Journal of Nonlinear Mathematical Physics Volume 10, Supplement 2 (2003), 1–12 SIDE V

Comparisons Between Vector and Matrix Padé Approximants

Claude BREZINSKI

Laboratoire de Mathématiques Appliquées, FRE CNRS 2222, Université des Sciences et Technologies de Lille, 59655–Villeneuve d'Ascq cedex, France
E-mail: Claude.Brezinski@univ-lille1.fr

This paper is part of the Proceedings of **SIDE V***;*
Giens, June 21-26, 2002

Abstract

In this paper, we compare the degrees and the orders of approximation of vector and matrix Padé approximants for series with matrix coefficients. It is shown that, in this respect, vector Padé approximants have better properties. Then, matrix–vector Padé approximants are defined and constructed. Finally, matrix Padé approximants are related to the method of moments.

1 Introduction

Padé approximants are rational functions whose series expansion in ascending powers of the variable matches a given series as far as possible, that is up to the sum of the degrees of its numerator and its denominator inclusively. This characteristic property defines completely a Padé approximant. Padé–type approximants are also rational functions, with an arbitrarily chosen denominator, and a numerator computed so that the power series expansion of the approximant matches the original series as far as possible, that is up to the degree of the numerator inclusively. Such approximants have received many important applications in numerical analysis and in various branches of applied mathematics, physics, chemistry, etc. They have been generalized to the case of series with nonscalar coefficients. So, a power series with vector coefficients can be approximated by vector Padé approximants and a series with matrix coefficients by matrix Padé approximants (see, for example, [8, 9]). In the matrix case, if the columns of each matrix coefficient are written one after the other into a vector, then vector Padé approximants can also be used. The aim of this paper is to compare the efficiency of vector and matrix Padé approximants in the case of series with matrix coefficients from the point of view of the degrees of the Padé approximant needed to achieve a certain order of approximation, and from the point of view of the order of approximation obtained with certain degrees. We will see that, with respect to these two criteria, vector Padé approximants are more efficient. Then, we will define and construct matrix–vector Padé approximants. Finally, matrix Padé approximants will be related to a generalization of the method of moments.

Let us first give some properties that will be useful in the sequel. Matrices will be designated by bold letters, while vectors will have an arrow above them.

Let $\mathbf{A}$ be a square $n \times n$ nonsingular matrix, we remind that

$$\mathbf{A}^{-1} = \frac{\text{adj } \mathbf{A}}{\det \mathbf{A}}$$

where adj $\mathbf{A}$ is the *adjunct matrix* of $\mathbf{A}$, that is the $n \times n$ matrix whose element (j,i) is the determinant of the submatrix obtained by deleting the ith row and the jth column of $\mathbf{A}$, multiplied by $(-1)^{i+j}$.

The characteristic polynomial of $\mathbf{A}$, $P_n(s) = \det(s\mathbf{I} - \mathbf{A})$, has degree n. Its zeros are the eigenvalues of $\mathbf{A}$. If we set $P_n(s) = a_n s^n + a_{n-1}s^{n-1} + \cdots + a_0$, with $a_n = 1$, then

$$\begin{aligned} \text{adj } (s\mathbf{I} - \mathbf{A}) &= \mathbf{A}^{n-1} + (s + a_{n-1})\mathbf{A}^{n-2} + \cdots + (s^{n-1} + a_{n-1}s^{n-2} + \cdots + a_1)\mathbf{I} \\ &= s^{n-1}\mathbf{I} + s^{n-2}(\mathbf{A} + a_{n-1}\mathbf{I}) + \cdots + (\mathbf{A}^{n-1} + a_{n-1}\mathbf{A}^{n-2} + \cdots + a_1\mathbf{I}) \end{aligned}$$

and we finally obtain

$$(s\mathbf{I} - \mathbf{A})^{-1} = \frac{\mathbf{Q}_{n-1}(s)}{P_n(s)} = \frac{\mathbf{B}_{n-1}s^{n-1} + \mathbf{B}_{n-2}s^{n-2} + \cdots + \mathbf{B}_0}{\det(s\mathbf{I} - \mathbf{A})} \tag{1.1}$$

where the $\mathbf{B}_i$ are $n \times n$ matrices.

The matrices $\mathbf{B}_i$ and the coefficients a_i of P_n can be computed by the Leverrier–Faddeev–Souriau formulae (which are not suitable for numerical computation when n becomes large)

$$\begin{aligned} \mathbf{B}_{n-1} &= \mathbf{I} \\ \mathbf{B}_{n-i-1} &= \mathbf{A}B_{n-i} + a_{n-i}\mathbf{I}, \quad i = 1, \ldots, n-1 \end{aligned}$$

with $a_{n-i} = -tr(\mathbf{A}B_{n-i})/i$ for $i = 1, \ldots, n-1$, where tr indicates the *trace* of a matrix, that is the sum of its diagonal elements (on this topic, see, for example, [7, pp. 19–20]).

2 Vector and matrix Padé approximants

In this Section, we will give the definitions and some properties of vector and matrix Padé–type and Padé approximants.

Let us consider a formal vector series

$$\vec{G}(s) = \sum_{i=0}^{\infty} \vec{C}_i s^i, \quad \vec{C}_i \in \mathbb{R}^d,$$

a vector polynomial $\vec{Q}_{k+h-1} \in \mathbb{R}^d$ of degree $k + h - 1 \geq 0$ and a scalar polynomial P_k of degree k. We assume that

$$\vec{Q}_{k+h-1}(s) - \vec{G}(s)P_k(s) = \mathcal{O}(s^q).$$

The integer q is called the *order of approximation*. If $q = k + h$, $\vec{Q}_{k+h-1}(s)/P_k(s)$ is called a *vector Padé–type approximant* (of *dimension* d to be more precise) of $\vec{G}$ and it is denoted by $(k + h - 1/k)_{\vec{G}}$. The construction of this approximant needs the knowledge of $\vec{C}_0, \ldots, \vec{C}_{k+h-1}$. By a convenient choice of P_k, an order $q = k + h + \alpha$, where α is the integer part of k/d, can be obtained. The approximants are then denoted $[k + h - 1/k]_{\vec{G}}$ and called *vector Padé approximants* (of dimension d). Assuming for simplicity that k/d is an integer (otherwise, vector approximants with an order of approximation $q + 1$ for the first components and q for the others can be defined), their computation needs the knowledge of $\vec{C}_0, \ldots, \vec{C}_{k+h+\alpha-1}$. So, the number of coefficients required is always equal to the order of approximation. This kind of vector Padé–type and Padé approximants was introduced by van Iseghem [18] (see [8, pp. 169–176] and [9, pp. 81–85] for a survey). Vector Padé approximants are built from formal vector orthogonal polynomials [19].

Let us now consider a formal matrix series

$$\mathbf{G}(s) = \sum_{i=0}^{\infty} \mathbf{C}_i s^i, \quad \mathbf{C}_i \in \mathbb{R}^{p\times m}$$

and the matrix polynomials $\mathbf{Q}_{k+h-1} \in \mathbb{R}^{p\times m}$ of degree $k + h - 1 \geq 0$ and $\mathbf{P}_k$ of degree k.

We assume that these polynomials satisfy one of the following relations

$$\begin{aligned} \mathbf{Q}_{k+h-1}(s) - \mathbf{G}(s)\mathbf{P}_k(s) &= \mathcal{O}(s^q), \quad \mathbf{P}_k \in \mathbb{R}^{m\times m} \\ \mathbf{Q}_{k+h-1}(s) - \mathbf{P}_k(s)\mathbf{G}(s) &= \mathcal{O}(s^q), \quad \mathbf{P}_k \in \mathbb{R}^{p\times p} \end{aligned}$$

where the integer q is the order of approximation. The left *matrix Padé–type approximants* $\mathbf{Q}_{k+h-1}(s)[\mathbf{P}_k(s)]^{-1}$ of $\mathbf{G}$ and the right matrix Padé–type approximants $[\mathbf{P}_k(s)]^{-1}\mathbf{Q}_{k+h-1}(s)$ correspond, respectively, to $q = k + h$ in these relations. They both need the knowledge of $\mathbf{C}_0, \ldots, \mathbf{C}_{k+h-1}$. Usually, the left approximants and the right ones are not identical. In the sequel, we will only use the left approximants and denote them by $(k + h - 1/k)_{\mathbf{G}}$. Square matrix Padé–type approximants were introduced by Draux [10] (see also [8, pp. 176–180]).

The polynomial $\mathbf{P}_k$ can be chosen in order to obtain an improved order of approximation $q > k + h$. Indeed, let us set

$$\mathbf{P}_k(s) = \mathbf{A}_0 + \cdots + \mathbf{A}_{k-1}s^{k-1} + \mathbf{A}_k s^k, \quad \mathbf{A}_k = \mathbf{I}.$$

Thus, the order of approximation will be $q = k + h + \beta$ if the matrices $\mathbf{A}_i \in \mathbb{R}^{m\times m}$ satisfy

$$\mathbf{C}_{k+h+i}\mathbf{A}_0 + \cdots + \mathbf{C}_{h+i}\mathbf{A}_k = \mathbf{0}, \quad i = 0, \ldots, \beta - 1.$$

Each of these matrix equations is equivalent to pm scalar equations with km^2 unknowns. Thus, β must be an integer satisfying $\beta pm = km^2$, that is

$$\beta p = km. \tag{2.1}$$

In this case, the approximants are called *matrix Padé approximants* and they are denoted by $[k+h-1/k]_{\mathbf{G}}$. They need the knowledge of $\mathbf{C}_0, \ldots, \mathbf{C}_{k+h+\beta-1}$. We see that the number of coefficients used in matrix approximants is again equal to the order of approximation. When $p = m$, the left and right matrix Padé approximants are identical. Results on matrix Padé approximants can be found in [1, pp. 429–466]. It is well–known that matrix Padé approximants are related to formal matrix orthogonal polynomials [10, 11, 17].

We have $[\mathbf{P}_k(s)]^{-1} = \text{ adj } \mathbf{P}_k(s)/\det \mathbf{P}_k(s)$. But $\det \mathbf{P}_k(s)$ is a polynomial of degree km in s and each element of adj $\mathbf{P}_k(s)$ is a polynomial of degree $k(m-1)$ in s. Since each element of $\mathbf{Q}_{k+h-1}$ is a polynomial of degree $k+h-1$, then $\mathbf{Q}_{k+h-1}(s)$ adj $\mathbf{P}_k(s)$ has degree $k(m-1)+k+h-1 = km+h-1$. Thus $\mathbf{Q}_{k+h-1}(s)[\mathbf{P}_k(s)]^{-1}$ is a rational function in s with a matrix numerator of degree $km+h-1$ and a scalar denominator of degree km.

Vector approximants can be used for treating the matrix case by forming vectors of dimension $d = pm$ where the columns of the matrices $\mathbf{C}_i$ arranged consecutively. Thus, in the vector case, we will take $k = \alpha d = \alpha pm$, where α is an integer.

Vector and matrix approximants can be compared according to two different criteria: both with numerator (and denominator) polynomials of the same degree or achieving the same order of approximation (that is, using the same number of coefficients of the series $\mathbf{G}$).

Let us begin by comparing approximants having the same degrees. Taking $k = \alpha p$ in the matrix case, the vector approximants $(\alpha pm + h - 1/\alpha pm)_{\vec{G}}$ and the matrix approximants $(\alpha p + h - 1/\alpha p)_{\mathbf{G}}$ both have a numerator of degree $\alpha pm + h - 1$ and a denominator of degree αpm. The order of approximation of the Padé–type approximants is $\alpha pm + h$ in the vector case and only $\alpha p + h$ in the matrix case. For Padé approximants, this order is $\alpha pm + \alpha + h$ in the vector case and $\alpha p + \beta + h = \alpha(p+m) + h$ in the matrix case. Thus, matrix Padé approximants have an order of approximation strictly smaller than the order of the vector Padé approximants if m and p are both strictly greater than 1. The order is the same if and only if p or m is equal to 1. This remark explains the results on the order of approximation given in Section 3 which are exactly recovered.

Let us now compare these approximants when they have the same order of approximation. In the Padé–type case, the vector and the matrix approximants $(\alpha pm + h - 1/\alpha pm)$ both have the order $\alpha pm + h$ but the matrix approximants have a numerator of degree $\alpha pm^2 + h - 1$ and a denominator of degree αpm^2. In the Padé case, assuming that $\hat{k} = \alpha p(pm+1)/(m+p)$ is an integer, the vector approximants $[\alpha pm + h - 1/\alpha pm]_{\vec{G}}$ and the matrix approximants $[\hat{k}+h-1/\hat{k}]_{\mathbf{G}}$ both have the order of approximation $\alpha(pm+1)+h$. However, the matrix approximants have a numerator of degree $\hat{k}m + h - 1$ and a denominator of degree $\hat{k}m > \alpha pm$.

Thus, in all cases, matrix Padé–type and Padé approximants have less interesting properties than the corresponding vector approximants since, with the same degrees, a lower order of approximation is obtained and, for achieving the same order of approximation, higher degrees have to be used.

3 Matrix–vector Padé approximants

As explained in Section 2, the matrix Padé approximant $[k-1/k]_{\mathbf{G}}$ is the product of a matrix polynomial of degree $k-1$ by the inverse of a second matrix polynomial of degree k with coefficients of dimension $m \times m$. So, the denominator of $[k-1/k]_{\mathbf{G}}$ is a polynomial of degree km, and it is the denominator of each entry of the numerator.

The generalization of scalar Padé approximants which allows to construct directly several approximants with the same denominator is that of vector Padé approximants. The power series expansion of a vector Padé approximant in ascending powers of s agrees with the series $\vec{G}$ as far as possible. So, we will make use of vector Padé approximants. However, since our problem is formulated in term of matrices, it seems that we first have to transform it into a vector form by using vectors of dimension $d = pm$ containing each column (or row) of the corresponding matrices consecutively. However, such a transformation is not necessary since the vectors involved will only be multiplied by scalars, an operation equivalent to the multiplication of the corresponding matrices by the scalar. Although they will appear in the sequel under a matrix form, these approximants are still the vector Padé approximants of Section 2 since they approximate simultaneously several series by rational functions with the same denominator, which is the property characterizing these approximants. To distinguish them, they will be called *matrix–vector* Padé (–type) approximants, and usual roman letters will be used.

The matrix power series $\mathbf{G}$ will be approximated by the rational function

$$R_k(s) = \sum_{i=0}^{k-1} A_i s^i / \sum_{i=0}^{k} b_i s^i$$

where $A_i \in \mathbb{R}^{p\times m}$ and $b_i \in \mathbb{R}$. The matrices A_i and the coefficients b_i are chosen so that

$$\sum_{i=0}^{k-1} A_i s^i - \left(\sum_{i=0}^{\infty} C_i s^i\right)\left(\sum_{i=0}^{k} b_i s^i\right) = \mathcal{O}(s^q)$$

where $q \in \mathbb{N}$ is the order of approximation.

Identifying the coefficients in both sides we see that, if the matrices A_i satisfy

$$\begin{array}{rcl} A_0 & = & b_0 C_0 \\ A_1 & = & b_0 C_1 + b_1 C_0 \\ & \vdots & \\ A_{k-1} & = & b_0 C_{k-1} + \cdots + b_{k-1} C_0, \end{array} \tag{3.1}$$

then the preceding approximation property holds with $q = k$.

Since a rational function is defined apart from a multiplying factor, we will take $b_0 = 1$. If the matrix coefficients A_i are given by (3.1), R_k is the vector Padé–type approximant $(k-1/k)_{\vec{G}}(s)$ and, for any choice of $b_1, \ldots, b_k$ (that is, for any choice of the denominator), we have

$$(k-1/k)_{\vec{G}}(s) = \vec{G}(s) + \mathcal{O}(s^k).$$

The denominator P_k of $(k-1/k)_{\vec{G}}$ has to be an approximation (in a certain sense) of the characteristic polynomial P_n of the matrix $\mathbf{A}$ (this point will be discussed in Section 4). For this purpose and for increasing the order q of approximation of the vector Padé–type approximant $(k-1/k)_{\vec{G}}$ from k to $k+\alpha$, as explained in Section 2, the following additional block equations must be satisfied

$$\begin{array}{rcl} 0 & = & b_0 C_k + \cdots + b_k C_0 \\ 0 & = & b_0 C_{k+1} + \cdots + b_k C_1 \\ & \vdots & \\ 0 & = & b_0 C_{k+\alpha-1} + \cdots + b_k C_{\alpha-1}. \end{array} \tag{3.2}$$

Each of these matrix equations corresponds, in fact, to pm scalar equations. So, in general, the system (3.2) has a solution only if $k \geq pm$. The maximum possible value for the integer α is $\alpha = \lfloor k/pm \rfloor$. If k is such that $(\alpha - 1)pm < k < \alpha pm$, that is if $k = (\alpha - 1)pm + \delta$ with $0 < \delta < pm$, only δ equations will be selected among the last block equation in (3.2). When these conditions are satisfied, $(k-1/k)_G$ becomes the vector Padé approximant $[k-1/k]_{\vec{G}}$ of $\vec{G}$. The order of approximation is $q = k + \lfloor k/pm \rfloor$ and that of the equations selected in the last block of (3.2) is $q+1$.

These remarks confirm the considerations on the degrees of the approximants and their orders of approximation given in Section 2. Indeed, when $k = \alpha pm$, the conclusions of this Section are recovered. It follows that Padé approximants for matrix series can be constructed by means of vector approximants. The algebraic aspects of matrix orthogonality for vector polynomials was already discussed in [17].

4 The generalized method of moments

As explained in [3], the method of moments of Vorobyev [20, pp. 14–23] is related to continued fractions and Padé approximants [20, pp. 54–60] and it falls into the framework of Galerkin's method. It is also related to Lanczos' methods for the biorthogonalization of two sets of vectors [15] and for the solution of a system of linear equations [16]; on these topics, see [2, pp. 79–81], [3, 6], and [4, pp. 154–164]. We will now generalize the method of moments to the matrix case and relate it to Padé–type and Padé approximants. A first account of this procedure was given in Vorobyev [20, pp. 128–134] where it was called the *generalized method of moments*. The results given below extend those of [6].

Let $\mathbf{v}_i \in \mathbb{R}^{n\times m}$ and $\mathbf{w}_i \in \mathbb{R}^{n\times p}$. We denote by E_r the subspace of $n \times m$ matrices of the form $\mathbf{v} = \mathbf{v}_0\mathbf{a}_0 + \cdots + \mathbf{v}_{r-1}\mathbf{a}_{r-1}$ where $\mathbf{a}_i \in \mathbb{R}^{m\times m}$. Similarly, let F_l be the subspace consisting of $n \times p$ matrices $\mathbf{w} = \mathbf{w}_0\mathbf{b}_0 + \cdots + \mathbf{w}_{l-1}\mathbf{b}_{l-1}$ with $\mathbf{b}_i \in \mathbb{R}^{p\times p}$. The columns of all matrices $\mathbf{v}_i$ are assumed to be linearly independent and a similar assumption for the columns of the matrices $\mathbf{w}_i$.

Let $\mathbf{H}_k \in \mathbb{R}^{n\times n}$ be the matrix representing the projection onto E_r orthogonally to F_l, which means that $\forall \mathbf{v} \in \mathbb{R}^{n\times m}$, $\mathbf{H}_k\mathbf{v} \in E_r$ and $\mathbf{v} - \mathbf{H}_k\mathbf{v} \perp F_l$. The generalized method of moments consists in constructing a linear operator $\mathbf{A}_k$ on E_r such that

$$\begin{aligned} \mathbf{v}_i &= \mathbf{A}_k\mathbf{v}_{i-1}, \quad i = 0, \ldots, r-1 \\ \mathbf{H}_k\mathbf{v}_r &= \mathbf{A}_k\mathbf{v}_{r-1}. \end{aligned}$$

Thus, $\mathbf{v}_i = \mathbf{A}_k^i\mathbf{v}_0$ for $i = 0, \ldots, r-1$. The meaning of the subscript k will be explained below.

These conditions define completely $\mathbf{A}_k$ which is represented by a $n \times n$ matrix. Indeed, for all $\mathbf{v} \in E_r$, we have $\mathbf{v} = \mathbf{v}_0\mathbf{a}_0 + \cdots + \mathbf{v}_{r-1}\mathbf{a}_{r-1}$, where $\mathbf{a}_i \in \mathbb{R}^{m\times m}$. Thus $\mathbf{A}_k\mathbf{v} = \mathbf{v}_1\mathbf{a}_0 + \cdots + \mathbf{v}_{r-1}\mathbf{a}_{r-2} + \mathbf{H}_k\mathbf{v}_r\mathbf{a}_{r-1} \in E_r$. Since $\mathbf{H}_k\mathbf{v}_r \in E_r$, we can write $\mathbf{H}_k\mathbf{v}_r = -\mathbf{v}_0\mathbf{h}_0 - \cdots - \mathbf{v}_{r-1}\mathbf{h}_{r-1}$, with $\mathbf{h}_i \in \mathbb{R}^{m\times m}$, and $\mathbf{v}_r - \mathbf{H}_k\mathbf{v}_r = \mathbf{v}_0\mathbf{h}_0 + \cdots + \mathbf{v}_{r-1}\mathbf{h}_{r-1} + \mathbf{v}_r$. The condition $\mathbf{v}_r - \mathbf{H}_k\mathbf{v}_r \perp F_l$ gives

$$\mathbf{w}_i^T\mathbf{v}_0\mathbf{h}_0 + \cdots + \mathbf{w}_i^T\mathbf{v}_{r-1}\mathbf{h}_{r-1} + \mathbf{w}_i^T\mathbf{v}_r = \mathbf{0}, \quad i = 0, \ldots, l-1. \tag{4.1}$$

Since $\mathbf{w}_i^T\mathbf{v}_j \in \mathbb{R}^{p\times m}$ and $\mathbf{h}_i \in \mathbb{R}^{m\times m}$, this is a system of lpm equations in rm^2 unknowns. Thus, in the sequel, we will assume that

$$rm = lp. \tag{4.2}$$

We set $k = rm = lp$. This condition is equivalent to the condition that E_r and F_l, considered as the subspaces spanned by the columns (which are vectors in $\mathbb{R}^n$) of the matrices v_i and w_i respectively, have the same dimension k. So, we must have $k \leq n$. The knowledge of k is sufficient to recover the integers r and l. This is the reason why k is used as a subscript in some places instead of the two subscripts r and l. Let $\mathbf{V}_k$, resp. $\mathbf{W}_k$, be the $n\times k$ matrix $[\mathbf{v}_0,\ldots,\mathbf{v}_{r-1}]$, resp. $[\mathbf{w}_0,\ldots,\mathbf{w}_{l-1}]$. Then

$$\mathbf{H}_k = \mathbf{V}_k(\mathbf{W}_k^T\mathbf{V}_k)^{-1}\mathbf{W}_k^T.$$

We have $\mathbf{H}_k^2 = \mathbf{H}_k$ which shows that $\mathbf{H}_k$ is an oblique projection. The projection is orthogonal if and only if $\mathbf{V}_k = \mathbf{W}_k$. On projections, see, for example, [4, pp. 18–23].

We assume that the system (4.1) is nonsingular and we consider the matrix polynomial

$$\mathbf{P}_r(t) = \mathbf{h}_0 + \cdots + t^{r-1}\mathbf{h}_{r-1} + t^r\mathbf{I}_m$$

where $\mathbf{I}_m$ is the $m\times m$ identity matrix. Thus $\mathbf{P}_r(\mathbf{M}) \in \mathbb{R}^{m\times m}$ if $\mathbf{M} \in \mathbb{R}^{m\times m}$ and $\mathbf{P}_r(t) = \mathbf{P}_r(t\mathbf{I}_m) \in \mathbb{R}^{m\times m}$. We will make use of the notation

$$\mathbf{P}_r(\mathbf{M})\circ\mathbf{v} = \mathbf{v}h_0 + \cdots + \mathbf{M}^{r-1}\mathbf{v}h_{r-1} + \mathbf{M}^r\mathbf{v}$$

where $\mathbf{M} \in \mathbb{R}^{n\times n}$ and $\mathbf{v} \in \mathbb{R}^{n\times m}$. So, $(\mathbf{P}_r(\mathbf{M})\circ\mathbf{v}) \in \mathbb{R}^{n\times m}$ (on matrix polynomials, see [13]). Such polynomials were already used in the multiparameter Lanczos method [5], where they were related to formal vector orthogonality, and in the block Lanczos method [12], where they were related to formal matrix orthogonality.

We have

$$\begin{aligned}\mathbf{P}_r(\mathbf{A}_k)\circ\mathbf{v}_0 &= \mathbf{v}_0\mathbf{h}_0 + \cdots + \mathbf{A}_k^{r-1}\mathbf{v}_0\mathbf{h}_{r-1} + \mathbf{A}_k^r\mathbf{v}_0\\ &= \mathbf{v}_0\mathbf{h}_0 + \cdots + \mathbf{v}_{r-1}\mathbf{h}_{r-1} + \mathbf{H}_k\mathbf{v}_r = 0.\end{aligned}$$

It follows that $\mathbf{P}_r(\mathbf{A}_k)\circ\mathbf{v}_i = \mathbf{0}$ for $i = 0,\ldots,r-1$. Indeed, from what precedes,

$$\begin{aligned}\mathbf{A}_k^i(\mathbf{P}_r(\mathbf{A}_k)\circ\mathbf{v}_0) &= \mathbf{0}\\ &= \mathbf{A}_k^i\mathbf{v}_0\mathbf{h}_0 + \cdots + \mathbf{A}_k^i\mathbf{A}_k^{r-1}\mathbf{v}_0\mathbf{h}_{r-1} + \mathbf{A}_k^i\mathbf{A}_k^r\mathbf{v}_0\\ &= \mathbf{v}_i\mathbf{h}_0 + \cdots + \mathbf{A}_k^{r-1}\mathbf{v}_i\mathbf{h}_{r-1} + \mathbf{A}_k^r\mathbf{v}_i\\ &= \mathbf{P}_r(\mathbf{A}_k)\circ\mathbf{v}_i, \quad i = 0,\ldots,r-1\end{aligned} \tag{4.3}$$

since $\mathbf{v}_i = \mathbf{A}_k^i\mathbf{v}_0$ for $i = 0,\ldots,r-1$. Therefore, by linear combination, $\forall\mathbf{v}\in E_r$, we have $\mathbf{P}_r(\mathbf{A}_k)\circ\mathbf{v} = \mathbf{0}$. In particular, if $\mathbf{v}$ is a matrix whose columns are eigenvectors of $\mathbf{A}_k$, then $\mathbf{P}_r(\mathbf{A}_k)\circ\mathbf{v} = \mathbf{v}\mathbf{P}_r(\mathbf{\Lambda}) = \mathbf{0}$, where $\mathbf{\Lambda}$ is the $m\times m$ diagonal matrix of the corresponding eigenvalues.

Now, we consider the $n\times n$ systems of linear equations with m right hand sides

$$s\mathbf{x} = \mathbf{A}\mathbf{x} + \mathbf{v}_0 \quad\text{and}\quad s\mathbf{x}_k = \mathbf{A}_k\mathbf{x}_k + \mathbf{v}_0$$

where s is a parameter, and $\mathbf{x}$ and $\mathbf{x}_k$ belong to $\mathbb{R}^{n\times m}$. We set $z = s^{-1}$. Since, formally, $(\mathbf{I}_n - z\mathbf{A})^{-1} = \mathbf{I}_n + z\mathbf{A} + z^2\mathbf{A}^2 + \cdots$ and a similar expression for $(\mathbf{I}_n - z\mathbf{A}_k)^{-1}$, then

$$\mathbf{x} = z(\mathbf{v}_0 + z\mathbf{A}v_0 + z^2\mathbf{A}^2\mathbf{v}_0 + \cdots) \quad \text{and} \quad \mathbf{x}_k = z(\mathbf{v}_0 + z\mathbf{A}_k\mathbf{v}_0 + z^2\mathbf{A}_k^2\mathbf{v}_0 + \cdots).$$

Setting $\mathbf{d}_i = \mathbf{y}^T\mathbf{A}_k^i\mathbf{v}_0$, where $\mathbf{y} \in \mathbb{R}^{n\times p}$, we have from (4.3)

$$\mathbf{d}_i\mathbf{h}_0 + \cdots + \mathbf{d}_{i+r-1}\mathbf{h}_{r-1} + \mathbf{d}_{i+r} = \mathbf{0}, \quad i = 0, 1, \ldots \tag{4.4}$$

This result shows that, formally, $\mathbf{d}_0 + \mathbf{d}_1 z + \mathbf{d}_2 z^2 + \cdots = \widetilde{\mathbf{Q}}_{r-1}(z)[\widetilde{\mathbf{P}}_r(z)]^{-1}$, where $\widetilde{\mathbf{Q}}_{r-1}$ is a polynomial of degree $r-1$ in z and $\widetilde{\mathbf{P}}_r$ a polynomial of degree r in z. Moreover, from (4.4), we see that $\widetilde{\mathbf{P}}_r(z) = z^r\mathbf{P}_r(z^{-1})$, where $\mathbf{P}_r$ is the polynomial obtained by the generalized method of moments. If we set $\widetilde{\mathbf{Q}}_{r-1}(z) = \mathbf{e}_0 z^{r-1} + \cdots + \mathbf{e}_{r-1}$, then

$$\begin{aligned}
\mathbf{e}_{r-1} &= \mathbf{d}_0 \\
\mathbf{e}_{r-2} &= \mathbf{d}_1 + \mathbf{d}_0\mathbf{h}_{r-1} \\
&\vdots \\
\mathbf{e}_0 &= \mathbf{d}_{r-1} + \mathbf{d}_{r-2}\mathbf{h}_{r-1} + \cdots + \mathbf{d}_0\mathbf{h}_1.
\end{aligned}$$

We set $\mathbf{g}_k(z) = \mathbf{y}^T\mathbf{x}_k = z(\mathbf{d}_0 + \mathbf{d}_1 z + \mathbf{d}_2 z^2 + \cdots)$ and $\mathbf{g}(z) = \mathbf{y}^T\mathbf{x} = z(\mathbf{c}_0 + \mathbf{c}_1 z + \mathbf{c}_2 z^2 + \cdots)$ with $\mathbf{c}_i = \mathbf{y}^T\mathbf{A}^i\mathbf{v}_0$. If $\forall i, \mathbf{v}_i = \mathbf{A}^i\mathbf{v}_0$, then $\mathbf{d}_i = \mathbf{c}_i$ for $i = 0, \ldots, r-1$ since $\mathbf{v}_i = \mathbf{A}_k^i\mathbf{v}_0$ for $i = 0, \ldots, r-1$. So, we see that

$$\mathbf{g}_k(z) = \mathbf{g}(z) + \mathcal{O}(z^r)$$

which shows that $\mathbf{g}_k$ is the $(r-1/r)_{\mathbf{g}}$ matrix Padé–type approximant of $\mathbf{g}$.

It must be remarked that

$$\begin{aligned}
\mathbf{g}_k(s^{-1}) &= s^{-1}\widetilde{\mathbf{Q}}_{r-1}(s^{-1})[\widetilde{\mathbf{P}}_r(s^{-1})]^{-1} \\
&= \mathbf{Q}_{r-1}(s)[\mathbf{P}_r(s)]^{-1}
\end{aligned}$$

with $\mathbf{Q}_{r-1}(s) = s^{r-1}\widetilde{\mathbf{Q}}_{r-1}(s^{-1})$. As already explained in Section 2, this result does not mean that g_k is the quotient of a polynomial of degree $r-1$ by a polynomial of degree r. In fact $\mathbf{Q}_{r-1}(s)[\mathbf{P}_r(s)]^{-1}$ is a rational function in s with a matrix numerator of degree $mr-1$ and a scalar denominator of degree mr (a result which can be seen directly from the equation satisfied by $\mathbf{x}_k$).

Let $\mathbf{P}$ and $\mathbf{Q}$ be two $m \times m$ matrix polynomials related by

$$s\mathbf{I}_m - \mathbf{P}(t) = (s-t)\mathbf{Q}(t). \tag{4.5}$$

Then, the equation

$$s\mathbf{x}_k = \mathbf{P}(\mathbf{A}_k) \circ \mathbf{x}_k + \mathbf{Q}(\mathbf{A}_k) \circ \mathbf{v}_0$$

has the same solution as $s\mathbf{x}_k = \mathbf{A}_k\mathbf{x}_k + \mathbf{v}_0$ since

$$(s\mathbf{I}_m - \mathbf{P}(\mathbf{A}_k)) \circ \mathbf{x}_k = \mathbf{Q}(\mathbf{A}_k) \circ \mathbf{v}_0 = \mathbf{Q}(\mathbf{A}_k) \circ ((s\mathbf{I}_n - \mathbf{A}_k)\mathbf{x}_k).$$

Since, from (4.5), the polynomial $\mathbf{P}$ must satisfy $\mathbf{P}(s) = s\mathbf{I}_m$ if $t = s$, let us take

$$\mathbf{P}(t) = s\mathbf{P}_r(t)[\mathbf{P}_r(s)]^{-1}.$$

But, $\mathbf{P}_r(s)\mathbf{P}_r(t) = \mathbf{P}_r(t)\mathbf{P}_r(s)$ and, multiplying on the left and on the right by the inverse of $\mathbf{P}_r(s)$, we see that $\mathbf{P}_r(t)$ commutes with it. Then

$$\begin{aligned} \mathbf{x}_k &= s^{-1}s[\mathbf{P}_r(s)]^{-1}(\mathbf{P}_r(\mathbf{A}_k)\circ\mathbf{x}_k) + s^{-1}\mathbf{Q}(\mathbf{A}_k)\circ\mathbf{v}_0 \\ &= s^{-1}\mathbf{Q}(\mathbf{A}_k)\circ\mathbf{v}_0 \end{aligned}$$

since $\mathbf{x}_k \in E_r$ and $\forall \mathbf{v} \in E_r, \mathbf{P}_r(\mathbf{A}_k)\circ\mathbf{v} = \mathbf{0}$.

Thus $\mathbf{g}_k(s^{-1}) = s^{-1}\mathbf{y}^T(\mathbf{Q}(\mathbf{A}_k)\circ\mathbf{v}_0)$. But, from (4.5),

$$\frac{\mathbf{P}_r(s) - \mathbf{P}_r(t)}{s-t} = s^{-1}\mathbf{Q}(t)\mathbf{P}_r(s).$$

Using the expression of $\mathbf{P}_r$, we get

$$\begin{aligned} \frac{\mathbf{P}_r(s) - \mathbf{P}_r(t)}{s-t} &= (\mathbf{h}_1 + \mathbf{h}_2 t + \cdots + \mathbf{h}_{r-1}t^{r-2} + t^{r-1}) \\ &\quad + s(\mathbf{h}_2 + \cdots + \mathbf{h}_{r-1}t^{r-3} + t^{r-2}) + \cdots + s^{r-2}(\mathbf{h}_{r-1} + t) + s^{r-1}. \end{aligned}$$

It follows

$$\begin{aligned} &\mathbf{y}^T(s\mathbf{I}_n - \mathbf{A}_k)^{-1}[(\mathbf{P}_r(s\mathbf{I}_n) - \mathbf{P}_r(\mathbf{A}_k))\circ\mathbf{v}_0] \\ &\quad = (\mathbf{d}_0\mathbf{h}_1 + \cdots + \mathbf{d}_{r-2}\mathbf{h}_{r-1} + \mathbf{d}_{r-1}) + \cdots + \mathbf{d}_0 s^{r-1} = \mathbf{Q}_{r-1}(s) \end{aligned}$$

and we finally recover

$$\mathbf{g}_k(s^{-1}) = \mathbf{Q}_{r-1}(s)[\mathbf{P}_r(s)]^{-1}.$$

The nonsingularity of the matrix $\mathbf{P}_r(s)$ remains to be discussed. Let λ_i, $i = 1, \ldots, m$, be an eigenvalue of $\mathbf{A}_k$ such that its associated eigenvector belongs to E_r. So, setting $\mathbf{\Lambda} = \text{diag}\ (\lambda_1, \ldots, \lambda_m)$, we have $\mathbf{A}_k v = v\mathbf{\Lambda}$ with $\mathbf{v} = \mathbf{v}_0\mathbf{a}_0 + \cdots + \mathbf{v}_{r-1}\mathbf{a}_{r-1} \in E_r$ with $\mathbf{a}_i \in \mathbb{R}^{m\times m}$ and

$$\begin{aligned} \mathbf{A}_k\mathbf{v} &= \mathbf{A}_k\mathbf{v}_0\mathbf{a}_0 + \cdots + \mathbf{A}_k\mathbf{v}_{r-2}\mathbf{a}_{r-2} + \mathbf{A}_k\mathbf{v}_{r-1}\mathbf{a}_{r-1} \\ &= \mathbf{v}_1\mathbf{a}_0 + \cdots + \mathbf{v}_{r-1}\mathbf{a}_{r-2} + \mathbf{H}_k\mathbf{v}_r\mathbf{a}_{r-1} \\ &= \mathbf{v}_1\mathbf{a}_0 + \cdots + \mathbf{v}_{r-1}\mathbf{a}_{r-2} + (-\mathbf{v}_0\mathbf{h}_0 - \cdots - \mathbf{v}_{r-1}\mathbf{h}_{r-1})\mathbf{a}_{r-1} \\ &= -\mathbf{v}_0\mathbf{h}_0\mathbf{a}_{r-1} + \mathbf{v}_1(\mathbf{a}_0 - \mathbf{h}_1\mathbf{a}_{r-1}) + \cdots + \mathbf{v}_{r-1}(\mathbf{a}_{r-2} - \mathbf{h}_{r-1}\mathbf{a}_{r-1}) \\ &= \mathbf{v}_0\mathbf{a}_0\mathbf{\Lambda} + \cdots + \mathbf{v}_{r-1}\mathbf{a}_{r-1}\mathbf{\Lambda}. \end{aligned}$$

Thus

$$\begin{aligned} -\mathbf{h}_0\mathbf{a}_{r-1} &= \mathbf{a}_0\mathbf{\Lambda} \\ \mathbf{a}_i - \mathbf{h}_{i+1}\mathbf{a}_{r-1} &= \mathbf{a}_{i+1}\mathbf{\Lambda}, \quad i = 0, \ldots, r-2 \end{aligned}$$

which shows that $\mathbf{a}_0, \ldots, \mathbf{a}_{r-1}$ are solution of the system

$$\begin{pmatrix} -\mathbf{\Lambda} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & -\mathbf{h}_0 \\ \mathbf{I}_m & -\mathbf{\Lambda} & \mathbf{0} & \cdots & \mathbf{0} & -\mathbf{h}_1 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{I}_m & -\mathbf{h}_{r-1} - \mathbf{\Lambda} \end{pmatrix} \begin{pmatrix} \mathbf{a}_0 \\ \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_{r-1} \end{pmatrix} = \mathbf{0}.$$

Since the solution of this system is not identically zero, then λ_i, $i = 1, \ldots, m$, must also be a zero of the determinant of this matrix. This matrix is called the *companion matrix* of $\mathbf{P}_r$ and its determinant is equal to $\det \mathbf{P}_r(\lambda)$; see [13, pp. 13–14] or [14, pp. 490–493]. So, $\det \mathbf{P}_r(s) \neq 0$ if and only if s is not an eigenvalue of $\mathbf{A}_k$ whose corresponding eigenvector belongs to E_r.

We will now choose the matrices $\mathbf{w}_i$ in order to improve the order of approximation. We have $\mathbf{P}_r(\mathbf{A}) \circ \mathbf{v}_0 - \mathbf{P}_r(\mathbf{A}_k) \circ \mathbf{v}_0 = \mathbf{v}_r - \mathbf{H}_k \mathbf{v}_r$. Thus

$$\mathbf{w}_i^T [\mathbf{P}_r(\mathbf{A}) \circ \mathbf{v}_0 - \mathbf{P}_r(\mathbf{A}_k) \circ \mathbf{v}_0] = \mathbf{0}, \quad i = 0, \ldots, l-1.$$

But $\mathbf{P}_r(\mathbf{A}_k) \circ \mathbf{v}_0 = \mathbf{0}$ and it follows

$$\mathbf{w}_i^T (\mathbf{P}_r(\mathbf{A}) \circ \mathbf{v}_0) = \mathbf{0}, \quad i = 0, \ldots, l-1.$$

Let us consider the choice $\mathbf{w}_i = (\mathbf{A}^T)^i \mathbf{y}$. Then

$$\mathbf{w}_i^T (\mathbf{P}_r(\mathbf{A}) \circ \mathbf{v}_0) = \mathbf{y}^T \mathbf{A}^i (\mathbf{P}_r(\mathbf{A}) \circ \mathbf{v}_0) = \mathbf{y}^T (\mathbf{P}_r(\mathbf{A}) \circ \mathbf{v}_i) = \mathbf{0}, \quad i = 0, \ldots, l-1$$

that is

$$\begin{aligned} \mathbf{0} &= \mathbf{y}^T \mathbf{v}_i \mathbf{h}_0 + \cdots + \mathbf{y}^T \mathbf{A}^{r-1} \mathbf{v}_i \mathbf{h}_{r-1} + \mathbf{y}^T \mathbf{A}^r \mathbf{v}_i \\ &= \mathbf{c}_i \mathbf{h}_0 + \cdots + \mathbf{c}_{i+r-1} \mathbf{h}_{r-1} + \mathbf{c}_{r+i}, \quad i = 0, \ldots, l-1 \end{aligned} \tag{4.6}$$

since $\mathbf{c}_{i+j} = \mathbf{y}^T \mathbf{A}^j \mathbf{v}_i$. Subtracting (4.6) from (4.4), we obtain

$$(\mathbf{c}_i - \mathbf{d}_i)\mathbf{h}_0 + \cdots + (\mathbf{c}_{i+r-1} - \mathbf{d}_{i+r-1})\mathbf{h}_{r-1} + \mathbf{c}_{i+r} - \mathbf{d}_{i+r} = \mathbf{0}, \quad i = 0, \ldots, l-1.$$

When $i = 0$, this relation shows that $\mathbf{d}_r = \mathbf{c}_r$ since $\mathbf{d}_i = \mathbf{c}_i$ for $i = 0, \ldots, r-1$. When $i = 1$, it leads to $\mathbf{d}_{r+1} = \mathbf{c}_{r+1}$, and so on. Thus, it follows that $\mathbf{d}_{r+i} = \mathbf{c}_{r+i}$ for $i = 0, \ldots, l-1$ and, finally, we have $\mathbf{d}_i = \mathbf{c}_i$, $i = 0, \ldots, r+l-1$ and

$$\mathbf{g}_k(z) = \mathbf{g}(z) + \mathcal{O}(z^{r+l}).$$

The condition (2.1) becomes (after replacing k by r) $\beta p = rm$ and, comparing with (4.2), we see that $\beta = l$. Thus, $\mathbf{g}_k$ is the $[r-1/r]_{\mathbf{g}}$ matrix Padé approximant of $\mathbf{g}$. Moreover, the relation (4.6) shows the well–known connection with formal matrix orthogonal polynomials [10].

Let us give an expression for $\mathbf{A}_k$. Let $\mathbf{v} = \mathbf{v}_0 \mathbf{c}_0 + \cdots + \mathbf{v}_{r-1} \mathbf{c}_{r-1} \in E_r$. Then

$$\begin{aligned} \mathbf{A} v &= \mathbf{A} v_0 \mathbf{c}_0 + \cdots + \mathbf{A} v_{r-1} \mathbf{c}_{r-1} \\ &= \mathbf{A} v_0 \mathbf{c}_0 + \cdots + \mathbf{A}^{r-1} \mathbf{v}_0 \mathbf{c}_{r-2} + \mathbf{A}^r \mathbf{v}_0 \mathbf{c}_{r-1} \\ &= \mathbf{A}_k \mathbf{v}_0 \mathbf{c}_0 + \cdots + \mathbf{A}_k^{r-1} \mathbf{v}_0 \mathbf{c}_{r-2} + \mathbf{A}^r \mathbf{v}_0 \mathbf{c}_{r-1}. \end{aligned}$$

It follows, since $\mathbf{H}_k \mathbf{v}_i = \mathbf{v}_i$ for $i = 0, \ldots, r-1$,

$$\begin{aligned} \mathbf{H}_k \mathbf{A} v &= \mathbf{A}_k \mathbf{v}_0 \mathbf{c}_0 + \cdots + \mathbf{A}_k^{r-1} \mathbf{v}_0 \mathbf{c}_{r-2} + \mathbf{H}_k \mathbf{v}_r \mathbf{c}_{r-1} \\ &= \mathbf{A}_k \mathbf{v}_0 \mathbf{c}_0 + \cdots + \mathbf{A}_k^{r-1} \mathbf{v}_0 \mathbf{c}_{r-2} + \mathbf{A}_k^r \mathbf{v}_0 \mathbf{c}_{r-1} \\ &= \mathbf{A}_k (\mathbf{v}_0 \mathbf{c}_0 + \cdots + \mathbf{v}_{r-1} \mathbf{c}_{r-1}) = \mathbf{A}_k \mathbf{v} \end{aligned}$$

which shows that $\mathbf{A}_k = \mathbf{H}_k\mathbf{A}$ on E_r. Since $\mathbf{H}_k\mathbf{v} \in E_r$ if $\mathbf{v} \in E_r$, the domain of $\mathbf{A}_k$ can be extended to the whole space by setting

$$\mathbf{A}_k = \mathbf{H}_k\mathbf{A}\mathbf{H}_k.$$

We set $\mathbf{V}_r = [\mathbf{v}_0, \ldots, \mathbf{v}_{r-1}]$ and $\mathbf{W}_l = [\mathbf{w}_0, \ldots, \mathbf{w}_{l-1}]$. By condition (4.2), these two matrices have the same dimension. We assume that $\mathbf{W}_l^T\mathbf{V}_r = \mathbf{I}$, which does not restrict the generality. We have $\mathbf{H}_k = \mathbf{V}_r\mathbf{W}_l^T$ and $\mathbf{A}_k = \mathbf{V}_r\mathbf{W}_l^T\mathbf{A}\mathbf{V}_r\mathbf{W}_l^T = \mathbf{V}_r\mathbf{J}_k\mathbf{W}_l^T$ with $\mathbf{J}_k = \mathbf{W}_l^T\mathbf{A}V_r$. It also holds $\mathbf{A}_k^i = \mathbf{V}_r\mathbf{J}_k^i\mathbf{W}_l^T$ and we have

$$\mathbf{g}_k(s) = \mathbf{y}^T\mathbf{V}_r(s\mathbf{I} - \mathbf{J}_k)^{-1}\mathbf{W}_l^T\mathbf{v}_0.$$

These results generalize those given in [2, pp. 75–78]. The case of systems of the form $A\mathbf{x} = \mathbf{v}_0$ and $\mathbf{A}_k\mathbf{x}_k = \mathbf{v}_0$ can be considered in a similar way; see [6].

References

[1] Baker Jr G A, Graves–Morris P R, Padé Approximants, 2nd ed., Cambridge University Press, Cambridge, 1996.

[2] Brezinski C, Padé–Type Approximation and General Orthogonal Polynomials, ISNM Vol. 50, Birkhäuser Verlag, Basel, 1980.

[3] Brezinski C, The methods of Vorobyev and Lanczos, *Linear Algebra Appl.* **234** (1996) 21–41.

[4] Brezinski C, Projection Methods for Systems of Equations, North–Holland, Amsterdam, 1997.

[5] Brezinski C, Multiparameter descent methods, *Linear Algebra Appl.* **296** (1999) 113–141.

[6] Brezinski C, The block Lanczos and Vorobyev methods, *C.R. Acad. Sci. Paris, Sér. I* **331** (2000) 137–142.

[7] Brezinski C, Computational Aspects of Linear Control, Kluwer, Dordrecht, 2002.

[8] Brezinski C, van Iseghem J, Padé approximations, in *Handbook of Numerical Analysis, vol. III*, P.G. Ciarlet and J.L. Lions eds., North–Holland, Amsterdam, 1994, 47–222.

[9] Brezinski C, van Iseghem J, A taste of Padé approximation, in *Acta Numerica 1995*, A Iserles ed., Cambridge University Press, Cambridge, 1995, 53–103.

[10] Draux A, Formal orthogonal polynomials and Padé approximants in a non–commutative algebra, in *Mathematical Theory of Networks and Systems, Beer Sheva, Israel, June 20–24, 1983*, P A Fuhrmann ed., *Lecture Notes in Control and Information Sciences*, vol. 58, Springer–Verlag, Berlin, 1984, 278–292.

[11] Duran A J, van Assche W, Orthogonal matrix polynomials and higher recurrence relations, Linear Algebra Appl. **219** (1995) 261–280.

[12] El Guennouni A, Mise en Œuvre et Variantes par Bloc des Méthodes de Type Lanczos, Thèse, Université des Sciences et Technologies de Lille, 2000.

[13] Gohberg I, Lancaster I P, Rodman L, Matrix Polynomials, Academic Press, New York, 1982.

[14] Lancaster P, Tismenetsky M, The Theory of Matrices, 2nd ed., Academic Press, New York, 1985.

[15] Lanczos C, An iteration method for the solution of the eigenvalue problem of linear differential and integral operators, *J. Res. Nat. Bur. Standards* **45** (1950) 255–282.

[16] Lanczos C, Solution of systems of linear equations by minimized iterations, *J. Res. Natl. Bur. Stand.* **49** (1952) 33–53.

[17] Sorokin V N, van Iseghem J, Algebraic aspects of matrix orthogonality for vector polynomials, *J. Approximation Theory* **90** (1997) 97–116.

[18] van Iseghem J, Vector Padé approximants, in *Numerical Mathematics and Applications*, R. Vichnevetsky and J. Vignes eds., North–Holland, Amsterdam, 1985, 73–77.

[19] van Iseghem J, Vector orthogonal relations, vector QD–algorithm, *J. Comput. Appl. Math.* **19** (1987) 141–150.

[20] Vorobyev Yu V, Method of Moments in Applied Mathematics, Gordon and Breach, New York, 1965.