Journal of Computational and Applied Mathematics 32 (1990) 27–37
North-Holland

Minimal vector Padé approximation

Adhemar BULTHEEL and Marc VAN BAREL
Department of Computer Science, Katholieke Universiteit Leuven, Celestijnenlaan 200A, B-3030 Heverlee, Belgium

Received 24 September 1989
Revised 6 November 1989

Abstract: We describe the minimal vector Padé approximation problem, which consists in finding Padé approximants with a common denominator for a number of series. These approximants are minimal in the sense that for a given order of approximation and a given discrepancy in numerator and denominator degrees, the degree of the rational approximant is minimal. Properties and solution methods are derived from an associated minimal partial realization problem.

Keywords: Minimal Padé, vector Padé, simultaneous approximation, minimal partial realization.

1. Motivation

Suppose you have a scalar normal Padé table (without blocks) and you want for a given order of approximation τ the "simplest" Padé approximant (PA) from the table. The answer depends of course on the criterion you choose to measure simplicity. Suppose you mean by it the one with the least degree (maximum of numerator and denominator degree). Then you get a PA which is on or near the main diagonal of the table. It is of course also possible to get other entries of the Padé table by allowing a discrepancy in numerator and denominator degree. We can, e.g., minimize the maximum of the numerator degree minus some shift s and the denominator degree. We shall call this the s-degree. For $s = 0$ we get solutions near the main diagonal. For s positive (negative) you select some diagonal below (above) the main diagonal of the table. We now get a problem that is parametrized in r and s, rather than in α and β (the numerator and denominator degrees). The (r, s)-net places a grid over the Padé table along the diagonals while the (α, β)-net was a row–column grid. The problem is somewhat generalized if we do not require the approximants to be PAs, but only rational approximants of order at least r, for which the s-degree is minimized. Not every solution of this problem is a PA, but among all possible solutions, there is always a Padé form which can be uniquely selected by imposing extra conditions on the order of approximation, on the denominator degree or on the degree of the numerator. We call the problem as described above (parametrized in r and s) a *minimal Padé approximation problem* (mPA problem). In the scalar case, this does not add much new to the theory of Padé approximation. With some extra conditions, all the solutions of the mPA problem are also Padé approximants and conversely. In the (r, s)-table of (the uniquely defined) minimal Padé approximants, one can find essentially the same block structure as in the ordinary Padé

table (see [4]). Other properties of PAs are transportable to mPAs as well. Our interest in the mPA problem stems from its usefulness to study generalizations of the (m)PA problem to the matrix case. Rational approximations to matrix sequences is a topic which has been studied extensively as the minimal partial realization (MPR) problem in linear system theory. This problem can be translated quite naturally into what we called the mPA problem, so that a number of results that are well known in system theory can be used to get corresponding results in the theory of (minimal) matrix Padé approximation (mMPA). Some results in this sense were obtained in [3] and in [14] a matrix two-point Padé problem was considered. In this paper we want to give a slightly different approach to the problem of minimal vector Padé approximation (mVPA).

In this paper we address the general problem of specifying uniquely a minimal vector Padé approximant. The algebraic theory of simultaneous Padé approximation (which is virtually identical to one type of vector Padé approximation) has been investigated by Mahler [12], Coates [5], Jager [10], de Bruin [6] and Nikisin [13]. In particular, Mahler noted the connection between the problems of simultaneous Padé approximation and Hermite–Padé approximation. The connection between Kalman's partial realization problem for SIMO (single-input-multi-output) systems and vector Padé approximation was noted by Graves-Morris [8]; some algorithms applicable to nondegenerate cases are given by Graves-Morris and Wilkins [9]. A more robust approach was adopted by Bultheel and Van Barel [3] in a more general matrix setting. These results were based upon earlier results of Dickinson et al. [7] and Anderson et al. [1] who described implicitly an algorithm for the general minimal partial realization problem for MIMO (multi-input-multi-output) systems. These ideas were clarified and an explicit algorithm, which generalizes the Euclidean algorithm to the case of matrix polynomials, was given in [15]. Similar results are also given by Antoulas in [2]. In [15] it is also shown how a unique (canonical) solution may be obtained by the algorithm.

A quite different approach to vector Padé approximation was given by Wynn [17] and developed by Graves-Morris and Jenkins [8]. The approximants they consider are almost unrelated to the approximants we shall consider here. The definition given by van Iseghem [16], however, is much more related. It turns out that her approximants are a special case of the ones defined here. The generalization being that we consider a different shift for each of the components separately, which means that we consider vectors with polynomial entries whereas considering polynomials with vector coefficients results naturally in an (almost) uniform shift for each of the components. To put it in yet another way: if, like in [16], we have a uniform degree for the numerator entries, then to match the degrees of freedom with the number of interpolation conditions, we can distribute these conditions equally over all components with possibly some extra conditions on the initial components. In our approach however, we shall be free to distribute the degrees of freedom (numerator degrees) and the interpolation conditions independently over whatever components we like.

2. Definitions and notations

Let $\boldsymbol{F}(z) = \sum_0^\infty \boldsymbol{F}_k z^k \in K^m[[z]]$ be a given formal power series. It has coefficients $\boldsymbol{F}_k \in K^m$. These are $1 \times m$ row vectors with entries in a field K. The ith component of $\boldsymbol{F}_k$ is denoted as $F_{k,i}$.

We shall need a special notation to denote vectors of polynomials with a different degree for each of their components. Therefore we introduce the following constructs. Let $\boldsymbol{k} = [k_1, \ldots, k_m] \in \mathbb{N}^m$. By $z^{\boldsymbol{k}}$ we mean $\operatorname{diag}(z^{k_1}, \ldots, z^{k_m})$. $\boldsymbol{F}_{\boldsymbol{k}}$ means the m-vector whose ith component is $F_{k_i,i}$.

For any $\boldsymbol{t} \in \mathbb{N}^m$, $\sum_{\boldsymbol{k}=\boldsymbol{0}}^{\boldsymbol{t}} \boldsymbol{F}_{\boldsymbol{k}} z^{\boldsymbol{k}}$ denotes the polynomial vector with ith component $\sum_{k_i=0}^{t_i} F_{k_i,i} z^{k_i}$. In general, a summation for the m-vector $\boldsymbol{k}$ ranging from some lower bound $\boldsymbol{l}$ to an upper bound $\boldsymbol{u}$ like in $\sum_{\boldsymbol{k}=\boldsymbol{l}}^{\boldsymbol{u}} \boldsymbol{v}_{\boldsymbol{k}}$ with $\boldsymbol{v}_{\boldsymbol{k}}$ some m-vectors with components $v_{k,i}$, $i = 1, \ldots, m$, means a componentwise summation so that the ith component of the result is $\sum_{k_i=l_i}^{u_i} v_{k_i,i}$. Inequalities between m-vectors like $\boldsymbol{k} \leqslant \boldsymbol{t}$ also mean componentwise inequalities: $k_i \leqslant t_i$ for all $i = 1, \ldots, m$.

Addition of a scalar and a vector means the addition of the scalar to each component of the vector. Also $\leqslant$ between a vector and a scalar means $\leqslant$ for each component.

A polynomial $\boldsymbol{P}(z) \in K^m[z]$, is said to have degree $\boldsymbol{s} = [s_1, \ldots, s_m] \in \overline{\mathbb{N}}^m$ if its ith entry $P_i(z)$ has degree s_i. We have set $\overline{\mathbb{N}} = \mathbb{N} \cup \{-\infty\}$ and define $\deg 0 = -\infty$. $\deg \boldsymbol{P}(z) = \boldsymbol{s}$ means that $\boldsymbol{s}$ is the minimal index vector for which $\boldsymbol{P}(z) = \sum_{\boldsymbol{k}=\boldsymbol{0}}^{\boldsymbol{s}} \boldsymbol{P}_{\boldsymbol{k}} z^{\boldsymbol{k}}$.

Next we introduce the idea of the $\boldsymbol{s}$-degree of a rational form which will play an important role in this paper. For $\boldsymbol{s} \in \mathbb{Z}^m$ we say that a rational form $Q(z)^{-1}\boldsymbol{P}(z)$ with $\boldsymbol{P}(z) \in K^m[z]$ and $Q(z) \in K[z]$ has $\boldsymbol{s}$-degree $v = \max\{\deg Q(z), \deg P_i(z) - s_i,\ i = 1, \ldots, m\}$.

We say that such a rational form is an approximant of order (at least) $\boldsymbol{r} \in \mathbb{N}^m$ for $\boldsymbol{F}(z)$ if

$$\boldsymbol{F}(z) - Q(z)^{-1}\boldsymbol{P}(z) = \tilde{\boldsymbol{F}}_{r+1} z^{r+1} + \tilde{\boldsymbol{F}}_{r+2} z^{r+2} + \cdots . \tag{2.1}$$

The minimal vector Padé approximation (mVPA) problem can now be formulated as follows: given $\boldsymbol{F}(z) \in K^m[[z]]$, $\boldsymbol{r} \in \mathbb{N}^m$ and $\boldsymbol{s} \in \mathbb{Z}^m$, find some $\boldsymbol{P}(z) \in K^m[z]$ and $Q(z) \in K[z]$ such that $Q(z)^{-1}\boldsymbol{P}(z)$ has minimal $\boldsymbol{s}$-degree among the approximants for $\boldsymbol{F}(z)$ of order at least $\boldsymbol{r}$.

Note that $Q(0)$ has to be nonzero because $Q(0) = 0$ implies $\boldsymbol{P}(0) = \boldsymbol{0}$ and then we do not have a minimal approximant since it can be reduced. In general $\boldsymbol{P}(z)$ and $Q(z)$ should be coprime (i.e., the m components $P_i(z)$ of $\boldsymbol{P}(z)$ and $Q(z)$ should have no common factors) for the same reason. If $Q(z) \neq 0$, then $Q(z)^{-1}\boldsymbol{P}(z)$ is an approximant of order at least $\boldsymbol{r}$ if and only if

$$Q(z)\boldsymbol{F}(z) - \boldsymbol{P}(z) = \boldsymbol{R}_{r+1} z^{r+1} + \boldsymbol{R}_{r+2} z^{r+2} + \cdots .$$

It is clear that $(\boldsymbol{P}(z), Q(z)) = (\sum_{\boldsymbol{k}=\boldsymbol{0}}^{\boldsymbol{r}} \boldsymbol{F}_{\boldsymbol{k}} z^{\boldsymbol{k}}, 1)$ always satisfies the order condition and therefore the set of order $\boldsymbol{r}$ approximants is not empty. For given $\boldsymbol{s}$ it is then possible to find in this set an approximant of minimal $\boldsymbol{s}$-degree. Thus the mVPA problem always has a solution.

For the sake of completeness, we recall here what we call a vector Padé approximation (VPA) problem. The problem is that of finding some $\boldsymbol{P}(z) \in K^m[z]$ and $Q(z) \in K[z]$ such that for given $\boldsymbol{\alpha} \in \mathbb{N}^m$ and $\beta \in \mathbb{N}$ and $\boldsymbol{F}(z) \in K^m[[z]]$, the rational form $Q(z)^{-1}\boldsymbol{P}(z)$ which approximates $\boldsymbol{F}(z)$ up to order $\boldsymbol{r}$ with $|\boldsymbol{r} + 1| = \Sigma(r_i + 1) \geqslant \Sigma(\alpha_i + 1) + \beta = |\boldsymbol{\alpha} + 1| + \beta$ while $\deg \boldsymbol{P}(z) \leqslant \boldsymbol{\alpha}$ and $\deg Q(z) \leqslant \beta$.

The readers who are not familiar with the MPR problem, should consult, e.g., [11,1,2,7,9,15]. To give a feeling for the work to be done in translating a MPR result into a mVPA result, we give a brief account of this. The MPR problem is defined as follows. Let $\boldsymbol{M}(z) = \sum_1^\infty \boldsymbol{M}_k z^{-1} \in K^m[[z^{-1}]]$ be a given formal power series in z^{-1}. Let also $\boldsymbol{r} \in \mathbb{N}^m$ be given. The problem is to find $\boldsymbol{C}(z) \in K^m[z]$ and $A(z) \in K[z]$ such that

$$\boldsymbol{M}(z) - A(z)^{-1}\boldsymbol{C}(z) = \boldsymbol{M}^*_{r+1} z^{-r-1} + \boldsymbol{M}^*_{r+2} z^{-r-2} + \cdots .$$

Since $\deg C_i(z)$ is less than $\deg A(z)$, minimizing the (-1)-degree of $A(z)^{-1}\boldsymbol{C}(z)$ is the same as minimizing $\deg A(z)$. Note that again, because of minimality, $A(z)$ and $\boldsymbol{C}(z)$ are coprime. In

particular $A(0) \neq 0$ or $\boldsymbol{C}(z) \neq \boldsymbol{0}$. Now, we replace z by z^{-1} to transform this into a mVPA problem. As we can easily check, the problem turns into a mVPA problem for $\boldsymbol{F}(z) = z^{-1}\boldsymbol{M}(z^{-1})$ with a solution $Q(z) = z^{\nu}A(z^{-1})$ and $\boldsymbol{P}(z) = z^{\nu-1}\boldsymbol{C}(z^{-1})$, where ν is the degree of $A(z)$. Note however that ν need not be equal to deg $Q(z)$, but it is equal to the (-1)-degree of the rational form $Q(z)^{-1}\boldsymbol{P}(z)$. As we mentioned in the introduction, the shift $\boldsymbol{s}$ is introduced to get other "nondiagonal" mVPAs. More details can be found in the proof of Theorem 3.1. This construction is slightly different from the one used in [3] as will be explained in Section 3.

3. Relation with minimal partial realization

Suppose that $Q(z)^{-1}\boldsymbol{P}(z)$ is an mVPA of $\boldsymbol{F}(z)$ for some given $\boldsymbol{s}$ and $\boldsymbol{r}$ and let its minimal $\boldsymbol{s}$-degree be given by d. Then, for the ith component, we may write ($\boldsymbol{F}_k = \boldsymbol{0}$ for $k < 0$)

$$\sum_{k=0}^{d} Q_k F_{j-k,i} = \begin{cases} P_{j,i}, & j = 0, \ldots, d + s_i, \\ 0, & j = d + s_i + 1, \ldots, r_i, \end{cases} \tag{3.1}$$

where

$$P_i(z) = \sum_{j=0}^{d+s_i} P_{j,i} z^j, \qquad Q(z) = \sum_{k=0}^{d} Q_k z^k, \qquad F_i(z) = \sum_{j=0}^{\infty} F_{j,i} z^j.$$

If $d + s_i > r_i$, (3.1) represents only a possible choice for the numerator $P_i(z)$. It does not contain conditions for the denominator. For simplicity, we restrict ourselves for the time being to the case where $\boldsymbol{r} = \boldsymbol{s} + n$ with $n \in \mathbb{N}$. This means that the numerator $\boldsymbol{P}(z)$ can be computed from $Q(z)$ and $\boldsymbol{F}(z)$, while the denominator has to satisfy the homogeneous block Hankel system

$$[Q_d \ \cdots \ Q_0] \begin{bmatrix} \boldsymbol{M}_1 & \boldsymbol{M}_2 & \cdots & \boldsymbol{M}_{n-d} \\ \boldsymbol{M}_2 & & & \\ \vdots & & & \vdots \\ \boldsymbol{M}_{d+1} & \cdots & \cdots & \boldsymbol{M}_n \end{bmatrix} = [0 \ 0 \ \cdots \ 0], \quad Q_0 \neq 0, \ d \text{ minimal}, \tag{3.2}$$

where

$$\boldsymbol{M}_k = \boldsymbol{F}_{s+k}. \tag{3.3}$$

Thus we have $\boldsymbol{M}_k = [M_{k,i}]_{i=1}^{m} \in K^m$ and $M_{k,i} = F_{s_i+k,i}$.

At this point, the mVPA problem is reduced to something which is formally the same as an nth-order minimal partial realization (MPR) problem for the vector sequence $\boldsymbol{M}_k$, $k = 1, 2, \ldots$. This problem has been intensively studied in the system literature for the more general $p \times m$ matrix case; references [1,2,7] are representative. Recently, a canonical solution for this problem was proposed which had a number of interesting properties [15]. In the analysis of the MPR problem, the Kronecker and dual Kronecker indices play an important role. For the vector case, there is only one dual Kronecker index $\nu(n)$ associated with the sequence $\boldsymbol{M}_1, \ldots, \boldsymbol{M}_n$. It is the smallest d for which the system (3.2) has a solution with $Q_0 \neq 0$. By definition of the mVPA problem, it is the minimal $\boldsymbol{s}$-degree of the mVPA of order $\boldsymbol{s} + n$ for the series $\boldsymbol{F}(z)$ whose

coefficients are related with the M_k sequence by (3.3). The Kronecker indices $\kappa(n) = [\kappa_1(n) \cdots \kappa_m(n)] \in \mathbb{N}^m$ for the same sequence are defined as follows: $\kappa_i(n)$ is the smallest d such that

$$\begin{bmatrix} M_1 & M_2 & \cdots & M_{d+1} \\ M_2 & & & \\ \vdots & & & \vdots \\ M_{n-d} & \cdots & \cdots & M_n \end{bmatrix} \begin{bmatrix} V_d \\ \vdots \\ V_0 \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

has a solution with $V_0^{\mathrm{T}} = [V_{0,1} \cdots V_{0,i-1}\ 1\ 0 \cdots 0] \in K^m$. Since it will be clear from the context whether we mean Kronecker or dual Kronecker indices, we shall call $\nu(n)$ as well as $\kappa(n)$ Kronecker indices for short. It is a well-known property of Kronecker indices that $|\kappa(n)| = \sum_i \kappa_i(n) = \nu(n)$ (see, e.g., [2]).

Now we are ready to formulate the following theorem.

Theorem 3.1. *Let $F(z) = \sum_0^\infty F_k z^k \in K^m[[z]]$. For a given integer $n \geqslant 0$, $-n \leqslant s \in \mathbb{Z}^m$ and $r = s + n \geqslant 0$ (componentwise inequalities), let $\nu = \nu(n)$ and $\kappa = \kappa(n)$ be the Kronecker indices of the sequence $M_1, \ldots, M_n$ with $M_k = F_{s+k}$ ($F_i = 0$ for $i < 0$). Then there exists a unique mVPA $Q(z)^{-1}P(z)$ with $Q(0) = 1$ of order r and minimal s-degree ν which satisfies*

$$F(z) - Q(z)^{-1}P(z) = R_{\kappa+\nu+1} z^{\kappa+s+\nu+1} + \cdots.$$

This mVPA is also a VPA of type $[\alpha, \beta]$ with $\alpha = s + \nu$ and $\beta = \nu$.

Proof. The present proof supplements that of a corresponding theorem in [15]. It clearly shows the mechanism which we mentioned at the end of the previous section relating MPR problems and mVPA problems for a general shift s.

The results of [15] are for a right MPR problem, but the results can be easily translated for a left MPR problem as we have here. Computing a left MPR of order n for the sequence $M_1, M_2, \ldots$ is essentially the same as computing a right MPR for the sequence $M_1^{\mathrm{T}}, M_2^{\mathrm{T}}, \ldots$. Applied to the present situation, the result then reads: define for $M_k = F_{s+k}$ the formal series

$$M(z) = M_1 z^{-1} + \cdots + M_n z^{-n} + M_{n+1} z^{-n-1} + \cdots.$$

Then there exists a unique (canonical) MPR $A(z)^{-1}C(z)$ of the sequence $M_1, \ldots, M_n$ which satisfies

$$M(z) - A(z)^{-1}C(z) = \tilde{M}_{\kappa+\nu+1} z^{-(\kappa+\nu+1)} + \cdots,$$

with $A(z)$ monic of degree ν and $C(z) \in K^m[z]$. Multiply with z^s from the right, after z is replaced by z^{-1}. Then you get

$$M(1/z)z^s - A(1/z)^{-1}C(1/z)z^s = \left(\tilde{M}_{\kappa+\nu+1} z^{\kappa+\nu+1} + \cdots\right) z^s,$$

which implies

$$F(z) - Q(z)^{-1}P(z) = \tilde{M}_{\kappa+\nu+1} z^{\kappa+s+\nu+1} + \cdots,$$

with

$$\begin{aligned} F(z) &= \sum_{k=0}^{s} F_k z^k + M(1/z) z^s, \\ Q(z) &= A(1/z) z^\nu, \quad Q(0) = 1, \\ P(z) &= Q(z) \left(\sum_{k=0}^{s} F_k z^k \right) + C(1/z) z^{s+\nu}. \end{aligned}$$

The ith component $P_i(z)$ of $P(z)$ is a polynomial of degree at most $s_i + \nu$. This is surely true if $s_i \geqslant 0$. If $s_i < 0$, then, because we agreed that $F_{k,i} = 0$ for $k < 0$, we also get $M_{k,i} = 0$ for $k < -s_i$. Hence, since $C_i(z)$ is given as the polynomial part of $A(z) M_i(z)$, it follows that $C_i(z)$ is of degree at most $\nu + s_i$. Thus $P_i(z)$ will be given by $C_i(1/z) z^{s_i+\nu}$ which is a polynomial of degree at most $s_i + \nu$.

To show that the mVPA is also a VPA, we note that the number of degrees of freedom in the pair $(P(z), Q(z))$, taking $Q(0) = 1$ into account, is equal to

$$\sum_{i=1}^{m} (s_i + \nu + 1) + \nu = |s| + m(\nu + 1) + \nu$$

and the number of coefficients fitted is

$$\sum_{i=1}^{m} (\nu + \kappa_i + s_i + 1) = |s| + m(\nu + 1) + |\kappa| = |s| + m(\nu + 1) + \nu.$$

With $\alpha = s + \nu$, $\beta = \nu$ and the order $r \geqslant \kappa + s + \nu$, we have shown that $|r + 1| \geqslant |\alpha + 1| + \beta$, so that the canonical mVPA solution is also a VPA. □

It is also possible to construct a matrix continued fraction whose convergents are the successive mVPAs for successive values of n.

Corollary. *Under the conditions of the previous theorem, apply the algorithm CMPR described in* [15] *to the data* $M_1^{\mathrm{T}}, \ldots, M_n^{\mathrm{T}}$; *then it will construct unimodular matrices*

$$V_k(z) = \begin{bmatrix} Y_k(z) & C_k(z) \\ X_k(z) & A_k(z) \end{bmatrix} \in K^{(m+1)\times(m+1)}[z]$$

such that the matrix continued fraction

$$\left\{ \left. \frac{\dfrac{\dfrac{\cdots}{\cdots} Y_2^{\mathrm{T}} + C_2^{\mathrm{T}}}{\dfrac{\cdots}{\cdots} X_2^{\mathrm{T}} + A_2} Y_1^{\mathrm{T}} + C_1^{\mathrm{T}}}{\dfrac{\dfrac{\cdots}{\cdots} Y_2^{\mathrm{T}} + C_2^{\mathrm{T}}}{\dfrac{\cdots}{\cdots} X_2^{\mathrm{T}} + A_2} X_1^{\mathrm{T}} + A_1} \right|_{z \to 1/z} \right\} z^s + \sum_{k=0}^{s} F_k z^k$$

has as successive convergents the uniquely defined mVPAs of the previous theorem for $n = 0, 1, 2, \ldots$.

Proof. Using the correspondence used in the previous theorem, this result is easily obtained from [15]. □

4. Another approach

In the previous section, we have reduced the problem of mVPA for given s and $r = s + n$, to a problem of left MPR computation for the vector sequence $F_{s+1}, \ldots, F_{s+n}$, which is the same as computing a right MPR for the transposed sequence. The algorithm of [15] then gives a recursive method to compute the solutions for $n = 0, 1, \ldots$. This means that during these computations s is kept constant while r increases in each step by 1 in all its components. In the scalar case, this corresponds to a recursion along a downward sloping diagonal in the Padé table. The algorithm CMPR of [15] performs on (formal) series some operations which are basically comparable to those performed in the Euclidean algorithm [3]. In the scalar case, it is also possible to use formally the same algorithm to move along an antidiagonal in the Padé table. This path of an antidiagonal corresponds for the vector case, just as in the scalar case, to choosing r constant and letting $s = r - n$ vary for $n = 0, 1, \ldots$. The systems (3.1) can then be reduced to

$$[Q_0 \ \cdots \ Q_d] H(n, d) = 0, \quad Q_0 \neq 0, \ d \text{ minimal},$$

where $H(n, d)$ is a block Hankel of the same form as in (3.2), but now the M_k are defined by $M_k = F_{r-k+1}$ with the same short-hand notation as in (3.3). As it is formulated now, we do not have exactly an MPR problem. It were, if only we required $Q_d \neq 0$ instead of $Q_0 \neq 0$. This means that we have to give up the reducedness condition of the approximant. We can only be sure that the linearized interpolation conditions (2.1) will be satisfied. This corresponds to the fact that in the scalar case when moving on an antidiagonal in the Padé table, it is possible to hit a singular block below its main antidiagonal. Here the reduced Padé approximants do not have the appropriate order of contact and it is only possible to satisfy the linearized conditions. This was the approach followed in [3] in the more general case of (right) matrix Padé approximation. We shall not repeat it here for the vector case.

5. The nice problem

Instead of choosing $r = s + n$ with n integer, we shall now come to the more general situation where $r = s + t$, $t = [t_1, \ldots, t_m] \in \mathbb{N}^m$, where the t_i can now be different from each other. Also in this case, the reduction to the MPR problem can be done as before. Define again $M_k = F_{s+k}$ as in (3.3) and let n be $n = t_{\max} = \max\{t_i, i = 1, \ldots, m\}$. The system (3.2) based on the sequence $M_1, \ldots, M_n$, however, contains too many restrictions. Indeed the solution of that system gives orders of approximation $s + t_{\max} = s + n \geqslant s + t = r$. Some of the entries in $H(n, d)$ are specified and used in the determination of the solution whereas others are not specified. We refer to the latter as "don't-care-entries", which we shall indicate by "?". These unspecified elements are given in the matrix $M^{\mathrm{T}} = [M_1^{\mathrm{T}} \ \cdots \ M_n^{\mathrm{T}}]$ which looks like

$$M^{\mathrm{T}} = \begin{bmatrix} M_{11} & \cdots & M_{t_1,1} & ? & ? & ? & ? \\ M_{12} & \cdots & \cdots & \cdots & M_{t_2,2} & ? & ? \\ & & & \vdots & & & \\ M_{1i} & \cdots & \cdots & \cdots & \cdots & & M_{t_{\max},i} \\ & & & \vdots & & & \\ M_{1m} & \cdots & \cdots & M_{t_m,m} & ? & ? & ? \end{bmatrix},$$

where we supposed $n = t_{\max} = t_i$. Such a problem is the vector case of what is known in the theory of MPR as a nice problem. It is called a nice problem because some of the elements of M^T are left unspecified, and this is always done so that $M_{k,i} = ?$ implies $M_{k+1,i} = ?$ for $i = 1, \ldots, m$. It is possible to construct a recursive algorithm similar to the algorithm CMPR of [15] to solve such a nice problem, but due to space limitations, we shall not do it here.

For nice problems, the role of the Kronecker indices is naturally played by indices which are known as nice indices. In fact there are both nice and dual nice indices, but as in the case of Kronecker indices, we shall drop the adjective dual most of the time. Since we have here a row vector, there is only one (dual) nice index and this then coincides with the (dual) Kronecker index. These should now be defined for matrices that are not completely specified (i.e., one that contains ?-elements. See, e.g., [11]). Consider the block Hankel matrix $H(n) = [\boldsymbol{M}_{i+j-1},\ i,\ j = 1, \ldots, n]$ where $\boldsymbol{M}_n$ is the last of the $\boldsymbol{M}_k$ that is not completely unspecified. The entries $\boldsymbol{M}_j$ for $j > n$ contain only ?-elements. Then $\nu(n) + 1$ is the number of the first row in $H(n)$ that is linearly dependent on the previous ones where the ?-elements are supposed to take the values that keep the rank of $H(n)$ as low as possible. Note, however, that there may exist many choices for these elements that will give the minimal rank. Because of the Hankel structure and the nice specification scheme, the linear dependency of row $\nu(n)$ implies the linear dependency of all the following rows. A similar convention is used to define the nice indices (and also Kronecker indices) corresponding to the columns. We call the integers $\boldsymbol{\kappa}(n) = [\kappa_1(n), \ldots, \kappa_m(n)]$ nice indices for the sequence $\boldsymbol{M}_1, \ldots, \boldsymbol{M}_n$ if in the matrix $H(n)$ all the columns $km + i$, $k = 0, 1, \ldots, \kappa_i(n) - 1$, $i = 1, \ldots, m$, are linearly independent and the others are linearly dependent on these. Again here this dependency is defined for a choice of the ?-elements that keeps the rank of $H(n)$ as low as possible. One clearly has rank $H(n) = \nu(n) = |\boldsymbol{\kappa}(n)| = \Sigma\kappa_i(n)$. The choice of the nice indices $\boldsymbol{\kappa}(n)$ is not unique. One can take any set of $\nu(n)$ independent columns as long as they satisfy the nice property, i.e., column $km + i$ dependent implies $(k+1)m + i$ dependent. Even for a fixed choice of the ?-elements there may be different sets of nice indices $\boldsymbol{\kappa}(n)$. On the other hand, the Kronecker indices are uniquely defined for a fixed choice of the ?-elements since they choose the first $\nu(n)$ independent columns that can be found. We could qualify them as minimal nice indices. A different minimal extension may or may not give different Kronecker indices. It is clear that if we extend the sequence $\boldsymbol{M}_k$ for k to infinity with completely unspecified elements, then the definition of the nice indices will not change when they are obtained from matrices $H(n+i)$, $i \geqslant 0$. For this reason we could drop the argument n from the notation and we shall speak of the nice indices ν and $\boldsymbol{\kappa}$ of the (infinite) sequence $\boldsymbol{M}_k$ or of the Hankel matrix $H = H(\infty)$.

As with the Kronecker indices, it is possible to prove the following theorem.

Theorem 5.1. *Let the sequence $\boldsymbol{M}_k = \boldsymbol{F}_{s+k}$ for $\boldsymbol{s} \in \mathbb{Z}^m$ be nicely specified up to $\boldsymbol{M}_t$ and $n = \max\{t_i\}$. Let ν and $\boldsymbol{\kappa}$ be a choice for the nice indices of this sequence. Then there exists a unique mVPA $Q(z)^{-1}\boldsymbol{P}(z)$ of order $\boldsymbol{r} = \boldsymbol{s} + \boldsymbol{t}$ which has minimal s-degree ν and which satisfies*

$$\boldsymbol{F}(z) - Q(z)^{-1}\boldsymbol{P}(z) = \tilde{\boldsymbol{F}}_{\kappa+\nu+1} z^{\kappa+s+\nu+1} + \cdots,$$

where $\boldsymbol{F}(z) = \Sigma_0^\infty \boldsymbol{F}_k z^k$. The couple $(\boldsymbol{P}(z), Q(z))$ can then be made unique by choosing $Q(0) = 1$.

Proof. This result can be derived from the corresponding result in MPR literature. The proof is essentially the same as for the Kronecker indices. See, e.g., [2]. □

Notes. (1) For the Kronecker indices, it is guaranteed that the required order of approximation is met, but one may not conclude that for all i one has $\kappa_i + \nu \geqslant t_i$. This means that sometimes, to fix the unique solution, more of the coefficients $\boldsymbol{F}_k$ are used and sometimes less than actually needed to meet the order condition.

(2) The previous theorem says that for given $\boldsymbol{s}$ and $\boldsymbol{t}$, we can find a unique mVPA for each choice of the nice indices (possibly depending on the choice of the extension made). It does not say anything about these mVPAs being the same or not.

(3) It is possible to parametrize all possible mVPAs for given $\boldsymbol{s}$ and $\boldsymbol{t}$ with a minimum number of free parameters. We shall not do this here.

We shall now describe how the nice indices change if one of the unspecified elements becomes specified. The jumps in ν correspond to the jumps in the denominator degree of the scalar PAs when one jumps over a (possibly trivial 1×1) block while moving on a downward sloping diagonal in the Padé table.

To obtain this, we follow the approach of Kalman [11]. For the scalar case, Kalman uses the following simple but powerful lemma.

Lemma 5.2. *Consider the matrix* $D(?) = \left[\begin{smallmatrix} A & B \\ C & ? \end{smallmatrix}\right]$ *with one unspecified element* ?. (1) *If* $\operatorname{rank} A = \operatorname{rank}[A\ B] = \operatorname{rank}[A^{\mathrm{T}}\ C^{\mathrm{T}}]$, *then there exists exactly one value* V *of* ? *for which* $\operatorname{rank} A = \operatorname{rank} D(V)$. *For any other value of* ?, $\operatorname{rank} D(?) = \operatorname{rank} A + 1$.

(2) *If* $\operatorname{rank}[A\ B] > \operatorname{rank} A$ *or* $\operatorname{rank}[A^{\mathrm{T}}\ C^{\mathrm{T}}] > \operatorname{rank} A$, *then* $\operatorname{rank} D(?)$ *is independent of the value of* ?.

This lemma, together with properties of Hankel matrices led Kalman to the conclusion in the scalar case ($m = 1$ and hence $\nu = \kappa$) that by specifying one more element in the sequence $M_1, \ldots, M_n$, we can come into different possible situations. If $2\nu > n$, then situation (2) of Lemma 5.2 applies to the Hankel matrix $H = [M_{i+j-1}]$, with entries $M_1, \ldots, M_n, ?, \ldots$. Therefore the minimal degree ν shall not change. This is plausible since to fix the unique solution, we used $M_1, \ldots, M_{2\nu}$ (see Theorem 5.1 with $s = -1$) which is more than required for an approximant of order n. If $2\nu \leqslant n$, then M_{n+1} shall appear in H in at least one position where situation (1) of Lemma 5.2 applies. The minimal degree will not increase if M_{n+1} is the specific value alluded to in (1) of the lemma. In that case, the sequence is extended without increasing the rank of the Hankel matrix H for this sequence. We shall call this value the minimal extension value. This means that for a minimal extension value of M_{n+1} we have $\nu(n+1) = \nu(n)$. For any other value of M_{n+1} one gets $\nu(n+1) = n + 1 - \nu(n)$. Thus, in general, it holds that $\nu(n+1) = \max\{\nu(n),\ n + 1 - \nu(n)\}$. See, e.g., [11]. In terms of Padé approximants the choice of the minimal extension value corresponds to the fact that a Padé approximant based on the data $M_1, \ldots, M_n$ has an expansion which defines an extension $M_1, \ldots, M_n, \tilde{M}_{n+1}, \tilde{M}_{n+2}, \ldots$ of this sequence. This extension is chosen such that it keeps the minimal s-degree of the approximant. It is a minimal extension. If it happens that $M_{n+1} = \tilde{M}_{n+1}$, then the nth PA will be equal to the $(n+1)$th if M_{n+1} is given. This stays like that until some M_k turns out to be different from the minimal extension value. In that case the denominator degree increases as indicated.

The following theorem generalizes this property to the vector case.

Theorem 5.3. *Consider the block Hankel matrix* $H = [\boldsymbol{M}_{i+j-1}]$ *with* $\boldsymbol{M}_k \in \overline{K}^m$, $\overline{K} = K \cup \{?\}$, *a nicely specified sequence up to* $\boldsymbol{M}_t$. *Let* $(\nu, \boldsymbol{\kappa})$ *be a choice for the nice indices of* H. *Mark with a* $\times$, *all the rows and columns that are chosen independent according to this choice of the nice indices. That are the rows* $1, \ldots, \nu$ *and the columns* $km + i$, $k = 0, \ldots, \kappa_i - 1$, $i = 1, \ldots, m$. *The remaining rows and columns are all marked with a* $\bullet$. *Let* $u = t_i + 1$; *then* $M_{u,i}$ *is the first unspecified element in column* i. *This element appears at several positions in the matrix* H, *viz. the positions* $(u - j, jm + i)$, $j = 0, \ldots, u - 1$. *If we specify this element, then we have one of the following cases*:

(1) *None of the positions where this element appears in* H *is of type* $(\bullet, \bullet)$. *Then the nice indices can be left unchanged whatever the specification of* $M_{u,i}$ *is.*

(2) *This element appears on at least one position which is of type* $(\bullet, \bullet)$. *Then either it takes a uniquely defined minimal extension value, in which case the nice indices can be left unchanged, or it is different from this value, in which case* ν *and* κ_i *increase with an amount* $u - \nu - \kappa_i$. *This increase is precisely equal to the number of times the element* $M_{u,i}$ *appears on a* $(\bullet, \bullet)$*-position.*

Proof. The elements $M_{u,i}$ appears in precisely u positions of the matrix H. Of these, only $u - \kappa_i$ appear on $\bullet$-columns, by definition of nice index κ_i. These are in the first $u - \kappa_i$ rows. Since the first ν rows are $\times$-rows, there are at most $\max\{0, u - \nu - \kappa_i\}$ elements on $(\bullet, \bullet)$-positions. If the element is not on a $(\bullet, \bullet)$-position, then, according to the previous lemma, the rank is independent of the choice of $M_{u,i}$ and this means that ν does not change, and hence also $\boldsymbol{\kappa}$ can be left unchanged. If the element appears on a $(\bullet, \bullet)$-position, then part (1) of Lemma 5.2 applies. This means that either $M_{u,i}$ takes the minimal extension value which keeps the rank down, in which case the nice indices need not be changed, or $M_{u,i}$ is not this value, and then all the rows and columns corresponding to these $(\bullet, \bullet)$-positions have to be changed into $\times$-rows and -columns, which means that the nice indices change as indicated. One only needs to exploit the Hankel structure to verify that the uniquely defined extension value is the same for each of the $(\bullet, \bullet)$-positions. The theorem is then proved. □

6. Conclusion

In this short note we have formulated the minimal vector Padé approximation problem. We reduced this problem to a minimal partial realization problem for a single-input-multi-output system. Such MPRs for SIMO systems were also considered in [9]. We also considered a nice problem which corresponds to different numerator degrees and different orders of approximation per component. We generalized a property about the jumps in numerator degree of Padé approximants to the vector case. Similar results hold for the MPR problem, even in the matrix case. See, e.g., [2].

Besides all the previously mentioned definitions of vector Padé approximants, this paper contains yet another approach to possible extensions of the notion of scalar Padé approximants to the vector case.

Acknowledgement

The authors thank Peter Graves-Morris for making many valuable suggestions to make the paper more readable.

References

[1] B.M. Anderson, F.M. Brasch Jr and P.V. Lopresti, The sequential construction of minimal partial realizations from finite input–ouput data, *SIAM J. Control* **13** (1975) 552–571.

[2] A.C. Antoulas, On recursiveness and related topics in linear systems, *IEEE Trans. Automat. Control* **31** (1986) 1121–1135.

[3] A. Bultheel and M. Van Barel, A matrix Euclidean algorithm and matrix Padé approximation, Report TW104, Dept. Comput. Sci., Kath. Univ. Leuven, 1988.

[4] A. Bultheel and G.-L. Xu, The structure of the minimal Padé approximation table, Report TW117, Dept. Comput. Sci., Kath. Univ. Leuven, 1988.

[5] J. Coates, On the algebraic computation of functions, *Indag. Math.* **28** (1966) 421–461.

[6] M. de Bruin, Generalized continued fractions and a multidimensional Padé table, Ph.D. Thesis, Amsterdam, 1974.

[7] B.W. Dickinson, M. Morf and T. Kailath, A minimal realization algorithm for matrix sequences, *IEEE Trans. Automat. Control* **19** (1974) 31–38.

[8] P.R. Graves-Morris, Vector valued interpolants I, *Numer. Math.* **42** (1983) 331–348; Vector valued interpolants II, *IMA J. Numer. Anal.* **4** (1984) 209–224; (with C.D. Jenkins) Vector valued interpolants III, *Constr. Approx.* **2** (1986) 263–289.

[9] P.R. Graves-Morris and J.M. Wilkins, A fast algorithm to solve Kalman's partial realisation problem for single-input, multi-output systems, in: E.B. Saff, Ed., *Approximation Theory, Tampa*, Lecture Notes in Math. **1287** (Springer, Berlin, 1987) 1–20.

[10] H. Jager, A multidimensional generalization of the Padé table, *Indag. Math.* **26** (1964) 193–249.

[11] R.E. Kalman, On partial realizations, transfer functions and canonical forms, *Acta Polytech. Scand. Math. Comput. Sci. Ser.* **31** (1979) 9–32.

[12] K. Mahler, Perfect systems, *Compositio Math.* **19** (1968) 95–166.

[13] E.M. Nikisin, On simultaneous Padé approximants, *Math. USSR Sb.* **41** (1982) 409–425.

[14] M. Van Barel and A. Bultheel, Minimal Padé sense matrix approximations around $s = 0$ and $s = \infty$, in: A. Cuyt, Ed., *Nonlinear Numerical Methods and Rational Approximation* (Reidel, Dordrecht, 1988) 143–154.

[15] M. Van Barel and A. Bultheel, A canonical matrix continued fraction solution of the minimal (partial) realization problem, *Linear Algebra Appl.* **122–124** (1989) 973–1002.

[16] J. van Iseghem, Approximants de Padé vectoriels, Ph.D. Thesis, Lille, 1987.

[17] P. Wynn, Continued fractions whose coefficients obey a non-commutative law of multiplication, *Arch. Rational Mech. Anal.* **12** (1963) 273–312.