Journal of Computational Mathematics, Vol.11, No.4, 1993, 339–349.

CONVERGENCE AND STABILITY PROPERTIES OF A VECTOR PADÉ EXTRAPOLATION METHOD*

Zhuang Guo-zhong
(Computing Center, Academia Sinica, Beijing, China)

Abstract

In this paper we introduce a new convergence accelerator for vector sequences — vector Padé approximation method(VPA), discuss its convergence and stability properties, and show that it is a bona fide acceleration method for some vector sequences.

1. Introduction

There are two well known families of convergence acceleration methods for vector sequences: polynomial methods and epsilon algorithms. Five classical methods have been discussed in [10]: the minimal polynomial extrapolation(MPE) of Cabay and Jackson[2], the reduced rank extrapolation(RRE) of Eddy[3] and Mešina[4], the scalar epsilon algorithm(SEA) and the vector epsilon algorithm(VEA) of Wynn[11,12], and the topological epsilon algorithm(TEA) of Brezinski[1]. In a recent paper [9], Sidi et al. improved the MPE and RRE methods and obtained a modified minimal polynomial extrapolation method(MMPE). Convergence and stability analyses of MPE, RRE, MMPE and TEA are discussed in [7] and [9] respectively. In [14] a rational acceleration method using vector Padé approximation is derived. This method also has the following important properties:

(1) It accelerates the convergence of a slowly converging vector sequence and makes a diverging sequence converge to an "anti–limit" which will be defined in the next section.

(2) It depends only upon the given vector sequence whose convergence is being accelerated; it does not depend on how the vector sequence is generated.

(3) It can use partial components of the vectors to accelerate convergence of the whole vectors.

We call this new method vector Padé approximation method(VPA). In [14] we also gave some properties of VPA, and obtained an algorithm quite similar to the H-algorithm[7]. In this paper, we first introduce vector Padé approximation and an associated acceleration method (i.e., VPA) in **2**. From the viewpoint of Shanks' transformation[6] we will explain the relation between VPA and MMPE. Adopting

* Received April 27, 1992.

the same technique as in [9], we will analyze the convergence and stability properties of VPA in **3** and **4** respectively. It is easy to see that our conclusions and even the remarks are quite similar to those of [7] and [9]. All the results show that VPA is a bona fide convergence acceleration method.

2. Notations and description of VPA method

Let C^p be the p-dimensional linear complex space, and the inner product $(\cdot,\cdot)$ and norm $\|\cdot\|$ be defined as usual.

In order to define vector Padé approximation, we introduce the following notations:

$$H_k := \{p(z) : p(z) = \sum_{i=0}^{k} a_i z^i, a_i \in C\},$$

$$E_k := \{e(z) : e(z) = \sum_{i=k+1}^{\infty} a_i z^i, a_i \in C\},$$

$$Z_+^p := \{\vec{l} : \vec{l} = (l_1, \cdots, l_p)^T, \quad l_i \in Z_+, \quad i = 1, \cdots, p\},$$

where p is a given positive integer, and Z_+ is the set of all nonnegative integers. Define $|\vec{l}\,| = \sum_{i=1}^{p} l_i$, for $\vec{l} \in Z_+^p$,

$$H_n^p := (H_n, \cdots, H_n)^T, \quad E_{\vec{w}}^p := (E_{w_1}, \cdots, E_{w_p})^T.$$

If $g(z) = \sum_{i=0}^{\infty} c_i z^i, c_i \in C$, we denote

$$T_{m,n}^l(g) = \begin{bmatrix} c_l & c_{l-1} & \cdots & c_{l-n+1} \\ c_{l+1} & c_l & \cdots & c_{l-n+2} \\ \cdots & \cdots & \cdots & \cdots \\ c_{l+m-1} & c_{l+m} & \cdots & c_{l+m-n} \end{bmatrix}. \tag{2.1}$$

Here we define $c_i = 0$, if $i < 0$.

Definition. *Let $f(z) = \sum_{i=0}^{\infty} c_i z^i, c_i \in C^p$ be a given power series, $n \in Z_+$ and $\vec{w} = (w_1, \cdots, w_p)^T \in Z_+^p$ be a given integer vector such that*

$$\vec{e} := \vec{w} - n := (w_1 - n, \cdots, w_p - n)^T \in Z_+^p, \qquad |\vec{w}| = p \cdot n + k$$

where k is a given integer. If we can find a vector polynomial $N(z) \in H_n^p$ and a scalar polynomial $M(z) \in H_k$ such that

$$f(z)M(z) - N(z) \in E_{\vec{w}}^p \quad \text{and} \quad M(0) = 1\ , \tag{2.2}$$

then we call $N(z)M(z)^{-1}$ the $[n, k, \vec{w}]$ vector Padé approximation of f. We denote it as $[n, k, \vec{w}]_f$.

If $H(n, k, \vec{w})$ is nonsingular, then we have (see [13], [14])

$$M(z) = \frac{1}{\det\ H(n, k, \vec{w})} \begin{vmatrix} 1 & z & \cdots & z^k \\ B(n, k, \vec{w}) & & H(n, k, \vec{w}) & \end{vmatrix}, \tag{2.3}$$

$$N(z) = \frac{1}{\det H(n, k, \vec{w})} \begin{vmatrix} f^{(n)} & z f^{(n-1)} & \cdots & z^k f^{(n-k)} \\ B(n, k, \vec{w}) & & H(n, k, \vec{w}) & \end{vmatrix}, \tag{2.4}$$

where for $i = 1, \cdots, p$

$$H(n,k,\vec{w}) = \begin{bmatrix} T^n_{e_1,k}(f_1) \\ \vdots \\ T^n_{e_p,k}(f_p) \end{bmatrix}, \quad B(n,k,\vec{w}) = \begin{bmatrix} T^{n+1}_{e_1,1}(f_1) \\ \vdots \\ T^{n+1}_{e_p,1}(f_p) \end{bmatrix},$$

$$f^{(n)}(z) = [f_1^{(n)}(z), \cdots, f_p^{(n)}(z)]^T, \qquad f_i^{(l)}(z) = \begin{cases} \sum_{j=0}^{l} (c_j, q_i) z^j, & \text{for} \quad l \geq 0, \\ 0, & \text{for} \quad l < 0 . \end{cases}$$

Here $q_i (i = 1, \cdots, p)$ is the i-th unit coordinate vector.

Now we will consider a sequence of vectors, x_m, $m = 0, 1, \cdots$, in C^p, satisfying

$$x_m \sim s + \sum_{i=1}^{\infty} v_i \lambda_i^m \quad \text{as} \quad m \to \infty \tag{2.5}$$

where s and v_i are vectors in C^p, and λ_i are scalars, independent of m, $\lambda_i \neq 1, i = 1, 2, \cdots$, $\lambda_i \neq \lambda_j$ for all $i \neq j$, and $|\lambda_1| \geq |\lambda_2| \geq \cdots$. We also assume that there can be at most p λ_i having the same modolus. An example of (2.5) is described in [8]. It is easy to see that, if $|\lambda_1| < 1$, $\{x_m\}$ converges to s; otherwise, it diverges, in which case we call s the anti-limit of the sequence $\{x_m\}$.

For the given vector sequence $\{x_m\}$, we construct $f(z)$ by the following formal power series

$$f(z) = \sum_{i=0}^{\infty} c_i z^i, \quad c_i \in C^p , \tag{2.6}$$

where

$$c_0 = x_0, \quad c_i = \Delta x_{i-1} = x_i - x_{i-1}, \quad i = 1, 2, \cdots . \tag{2.7}$$

Then by the determinant expression (2.3) and (2.4), we define the acceleration formula as

$$s_{n,k,\vec{w}} = [n+k, k, \vec{w}]_f(1) = \frac{\begin{vmatrix} x_n & x_{n+1} & \cdots & x_{n+k} \\ \Delta x_{n,\vec{e}} & \Delta x_{n+1,\vec{e}} & \cdots & \Delta x_{n+k,\vec{e}} \end{vmatrix}}{\begin{vmatrix} 1 & 1 & \cdots & 1 \\ \Delta x_{n,\vec{e}} & \Delta x_{n+1,\vec{e}} & \cdots & \Delta x_{n+k,\vec{e}} \end{vmatrix}}$$

where

$$\vec{e} = (e_1, \cdots, e_p)^T = \vec{w} - (n+k) ,$$

$$x_{n,\vec{e}} = [(x_n, q_1), \cdots, (x_{n+e_1-1}, q_1), \cdots, (x_n, q_p), \cdots, (x_{n+e_p-1}, q_p)]^T ,$$

$$\Delta x_{n,\vec{e}} = x_{n+1,\vec{e}} - x_{n,\vec{e}}$$

and

$$\vec{w} - (n+k) \geq 0, \quad |\vec{w}| = p \cdot (n+k) + k .$$

It has been pointed out in [14] that (2.8) is an acceleration formula that is a generalization of the Henrici transformation [5]. In order to implement this VPA method, an algorithm similar to the H-algorithm [5] is provided in [14]. We will not state it here.

For the simplicity of discussing the convergence and stability properties of VPA, we give another expression of $s_{n,k,\vec{w}}$ here. Let the vector sequence be generated as in (2.5). We write $u_i = \Delta x_i = x_{i+1} - x_i$, $i = 0, 1, \cdots$. Without loss of generality we will assume that $\lambda_i \neq 0$, and $v_i \neq 0$ for all $i \geq 1$. Then

$$u_m \sim \sum_{i=1}^{\infty} z_i \lambda_i^m \quad \text{as} \quad m \to \infty, \tag{2.9}$$

where $z_i = (\lambda_i - 1)v_i, i = 1, 2, \cdots$. Denote $u_{i,j} = (u_i, q_j)$, $z_{i,j} = (z_i, q_j)$, $i \geq 1$, $1 \leq j \leq p$. Then

$$u_{m,j} \sim \sum_{i=1}^{\infty} z_{i,j} \lambda_i^m \quad \text{as} \quad m \to \infty. \tag{2.10}$$

With the above notations we define $D(\sigma_0, \sigma_1, \cdots, \sigma_k)$ to be the determinant

$$D(\sigma_0, \sigma_1, \cdots, \sigma_k) = \begin{vmatrix} \sigma_0 & \sigma_1 & \cdots & \sigma_k \\ u_{n,1} & u_{n+1,1} & \cdots & u_{n+k,1} \\ u_{n,2} & u_{n+1,2} & \cdots & u_{n+k,2} \\ \vdots & \vdots & & \vdots \\ u_{n,k} & u_{n+1,k} & \cdots & u_{n+k,k} \end{vmatrix}, \tag{2.11}$$

where σ_i are scalars. Let N_i be the cofactor of σ_i in the first row expansion of this determinant. Then

$$D(\sigma_0, \sigma_1, \cdots, \sigma_k) = \sum_{i=0}^{k} \sigma_i N_i\ . \tag{2.12}$$

When σ_i are vectors, we again let $D(\sigma_0, \sigma_1, \cdots, \sigma_k)$ be defined by the determinant in (2.11), and take (2.12) as the interpretation of this determinant.

Obviously,

$$s_{n,k,\vec{w}} = \frac{D(x_n, x_{n+1}, \cdots, x_{n+k})}{D(1, 1, \cdots, 1)}. \tag{2.13}$$

It is easy to prove that

$$s_{n,k,\vec{w}} - s = \frac{D(x_n - s, x_{n+1} - s, \cdots, x_{n+k} - s)}{D(1, 1, \cdots, 1)}. \tag{2.14}$$

Let

$$\gamma_i = \frac{N_i}{D(1, 1, \cdots, 1)} = \frac{N_i}{\sum_{j=0}^{k} N_j}, \quad 0 \leq i \leq k. \tag{2.15}$$

Then we can also write $s_{n,k,\vec{w}}$ as

$$s_{n,k,\vec{w}} = \sum_{i=0}^{k} \gamma_i x_{n+i}. \tag{2.16}$$

As [9] does, we consider the overdetermined (and in general inconsistent) linear system

$$\sum_{i=0}^{k-1} c_i \Delta x_{m+i} = -\Delta x_{m+k}, \quad n \leq m \leq n + k - 1 \tag{2.17}$$

for c_i. Once $c_i, i = 0, 1, \cdots, k-1$, have been determined "in some sense", set

$$c_k = 1, \quad \gamma_i = \frac{c_i}{\sum_{j=0}^{k} c_j}, \quad 0 \le i \le k, \tag{2.18}$$

provided that $\sum_{j=0}^{k} c_j \neq 0$. Then $s_{n,k,\vec{w}}$ can be computed by (2.16).

Now the problem is how to decide on $c_i, i = 0, 1, \cdots, k-1$. MMPE considers only one of the equations in (2.17), namely that with $m = n$, and obtains c_i by solving the system of k equations

$$\sum_{i=0}^{k-1} c_i Q_j(\Delta x_{n+i}) = -Q_j(\Delta x_{n+k}), \quad j = 1, \cdots, k, \tag{2.19}$$

where Q_j are linearly independent bounded functionals. For VPA method, if we define linear functionals $Q_j (j = 1, \cdots, p)$ by

$$Q_j(y) = (y, q_j), \quad y \in C^p, \tag{2.20}$$

then we determine $c_i (i = 0, 1, \cdots, k-1)$ by the following equations:

$$\sum_{i=0}^{k-1} c_i Q_j(\Delta x_{n+l+i}) = -Q_j(\Delta x_{n+l+k}), \quad l = 1, \cdots, e_j; \ \ j = 1, \cdots, p\,. \tag{2.21}$$

Clearly, if $e_1 = e_2 = \cdots = e_p = 1$, the systems (2.19) and (2.21) are the same if the linear functionals Q_j $(j = 1, \cdots, p)$ in (2.19) are defined by (2.20). Therefore MMPE and VPA are related. However, they do not include each other in general.

3. Convergence analysis of VPA

In the following we shall assume that $k \le p$, $e := \vec{w} - (n+k) \le 1$ which means $0 \le e_j \le 1$ for all $j = 1, \cdots, p$. Without loss of generality, we will assume that $e_1 = \cdots = e_k = 1$, $e_{k+1} = \cdots = e_p = 0$. Moreover, later we will always assume that the vector sequence $\{x_m\}$ is generated as in (2.5). Before stating the main theorem, we first introduce a lemma given in [9], which will be used many times in this paper.

Lemma 3.1. *Let $i_0, i_1, \cdots, i_k$ be integers greater than or equal to 1, and assume that the scalars $v_{i_0,i_1,\cdots,i_k}$ are odd under an interchange of any two indices of $i_0, i_1, \cdots, i_k$. Let $\sigma_i, i \ge 1$, be scalars (or vectors), and let $t_{ij}, i \ge 1, 1 \le j \le k$, be scalars. Define*

$$I_{k,N} = \sum_{i_0=1}^{N} \cdots \sum_{i_k=1}^{N} \sigma_{i_0} \left(\prod_{j=0}^{k} t_{i_j,j} \right) v_{i_0,\cdots,i_k} \tag{3.1}$$

and

$$J_{k,N} = \sum_{1\le i_0<i_1<\cdots<i_k\le N} \begin{vmatrix} \sigma_{i_0} & \sigma_{i_1} & \cdots & \sigma_{i_k} \\ t_{i_0,1} & t_{i_1,1} & \cdots & t_{i_k,1} \\ t_{i_0,2} & t_{i_1,2} & \cdots & t_{i_k,2} \\ \vdots & \vdots & & \vdots \\ t_{i_0,k} & t_{i_1,k} & \cdots & t_{i_k,k} \end{vmatrix} v_{i_0,\cdots,i_k}, \tag{3.2}$$

where the determinant in (3.2) *is to be interpreted in the same way as* $D(\sigma_0,\cdots,\sigma_k)$ *in* (2.11). *Then*

$$I_{k,N} = J_{k,N}. \tag{3.3}$$

Theorem 3.2. *Define* $(v_i, q_j) = v_{ij}, i \ge 1, 1 \le j \le p$, *and let*

$$F = \begin{vmatrix} v_{1,1} & v_{2,1} & \cdots & v_{k,1} \\ v_{1,2} & v_{2,2} & \cdots & v_{k,2} \\ \vdots & \vdots & & \vdots \\ v_{1,k} & v_{2,k} & \cdots & v_{k,k} \end{vmatrix} \neq 0 . \tag{3.4}$$

Assume that $|\lambda_i| = \cdots = |\lambda_{i+r-1}|$ *implies the vectors* $v_i, v_{i+1}, \cdots, v_{i+r-1}$ *are linearly independent, for all* $i = 1, 2, \cdots; r \le p$, *and that* λ_i *satisfy*

$$|\lambda_1| \ge \cdots \ge |\lambda_k| > |\lambda_{k+1}| \ge |\lambda_{k+2}| \ge \cdots. \tag{3.5}$$

Then, for all sufficiently large n, $D(1,1,\cdots,1) \neq 0$; *hence,* $s_{n,k,\vec{w}}$, *as given in* (2.13) *exists.*

Furthermore,

$$s_{n,k,\vec{w}} - s = \Gamma(n)\lambda_{k+1}^n[1+o(1)] \quad as \quad n \to \infty , \tag{3.6}$$

where the vector $\Gamma(n)$ *is nonzero and bounded for all sufficiently large* n. *If, in addition,* $|\lambda_{k+1}| > |\lambda_{k+2}|$, *then*

$$\Gamma(n) = \frac{1}{F} \begin{vmatrix} v_1 & v_2 & \cdots & v_{k+1} \\ z_{1,1} & z_{2,1} & \cdots & z_{k+1,1} \\ z_{1,2} & z_{2,2} & \cdots & z_{k+1,2} \\ \vdots & \vdots & & \vdots \\ z_{1,k} & z_{2,k} & \cdots & z_{k+1,k} \end{vmatrix} \prod_{i=1}^{k} \frac{(\lambda_{k+1}-\lambda_i)}{(\lambda_i - 1)^2} . \tag{3.7}$$

Proof. For simplicity of notations we shall sometimes denote $G_{n,k,\vec{w}} = D(x_n - s, x_{n+1}-s,\cdots,x_{n+k}-s)$ and $H_{n,k,\vec{w}} = D(1,\cdots,1)$.

From (2.11) and (2.10) we have

$$H_{n,k,\vec{w}} \sim \begin{vmatrix} 1 & 1 & \cdots & 1 \\ \sum_{i_1=1}^{\infty} z_{i_1,1}\lambda_{i_1}^n & \sum_{i_1=1}^{\infty} z_{i_1,1}\lambda_{i_1}^{n+1} & \cdots & \sum_{i_1=1}^{\infty} z_{i_1,1}\lambda_{i_1}^{n+k} \\ \sum_{i_2=1}^{\infty} z_{i_2,2}\lambda_{i_2}^n & \sum_{i_2=1}^{\infty} z_{i_2,2}\lambda_{i_2}^{n+1} & \cdots & \sum_{i_2=1}^{\infty} z_{i_2,2}\lambda_{i_2}^{n+k} \\ \vdots & \vdots & & \vdots \\ \sum_{i_k=1}^{\infty} z_{i_k,k}\lambda_{i_k}^n & \sum_{i_k=1}^{\infty} z_{i_k,k}\lambda_{i_k}^{n+1} & \cdots & \sum_{i_k=1}^{\infty} z_{i_k,k}\lambda_{i_k}^{n+k} \end{vmatrix} . \tag{3.8}$$

By the multilinearity property of determinant, (3.8) is equivalent to

$$H_{n,k,\vec{w}} \sim \sum_{i_1=1}^{\infty} \cdots \sum_{i_k=1}^{\infty} \left(\prod_{j=1}^{k} z_{i_j,j} \right) \left(\prod_{j=1}^{k} \lambda_{i_j}^{n} \right) V(1, \lambda_{i_1}, \cdots, \lambda_{i_k}) \,, \tag{3.9}$$

where $V(\xi_0, \xi_1, \cdots, \xi_k)$ is the Vandermonde determinant defined by

$$V(\xi_0, \xi_1, \cdots, \xi_k) = \begin{vmatrix} 1 & 1 & \cdots & 1 \\ \xi_0 & \xi_1 & \cdots & \xi_k \\ \xi_0^2 & \xi_1^2 & \cdots & \xi_k^2 \\ \vdots & \vdots & & \vdots \\ \xi_0^k & \xi_1^k & \cdots & \xi_k^k \end{vmatrix} \,. \tag{3.10}$$

Since $V(\xi_0, \xi_1, \cdots, \xi_k)$ is odd under an interchange of the indices $0, 1, \cdots, k$, by Lemma 3.1, (3.9) can be expressed as

$$H_{n,k,\vec{w}} \sim \sum_{1 \leq i_1 < \cdots < i_k}^{\infty} \begin{vmatrix} z_{i_1,1} & z_{i_2,1} & \cdots & z_{i_k,1} \\ z_{i_1,2} & z_{i_2,2} & \cdots & z_{i_k,2} \\ \vdots & \vdots & & \vdots \\ z_{i_1,k} & z_{i_2,k} & \cdots & z_{i_k,k} \end{vmatrix} \left(\prod_{j=1}^{k} \lambda_{i_j}^{n} \right) V(1, \lambda_{i_1}, \cdots, \lambda_{i_k}). \tag{3.11}$$

By (3.5), if

$$\tilde{F} = \begin{vmatrix} z_{1,1} & z_{2,1} & \cdots & z_{k,1} \\ z_{1,2} & z_{2,2} & \cdots & z_{k,2} \\ \vdots & \vdots & & \vdots \\ z_{1,k} & z_{2,k} & \cdots & z_{k,k} \end{vmatrix} , \tag{3.12}$$

then as $n \to \infty$ the most dominant term in the summation on the left side of (3.11) is that for which $i_1 = 1, \cdots, i_k = k$. Since $z_i = (\lambda_i - 1)v_i$, $i \geq 1$, we have

$$\tilde{F} = \left[\prod_{i=1}^{k} (\lambda_i - 1) \right] F. \tag{3.13}$$

Because $F \neq 0$ by assumption, $\tilde{F} \neq 0$. Thus the first part of the theorem follows, with

$$D(1, \cdots, 1) = \left[\prod_{i=0}^{k} (\lambda_i - 1) \right] F \left(\prod_{j=1}^{k} \lambda_j^{n} \right) V(1, \lambda_1, \cdots, \lambda_k)[1 + o(1)] \quad \text{as } n \to \infty \,. \tag{3.14}$$

For the proof of the second part we can proceed similarly. By (2.5), (2.10) and (2.11) we have

$$G_{n,k,\vec{w}} \sim \begin{vmatrix} \sum_{i_0=1}^{\infty} v_{i_0} \lambda_{i_0}^{n} & \sum_{i_0=1}^{\infty} v_{i_0} \lambda_{i_0}^{n+1} & \cdots & \sum_{i_0=1}^{\infty} v_{i_0} \lambda_{i_0}^{n+k} \\ \sum_{i_1=1}^{\infty} z_{i_1,1} \lambda_{i_1}^{n} & \sum_{i_1=1}^{\infty} z_{i_1,1} \lambda_{i_1}^{n+1} & \cdots & \sum_{i_1=1}^{\infty} z_{i_1,1} \lambda_{i_1}^{n+k} \\ \sum_{i_2=1}^{\infty} z_{i_2,2} \lambda_{i_2}^{n} & \sum_{i_2=1}^{\infty} z_{i_2,2} \lambda_{i_2}^{n+1} & \cdots & \sum_{i_2=1}^{\infty} z_{i_2,2} \lambda_{i_2}^{n+k} \\ \vdots & \vdots & & \vdots \\ \sum_{i_k=1}^{\infty} z_{i_k,k} \lambda_{i_k}^{n} & \sum_{i_k=1}^{\infty} z_{i_k,k} \lambda_{i_k}^{n+1} & \cdots & \sum_{i_k=1}^{\infty} z_{i_k,k} \lambda_{i_k}^{n+k} \end{vmatrix} \,. \tag{3.15}$$

Again from the multilinearity property of the determinant,

$$G_{n,k,\vec{w}} \sim \sum_{i_0=0}^{\infty}\sum_{i_1=1}^{\infty}\cdots\sum_{i_k=1}^{\infty} v_{i_0}\left(\prod_{j=1}^{k} z_{i_j,j}\right)\left(\prod_{j=1}^{k}\lambda_{i_j}^n\right)V(\lambda_{i_0},\lambda_{i_1},\cdots,\lambda_{i_k})\,. \tag{3.16}$$

By Lemma 3.1, (3.16) can be expressed as

$$G_{n,k,\vec{w}} \sim \sum_{1\le i_0<i_1<\cdots<i_k}^{\infty} \begin{vmatrix} v_{i_0} & v_{i_1} & \cdots & v_{i_k} \\ z_{i_0,1} & z_{i_1,1} & \cdots & z_{i_k,1} \\ z_{i_0,2} & z_{i_1,2} & \cdots & z_{i_k,2} \\ \vdots & \vdots & & \vdots \\ z_{i_0,k} & z_{i_1,k} & \cdots & z_{i_k,k} \end{vmatrix} \lambda_{i_0}^n\left(\prod_{j=1}^{k}\lambda_{i_j}^n\right)V(1,\lambda_{i_1},\cdots,\lambda_{i_k}). \tag{3.17}$$

By the assumption made following (2.5), there are at most p λ_i with modulus equal to $|\lambda_{k+1}|$. Let $|\lambda_{k+1}| = |\lambda_{k+2}| = \cdots = |\lambda_{k+r}|$ (clearly $r \le p$). From this and (3.5), it follows that the dominant term on the right side of (3.17) is the sum of those terms with indices $i_0 = 1, i_1 = 2, \cdots, i_{k-1} = k$, $i_k = k + l$, $l = 1, \cdots, r$. That is,

$$\begin{aligned} G_{n,k,\vec{w}} &= D(x_n - s, x_{n+1} - s, \cdots, x_{n+k} - s) \\ &= \left(\prod_{j=1}^{k}\lambda_j^n\right)\sum_{l=1}^{r}\lambda_{k+l}^n V(\lambda_1,\lambda_2,\cdots,\lambda_k,\lambda_{k+l}) \\ &\times \begin{vmatrix} v_1 & \cdots & v_k & v_{k+l} \\ z_{1,1} & \cdots & z_{k,1} & z_{k+l,1} \\ z_{1,2} & \cdots & z_{k,2} & z_{k+l,2} \\ \vdots & & \vdots & \vdots \\ z_{1,k} & \cdots & z_{k,k} & z_{k+l,k} \end{vmatrix} [1 + o(1)] \quad \text{as} \quad n \to \infty\,. \end{aligned} \tag{3.18}$$

Now the cofactor of v_{k+l} in the determinant in (3.1) is $\tilde{F}$, which is nonzero since $F \neq 0$. Thus, for n sufficiently large, the coefficients of $v_{k+1}, \cdots, v_{k+r}$ are nonzero. By the assumption $v_{k+1}, \cdots, v_{k+r}$ are linearly independent. Therefore the summation in (3.15) is never zero. Combining (3.14) and (3.18) in (2.14) results in (3.6).

If $|\lambda_{k+1}| > |\lambda_{k+2}|$, then $r = 1$. In this case, because

$$V(\xi_0, \xi_1, \cdots, \xi_k) = \prod_{0\le i<j\le k}(\xi_i - \xi_j)\,,$$

from (3.14) and (3.18), we have (3.7).

The asymptotic error analysis of VPA leads to the following important conclusions:

(1) Under the conditions stated in Theorem 3.2, VPA is a bona fide vector accelerator in the sense that

$$\frac{\|s_{n,k,\vec{w}} - s\|}{\|x_{n+k+1} - s\|} = O\left[\left(\frac{\lambda_{k+1}}{\lambda_1}\right)^n\right] \quad \text{as} \quad n \to \infty\,. \tag{3.19}$$

(2) The result in (3.6) shows that, when the VPA is applied to a vector sequence generated by using the matrix iterative method described in [8], it will be especially

effective if the iteration matrix A has a small number of large eigenvalues (the number can not be larger than k when $s_{n,k,\vec{w}}$ is used) that are well separated from the small eigenvalues.

(3) Bearing $\Gamma(n)$ in (3.6) and (3.7) in mind, we can see that a loss of accuracy will take place in $s_{n,k,\vec{w}}$ when $\lambda_1, \cdots, \lambda_k$ are close to 1, since $\Gamma(n)$ becomes large in this case. When the vector sequences are obtained by solving a linear system of equations $(I - A)X = b$ by an iterative technique, which means if A has large eigenvalues near 1, there will be a loss of accuracy in $s_{n,k,\vec{w}}$. In fact, eigenvalues near 1 would cause the system to be nearly singular.

4. Stability analysis of VPA

With the assumption at the beginning of the last section, we will consider stability of VPA for vector sequences generated by (2.5). Let us denote γ_i in (2.15) by $\gamma_i^{(n,k)}$. Then the propagation of errors introduced in $\{x_m\}$ will be controlled, to some extent, by $\sum_{j=0}^{k} |\gamma_j^{(n,k)}|$; the larger this quantity, the worse the error propagation is expected to be. With this in mind, we say that $s_{n,k,\vec{w}}$ is asymptotically stable if

$$\sup_n \sum_{j=0}^{k} |\gamma_j^{(n,k)}| < \infty . \tag{4.1}$$

Since $\sum_{j=0}^{k} \gamma_j^{(n,k)} = 1$ by (2.15), it follows that $\sum_{j=0}^{k} |\gamma_j^{(n,k)}| \geq 1$. Thus the most ideal situation is that $\gamma_j^{(n,k)} \geq 0$, $0 \leq j \leq k$, for n sufficiently large.

Theorem 4.1. *Under the conditions of Theorem 3.2, $s_{n,k,\vec{w}}$ is asymptotically stable.*

Proof. We need only to show that $\gamma_j^{(n,k)}, 0 \leq j \leq k$, remains bounded for $n \to \infty$.

From (2.11), we have

$$N_j = (-1)^j \begin{vmatrix} u_{n,1} & \cdots & u_{n+j-1,1} & u_{n+j,1} & \cdots & u_{n+k,1} \\ u_{n,2} & \cdots & u_{n+j-1,2} & u_{n+j,2} & \cdots & u_{n+k,2} \\ \vdots & & \vdots & \vdots & & \vdots \\ u_{n,k} & \cdots & u_{n+j-1,k} & u_{n+j,k} & \cdots & u_{n+k,k} \end{vmatrix} . \tag{4.2}$$

Substituting the asymptotic expansion of $u_{m,j}$ as given in (2.10), and utilizing the multilinearity property of determinants, we obtain

$$N_j \sim \sum_{i_1=1}^{\infty} \cdots \sum_{i_k=1}^{\infty} \prod_{j=1}^{k} \lambda_{i_j}^n \prod_{j=1}^{k} z_{i_j,l} C_j(\lambda_{i_1}, \cdots, \lambda_{i_k}) \tag{4.3}$$

where

$$C_j(\xi_1, \cdots, \xi_k) = (-1)^j \begin{vmatrix} 1 & \xi_1 & \cdots & \xi_1^{j-1} & \xi_1^{j+1} & \cdots & \xi_1^k \\ 1 & \xi_2 & \cdots & \xi_2^{j-1} & \xi_2^{j+1} & \cdots & \xi_2^k \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 1 & \xi_k & \cdots & \xi_k^{j-1} & \xi_k^{j+1} & \cdots & \xi_k^k \end{vmatrix} . \tag{4.4}$$

Using Lemma 3.1 again, we have

$$N_j \sim \sum_{1\le i_1<\cdots<i_k}^{\infty} \begin{vmatrix} z_{i_1,1} & z_{i_2,1} & \cdots & z_{i_k,1} \\ z_{i_1,2} & z_{i_2,2} & \cdots & z_{i_k,2} \\ \vdots & \vdots & & \vdots \\ z_{i_1,k} & z_{i_2,k} & \cdots & z_{i_k,k} \end{vmatrix} \left(\prod_{j=1}^{k} \lambda_{i_j}^n\right) C_j(\lambda_{i_1},\cdots,\lambda_{i_k}) . \tag{4.5}$$

By (3.4), (3.12), (3.13) and (3.15),

$$N_j = \left(\prod_{j=1}^{k} \lambda_j^n\right) \tilde{F}[C_j(\lambda_1,\cdots,\lambda_k) + o(1)] \quad \text{as} \quad n \to \infty . \tag{4.6}$$

Combining (4.6) and (3.14) in (2.15), and using (3.13), we obtain

$$\gamma_j^{(n,k)} = \frac{C_j(\lambda_1,\cdots,\lambda_k)}{V(1,\lambda_1,\cdots,\lambda_k)} + o(1) \quad \text{as} \quad n \to \infty . \tag{4.7}$$

Obviously, (4.7) implies $|\gamma_j^{(n,k)}| < \infty$ for n sufficiently large. Thus (4.1) holds.

Remark. If $\lambda_i, 1 \le i \le k$, are real and negative, then $\gamma_j^{(n,k)} > 0, 0 \le j \le k$, for n sufficiently large. Detailed discussions are given in [9].

5. Conclusions

In this paper, we have only discussed convergence acceleration of vector sequences in the finite dimensional linear normed space C^p. However, C^p can be replaced by any general linear normed space B, since our vector accelerator (2.13) still holds if we modify $u_{m,j}$ in (2.10) by appropriate linear independent bounded functionals on the space B. Sidi[9] has dealt with this case for MMPE.

References

[1] C. Brezinski, Accélération de la convergence en analyse numérique, Lect. Notes in Math. **584**, Springer–Verlag, Berlin, 1977.

[2] S. Cabay and L. W. Jackson, A polynomial extrapolation method for finding limits and antilimits of vector sequences, *SIAM J. Numer. Anal.*, **13** (1976), 734–752.

[3] R. P. Eddy, Extrapolation to the limit of a vector sequence, in Information Linkage between Applied Mathematics and Industry, P. C. C. Wang, ed., Academic Press, New York, 1979, 287–396.

[4] M. Mešina, Convergence acceleration for the iterative solution of the equations $X = AX + f$, *Comp. Meth. Appl. Mech. Eng.*, **10** (1977), 165–173.

[5] H. Sadok, About Henrici's transformation for accelerating vector sequences, *J. Comp. Appl. Math.*, **29** (1990), 101–110.

[6] D. Shanks, Non-linear transformations of divergent and slowly convergent sequence, *J. Math. Phys.*, **34** (1955), 1–42.

[7] A. Sidi, Convergence and stability properties of minimal polynoimal and reduced rank extrapolation algorithms, *SIAM J. Numer. Anal.*, **23** (1986), 197–209.

[8] A. Sidi and J. Bridger, Convergence and stability analyses for some vector extrapolation methods in the presence of defective iteration matrices, *J. Comp. Appl. Math.*, **22** (1988), 35–61.

[9] A. Sidi, W. F. Ford and D. A. Smith, Acceleration of convergence of vector sequences, *SIAM J. Numer. Anal.*, **23** (1986), 178–196.

[10] D.A. Smith, W.F. Ford and A. Sidi, Extrapolation methods for vector sequences, *SIAM Rev.*, **29** (1987), 199–233.

[11] P. Wynn, On a device for computing the $e_m(S_n)$ transformation, *MTAC*, **10** (1956), 91–96.

[12] P. Wynn, Acceleration techniques for iterated vector and matrix problem, *Math. Comp.*, **16** (1962), 301–322.

[13] G. Xu and A. Bultheel, Matrix Padé approximation: Definitions and properties, *Lin. Alg. Appl.*, **137/138** (1990), 67–136.

[14] G. Xu and G. Zhuang, Convergence acceleration of sequences by vector Padé approximation , 1991, submitted to *J. Comp. Math.*