CONVERGENCE AND STABILITY ANALYSES FOR SOME VECTOR EXTRAPOLATION METHODS IN THE PRESENCE OF DEFECTIVE ITERATION MATRICES

by

A. Sidi and J. Bridger

Technical Report #450
April 1987
CONVERGENCE AND STABILITY ANALYSES FOR SOME VECTOR EXTRAPOLATION METHODS IN THE PRESENCE OF DEFECTIVE ITERATION MATRICES

Avram Sidi  
Computer Science Department  
Technion - Israel Institute of Technology  
Haifa 32000, Israel

Jacob Bridger  
Mathematics Department  
Technion - Israel Institute of Technology  
Haifa 32000, Israel

ABSTRACT

In two previous papers [8,9] convergence and stability results for the following vector extrapolation methods were presented: Minimal Polynomial Extrapolation, Reduced Rank Extrapolation, Modified Minimal Polynomial Extrapolation, and Topological Epsilon Algorithm. The analyses were carried out for vector sequences that include those arising from iterative methods for linear systems of equations having diagonalizable iteration matrices. In this paper the analyses of [8,9] are extended to vector sequences that include those arising from iterative methods for linear systems having defective iteration matrices. The results are illustrated with numerical examples. Based on the analyses above, extensions of the well known power method are suggested enabling one to obtain estimates for several dominant eigenvalues of a general matrix.

* Work of this author was performed in partial fulfillment of the requirements for the degree of Master of Science at the Technion-Israel Institute of Technology, Haifa, Israel.
1. INTRODUCTION

Let $B$ be a normed linear space over the field of complex numbers, and denote the norm associated with $B$ by $\| \cdot \|$. In case $B$ is also an inner product space, we adopt the following convention for the homogeneity property of the inner product: For $y, z \in B$ and $\alpha, \beta$ complex numbers, the inner product $(\cdot, \cdot)$ satisfies $(\alpha y, \beta z) = \bar{\alpha} \beta (y, z)$. The norm in this case is the one induced by the inner product, i.e., if $x \in B$, $\| x \| = \sqrt{(x, x)}$.

Let $x_i, i = 0, 1, \ldots$, be a sequence in $B$. We shall assume that

$$x_m \to s + \sum_{i=1}^{\infty} P_i(m) \lambda_i^m \text{ as } m \to \infty. \quad (1.1)$$

Here $s$ is a vector in $B$ and $\lambda_i$ are scalars ordered such that

$$|\lambda_1| \geq |\lambda_2| \geq |\lambda_3| \geq \cdots, \quad (1.2)$$

and satisfying $\lambda_i \neq 0$ and $\lambda_i \neq 1$, $i = 1, 2, \ldots$, and $\lambda_i \neq \lambda_j$ if $i \neq j$. In addition, we assume that there can be only a finite number of $\lambda_i$ having the same modulus. $P_i(m)$ are polynomials in $m$ with vector-coefficients (thus $P_i(m)$ are vectors in $B$), which we write in the form

$$P_i(m) = \sum_{l=0}^{P_i} y_{il} \binom{m}{l}, \quad (1.3)$$

where $y_{il}$ are binomial coefficients and $y_{il}, i = 0, \ldots, P_i, l = 1, 2, \ldots$, form a linearly independent set of vectors. We agree to order the $\lambda_i$ such that if $|\lambda_j| = |\lambda_{j+1}|$, then $p_j \geq p_{j+1}$. The meaning of (1.1) is that for any positive integer $N$ there exist a positive constant $K$ and a positive integer $m_0$ that depend only on $N$, such that for every $m \geq m_0$,

$$\| x_m - s - \sum_{i=1}^{N-1} P_i(m) \lambda_i^m \| \leq K \lambda_1^m m^P. \quad (1.4)$$

If $|\lambda_1| < 1$, then $\lim_{m \to \infty} x_m$ exists and is simply $s$. If $|\lambda_1| \geq 1$, then $\lim_{m \to \infty} x_m$ does not exist, and $s$, in this case, is said to be the anti-limit of the sequence $x_m, m = 0, 1, \ldots$.

As will be shown in Section 2, sequences of vectors generated by iterative solution of linear systems of equations having defective iteration matrices are in general of the form described above. In fact, this has been the source of motivation for the assumptions above.
Our aim is to find a good approximation to \( s \), from a small number of terms of the sequence \( x_m, m = 0, 1, \ldots \), whether \( s \) is the limit or the anti-limit of this sequence. To this effect several vector extrapolation methods have been proposed. In a recent paper by Smith, Ford, and Sidi [10] some of these methods have been surveyed and tested numerically. The methods that have been considered in [10] are the Minimal Polynomial Extrapolation (MPE) of Cabay and Jackson [3], the Reduced Rank Extrapolation (RRE) of Eddy [4] and Mesina [7], the Scalar Epsilon Algorithm (SEA) of Wynn [12], the Vector Epsilon Algorithm (VEA) of Wynn [13], and the Topological Epsilon Algorithm (TEA) of Brezinski [1]. In yet another work by Sidi, Ford, and Smith [9] a new method designated the Modified MPE (MMPE) has been proposed.

Four of the methods above, namely, MPE, RRE, MMPE, and TEA have been analyzed in Sidi [8] and in [9] for their convergence and stability properties. Their analyses have been carried out for sequences of the form (1.1) with \( p_i = \deg P_i (m) = 0 \) for all \( i \).

For future reference we will now give a brief description of the above mentioned extrapolation methods based on the developments in [9] for MMPE and TEA and in [8] for MPE and RRE.

Below \( k \) denotes a positive integer less than or equal to the dimension of the space \( B \) and

\[
\mu_m = \Delta x_m = x_{m+1} - x_m, \quad w_m = \Delta \mu_m = \mu_{m+1} - \mu_m, \quad m = 0, 1, \ldots
\]

Also \( s_{n,k} \) denotes the approximation to \( s \) obtained by applying any of the methods above to the vector sequence \( x_m, m = 0, 1, \ldots \). Clearly, \( s_{n,k} \) will be different for each method. For each method \( s_{n,k} \) can be shown to be of the form

\[
s_{n,k} = \sum_{i=0}^{k} \gamma_i x_{n+i}
\]

subject to

\[
\sum_{i=0}^{k} \gamma_i = 1.
\]

It can also be shown that \( s_{n,k} \) has the determinant representation

\[
s_{n,k} = \frac{D(x_n, x_{n+1}, \ldots, x_{n+k})}{D(1, 1, \ldots, 1)}
\]

where
with $u_{ij}$ scalars dependent on the extrapolation method being used. If we let $N_i$ be the cofactor of $\sigma_i$ in the first row of $D(\sigma_0, \ldots, \sigma_k)$, then

$$D(\sigma_0, \ldots, \sigma_k) = \sum_{i=0}^{k} \sigma_i N_i$$

(1.9)

when $\sigma_i$ are scalars. When $\sigma_i$ are vectors, (1.9) is taken to be the interpretation of $D(\sigma_0, \ldots, \sigma_k)$. The computation of the $u_{ij}$ for the different methods is explained below.

1. For MMPE

$$u_{ij} = Q_{i+1}(u_{n+j}),$$

(1.10)

where $Q_1, \ldots, Q_k$ form a linearly independent set of bounded linear functionals over $B$.

2. For TEA

$$u_{ij} = Q(u_{n+i+j}),$$

(1.11)

where $Q$ is a bounded linear functional over $B$.

3. For MPE

$$u_{ij} = (w_{n+i}, w_{n+j}).$$

(1.12)

4. For RRE

$$u_{ij} = (w_{n+i}, w_{n+j}).$$

(1.13)

For all four methods the $\gamma_j$ satisfy the system of linear equations consisting of (1.6) and

$$\sum_{j=0}^{k} \gamma_j u_{ij} = 0, \quad 0 \leq i \leq k-1.$$

(1.14)

For more details the reader is referred to [8] and [9].

An extrapolation method almost identical to RRE has been proposed by Kaniel and Stein [6]. In this method
\[ s_{n,k} = \sum_{i=0}^{k} \gamma_i x_{n+i+1}, \]  
where the \( \gamma_i \) are determined exactly as for RRE. Actually, as suggested in [9], one can consider applying all the methods above in the form

\[ s_{n,k} = \sum_{i=0}^{k} \gamma_i x_{n+i+q}, \]  
where \( \gamma_i \) are determined exactly as before, i.e., \( s_{n,k} \) has the determinant representation

\[ s_{n,k} = \frac{D(x_{n+q}, x_{n+q+1}, \ldots, x_{n+q+2k})}{D(1, 1, \ldots, 1)}. \]

c.f. (1.5) and (1.7). Note that the determination of the \( \gamma_i \) for MMPE, MPE, and RRE involves \( x_n, x_{n+1}, \ldots, x_{n+2k} \). This suggests that for computational economy \( q \leq 1 \). For TEA, on the other hand, the \( \gamma_i \) are determined from \( x_n, x_{n+1}, \ldots, x_{n+2k} \), which suggests that \( q \leq k \).

In [8] and [9] it was shown, under the assumption that \( p_i = \deg P_i(m) = 0 \) and \( |\lambda_i| > |\lambda_{i+1}| \) and under additional mild assumptions, that

\[ \| s_{n,k} \| = 0 \left[ \lambda_{k+1}^n \right] \]  
as \( n \to \infty \),  
(1.18)

where \( s_{n,k} \) for all four methods is as given in (1.5). In addition, it was shown that the methods are asymptotically stable in the sense that

\[ \sup_n \sum_{i=0}^{k} |\gamma_i(s, k)| < \infty \]  
(1.19)

(we have denoted the \( \gamma_i \) by \( \gamma_i(s, k) \) to show their dependence on \( n \) and \( k \)). In fact, it was shown that

\[ \lim_{n \to \infty} \sum_{i=0}^{k} \gamma_i(s, k) \lambda_i = \prod_{i=1}^{k} \frac{\lambda_i - \lambda_i}{1 - \lambda_i} \]  
(1.20)

In Section 3 of this work we state the extensions of the results (1.18)-(1.20) to the case of arbitrary \( p_i \). The proofs are carried out in Section 5. In Section 4 we illustrate the results of Section 3 with numerical examples. Finally, in Section 6 we propose extensions of the well known power method that enable one to obtain estimates for several dominant eigenvalues of a general matrix.
2. EXAMPLE: LINEAR ITERATIVE METHODS WITH DEFECTIVE MATRICES

Let us consider a vector sequence generated by a matrix iterative technique used in solving the linear system of equations

\[ x = Ax + b, \]

where \( A \) is a general and possibly defective \( M \times M \) (complex) matrix and \( b \) and \( x \) are \( M \) dimensional (complex) column vectors. We note that if \( \lambda = 1 \) is an eigenvalue of \( A \), then the system (2.1) does not have a unique solution. Thus, we assume that all eigenvalues of \( A \) are different than 1. Under this assumption, the unique solution vector \( s \) satisfies

\[ s = As + b. \]

(2.2)

For a given vector \( x_0 \), we generate the vectors \( x_m \) by

\[ x_{m+1} = Ax_m + b, \quad m = 0, 1, \ldots \]

(2.3)

From (2.2) and (2.3) we obtain

\[ x_m - s = A^m (x_0 - s). \]

(2.4)

For any \( M \times M \) matrix \( A \) we can find a non-singular matrix \( V \) such that

\[ V^{-1} AV = J = \begin{bmatrix} J_1 & J_2 \\ & \ddots & \ddots \\ & & J_v \end{bmatrix}, \]

(2.5)

where the Jordan blocks \( J_i \) are of dimension \( r_i \) and have the form

\[ J_i = \begin{bmatrix} \lambda_i & 1 \\ & \ddots & \ddots \\ & & \lambda_i \end{bmatrix}, \quad \lambda_i \text{ eigenvalue.} \]

(2.6)

It can be shown (see Varga [11]) that
where by convention \([m_j] = 0\) if \(j > m\) or \(j < 0\). If we denote the columns of the matrix \(V\) by

\[ v_{11}, v_{12}, \ldots, v_{1r_1}, v_{21}, v_{22}, \ldots, v_{2r_2}, \ldots, v_{s1}, v_{s2}, \ldots, v_{sr}\],

then \(v_{j1}\) is the eigenvector corresponding to the eigenvalue \(\lambda_j\) and \(v_{ji}, i = 2, \ldots, r_j\), are the principal vectors corresponding to the same eigenvalue. As is known, the set of the eigenvectors and principal vectors is linearly independent and forms a basis for \(\mathbb{C}^M\).

For the initial error vector there exist scalars \(a_{ji}\) such that

\[ x_0 = \sum_{j=1}^{w} \sum_{i=1}^{p_j} a_{ji} v_{ji}. \]  

Here we have assumed that Jordan blocks that do not contribute to (2.8) have been removed, and the remaining blocks renumbered.

Define \(p_j\) to be the largest nonnegative integer for which \(a_{j,p_j+1} \neq 0\). Consequently, \(p_j+1 \leq r_j\). By (2.4), (2.5), (2.7), (2.8) and the fact that \(A^mV = VJ^m\), it follows that

\[ x_{m-s} = \sum_{j=1}^{w} \sum_{i=1}^{p_j+1} a_{ji} \sum_{l=1}^{m} \lambda_j^{m-i+l} v_{ji}. \]  

We first observe that if 0 is an eigenvalue of \(A\), then for \(m\) sufficiently large, this eigenvalue does not contribute to the triple sum in (2.9). Thus it can be assumed that \(\lambda_j \neq 0\) for all \(j\).

Changing the index \(l\) to \(q = q-l\), and interchanging the summations over \(i\) and \(q\), (2.9) becomes

\[ x_{m-s} = \sum_{j=1}^{w} \left[ \sum_{q=0}^{p_j} y_{jq} \left( \frac{m}{q} \right) \right] \lambda_j^m, \]  

where
\[
y_{jq} = \sum_{i=q+1}^{p_{j+1}} a_{ji} v_{j,i-1} \lambda_j^{-q}, \quad 0 \leq q \leq p_j. \tag{2.11}
\]

By the fact that \(a_{ji} v_{j,i+1} \neq 0\) and the linear independence of the set of vectors \(v_{ji}, 1 \leq i \leq p_{j+1}\), it follows that \(y_{jq}, 0 \leq q \leq p_j\), form a linearly independent set as well. Thus, the set of vectors \(y_{jq}, 0 \leq q \leq p_j, 1 \leq j \leq \nu\), is linearly independent.

Assume now that there are several Jordan blocks that have the same eigenvalue. For the sake of argument suppose that \(\lambda = \lambda_1 = \lambda_2 = \cdots = \lambda_N \neq \lambda_{N+1}\), and that \(p = p_1 \geq p_2 \geq \cdots \geq p_N\). Then the part of the double sum in (2.10) having \(j = 1, 2, \ldots, N\), can be expressed as

\[
\sum_{j=1}^{N} \sum_{q=0}^{p_j} y_{jq} \left( \begin{array}{c} m \\ q \end{array} \right) \lambda_j^m = \sum_{q=0}^{p} \sum_{j=q}^{N} y_{jq} \left( \begin{array}{c} m \\ q \end{array} \right) \lambda_j^m, \tag{2.12}
\]

where

\[
y_q = \sum_{j=1}^{N} y_{jq}, \quad 0 \leq q \leq p. \tag{2.13}
\]

and \(y_{jq} = 0\) when \(q > p_j\). It is easy to see that the vectors \(y_q, 0 \leq q \leq p\), form a linearly independent set.

We thus have shown that Jordan blocks having the same eigenvalue can be combined into a single block, enabling us to assume that all \(\lambda_j\) in (2.10) are distinct and that the \(y_{jq}\) form a linearly independent set.

In summary, we have shown that if the vector sequence \(x_m, m=0,1,2,\ldots\), is generated by the matrix iterative process (2.3) and if \((I-A)^{-1}\) exists, then it automatically obeys (1.1) in conjunction with all the conditions imposed on the scalars \(\lambda_j\) and the vectors \(y_{jq}\).

3. STATEMENT OF CONVERGENCE AND STABILITY RESULTS

In accordance with the assumptions of Section 1, let the positive integers \(t\) and \(r\) be such that

\[
|\lambda_1| > |\lambda_{t+1}| > \cdots > |\lambda_{r+1}| > |\lambda_{t+r+1}|. \tag{3.1}
\]

Now from (3.1) and the ordering \(p_{t+1} \geq \cdots \geq p_{t+r}\), it follows that there is a greatest integer \(r' \leq r\), for which

\[
p_{t+1} = \cdots = p_{t+r}. \tag{3.2}
\]

Obviously, \(r' = r\) when \(r = 1\) or \(p_{t+r} = p_{t+1}\). Let
Theorem 3.1: If \( s_{n,k} \) is as given in (1.16) (equivalently (1.17)), then

\[
k = \sum_{j=1}^{t} (p_j+1). \tag{3.3}
\]

where

\[
s_{n,k} = \sup \| \Gamma(n) \| < \infty, \tag{3.4}
\]

provided

\[
\begin{bmatrix}
Q_1(y_{10}) & Q_1(y_{1p}) & \cdots & Q_1(y_{pt}) \\
Q_2(y_{10}) & Q_2(y_{1p}) & \cdots & Q_2(y_{pt}) \\
\vdots & \vdots & \ddots & \vdots \\
Q_k(y_{10}) & Q_k(y_{1p}) & \cdots & Q_k(y_{pt})
\end{bmatrix} \neq 0 \tag{3.5}
\]

for MMPE, or

\[
\prod_{j=1}^{t} Q_j(y_{jp}) \neq 0 \tag{3.6}
\]

for TEA. For MPE and RRE there are no additional restrictions.

The reason that MPE and RRE need no additional restrictions arises from the fact that the Gram determinant of the vectors \( y_{jp}, 0 \leq p \leq p_j, 1 \leq j \leq t \), is nonzero, which follows from their linear independence.

The vector \( \Gamma(n) \) for each method has the asymptotic form

\[
\Gamma(n) = \sum_{j=1}^{t} \Gamma_j e^{i \text{arg} \lambda_{j+1}} + o(1) \text{ as } n \to \infty,
\]

where \( \Gamma_j \) are vectors independent of \( n \) and dependent on the method used. If we denote

\[
\mu = \max \left| \begin{array}{c}
\lambda_{j+1} \\
\lambda_j \\
\vdots \\
\lambda_{t+1}
\end{array} \right| /
\]

then the \( o(1) \) term in the expression for \( \Gamma(n) \) is actually \( O(\mu^t) \) if \( p_{t+1} = 0 \) and \( O(n^{-1}) \) otherwise. We do not dwell on this point in the proofs given in Section 5, but they can be obtained easily from the proofs themselves.
For MPE and RRE the vectors \( \Gamma_j, 1 \leq j \leq r' \), turn out to be identical. For complete details see Section 5.

**Theorem 3.2:** If \( s_{n,k} \) is as given in (1.16) (equivalently (1.17)), then, under the conditions of Theorem 3.1, all four methods are asymptotically stable in the sense that

\[
\sup_n \sum_{i=0}^{k} | \gamma^{(n,k)}_i | < \infty. 
\]

(3.8)

In fact,

\[
\lim_{n \to \infty} \sum_{i=0}^{k} \gamma^{(n,k)}_i \lambda_i^j = \prod_{i=1}^{t} \left( \frac{\lambda_i - \lambda_i^j}{1 - \lambda_i} \right)^{p_{i+1}} 
\]

(3.9)

and (3.8) is a consequence of (3.9).

Actually, it is true that \( \gamma^{(n,k)}_i = \delta_i + O(1/\lambda_{t+1}^j \lambda_i^n) \) as \( n \to \infty \), where \( \sum_{i=0}^{k} \delta_i \lambda_i = \prod_{i=1}^{t} \left( \frac{\lambda_i - \lambda_i^j}{1 - \lambda_i} \right)^{p_{i+1}} \)

We do not give the details of the proof of this statement, although they can be extracted from the proofs of Section 5.

**Remarks**

1) As can be seen from (3.4) and (3.5), all four methods are bona fide acceleration methods in the sense that

\[
\| s_{n,k} - s \| = 0 \left[ \frac{\lambda_{t+1}}{\lambda_1} \right]^n \] as \( n \to \infty \). 

(3.10)

Consequently, if \( \lambda_{t+1} \) is well separated from \( \lambda_1 \), and \( | \lambda_{t+1} | < 1 \), then \( s_{n,k} \) converges to \( s \) much more quickly than \( x_n \) itself as \( n \to \infty \), irrespective of whether the sequence \( x_0,x_1,x_2, \ldots \) converges or diverges.

2) As is suggested by (3.4), if \( p_{t+1} > 0 \), then the quality of \( s_{n,k} \) will deteriorate initially, but will improve for increasing \( n \), provided \( | \lambda_{t+1} | < 1 \).

3) Inspection of the leading asymptotic behavior of \( \Gamma(n) \) for the different methods shows that if some of \( \lambda_1, \ldots, \lambda_t \) are close to 1, then \( \| \Gamma(n) \| \) is large; and this has an adverse effect on the accuracy of
When the sequence $x_0, x_1, x_2, \ldots$ is obtained from iterative solution of a linear system of equations, the closeness of some of $\lambda_1, \ldots, \lambda_q$, to 1 means that the matrix of the system is nearly singular.

4) When some of $\lambda_1, \ldots, \lambda_q$ are close to 1, Theorem 3.2 implies that the $\gamma_i^{(n,k)}$ will be large in modulus although $\sum_{i=0}^{k} \gamma_i^{(n,k)} = 1$. This causes $\sum_{i=0}^{k} |\gamma_i^{(n,k)}|$ to be very large, which, in turn, causes errors in the $x_i$ to be magnified severely.

The conclusions above are identical to those derived in [8,9] for the case $p_i = 0$, $i = 1, 2, \ldots$.

4. NUMERICAL EXAMPLES

In this section we illustrate the convergence results of Section 3 for MPE and MMPE with two examples. In both of the examples the space $B$ is the Euclidean space dimension $M=12$, and the sequence $x_0, x_1, \ldots$, is obtained by the matrix iterative method described in the first paragraph of Section 2 with the notation therein. To simplify matters the solution $s$ to the system (2.1) is taken to be $s = (1,1,\ldots,1)^T$ and $b$ is determined by $b = s - At$. The initial vector $x_0$ is taken to be zero. The matrix $A$ is defective and is determined from another defective matrix $J$ that has a simple form by the similarity transformation

$$ A = W^{-1} J W, $$

where the matrix $W = (w_{ij})$ for both examples is given by

$$ w_{ij} = \begin{cases} i+j-1 & \text{if } i \neq j \\ 100 + 10j & \text{if } i = j \end{cases} $$

As such, $W$ is strictly diagonally dominant so that $W^{-1}$ can be computed numerically to very high accuracy.

Example 1: The matrix $J$ is given as the block diagonal matrix

$$ J = \begin{bmatrix} C_1 & 0 \\ 0 & C_2 \\ 0 & C_3 \end{bmatrix} $$

where
Each of the matrices $C_1, C_2,$ and $C_3$ is defective. $C_1$ is a Frobenius matrix with eigenvalues $\lambda_1 = -1 + 0.6i$ and $\lambda_2 = \bar{\lambda}_1, \bar{\lambda}_1,$ and $\lambda_2$ having algebraic multiplicity 2 and geometric multiplicity 1. $C_2$ is a Jordan matrix with eigenvalue $\lambda_2 = 0.6$ having algebraic multiplicity 4 and geometric multiplicity 1. Similarly, $C_3$ is a Jordan matrix with eigenvalue $\lambda_4 = 0.1$ having algebraic multiplicity 4 and geometric multiplicity 1. Combining this information, we have, for an arbitrary initial vector $x_0$, the expansion

$$x_m = s + \sum_{i=1}^{4} P_i(m) \lambda_i^m,$$

with $\lambda_i$ as above and $P_1 = P_2 = 1$ and $P_3 = P_4 = 3$.

In Figures 1 and 2 we give the results of the computations for $|| s_n, k \rightarrow s ||_\infty$ using both MPE and MMPE with $k=4$ and $k=8$, respectively. We do not include $|| x_n - s ||_\infty$ as the sequence $x_0, x_1, \ldots$, diverges by $p(A) = |\lambda_2| > 1$. According to the theory of Section 3 we should have

$$-\log_{10} || s_n, 4 \rightarrow s ||_\infty = (-\log_{10} 0.6)n - 3 \log_{10} n + O(1) \text{ as } n \rightarrow \infty,$$

and

$$-\log_{10} || s_n, 8 \rightarrow s ||_\infty = (-\log_{10} 0.1)n - 3 \log_{10} n + O(1) \text{ as } n \rightarrow \infty.$$

These results are indeed born out by Figures 1 and 2.

Example 2: The matrix $J$, this time is given as the block diagonal matrix

$$J = \begin{bmatrix} C_1' & 0 & 0 \\ 0 & C_2' & 0 \\ 0 & 0 & C_3' \end{bmatrix}$$

where $C_1$ is exactly as in (4.4) and
\( C_2 = \begin{bmatrix} 0.6 & 1 \\ 0 & 0.6 \end{bmatrix}, \quad C_3 = \begin{bmatrix} 0.1 & 1 \\ 0 & 0.1 \end{bmatrix} \)  
(4.10)

Again the eigenvalues are \( \lambda_1 = -1+0.6i, \lambda_2 = \lambda_1, \lambda_3 = 0.6, \) and \( \lambda_4 = 0.1, \) and, for an arbitrary initial vector \( x_0, x_m \) has an expansion of the form given in (4.6) with \( p_1 = p_2 = 1, \) as in Example 1 but with \( p_3 = p_4 = 1 \) unlike Example 1, as explained at the end of Section 2.

In Figures 3 and 4 we give the results of the computations for \( \| s_{x,4} \| \to \infty \) using MPE and MMPE with \( k=4 \) and \( k=6, \) respectively. We again do not include \( \| x_n \| \to \infty \) as the sequence \( x_0, x_1, \ldots \) diverges by \( \rho(A) = |\lambda_1| > 1. \) This time we should have

\[
-\log_{10} \| s_{x,4} \| \to \infty = (\log_{10} 0.6 n + \log_{10} 10 + 0(1)) \text{ as } n \to \infty
\]

and

\[
-\log_{10} \| s_{x,6} \| \to \infty = (\log_{10} 0.1 n + \log_{10} 10 + 0(1)) \text{ as } n \to \infty
\]

These results are born out by Figures 3 and 4.

As functions of \( \eta, -\log_{10} \| s_{x,4} \| \to \infty \) in the examples above exhibit almost a straight line behavior, which is slightly distorted due to the presence of the terms \( -3 \log_{10} n \) in (4.7) and (4.8) and of \( -\log_{10} n \) in (4.11) and (4.12). The source of these terms is of course in the Jordan blocks \( C_2 \) and \( C_3 \) in Example 1 and \( C_2 \) and \( C_3 \) in Example 2. We see that the behavior of \( -\log_{10} \| s_{x,4} \| \to \infty \) is closer to that of a straight line in Figures 3 and 4 than in Figures 1 and 2 since the Jordan blocks \( C_2 \) and \( C_3 \) have smaller sizes than \( C_2 \) and \( C_3. \) Also, by the same reason, \( s_{x,4} \) and \( s_{x,8} \) in Example 2 achieve the same accuracies as \( s_{x,4} \) and \( s_{x,8} \) respectively in Example 1, with fewer iterations. Also recall that \( s_{x,8} \) is obtained with less labor than \( s_{x,8}. \)

5. PROOFS OF MAIN RESULTS

**Definition 5.1:** Let \( \lambda \) be a scalar and let \( m,j, \) and \( q \) be integers. Then the linear operator \( A \) is defined via

\[
A^m \left[ \begin{array}{c} m \\ j \end{array} \right] \lambda^q = \left[ \begin{array}{c} m+1 \\ j \end{array} \right] \lambda^{q+1} - \left[ \begin{array}{c} m \\ j \end{array} \right] \lambda^q
\]

(5.1)

**Definition 5.2:** Let
where \( b_{ij} \) and \( \lambda_{ij} \) are scalars and \( n_{ij}, l_{ij} \) and \( m_{ij} \) are integers. Define the \( N \times N \) matrix \( W \) by

\[
(W)_{ip} = \sum_{j=1}^{k} b_{ij} \left( \begin{array}{c} n_{ij} \\ l_{ij} \end{array} \right) \lambda_{ij}^{m_{ij}p-1}, \quad 1 \leq i, p \leq N,
\]

and denote

\[
Y(g_1, g_2, \ldots, g_N) = \det W.
\]

When \( g_i = 1 \) for \( i = i_0 \), we shall take \((W)_{i_0 p} = 1, 1 \leq p \leq N\). Note that the first column of \( W \) is composed of the transpose of the vector \((g_1, g_2, \ldots, g_N)\). Note also that if \( g_i = g_j \) for some \( i, j \) \((i \neq j)\), then \( Y(g_1, g_2, \ldots, g_N) = 0 \) since in this case \( W \) has two identical rows.

Example 5.1:

\[
Y\left(\begin{array}{c} n \\ p \end{array}\right) \lambda_1^0 \Delta\left(\begin{array}{c} n \\ p \end{array}\right) = \begin{vmatrix} 1 & 1 & 1 \\ \frac{n+1}{p} \lambda_1 & \frac{n+2}{p} \lambda_1^2 & \frac{n+3}{p} \lambda_2^3 \\ \frac{n+1}{p} \lambda_2 & \frac{n+2}{p} \lambda_2^2 & \frac{n+3}{p} \lambda_2^3 \end{vmatrix}.
\]

Example 5.2: Let \( a_1, \ldots, a_h \) be arbitrary scalars, and let \( q_1, \ldots, q_h \) be non-negative integers. Then,

\[
Y(a_1 q_1, a_2 q_2, \ldots, a_h q_h) = \prod_{1 \leq i < j \leq h} (a_j - a_i) (q_{i+1} q_{k+1}) \cdot
\]

Actually, the determinant \( Y \) in (5.5) can be shown to be the generalized Vandermonde determinant of \( a_1, \ldots, a_h \). For details see, for example, [5].

Lemma 5.1: For arbitrary \( g_i \) and \( \lambda \) let

\[
R = Y(g_1, \ldots, g_h, \left[ \begin{array}{c} n \\ 0 \end{array} \right] \lambda^m, \ldots, \left[ \begin{array}{c} n \\ j-1 \end{array} \right] \lambda^m \Delta \left[ \begin{array}{c} n \\ j \end{array} \right] \lambda^m, \left[ \begin{array}{c} n \\ j+1 \end{array} \right] \lambda^m, \ldots, \left[ \begin{array}{c} n \\ p \end{array} \right] \lambda^m).
\]

Then

\[
R = (\lambda - 1) Y(g_1, \ldots, g_h, \left[ \begin{array}{c} n \\ 0 \end{array} \right] \lambda^m, \ldots, \left[ \begin{array}{c} n \\ j \end{array} \right] \lambda^m, \ldots, \left[ \begin{array}{c} n \\ p \end{array} \right] \lambda^m).
\]

Proof: That (5.7) holds for \( j = 0 \) is easily seen since by (5.1)
Consider now \( j > 0 \). Again by virtue of (5.1) we have
\[
\Delta \left[ \begin{array}{c} n \\ j \end{array} \right] \lambda^m = (\lambda-1) \lambda^m = (\lambda-1) \left[ \begin{array}{c} n \\ 1 \end{array} \right] \lambda^m. 
\]
(5.8)

Substituting (5.9) in (5.6), and using the fact that determinants are multilinear in their rows, we have
\[
\begin{align*}
\hat{R} &= \lambda Y(g_1, \ldots, g_k, \left[ \begin{array}{c} n \\ 0 \end{array} \right] \lambda^m, \ldots, \left[ \begin{array}{c} n \\ j-1 \end{array} \right] \lambda^m, \left[ \begin{array}{c} n \\ j \end{array} \right] \lambda^m, \left[ \begin{array}{c} n \\ j+1 \end{array} \right] \lambda^m, \ldots, \left[ \begin{array}{c} n \\ p \end{array} \right] \lambda^m) + \\
&\quad + (\lambda-1) Y(g_1, \ldots, g_k, \left[ \begin{array}{c} n \\ 0 \end{array} \right] \lambda^m, \ldots, \left[ \begin{array}{c} n \\ j \end{array} \right] \lambda^m, \ldots, \left[ \begin{array}{c} n \\ p \end{array} \right] \lambda^m). 
\end{align*}
\]
(5.10)

The first of the determinants \( Y' \) on the right hand side of (5.10) vanishes since it has two identical rows, and what remains is (5.7). \( \square \)

**Corollary:** If in (5.6) there are \( l \) terms \( (l \leq p + 1) \) of the form \( \Delta \left[ \begin{array}{c} n \\ j \end{array} \right] \lambda^m, \ i = 1, \ldots, l \), then any \( \Delta \) can be removed provided the resulting determinant is multiplied by \( (\lambda-1) \). In particular,
\[
Y(g_1, \ldots, g_k, \Delta \left[ \begin{array}{c} n \\ 0 \end{array} \right] \lambda^m, \ldots, \Delta \left[ \begin{array}{c} n \\ p \end{array} \right] \lambda^m) = (\lambda-1)^{p+1} Y(g_1, \ldots, g_k, \left[ \begin{array}{c} n \\ 0 \end{array} \right] \lambda^m, \ldots, \left[ \begin{array}{c} n \\ p \end{array} \right] \lambda^m). 
\]
(5.11)

**Lemma 5.2:** For arbitrary \( g_j \) and \( \lambda \), and non-negative integers \( q \), let
\[
\hat{R}_{q_1, \ldots, q_p} = Y(g_1, \ldots, g_k, \left[ \begin{array}{c} n \\ 0 \end{array} \right] \lambda^m, \left[ \begin{array}{c} n+q_1 \\ 1 \end{array} \right] \lambda^m, \ldots, \left[ \begin{array}{c} n+q_p \\ p \end{array} \right] \lambda^m). 
\]
(5.12)

Then \( \hat{R}_{q_1, \ldots, q_p} \) is independent of \( n \) and the \( q_i \) thus
\[
\hat{R}_{q_1, \ldots, q_p} = Y(g_1, \ldots, g_k, \left[ \begin{array}{c} 0 \\ 0 \end{array} \right] \lambda^m, \left[ \begin{array}{c} 0 \\ 1 \end{array} \right] \lambda^m, \ldots, \left[ \begin{array}{c} 0 \\ p \end{array} \right] \lambda^m). 
\]
(5.13)

**Proof:** We first show that \( \hat{R}_{q_1, \ldots, q_p} \) is independent of the \( q_i \). We shall prove this assertion by proving that
\[
\hat{R}_{q_1, \ldots, q_p} = Y(g_1, \ldots, g_k, \left[ \begin{array}{c} n \\ 0 \end{array} \right] \lambda^m, \ldots, \left[ \begin{array}{c} n \\ j \end{array} \right] \lambda^m, \left[ \begin{array}{c} n+q_j+1 \\ j+1 \end{array} \right] \lambda^m, \ldots, \left[ \begin{array}{c} n+q_p \\ p \end{array} \right] \lambda^m). 
\]
(5.14)

First, (5.14) holds for \( j = 0 \) since
\[
\left[ \begin{array}{c} n+q_j \\ 0 \end{array} \right] = \left[ \begin{array}{c} n \\ 0 \end{array} \right] = 1, \quad \forall q = 0, 1, 2, \ldots . 
\]
(5.15)

We now assume that (5.14) holds for \( j \). Substituting the identity...
in (5.14) with \( q = q_{j+1} \) and \( l = j+1 \), and using the fact that \( Y \) is multilinear in its arguments, we have

\[
R_{q_0 \ldots q_{j+1}} = \sum_{i=0}^{j+1} \left( q_{j+1} \right) Y(\underbrace{g_1, \ldots, g_k}_{i=0}, \ldots, \underbrace{g_1, \ldots, g_k}_{i}) \lambda^n, \ldots, \underbrace{\lambda^n}_{i+1}, \underbrace{\lambda^n}_{j+2}, \ldots, \underbrace{\lambda^n}_{p+1}, \lambda^m, \ldots, \underbrace{\lambda^m}_{i+1}, \underbrace{\lambda^m}_{j+1}, \ldots, \underbrace{\lambda^m}_{p+1}, \lambda^m).
\]  

(5.17)

One can see that all terms with \( i \leq j \) in (5.17) vanish, from which one concludes that \( R_{q_0 \ldots q_{j+1}} \) is independent of \( q_{j+1} \) also.

Now that we have proved \( R_{q_0 \ldots q_{j+1}} \) to be independent of \( q_{0} \ldots q_{j} \), we can write

\[
R_{q_0 \ldots q_{j+1}} = R_{l_{j+1}}.
\]  

(5.18)

But by definition of \( R_{q_0 \ldots q_{j+1}} \),

\[
R_{l_{j+1}} = R_{0_{j+1}},
\]  

(5.19)

thus proving the lemma.

We now state a lemma whose proof can be found in [9].

**Lemma 5.3:** Let \( i_0, i_1, \ldots, i_k \) be integers greater than or equal to 1, and assume that the scalars \( v_{i_0 \ldots i_k} \) are odd under an interchange of any two indices, \( i_0, \ldots, i_k \). Let \( \sigma_i, i \geq 1 \), be scalars (or vectors), and let \( t_{i,j}, i \geq 1, 1 \leq j \leq k \) be scalars. Define

\[
I_{k,N} = \sum_{i=1}^{N} \left( \sum_{i_0=1}^{N} \sigma_{i_0} \prod_{p=1}^{k} \prod_{i_p, i_{p+1}}^{i_{p+1}} v_{i_0 \ldots i_k} \right)
\]

(5.20)

and

\[
J_{k,N} = \sum_{1 \leq i_0 < i_1 < \ldots < i_k \leq N} \begin{vmatrix}
\sigma_{i_0} & \sigma_{i_1} & \ldots & \sigma_{i_k} \\
t_{i_0,1} & t_{i_1,1} & \ldots & t_{i_k,1} \\
t_{i_0,2} & t_{i_1,2} & \ldots & t_{i_k,2} \\
\vdots & \vdots & \ddots & \vdots \\
t_{i_0,k} & t_{i_1,k} & \ldots & t_{i_k,k}
\end{vmatrix} v_{i_0 \ldots i_k}
\]

(5.21)

where the determinant in (5.21) is to be interpreted in the same way as \( D(\sigma_0, \ldots, \sigma_k) \) in (1.9). Then-
Definition 5.3: Let $j_h$, $l_h$, $j_p$, $l_p$ be non-negative integers. We will write

\[ j_h l_h < j_p l_p \quad \text{if} \quad j_h < j_p \quad \text{or if} \quad j_h = j_p \quad \text{and} \quad l_h < l_p; \]

\[ j_h l_h = j_p l_p \quad \text{if} \quad j_h = j_p \quad \text{and} \quad l_h = l_p; \]

\[ j_h l_h \leq j_p l_p \quad \text{if either} \ (5.23) \quad \text{or} \ (5.24) \quad \text{holds.} \quad (5.25) \]

Note that Definition 5.3 is equivalent to ordering the set of pairs of non-negative integers lexicographically.

For brevity, in the sequel we shall denote

\[
\sum_{j} = \sum_{j=1}^{p_j} \sum_{i=0}^{j} \sum_{l_1 < l_2 < \cdots < l_h} \sum_{l_1 = 1}^{j_1} \sum_{j_2 = j_1}^{j_2} \cdots \sum_{l_h = 1}^{j_h} \quad \text{if} \ (k_h \geq j_h - 1). \quad (5.26)
\]

In convergence and stability analyses below, we will make use of the three lemmas above, as well as of the following asymptotic expansion for the vectors $u_m$, which follows from (1.1).

\[
u_m = x_{m+1} - x_m \approx \sum_{j=1}^{p} \sum_{l=0}^{j} y_{jl} \Delta \left( \frac{m}{j} \right) \lambda_j^m, \quad \text{as} \quad m \to \infty. \quad (5.27)
\]

In addition, for notational brevity, let us agree that "$\alpha_n - \beta_n$" is equivalent to "$\alpha_n - \beta_n$ as $n \to \infty".\)

Finally, the relation

\[
\frac{\tilde{s}_n^{\lambda - \xi}}{\tilde{s}_n^{\lambda - \xi}} = \frac{D(\tilde{x}_{n+q-\xi}^{-1}, \tilde{x}_{n+q+1-\xi}^{-1}, \ldots, \tilde{x}_{n+q+k-\xi}^{-1})}{D(1, 1, \ldots, 1)} \quad (5.28)
\]

and

\[
\sum_{j=0}^{k} y_{j}^{\lambda} \lambda_j = \frac{D(\tilde{x}_{1}, \ldots, \tilde{x}_{k})}{D(1, 1, \ldots, 1)} \quad (5.29)
\]

will be of use in the proofs below. Both (5.28) and (5.29) are consequences of (1.6); (1.8), and (1.9).

5.1 Convergence and stability. Proofs for MMPE

From (1.10) and (5.27) it follows that

\[
u_n = \sum_{j=1}^{p} \sum_{l=0}^{j} \Delta \left( \frac{n}{l} \right) \lambda_j \lambda_j^m, \quad (5.30)
\]

where
Lemma 5.4: Define

\[ Z_{j_1, j_2, \ldots, j_k} = \begin{vmatrix} z_{j_1,1} & z_{j_1,2} & \cdots & z_{j_1,1} \\ z_{j_2,1} & z_{j_2,2} & \cdots & z_{j_2,1} \\ \vdots & \vdots & \ddots & \vdots \\ z_{j_k,1} & z_{j_k,2} & \cdots & z_{j_k,1} \end{vmatrix} \]  

(5.32)

Then \( H_n(\lambda) = D(1, \lambda, \ldots, \lambda^k) \) has the asymptotic behavior

\[ H_n(\lambda) \sim \sum_{j_1, \ldots, j_k} Z_{j_1, j_2, \ldots, j_k} \left( \prod_{k=1}^n \lambda_{j_k} \right) Y \left( \begin{bmatrix} n \\ 0 \end{bmatrix} \lambda_{j_1}, \Delta \left( \begin{bmatrix} n \\ 1 \end{bmatrix} \lambda_{j_1} \right), \ldots, \Delta \left( \begin{bmatrix} n \\ k \end{bmatrix} \lambda_{j_k} \right) \right). \]  

(5.33)

Proof: Substituting (5.30) into the determinant expression (1.8) for \( D(1, \lambda, \ldots, \lambda^k) \), we obtain

\[ H_n(\lambda) \sim \sum_{j_1, j_2, \ldots, j_k} Z_{j_1, j_2, \ldots, j_k} \left( \prod_{k=1}^n \lambda_{j_k} \right) Y \left( \begin{bmatrix} n \\ 1 \end{bmatrix} \lambda_{j_1}, \Delta \left( \begin{bmatrix} n \\ 1 \end{bmatrix} \lambda_{j_1} \right), \ldots, \Delta \left( \begin{bmatrix} n \\ k \end{bmatrix} \lambda_{j_k} \right) \right). \]  

(5.34)

Using the multilinearity property of determinants, and removing common factors from each row, we can express (5.34) in the form

\[ H_n(\lambda) \sim \sum_{j_1, j_2, \ldots, j_k} \prod_{k=1}^n \lambda_{j_k} \left( \prod_{k=1}^n \lambda_{j_k} \right) Y \left( \begin{bmatrix} n \\ 0 \end{bmatrix} \lambda_{j_1}, \Delta \left( \begin{bmatrix} n \\ 1 \end{bmatrix} \lambda_{j_1} \right), \ldots, \Delta \left( \begin{bmatrix} n \\ k \end{bmatrix} \lambda_{j_k} \right) \right). \]  

(5.35)

Observing that the product \( \prod_{k=1}^n \lambda_{j_k} \) is odd under an interchange of
two pairs of indices \( j_k, k \), we can invoke Lemma 5.3 to obtain (5.33).

Theorem 5.1: Provided \( \lambda \neq \lambda_i, i = 1, \ldots, t \),

\[
D(1, \lambda, \ldots, \lambda^t) = Z \left[ \prod_{k=1}^t \lambda_k^p \right] \left[ \prod_{k=1}^t (\lambda_k-1)^{p+1} \lambda_k^{p+1} \right]^{1/2} \\
\times Y(0;\lambda_1 p_1; \ldots; \lambda_t p_t) \left[ 1 + o(1) \right] \text{ as } n \rightarrow \infty.
\]

(5.36)

Note: (5.36) implies that \( D(1,1, \ldots, 1) \neq 0 \) as \( n \rightarrow \infty \), which guarantees the existence of \( s_{n,k} \) for large enough \( n \).

Proof: It can be shown that the dominant term in the expansion (5.33) is the one whose indices \( j_1, j_2, \ldots, j_k \) take on the values \( 10, 11, \ldots, 1p_1, 20, \ldots, 2p_2, \ldots, 0, \ldots, p_t \), respectively. That is,

\[
H_n(\lambda) = Z \left[ \prod_{k=1}^t \lambda_k^p \right] \left[ \prod_{k=1}^t (\lambda_k-1)^{p+1} \lambda_k^{p+1} \right]^{1/2} \\
\times Y \left( \left[ \begin{array}{c} n \cr 0 \end{array} \right] \lambda_0, \Delta \left[ \begin{array}{c} n \cr p_1 \end{array} \right] \lambda_1^p, \ldots, \Delta \left[ \begin{array}{c} n \cr p_t \end{array} \right] \lambda_t^p \right) \\
\text{ provided that this term is not zero. From (5.32), (5.31) and (3.5), we have } Z_{10,1p_1,0, \ldots, 0} = Z, \text{ which is non-zero by assumption. Applying the corollary to Lemma 5.1 and Lemma 5.2 to the determinant } Y \text{ in (5.37) shows this determinant to be a multiple of } Y(\lambda,0;\lambda_1 p_1; \ldots; \lambda_t p_t) \text{ which is non-zero by virtue of (5.5) and the assumption } \lambda \neq \lambda_i, 1 \leq i \leq t. \text{ This proves the lemma.}
\]

Lemma 5.5: \( G_n = D(x_{n+q} \ldots, x_{n+q+k} \ldots) \) has the asymptotic behavior

\[
G_n \sim \sum_j \lambda_j^p \sum \left[ \prod_{k=1}^q \lambda_k^p \right] Y \left( \left[ \begin{array}{c} n+q \cr 0 \end{array} \right] \lambda_0^p, \Delta \left[ \begin{array}{c} n \cr p_1 \end{array} \right] \lambda_1^p, \ldots, \Delta \left[ \begin{array}{c} n \cr p_t \end{array} \right] \lambda_t^p \right). \quad (5.38)
\]

Proof: Substituting (5.30) and (1.1) into the determinant expression (1.8) for \( D(x_{n+q} \ldots, x_{n+q+k} \ldots) \), and proceeding as in the proof of Lemma 5.4, we obtain

\[
G_n \sim \sum_j \lambda_j^p \sum \left[ \prod_{k=1}^q \lambda_k^p \right] Y \left( \left[ \begin{array}{c} n+q \cr 0 \end{array} \right] \lambda_0^p, \Delta \left[ \begin{array}{c} n \cr p_1 \end{array} \right] \lambda_1^p, \ldots, \Delta \left[ \begin{array}{c} n \cr p_t \end{array} \right] \lambda_t^p \right). \quad (5.39)
\]

The product \( \left[ \prod_{k=1}^q \lambda_k^p \right] Y \left( \left[ \begin{array}{c} n+q \cr 0 \end{array} \right] \lambda_0^p, \Delta \left[ \begin{array}{c} n \cr p_1 \end{array} \right] \lambda_1^p, \ldots, \Delta \left[ \begin{array}{c} n \cr p_t \end{array} \right] \lambda_t^p \right) \) is odd under an interchange of two pairs of indices \( j_k, k \), \( h=1, \ldots, k \). Thus Lemma 5.3 can be invoked, resulting in (5.38).
Theorem 5.2: Define the vectors $\vec{z}_{i,j}$ by

$$\vec{z}_{i,j} = \begin{bmatrix} d_{1}y_{10} & \cdots & d_{1}y_{1p} & \cdots & d_{1}y_{p}, d_{1}y_{1} \\ z_{10,1} & \cdots & z_{1p,1} & \cdots & z_{1p,1} z_{10,1} \\ \vdots & & \vdots & & \vdots \\ z_{10,k} & \cdots & z_{1p,k} & \cdots & z_{1p,k} z_{10,k} \end{bmatrix}$$

(5.40)

where $d_k = \lambda / (\lambda_k - 1)$. Then with $r'$ as defined in (3.2),

$$G_n = \prod_{k=1}^{t} \lambda_n^{(p+1)} \prod_{k=1}^{t} (\lambda_n - 1)^{y_{k}^{(p+1)}} \lambda_n^{(p+1)} \lambda_n^{(p+1)}$$

Proof: We shall treat the case in which $r=1$ first. In this case the dominant terms in (5.38) are those for which the pairs of indices $j_{10}, j_{1p}, \ldots, j_{l_0}, \ldots, j_{l_p}, (l+1)l$, $(0 \leq l \leq p+1)$, subject to the constraint $j_{10,1} < j_{20,2} < \cdots < j_{k0,k}$, and are the dominant terms in the asymptotic expansion (5.38) of $G_n$ thus becomes:

$$G_n \sim \sum_{l=0}^{p+1} \sum_{j_{0}, j_{1}, \ldots, j_{l}, t} \lambda_{l}^{j_{0}} \lambda_{l+1}^{j_{l+1}} \lambda_{l+2}^{j_{l+2}} \cdots \lambda_{l+1}^{j_{l+1}} \lambda_{l}^{j_{0} \cdot \Delta [l_0, l_1, \ldots, l_{t}] (5.43)$$

subject to the above mentioned constraints on the indices $j_{10}, j_{1p}, \ldots, j_{l_0}, \ldots, j_{l_p}, (l+1)l$. Here we have denoted

$$\sum_{l} = \begin{cases} \sum_{l} & \text{for } 1 \leq j_{0} \leq t \\ \sum_{l} & \text{for } j_{0} = r+1 \end{cases}$$

By rearranging rows in the determinant $Y$ in (5.43), and using the corollary to Lemma 5.1 and Lemma 5.2, in this order, we can show that this determinant is actually

$$(-1)^{\sum_{l}^{p+1}} (\lambda_{l} - 1)^{y_{l+1}} / (\lambda_{l} - 1)^{y_{l+1}} (\lambda_{l} - 1)^{y_{l+1}} (\lambda_{l} - 1)^{y_{l+1}}$$

for $1 \leq j_{0} \leq t$, and
\[-\frac{(n+q)}{l}\] for $j_0 = t+1$ and $l_0 = l$. Using the facts that $\Delta \left[ \frac{n}{l} \lambda^0_l \right] = \frac{nI}{l} (\lambda_{t+1}-1) + O(n^{-1})$ as $n \to \infty$ and

\[
\left[ \frac{n+q}{l} \right] = \frac{n}{l} + O(n^{-1}) \text{ as } n \to \infty,
\]
both cases can be combined in (5.43), to yield

\[
G_n = \sum_{i=0}^{P_{t+1}} \sum_{j=0}^{P_{t+1}} d_{ij} y_{j,i} \left[ (-1) \gamma^0 \left( \frac{n}{l} \right) x_{i,l} \right] (\lambda_{t+1}-1) \times \left( \prod_{i=1}^{P_{t+1}} \left[ \frac{n}{l} \lambda^0_l \right] \frac{n}{l} I \right)^{\frac{n}{l} - 1} \left( \prod_{i=1}^{P_{t+1}} \left[ \frac{0}{l} \lambda^0_l \right] \frac{0}{l} I \right)^{\frac{n}{l} + 1}.
\]

Using the fact that

\[
\sum_{j=1}^{l+1} (-1)^{\gamma^0_l} d_{ij} y_{j,i} z_{j+1,l} = z_{l+1,l},
\]

and noting that $l = p_{t+1}$ yields the dominant term, and exploiting the relation between the determinant $Y$ in (5.44) and the appropriate generalized Vandermonde determinant $S$, (5.44) reduces to (5.41) with $r' = 1$. It should now be clear that $G_n$ can be expressed via (5.41) for arbitrary $r'$ as well. What remains to be shown is that the dominant term in (5.41) is non-zero. For this it is enough to show that the summation in (5.41) does not vanish. By (5.45) we can express this summation as a linear combination of the linearly independent vectors $y_{j,i}$, $10 \leq j, p \leq p_{t+1}$, and $y_{i+1,j}$, $1 \leq i \leq r'$. Now the coefficient multiplying the vector $y_{i+1,j}$ in this linear combination is the product

\[
\left[ (-1)^{\gamma^0_l} \left[ \frac{n}{l} \lambda^0_l \right] \frac{n}{l} I \right] \left( \frac{n}{l} \lambda^0_l \right) (\lambda_{t+1}-1) \times \left( \prod_{i=1}^{P_{t+1}} \left[ \frac{n}{l} \lambda^0_l \right] \frac{n}{l} I \right)^{\frac{n}{l} - 1} \left( \prod_{i=1}^{P_{t+1}} \left[ \frac{0}{l} \lambda^0_l \right] \frac{0}{l} I \right)^{\frac{n}{l} + 1},
\]

and this product is non-zero by $z_{l+1,l} = Z$ and the assumption that $Z \neq 0$, and the rest of the assumptions on the $\lambda_i$. \[\Box\]

Theorem 3.1 for MMPE can now be proved by dividing the asymptotic behavior of $G_n$ in (5.41) by that of $H_n(1)$ in (5.36), the vector $\Gamma(n)$ in (3.4) being identified as

\[
\Gamma(n) = \frac{1}{Z P_{t+1}!} \sum_{j=1}^{P_{t+1}} e^{-\lambda_j} \frac{n}{l} \left( \frac{\lambda_j}{\lambda_j - 1} \right)^{P_{t+1} - j} \left( \frac{\lambda_{t+1} - \lambda_j}{\lambda_j - 1} \right)^{\gamma^0_j} + O(1) \quad \text{as } n \to \infty.
\]

Clearly this $\Gamma(n)$ satisfies (4.5).
The proof of Theorem 3.2 can be achieved by combining the asymptotic behaviour of $H_n(\lambda)$ and $H_n(1)$ from (5.36) in (5.29).

We conclude this subsection by exploring the meaning of the constraint $Z \neq 0$ for the example described in Section 2. Substituting (2.11) in (3.6) we find after some elementary column transformations that $Z \neq 0$ is equivalent to

$$
\begin{bmatrix}
Q_1(v_{11}) & \ldots & Q_1(v_{1,p+1}) & \ldots & Q_1(v_{1,\alpha+1}) \\
\vdots & \ddots & \vdots & \ddots & \vdots \\
Q_k(v_{11}) & \ldots & Q_k(v_{1,p+1}) & \ldots & Q_k(v_{1,\alpha+1})
\end{bmatrix} \neq 0,
$$

(5.47)

where $v_{ij}$ are eigenvectors or principal vectors.

5.2 Convergence and Stability Proofs for TEA

From (1.11) and (5.27) it follows that

$$
u_{h,i} = Q(u_{h,i+1}) - \sum_{j=1}^{\infty} \sum_{p=0}^{\infty} Q(y_{jp}) \Delta \left[ \frac{n+h+i}{p} \right] \lambda_{j}^{p+h+i}.
$$

(5.48)

Using the binomial identity in (5.16) with $n$ and $q$ there replaced by $n+i$ and $h$ respectively, (5.48) becomes

$$
u_{h,i} = \sum_{j=1}^{\infty} \sum_{p=0}^{\infty} Q(y_{jp}) \sum_{l=0}^{h} \Delta \left[ \frac{n+i}{j} \right] \lambda_{j}^{p+h+i}.
$$

(5.49)

Interchanging the summations over $p$ and $l$, and letting

$$
z_{j,l,h} = \sum_{j=1}^{\infty} \sum_{p=0}^{\infty} Q(y_{jp}) \lambda_{j}^{p+h+i},
$$

(5.50)

we see that $u_{h,i}$ for TEA has the same form as that for MMPE given in (5.30). Therefore, the proofs for TEA are identical to those for MMPE, provided we replace $Q_{h+1}(y_h)$ in the definition of $Z$ in (3.6) with $z_{j,l,h}$ of (5.50). The only thing that remains to be shown is that (3.7) implies that $Z \neq 0$. By performing elementary column transformations on this new $Z$ we find that

$$
Z = (-1)^{p_j} \sum_{j=1}^{p_j} \prod_{j=1}^{p_j} Q(y_{jp_j}) \left[ \lambda_{j+1} \ldots \lambda_{p_j} \right] \Delta \left[ \frac{n+i}{p_j} \right] \lambda_{j}^{p_j+h+i},
$$

(5.51)

where $\sigma_j = p_j(p_j+1)/2$. The desired result now follows.
In the context of the example of Section 2, the condition (3.7) is equivalent to

$$\prod_{j=1}^{n} \Omega(v_{j}) \neq 0,$$

which can be seen by observing that $y_{j k} = \lambda_{j}^{\nu_{j}} a_{j}^{\nu_{j}+1} v_{j}$. It is interesting to note that (5.52) imposes no conditions on the operator $\Omega$ with respect to the principal vectors.

### 5.3 Convergence and Stability Proofs for MPE

From (1.12) and (5.27) we have

$$u_{v_{j}} = (u_{v_{j}} + u_{v_{n+1}}) - \sum_{m=1}^{N} \sum_{j=1}^{n} z_{j}^{m} \left[ \Delta \left( n+1 \right) \lambda_{n+1} \right] \left[ \Delta \left( n+1 \right) \lambda_{n+1} \right],$$

where

$$z_{j}^{m} = \left( y_{j}^{m} v_{j} \right).$$

**Theorem 5.3:** Let

$$Z_{j_{1}, \ldots, j_{l}}^{h_{1}, \ldots, h_{k}} =$$

$$
\begin{array}{cccc}
z_{j_{1}}^{h_{1}} & z_{j_{2}}^{h_{2}} & \cdots & z_{j_{l}}^{h_{l}} \\
\vdots & \vdots & \ddots & \vdots \\
z_{j_{1}}^{h_{1}} & z_{j_{2}}^{h_{2}} & \cdots & z_{j_{l}}^{h_{l}}
\end{array}
$$

Provided that $\lambda \neq \lambda_{i}$, $i=1, \ldots, l$,

$$\tilde{D}(1, \lambda_{1}, \ldots, \lambda_{k}) = Z_{10, \ldots, 0}^{10, \ldots, 0} \prod_{m=1}^{l} \lambda_{m}^{2(n+1)} \left[ \prod_{m=1}^{l} \left( \lambda_{m} - 1 \right)^{2(m+1)} \lambda_{m}^{m(m+1)} \right] \tilde{Y}^{(\lambda_{1}, \lambda_{2}, \ldots, \lambda_{k})}(\lambda, 0; \lambda_{1}, p_{1}; \lambda_{2}, p_{2}; \ldots, \lambda_{k}, p_{k}) \left[ 1 + o(1) \right] \text{ as } n \to \infty.$$  

**Proof:** The proof of this theorem is similar to that of Theorem 5.1. Therefore, we shall provide an outline for the reader. Substituting (5.53) into the determinant expression (1.8) for $\tilde{D}(1, \lambda_{1}, \ldots, \lambda_{k})$, and using the multilinearity property of determinants, we obtain an expansion similar to (5.35) involving the summation indices $j_{1}^{t_{1}} \ldots, j_{l}^{t_{l}}$ and $h_{1}^{t_{1}}, \ldots, h_{k}^{t_{k}}$. It turns out that Lemma 5.3 can be applied twice, first to the summation over $j_{1}^{t_{1}} \ldots, j_{l}^{t_{l}}$ and then to that over $h_{1}^{t_{1}}, \ldots, h_{k}^{t_{k}}$ - resulting in
which is the analogue of (5.33) in Lemma 5.4. (5.56) now follows from (5.57) in the same way (5.36) follows from (5.33). To complete the argument, we note that \( Z_{10}^{10} \ldots \) in (5.56) is the Gram determinant of the linearly independent vectors \( y_{jl} \). \( 10 \leq j \leq \theta_1 \), and is thus non-zero.

**Theorem 5.4:** Define the vectors

\[
\tilde{z}_{i,j} = \begin{vmatrix}
    d_1 y_{10} & \cdots & d_1 y_{\theta_1} & \cdots & d_1 y_{\theta_p} & d_1 y_d \\
    z_{10} & \cdots & z_{10} & \cdots & z_{10} & \cdots & z_{10} \\
    \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\
    z_{\theta_p} & \cdots & z_{\theta_p} & \cdots & z_{\theta_p} & \cdots & z_{\theta_p} \\
    z_{\theta_d} & \cdots & z_{\theta_d} & \cdots & z_{\theta_d} & \cdots & z_{\theta_d} \\
    \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\
    z_{\theta_d} & \cdots & z_{\theta_d} & \cdots & z_{\theta_d} & \cdots & z_{\theta_d} \\
\end{vmatrix}
\]

(5.58)

where \( d_\theta = \lambda_\theta/(\lambda_\theta - 1) \). Then with \( \Gamma' \) as defined in (3.2), \( G_n = D(x_{n+q-s} \ldots x_{n+q+k-s}) \) has the asymptotic behavior

\[
G_n = \left[ \prod_{m=1}^{n} \lambda_m^{2\epsilon_1(p+1)} \right] \left[ \prod_{m=1}^{n} (\lambda_m - 1)^{2\epsilon_1(p+1)} \lambda_m^{\epsilon_1(p+1)} \right] Y(\lambda_1, \ldots, \lambda_{p+1}) \frac{n^{p+1}}{p+1!} \times \sum_{i=1}^{\theta_1} \lambda_{i+1}^{\epsilon_1} \tilde{z}_{i,l}(\lambda_{i+1} - 1) \tilde{y}(\lambda_1, \ldots, \lambda_{p+1}; \lambda_{i+1}; 0).
\]

(5.59)

**Proof:** The proof of (5.59) can be achieved using arguments similar to those found in the proofs of Theorems 5.2 and 5.3. The details are left to the reader.

Theorem 3.1 for MPE can now be proved by dividing the asymptotic behavior of \( G_n \) in (5.59) by that of \( D(1,1,\ldots,1) \) in (5.56), where the vector \( \Gamma(n) \) of (3.4) is exactly of the form (5.46) with \( Z' \) and \( \tilde{z}_{i,j} \) of (5.46) being replaced by \( Z_{10}^{10} \ldots \) of (5.55) and \( \tilde{z}_{i,j} \) of (5.58) respectively.

Theorem 3.2 for MPE can be proved by considering the asymptotic behavior of \( H_n(\lambda) \) divided by the asymptotic behavior of \( H_n(1) \).
5.4 Convergence and Stability Proofs for RRE

From (1.13) and (5.27) we have

$$u_{h,i} = \langle w_{n+h,i}, u_{n+1,i} \rangle = \sum_{m,p,l} y_{m,p,l}^{(n)} \left[ \Delta^2 \left( \frac{n+h}{p} \right) \lambda_m \lambda_p \lambda_l \right] \left[ \Delta \left( \frac{n+h}{l} \right) \lambda_m \lambda_p \right].$$

(5.60)

where $y_{m,p,l}^{(n)}$ is as defined in (5.54) and $\Delta^2 = \Delta \Delta$.

The proofs of this section are nearly identical to those for MPE, the only difference being that in the determinants $Y$ having $\lambda_j$ as their arguments, the operator $\Delta$ is replaced by $\Delta^2$, c.f. (5.57). Ultimately, however, these determinants have no effect on the final results since they disappear from the dominant terms of $G_nH_n(1)$ and $H_n(\lambda)/H_n(1)$. Consequently, the vector $\Gamma(n)$ for RRE is asymptotically equivalent to $\Gamma(n)$ for MPE.

6. EXTENSIONS OF POWER METHOD.

From Theorem 3.2 it is apparent that, as $n \to \infty$, the zeros of the polynomial $\sum_{i=0}^{k} \gamma^{(n,k)}(n) \lambda_i$ approach the $\lambda_i$, $i = 1, \ldots, t$, with corresponding multiplicities $p_i+1$. Based on this observation, in this section we propose some extensions to the well-known power method that is used to estimate the largest eigenvalue (in modulus) of a matrix $A$. These extensions enable us to estimate the first few dominant eigenvalues of the matrix $A$.

Let $x_0, x_1, x_2, \ldots$ be a sequence of vectors in $B$ satisfying

$$x_{m+1} = \sum_{i=1}^{\infty} P_i(m) \lambda_i x_m$$

as $m \to \infty$.

(6.1)

where $P_i(m)$ and $\lambda_i$ are exactly as described in Section 1 with the notation therein, with the exception that $\lambda_i \neq 1, i = 1, 2, \ldots$ is not required.

A natural example for a sequence of this kind is one generated by the iterative procedure

$$x_{j+1} = Ax_j, \quad j = 0, 1, \ldots, x_0 \text{ given},$$

(6.2)

where $A$ is the matrix of Section 2, with no restrictions being imposed on its spectrum. In fact, (6.1) can be obtained for this example, beginning with (2.4) and deleting $s$ everywhere in Section 2.
Now a close look at the power method for the matrix $A$ above reveals that this method actually approximates $\lambda_i$ in (6.1) provided $p_1=0$ and $|\lambda_1| > |\lambda_2|$, by utilizing only the vector sequence $x_0, x_1, \ldots$ with any reference to the the matrix $A$ being indirectly through the vectors $x_0, x_1, \ldots$. With this in mind we now propose the following extensions to the power method for estimating the first few dominant $\lambda_i$ in (6.1) counting multiplicities:

Let the vector sequence $x_0, x_1, \ldots$ be as above. Construct the polynomial

$$P^{(n,A)}(\lambda) = \sum_{i=0}^{k-1} c_i^{(n,A)} \lambda^i, \quad c_k^{(n,A)} = 1,$$

where the coefficients $c_i = c_i^{(n,A)}, 0 \leq i \leq k-1,$ are determined in one of the following ways:

1) MMPE extension

$$\sum_{j=0}^{k-1} c_j Q_i(x_{n+j}) = -Q_i(x_{n+k}), \quad 1 \leq i \leq k,$$

2) TEA extension

$$\sum_{j=0}^{k-1} c_j Q(x_{n+i+j}) = -Q(x_{n+i+k}), \quad 0 \leq i \leq k-1,$$

3) MPE extension

$$\sum_{j=0}^{k-1} c_j (x_{n+i+j}) = -(x_{n+i+k}), \quad 0 \leq i \leq k-1,$$

4) RRE extension

$$\sum_{j=0}^{k-1} c_j (u_{n+i+j}) = -(u_{n+i+k}), \quad 0 \leq i \leq k-1.$$

Here $Q$, $Q_i$ are as described in Section 1. Finally, the zeros of the polynomial $P^{(n,A)}(\lambda)$ are taken to the estimates of the first $k$ most dominant $\lambda_i$ including their multiplicities.

In a finite dimensional space $W$, for every bounded linear functional $F$ there exists a unique vector $f \in W$ such that $F(x) = \langle f, x \rangle$ for every $x \in W$, where $\langle \cdot, \cdot \rangle$ is the inner product associated with $W$. $f$ is called the representor of $F$.

Taking $k=1$ and letting $Q_1 = Q$ be represented by the vector $q$, we see that the MMPE and TEA extensions above are equivalent to the standard power method. The MPE extension reduces to the Rayleigh quotient for $k=1$. The RRE extension, however, has no analogue that we know of.
The following theorem provides the justification for the extensions above.

**Theorem 6.1:** Let the vector sequence \( x_0, x_1, \ldots \) be as described in the beginning of this section. Assume in addition, that the \( \lambda_i \) satisfy (3.1) and let \( k \) be as in (3.3). Then

\[
\lim_{n \to \infty} P^{(n,k)}(\lambda) = \prod_{i=1}^r (\lambda - \lambda_i)^{p_{n+1}^{i+1}}
\]

provided (3.6) holds for MMPE, (3.7) for TEA. No additional conditions are needed for MPE and RRE.

**Proof:** In analogy to (5.29) it can be shown that \( P^{(n,k)}(\lambda) \) can be expressed as

\[
p^{(n,k)}(\lambda) = \frac{D(1, \lambda, \ldots, \lambda^k)}{D_k}
\]

where \( D(1, \lambda, \ldots, \lambda^k) \) is as defined in (1.8) with \( u_{i,j} \) there redefined as

\[
u_{i,j} = Q_{i+1}(x_{n+i}) \quad \text{for MMPE},
\]

\[
u_{i,j} = Q(x_{n+i}, x_{n+j}) \quad \text{for TEA},
\]

\[
u_{i,j} = (x_{n+i}, x_{n+j}) \quad \text{for MPE},
\]

\[
u_{i,j} = (u_{n+i}, x_{n+j}) \quad \text{for RRE},
\]

and \( D_k \) is the cofactor of \( \lambda^k \) in the expansion of \( D(1, \lambda, \ldots, \lambda^k) \) with respect to its first row.

The proof now proceeds along the same lines as the proofs of Theorems 3.1 and 3.2. A cursory look at the asymptotic expansions for \( D(1, \lambda, \ldots, \lambda^k) \) reveals that it is not necessary to require \( \lambda_i \neq 1 \). This is easily seen especially for the cases of MMPE, TEA, and MPE, and can be shown for RRE as well by using the tools developed in the beginning of Section 5.

Extensions of the power method based on the scalar and vector epsilon algorithms of Wynn [12,13] and TEA were given by Brezinski [2]. Brezinski shows convergence under the conditions that \( \lambda_i \neq 1 \) and \(|\lambda_1| > |\lambda_2| > \cdots \).
REFERENCES


[12] P. Wynn, On a device for computing the $e_m(S_n)$ transformation MTAC, 10 (1956), pp. 91-96.

FIGURE CAPTIONS

Fig. 1 - Results for Example 1 taking \( k=4 \).

\[ +: - \log_{10} |s_n - s| \quad \text{for MMPE} \]

\[ \Box: - \log_{10} |s_n - s| \quad \text{for MPE} \]

Fig. 2 - Results for Example 1 taking \( k=8 \).

\[ +: - \log_{10} |s_n - s| \quad \text{for MMPE} \]

\[ \Box: - \log_{10} |s_n - s| \quad \text{for MPE} \]

Fig. 3 - Results for Example 2 taking \( k=4 \).

\[ +: - \log_{10} |s_n - s| \quad \text{for MMPE} \]

\[ \Box: - \log_{10} |s_n - s| \quad \text{for MPE} \]

Fig. 4 - Results for Example 2 taking \( k=6 \).

\[ +: - \log_{10} |s_n - s| \quad \text{for MMPE} \]

\[ \Box: - \log_{10} |s_n - s| \quad \text{for MPE} \]
Figure 1 - RESULTS FOR EXAMPLE 1 WITH $k = 4$.
Figure 2 - RESULTS FOR EXAMPLE 1 WITH K = 8
Figure 3 - RESULTS FOR EXAMPLE 2 WITH $k = 4$
Figure 4 - Results for Example 2 with $k = 6$