1,U+1,W'+1]\,[I - Q^{-1}(V,U,W,W')D(V,U,W,W')]\,.\end{aligned}$$

Comparing this with formula (2.21), we get (ii).

(iii) Since

$$\sum_{i=1}^{p}(v_i+1) + \sum_{i=1}^{m}(u_i-1) = \sum_{i=1}^{p} w_i + p - m = \sum_{i=1}^{p} w_i'',$$

(1.2) holds for $(V+1,U-1,W'')$. In formula (2.21), take the parameter (V,U,W,W') to be $(V+1,U-1,W'',W+1)$, then the conditions on conclusion (iii) of the theorem hold. (2.21) now becomes

$$\begin{aligned}[V,U,W] &= [V+1,U-1,W'']T(V+1,U-1,W'',W+1) \\ &\quad +[V+1,U,W+1](I - T(V+1,U-1,W'',W+1)).\end{aligned}$$

Substituting it into (2.17), we have

$$\begin{aligned}&[V+1,U+1,W'+1] \\ &\quad = z[V+1,U-1,W'']T(V+1,U-1,W'',W+1)Q(V,U,W,W') \\ &\quad +z[V+1,U,W+1]\,[I - T(V+1,U-1,W'',W+1)]\,Q(V,U,W,W') \\ &\quad +[V+1,U,W+1].\end{aligned} \tag{2.29}$$

In Formula (2.23), changing the parameter (V,U,W,W') to $(V+2,U,W+2,W'+2)$, we have

$$\begin{aligned}&[V+1,U+1,W'+1] = [V+2,U,W+2] \\ &\quad +z[V+1,U,W+1]S(V+2,U,W+2,W'+2).\end{aligned} \tag{2.30}$$

Comparing the coefficient of z of the denominator on the right hand side of (2.29) and (2.30), we get

$$Q(V,U,W,W') + M^{(1)}(V+1,U,W+1) = M^{(1)}(V+2,U,W+2)$$
$$+S(V+2,U,W+2,W'+2), \tag{2.31}$$

where $M^{(1)}(V,U,W)$ denotes the coefficient matrix of z in the denominator. In formula (2.19), take the parameter (V,U,W,W') as $(V+1,U-1,W'',W+1)$, then one has

$$[V+2,U,W+2] = z[V+1,U-1,W'']D(V+1,U-1,W'',W+1) + [V+1,U,W+1].$$

It follows that

$$M^{(1)}(V+2,U,W+2) = D(V+1,U-1,W'',W+1) + M^{(1)}(V+1,U,W+1).$$

Combining it with (2.31), we get (2.27).

(iv) Comparing the leading coefficient of the denominator in (2.29) and (2.30), we have

$$M^{hc}(V+1,U,W+1)\,[I - T(V+1,U-1,W'',W+1)]\,Q(V,U,W,W')$$
$$= M^{hc}(V+1,U,W+1)S(V+2,U,W+2,W'+2).$$

Since $M^{hc}(V+1,U,W+1)$ is nonsingular,

$$S(V+2,U,W+2,W'+2) = [I - T(V+1,U-1,W'',W+1)]\,Q(V,U,W,W').$$

Using (2.27), we get (2.28). The proof is completed. □

If $p = m$ and $V,U,W,W',W'' \in \mathbf{Z}_+$, i.e., $V = v,\ U = u,\ W = w,\ W' = w',\ W'' = w''$, then $w = v + u, \quad w' = w + 1, \quad w'' = w$. Let

$$U_u^v = S(V+1,U-1,W,W+1), \qquad T_u^v = T(V+1,U-1,W,W+1),$$

$$E_u^v = -D(V-1,U-1,W-2,W-1), \qquad Q_u^v = -Q(V-1,U-1,W-2,W-1).$$

Then we have the following relations for these matrices

Corollary 2.6 *Suppose f is normal, then with the notation just introduced, we get the following qd-like relations :*

$$E_u^v = Q_u^v + U_u^{v-1}, \qquad v \geq 1, u \geq 1, \tag{2.32}$$

$$U_u^{v-1} = E_{u-1}^v - Q_u^{v-1}, \quad v \geq 1, u \geq 1, \tag{2.33}$$

$$Q_{u+1}^v = (T_u^{v-1})^{-1} E_u^{v+1}, \quad v \geq 1, u \geq 1, \tag{2.34}$$

$$T_u^{v-1} = (Q_u^v)^{-1} E_u^v, \qquad v \geq 1, u \geq 1, \tag{2.35}$$

where

$$E_0^v = 0, \qquad v = 1,2,\ldots, \tag{2.36}$$

$$Q_1^v = c_v^{-1} c_{v+1}, \quad v = 0,1,\ldots, \tag{2.37}$$

$$Q_u^0 = 0, \qquad u = 2,3,\ldots. \tag{2.38}$$

Proof. All the formulas in (2.32)–(2.35) are special cases of Theorem 2.5 except for $v = 1$ or $u = 1$ in (2.33), because the parameters Q_u^0 and E_0^v are not defined in Theorem 2.5. Therefore, the only thing to be verified is that the initial parameters given in (2.36)–(2.38) are correct indeed. From the definition of Q_1^v, we have

$$\begin{aligned} Q_1^v &= -Q(v-1,0,v-1,v) \\ &= E(v-1,0,v-1,v)^{-1}E(v,0,v,v+1) = c_v^{-1}c_{v+1}. \end{aligned}$$

i.e., (2.37) holds.

If $u = 1$ and $v \geq 1$, we get from the definition of U_1^{v-1},

$$\begin{aligned} U_1^{v-1} &= S(v,0,v,v+1) \\ &= -N(v-1,0,v-1,v-2)^{-1}N(v,0,v,v-1) = -c_{v-1}^{-1}c_v. \end{aligned}$$

i.e., $U_1^{v-1} = -Q_1^{v-1}$. Therefore, (2.36) is correct.

If $v = 1, u \geq 2$, from (2.19) and (2.23) we have

$$[1,u-1,u] = -z[0,u-2,u-2]E_{u-1}^1 + [0,u-1,u-1],$$

$$[0,u,u] = [1,u-1,u] + z[0,u-1,u-1]U_u^0.$$

Hence

$$[0,u,u] = -z[0,u-2,u-2]E_{u-1}^1 + [0,u-1,u-1] + z[0,u-1,u-1]U_u^0.$$

Equating the coefficient of z in the denomimator, we have

$$U_u^0 = E_{u-1}^1 + [M^{(1)}(0,u,u) - M^{(1)}(0,u-1,u-1)].$$

Since $M^{(1)}(0,u,u) - M^{(1)}(0,u-1,u-1) = 0$, (2.38) is correct. □

Using the formulas given in the corollary, all the parameters can be evaluated recursively from the initial parameters.

From the previous corollary, we immediately get the following one

Corollary 2.7 *Suppose f is normal, then*

$$E_u^v = Q_u^v + E_{u-1}^v - Q_u^{v-1}, \quad v \geq 1, u \geq 1,$$

$$Q_{u+1}^v = (E_u^v)^{-1}Q_u^v, E_u^{v+1}, \quad v \geq 1, u \geq 1,$$

where

$$E_0^v = 0, \qquad v = 1,2,\ldots,$$

$$Q_1^v = c_v^{-1}c_{v+1}, \quad v = 0,1,\ldots,$$

$$Q_u^0 = 0, \qquad u = 2,3,\ldots.$$

This is the analogue of the qd algorithm for the classical Padé approximants.

References

[1] G.L. Xu, A. Bultheel, Matrix Padé Approximants : Definitions and Properties, submitted. See also : G.L. Xu, A. Bultheel, The Problem of Matrix Padé Approximation, Report TW116, Dept. Computer Science, K.U. Leuven, November 1988, revised March 1989.