# Remarks on Some Newton and Chebyshev-type Methods for Approximation Eigenvalues and Eigenvectors of Matrices 

I. Pǎvǎloiu<br>E. Cǎtinaş

## 1 Introduction

It is well known that the Newton and the Chebyshev methods for nonlinear systems require solving of a linear system at each iteration step. In this note we shall study two modified methods which avoid solving of linear systems by using the Schultz method to approximate inverses of Fréchet derivatives. At the same time we shall use the particularities of nonlinear systems arising from eigenproblems, since the Fréchet derivatives of order higher than two are the null multilinear operators. Some numerical examples will be provided in the end of this note.

Denote $V=\mathbb{K}^{n}$ and let $A=\left(a_{i j}\right) \in \mathbb{K}^{n \times n}$, where $\mathbb{K}=\mathbb{R}$ or $\mathbb{C}$. We recall that the scalar $\lambda \in \mathbb{K}$ is an eigenvalue of $A$ if there exists $v \in V$, $v \neq 0$ such that

$$
\begin{equation*}
A v-\lambda v=0 . \tag{1}
\end{equation*}
$$

The vector $v$ is called the eigenvector corresponding to the eigenvalue $\lambda$. Since for an eigenvalue $\lambda$ the eigenpair $(v, \lambda)$ is not uniquely determined, it is necessary to impose a supplementary condition. Different Newton-type methods were studied in papers [1]-[4], [6], [9]-[11], [16]-[18], [20], [21]. It is worth mentioning that the Rayleigh quotient method is equivalent with a certain Newton method ("the scaled Newton method" [10]).

We shall consider a "norming" function $G: V \rightarrow \mathbb{K}, G(0) \neq 1$ and,

[^0]besides (1), the equation
$$
G(v)-1=0
$$

The function $G$ may be chosen in different ways (see [16] and [3]):

$$
\begin{aligned}
\text { I } & G(v)=\frac{1}{2}\|v\|_{2}^{2}, \\
\text { II } & G(v)=\frac{1}{2 n}\|v\|_{2}^{2} .
\end{aligned}
$$

We shall consider in the theoretical results hereafter the choice II.
Let $X=V \times \mathbb{K}\left(=\mathbb{K}^{n+1}\right)$ and for $x=\binom{v}{\lambda} \in X$ take

$$
\|x\|=\max \{\|v\|,|\lambda|\}
$$

where the norm on $V$ is one of usual norms.
Consider the system

$$
\begin{equation*}
F(x)=0 \quad\left(\text { here } 0 \in \mathbb{K}^{\mathrm{n}+1}\right) \tag{2}
\end{equation*}
$$

with the mapping $F$ given by

$$
F(x)=\binom{A v-\lambda v}{G v-1}, \quad x=\binom{v}{\lambda} \in X .
$$

Denoting $v=\left(x^{(1)}, x^{(2)}, \ldots, x^{(n)}\right)$ and $\lambda=x^{(n+1)}$ then the system (2) can be written explicitly

$$
\begin{aligned}
& F_{1}(x)=\left(a_{11}-x^{(n+1)}\right) x^{(1)}+a_{12} x^{(2)}+\cdots+a_{1 n} x^{(n)}=0 \\
& \vdots \\
& F_{n}(x)=a_{n 1} x^{(1)}+a_{n 2} x^{(2)}+\cdots+\left(a_{n n}-x^{(n+1)}\right) x^{(n)}=0 \\
& F_{n+1}(x)=\frac{1}{2 n}\left(x^{(1)}\right)^{2}+\frac{1}{2 n}\left(x^{(2)}\right)^{2}+\cdots+\frac{1}{2 n}\left(x^{(n)}\right)^{2}-1=0 .
\end{aligned}
$$

It can be easily seen that the Fréchet derivatives of $F$ are given by the following relations:

$$
F^{\prime}\left(x_{0}\right) h=\left(\begin{array}{cc}
A-\lambda_{0} I & -v_{0} \\
\frac{1}{n} v_{0}^{t} & 0
\end{array}\right)\binom{u}{\alpha}
$$

$$
F^{\prime \prime}\left(x_{0}\right) h k=\binom{-\alpha w-\beta u}{\frac{1}{n} w^{t} u}
$$

for all $x_{0}=\binom{v_{0}}{\lambda_{0}}, h=\binom{u}{\alpha}, k=\binom{w}{\beta} \in X$.
Since the Fréchet derivative of order 2 does not depend on $x_{0}$ it is obvious that the derivatives of order higher than 2 an the null multilinear operators, so for any fixed $x_{0} \in X$

$$
\begin{equation*}
F(x)=F\left(x_{0}\right)+F^{\prime}\left(x_{0}\right)\left(x-x_{0}\right)+\frac{1}{2} F^{\prime \prime}\left(x_{0}\right)\left(x-x_{0}\right)^{2} \quad \forall x \in X . \tag{3}
\end{equation*}
$$

It is easy to verify that when we use the max norm on $V$ and $G$ is given by the choice II

$$
\left\|F^{\prime \prime}(x)\right\|_{\infty}=2, \quad \forall x \in X
$$

The following result concerning the invertibility of $F^{\prime}$ at a solution hold.

Lemma 1 Let $x^{*}=\left(v^{*}, \lambda^{*}\right)$ be an eigenpair of a given matrix $A \in$ $\mathbb{K}^{n \times n}$. Then the eigenvalue $\lambda^{*}$ is simple if and only if the Jacobian $F^{\prime}\left(x^{*}\right)$ is nonsingular.

Proof. The corresponding result for the choice I of $G$ was proved by Yamamoto [21], for which

$$
F^{\prime}\left(x_{0}\right)=\left(\begin{array}{cc}
A-\lambda_{0} I & -v_{0} \\
-v_{0}^{t} & 0
\end{array}\right) .
$$

The stated affirmation follows immediately observing that the two matrices differ by a nonzero factor in the last row.

## 2 Some Newton and Chebyshev-type methods

We shall study first the convergence of the sequences $\left(x_{k}\right)_{k \geq 0} \subset X$ and $\left(\Gamma_{k}\right) \subset \mathcal{L}(X)\left(=\mathbb{K}^{(n+1) \times(n+1)}\right)$ generated by the following Newtontype process applied to the nonlinear system (2), initially proposed by

Ul'm [19] and studied by Diaconu and Pǎvăloiu [7]:

$$
\begin{align*}
& x_{k+1}=x_{k}-\Gamma_{k} F\left(x_{k}\right)  \tag{4}\\
& \Gamma_{k+1}=\Gamma_{k}\left(2 I-F^{\prime}\left(x_{k+1}\right) \Gamma_{k}\right), \quad k=0,1, . .
\end{align*}
$$

$x_{0} \in X$ and $\Gamma_{0} \in \mathcal{L}(X)$ being given.
We shall need the following preliminary result.
Lemma 2 [4] If the sequences $\left(\delta_{k}\right)_{k \geq 0}$ and $\left(\rho_{k}\right)_{k \geq 0}$ of real positive numbers satisfy

$$
\begin{align*}
& \delta_{k+1} \leq\left(\delta_{k}+2 \rho_{k}\right)^{2}  \tag{5}\\
& \rho_{k+1} \leq \rho_{k} \delta_{k}+\rho_{k}^{2}, \quad k=0,1, . .
\end{align*}
$$

with $\max \left(\delta_{0}, \rho_{0}\right) \leq \frac{1}{9} d$ for some $d \in(0,1)$, then the following inequalities are true:

$$
\max \left\{\delta_{k}, \rho_{k}\right\} \leq \frac{1}{9} d^{2^{k}}, \quad k=0,1, . .
$$

Denoting $\bar{B}_{r}\left(x_{0}\right)=\left\{x \in X:\left\|x-x_{0}\right\| \leq r\right\}$ we can state the following result.

Theorem 3 Assume that the operator $F$ and the elements $x_{0} \in X$, $\Gamma_{0} \in \mathcal{L}(X)$ and $r>0$ satisfy the following conditions
a) there exists $F^{\prime}\left(x_{0}\right)^{-1}$ and $\left\|F^{\prime}\left(x_{0}\right)^{-1}\right\| \leq \beta_{0}$;
b) $q=2 \beta_{0} r<1$;
c) denoting $\delta_{0}=\left\|I-F^{\prime}\left(x_{0}\right) \Gamma_{0}\right\|, \rho_{0}=\frac{100}{81} \beta^{2}\left\|F\left(x_{0}\right)\right\|$ and $\beta=\frac{\beta_{0}}{1-q}$, suppose

$$
\max \left\{\delta_{0}, \rho_{0}\right\} \leq \frac{1}{9} d \quad \text { for some } d \in(0,1) ;
$$

d) $\frac{d}{10(1-q)} \leq r$.

Then the sequences $\left(x_{k}\right)_{k \geq 0},\left(\Gamma_{k}\right)_{k \geq 0}$ generated by (4) converge and $\left(x_{k}\right)_{k \geq 0} \subset \bar{B}_{r}\left(x_{0}\right)$. Denoting $x^{*}=\lim _{k \rightarrow \infty} x_{k}$ and $\Gamma^{*}=\lim _{k \rightarrow \infty} \Gamma_{k}$, then $x^{*}$
is a solution of the nonlinear system (2) and $\Gamma^{*}=F^{\prime}\left(x^{*}\right)^{-1}$. Moreover, the following estimations hold:

$$
\begin{aligned}
& \left\|x^{*}-x_{k}\right\| \leq \frac{d^{2^{k}}}{10 \beta\left(1-d^{2^{k}}\right)}, \quad k=0,1, . . \\
& \left\|\Gamma^{*}-\Gamma_{k}\right\| \leq \frac{d^{2^{k}}}{3\left(1-d^{2^{k}}\right)}, \quad k=0,1, \ldots
\end{aligned}
$$

Proof. It can be easily seen that in our hypotheses the derivatives $F^{\prime}(x)$ are invertible for all $x \in \bar{B}_{r}\left(x_{0}\right)$.

Using the inequality

$$
\left\|I-F^{\prime}\left(x_{0}\right)^{-1} F^{\prime}(x)\right\| \leq 2 \beta_{0} r=q<1
$$

and applying the Banach lemma we get

$$
\left\|F^{\prime}(x)^{-1}\right\| \leq \frac{\beta_{0}}{1-q}=\beta .
$$

Taking into account b) and c) it follows that

$$
\begin{aligned}
\left\|\Gamma_{0}\right\| & \leq\left\|F^{\prime}\left(x_{0}\right)^{-1}\right\|\left(\left\|F^{\prime}\left(x_{0}\right) \Gamma_{0}-I\right\|+1\right) \\
& \leq \beta_{0}\left(1+\delta_{0}\right) \leq \frac{10}{9} \beta_{0} \leq \frac{10}{9} \beta
\end{aligned}
$$

which together with relation (4) imply

$$
\left\|x_{1}-x_{0}\right\| \leq\left\|\Gamma_{0}\right\|\left\|F\left(x_{0}\right)\right\|<\frac{d}{10 \beta(1-d)} \leq r
$$

i.e. $x_{1} \in \bar{B}_{r}\left(x_{0}\right)$.

Denote $\rho_{1}=\frac{100}{81} \beta^{2}\left\|F\left(x_{1}\right)\right\|$ and $\delta_{1}=\left\|I-F^{\prime}\left(x_{1}\right) \Gamma_{1}\right\|$. If we take $x=x_{1}$ in (3) then an elementary reasoning shows that

$$
\begin{aligned}
& \rho_{1} \leq \rho_{0}^{2}+\delta_{0} \rho_{0} \\
& \delta_{1} \leq\left(\delta_{0}+2 \rho_{0}\right)^{2}
\end{aligned}
$$

whence, by lemma 2 it follows that $\max \left\{\rho_{1}, \delta_{1}\right\} \leq \frac{1}{9} d^{2}$.
It can be easily proved by induction that the following relations hold for $k=0,1, .$.

- $x_{k} \in \bar{B}_{r}\left(x_{0}\right)$;
- $\delta_{k}:=\left\|I-F^{\prime}\left(x_{k}\right) \Gamma_{k}\right\| \leq \frac{1}{9} d^{2^{k}}$;
- $\rho_{k}:=\frac{100}{81} \beta^{2}\left\|F\left(x_{k}\right)\right\| \leq \frac{1}{9} d^{2^{k}}$;
- $\left\|x_{k+1}-x_{k}\right\| \leq \frac{d^{k^{k}}}{10 \beta}$.

From the above properties it results that the sequence $\left(x_{k}\right)_{k \geq 0}$ is a Cauchy one and therefore there exists $x^{*} \in \bar{B}_{r}\left(x_{0}\right)$ such that $\bar{x}^{*}=$ $\lim _{k \rightarrow \infty} x_{k}$. The last inequality above implies that for all $m \in \mathbb{N}$

$$
\left\|x_{k+m}-x_{k}\right\| \leq \sum_{i=k}^{k+m-1}\left\|x_{i+1}-x_{i}\right\| \leq \frac{d^{2^{k}}}{10 \beta\left(1-d^{2^{k}}\right)}, \quad k=0,1, . .
$$

which leads to the first estimation from the enounce.
The convergence of the sequence $\left(\Gamma_{k}\right)_{k \geq 0}$ is infered from the inequalities

$$
\begin{aligned}
\left\|\Gamma_{k+1}-\Gamma_{k}\right\| & =\left\|I-F^{\prime}\left(x_{k+1}\right) \Gamma_{k}\right\| \\
& \leq\left\|I-F^{\prime}\left(x_{k}\right) \Gamma_{k}\right\|+\left\|F^{\prime}\left(x_{k}\right)-F^{\prime}\left(x_{k+1}\right)\right\|\left\|\Gamma_{k}\right\| \\
& \leq \delta_{k}+2\left\|\Gamma_{k}\right\|^{2}\left\|F\left(x_{k}\right)\right\| \\
& \leq \delta_{k}+2 \rho_{k} \leq \frac{1}{3} d^{2^{k}}, \quad k=0,1, . .,
\end{aligned}
$$

which lead to the second stated estimation.
As we can see, the Newton-type method (4) has the $r$-convergence order at least 2. The conditions from the above theorem assure that the eigenvalue $\lambda^{*}=\lim _{k \rightarrow \infty} x_{k}^{(n+1)}$ is simple, according to lemma 1.

We shall consider now the following sequences given by the Chebyshev-type method, initially proposed by Diaconu [8]

$$
\begin{align*}
C_{k} & =B_{k}\left(2 I-F^{\prime}\left(x_{k}\right) B_{k}\right) \\
x_{k+1} & =x_{k}-C_{k} F\left(x_{k}\right)-\frac{1}{2} C_{k} F^{\prime \prime}\left(x_{k}\right)\left(C_{k} F\left(x_{k}\right)\right)^{2}  \tag{6}\\
B_{k+1} & =B_{k}\left[3 I-3 F^{\prime}\left(x_{k+1}\right) B_{k}+\left(F^{\prime}\left(x_{k+1}\right) B_{k}\right)^{2}\right], \quad k=0,1, . .,
\end{align*}
$$

where $x_{0} \in X$ and $B_{0} \in \mathcal{L}(X)$ are given.
For the study of the above method we need the following auxilliary result.

Lemma 4 [3] If the sequences of real positive numbers $\left(\delta_{k}\right)_{k \geq 0}$ and $\left(\rho_{k}\right)_{k \geq 0}$ satisfy

$$
\begin{aligned}
& \delta_{k+1} \leq\left(\delta_{k}+2 \rho_{k}+2 \rho_{k}^{2}\right)^{3} \\
& \rho_{k+1} \leq \rho_{k} \delta_{k}^{2}+\rho_{k}^{2} \delta_{k}^{2}+2 \rho_{k}^{3}+\rho_{k}^{4}, \quad k=0,1, . .
\end{aligned}
$$

where $\max \left\{\delta_{0}, \rho_{0}\right\} \leq \frac{1}{7} d$ for some $0<d<1$, then the following relation holds:

$$
\max \left\{\delta_{k}, \rho_{k}\right\} \leq \frac{1}{7} d^{3^{k}}, \quad k=0,1, \ldots
$$

As for the previous method, we shall consider the elements $x_{0} \in X$ and the ball $\bar{B}_{r}\left(x_{0}\right)$.

Theorem 5 Assume that the operator $F$ and the elements $x_{0} \in X$, $B_{0} \in \mathcal{L}(X)$ satisfy:
a) there exists $F^{\prime}\left(x_{0}\right)^{-1}$ and $\left\|F^{\prime}\left(x_{0}\right)^{-1}\right\| \leq \beta_{0}$;
b) $q=2 \beta_{0} r<1$;
c) denoting $\beta=\frac{\beta_{0}}{1-q}, a=\frac{64}{49} \beta, \delta_{0}=\left\|I-F^{\prime}\left(x_{0}\right) B_{0}\right\|$ and $\rho_{0}=$ $a^{2}\left\|F\left(x_{0}\right)\right\|$, suppose

$$
\max \left\{\delta_{0}, \rho_{0}\right\} \leq \frac{1}{7} d \quad \text { for some } d \in(0,1) ;
$$

d) $\frac{8 d}{49 a\left(1-d^{2}\right)} \leq r$.

Then the sequences $\left(x_{k}\right)_{k \geq 0},\left(B_{k}\right)_{k \geq 0},\left(C_{k}\right)_{k \geq 0} \quad$ converge and $\left(x_{k}\right)_{k \geq 0} \subset \bar{B}_{r}\left(x_{0}\right)$. Denoting $x^{*}=\lim x_{k}, B^{*}=\lim B_{k}, C^{*}=\lim C_{k}$,
then $F\left(x^{*}\right)=0$ and $B^{*}=C^{*}=F^{\prime}\left(x^{*}\right)^{-1}$. Moreover, the following estimations hold:

$$
\begin{aligned}
\left\|x^{*}-x_{k}\right\| & \leq \frac{8 d^{3^{k}}}{49 a\left(1-d^{2 \cdot 3 k}\right)} ; \\
\left\|F^{\prime}\left(x^{*}\right)^{-1}-B_{k}\right\| & \leq \frac{1656 a d^{3^{k}}}{2401\left(1-d^{2 \cdot 3^{k}}\right)}, \quad k=0,1, \ldots
\end{aligned}
$$

Proof. From a), b) and the Banach lemma it easily follows that for any $x \in \bar{B}_{r}\left(x_{0}\right)$ the Jacobian of $F$ is invertible and

$$
\left\|F^{\prime}(x)^{-1}\right\| \leq \frac{\beta_{0}}{1-q}=\beta .
$$

For the norms of $B_{0}$ and $C_{0}$, taking into account the hypotheses, we get

$$
\begin{aligned}
\left\|B_{0}\right\| & \leq\left\|B_{0}-F^{\prime}\left(x_{0}\right)^{-1}\right\|+\left\|F^{\prime}\left(x_{0}\right)^{-1}\right\| \\
& \leq\left\|F^{\prime}\left(x_{0}\right)^{-1}\right\|\left(1+\left\|I-F^{\prime}\left(x_{0}\right) B_{0}\right\|\right) \\
& \leq \beta_{0}\left(1+\delta_{0}\right) \leq \frac{8}{7} \beta_{0}<\frac{8}{7} \beta
\end{aligned}
$$

and

$$
\begin{aligned}
\left\|C_{0}\right\| & \leq\left\|B_{0}\right\|+\left\|I-F^{\prime}\left(x_{0}\right) B_{0}\right\| \cdot\left\|B_{0}\right\| \\
& \leq\left\|B_{0}\right\|\left(1+\delta_{0}\right) \leq \frac{64}{49} \beta=a
\end{aligned}
$$

so $\max \left\{\left\|B_{0}\right\|,\left\|C_{0}\right\|\right\} \leq a$.
From (6) we have that

$$
\begin{aligned}
\left\|x_{1}-x_{0}\right\| & \leq a\left(1+a^{2}\left\|F\left(x_{0}\right)\right\|\right)\left\|F\left(x_{0}\right)\right\| \\
& \leq a\left(1+\rho_{0}\right)\left\|F\left(x_{0}\right)\right\| \\
& <\frac{8}{7} a^{2}\left\|F\left(x_{0}\right)\right\| \\
& \leq \frac{\rho_{0}}{a^{2}} \cdot \frac{8}{7} a=\frac{8 d}{49 a},
\end{aligned}
$$

whence, taking into account d) it follows that $x_{1} \in \bar{B}_{r}\left(x_{0}\right)$.

Further, by the identity (3) and by (6) one obtains

$$
\begin{aligned}
\left\|F\left(x_{1}\right)\right\| & \leq\left\|I-F^{\prime}\left(x_{0}\right) C_{0}\right\|\left(1+\frac{1}{2}\left\|F^{\prime \prime}\left(x_{0}\right)\right\|\left\|C_{0}\right\|^{2}\left\|F^{\prime}\left(x_{0}\right)\right\|\right)\left\|F\left(x_{0}\right)\right\|+ \\
& +\frac{1}{2}\left\|F^{\prime \prime}\left(x_{0}\right)\right\|^{2}\left\|C_{0}\right\|^{4}\left\|F\left(x_{0}\right)\right\|^{3}+\frac{1}{8}\left\|F^{\prime \prime}\left(x_{0}\right)\right\|^{3}\left\|C_{0}\right\|^{6}\left\|F\left(x_{0}\right)\right\|^{4}
\end{aligned}
$$

whence

$$
\begin{aligned}
a^{2}\left\|F\left(x_{1}\right)\right\| & \leq a^{2}\left\|F\left(x_{0}\right)\right\| \cdot\left\|I-F^{\prime}\left(x_{0}\right) C_{0}\right\|\left(1+a^{2}\left\|F\left(x_{0}\right)\right\|\right)+ \\
& +2\left(a^{2}\left\|F\left(x_{0}\right)\right\|\right)^{3}+\left(a^{2}\left\|F\left(x_{0}\right)\right\|\right)^{4}
\end{aligned}
$$

Denoting $\rho_{1}=a^{2}\left\|F\left(x_{1}\right)\right\|$ and taking into account the inequality

$$
\left\|I-F^{\prime}\left(x_{0}\right) C_{0}\right\| \leq\left\|I-F^{\prime}\left(x_{0}\right) B_{0}\right\|^{2}=\delta_{0}^{2}
$$

it follows

$$
\rho_{1} \leq \rho_{0} \delta_{0}^{2}+\rho_{0}^{2} \delta_{0}^{2}+2 \rho_{0}^{3}+\rho_{0}^{4}
$$

From the third relation of (6) we get

$$
\begin{aligned}
\left\|I-F^{\prime}\left(x_{1}\right) B_{1}\right\| & =\left\|\left(I-F^{\prime}\left(x_{1}\right) B_{0}\right)^{3}\right\| \\
& \leq\left\|I-F^{\prime}\left(x_{1}\right) B_{0}\right\|^{3} \\
& \leq\left(\left\|I-F^{\prime}\left(x_{0}\right) B_{0}\right\|+2\left\|B_{0}\right\|\left\|x_{1}-x_{0}\right\|\right)^{3} \\
& \leq\left(\delta_{0}+2 \rho_{0}+2 \rho_{0}^{2}\right)^{3},
\end{aligned}
$$

i.e.,

$$
\delta_{1} \leq\left(\delta_{0}+2 \rho_{0}+2 \rho_{0}^{2}\right)^{3}
$$

By lemma 4 the above inequalities imply that $\max \left\{\rho_{1}, \delta_{1}\right\} \leq \frac{1}{7} d^{3}$. Assume now that the following properties hold:

- $x_{0}, x_{1}, . ., x_{k} \in \bar{B}_{r}\left(x_{0}\right)$;
- $\rho_{i}:=a^{2}\left\|F\left(x_{i}\right)\right\| \leq \frac{1}{7} d^{3^{i}}$ and $\delta_{i}:=\left\|I-F^{\prime}\left(x_{i}\right) B_{i}\right\| \leq \frac{1}{7} d^{3^{i}}, i=$ $0, . ., k$.

It easily follows that $\max \left\{\left\|B_{k}\right\|,\left\|C_{k}\right\|\right\} \leq a$ and

$$
\begin{aligned}
\left\|x_{k+1}-x_{k}\right\| & \leq a\left(1+a^{2}\left\|F\left(x_{k}\right)\right\|\right)\left\|F\left(x_{k}\right)\right\| \\
& \leq a\left(1+\rho_{k}\right)\left\|F\left(x_{k}\right)\right\| \\
& \leq \frac{8 \rho_{k}}{7 a} \leq \frac{8 d^{3}}{49 a} .
\end{aligned}
$$

From the above formula it follows that $x_{k+1} \in \bar{B}_{r}\left(x_{0}\right)$ :

$$
\left\|x_{k+1}-x_{0}\right\| \leq \frac{8 d}{49 a} \sum_{i=0}^{k} d^{3^{i}-1} \leq \frac{8 d}{49 a\left(1-d^{2}\right)} \leq r .
$$

Denoting $\rho_{k+1}=a^{2}\left\|F\left(x_{k+1}\right)\right\|$ and $\delta_{k+1}=\left\|I-F^{\prime}\left(x_{k+1}\right) B_{k+1}\right\|$, the following relations are obtained in the same manner as for $\rho_{1}$ and $\delta_{1}$ :

$$
\begin{aligned}
\rho_{k+1} & \leq \rho_{k} \delta_{k}^{2}+\rho_{k}^{2} \delta_{k}^{2}+2 \rho_{k}^{3}+\rho_{k}^{4} \\
\delta_{k+1} & \leq\left(\delta_{k}+2 \rho_{k}+2 \rho_{k}^{2}\right)^{3},
\end{aligned}
$$

whence, by lemma 4 , we get that $\max \left\{\rho_{k+1}, \delta_{k+1}\right\} \leq \frac{1}{7} d^{3^{k+1}}$ and the induction is proved.

We will show now that $\left(x_{k}\right)_{k \geq 0}$ is a Cauchy sequence. Indeed,

$$
\left\|x_{k+m}-x_{k}\right\| \leq \frac{8 d^{3^{k}}}{49 a} \sum_{i=k}^{k+m-1} d^{3^{i}} \leq \frac{8 d^{3^{k}}}{49 a\left(1-d^{2 \cdot 3^{k}}\right)},
$$

for all $k, m \in \mathbb{N}$, which implies that $\left(x_{k}\right)_{k \geq 0}$ converges. Denoting $x^{*}=$ $\lim _{k \rightarrow \infty} x_{k}$ we obtain

$$
\left\|x^{*}-x_{k}\right\| \leq \frac{8 d^{3^{k}}}{49 a\left(1-d^{2} \cdot 3^{k}\right)}, \quad k=0,1, \ldots
$$

The convergence of $\left(B_{k}\right)_{k \geq 0}$ is obtained from the third relation of (6):

$$
\begin{aligned}
\left\|B_{k+1}-B_{k}\right\| & \leq\left\|B_{k}\right\| \cdot\left\|2 I-F^{\prime}\left(x_{k+1}\right) B_{k}\right\| \cdot\left\|I-F^{\prime}\left(x_{k+1}\right) B_{k}\right\| \\
& \leq a\left(1+\delta_{k}+2 \rho_{k}+2 \rho_{k}^{2}\right)\left(\rho_{k}+2 \rho_{k}+2 \rho_{k}^{2}\right) \\
& \leq a \frac{1656}{2401} d^{3^{k}} .
\end{aligned}
$$

Denoting $B^{*}=\lim B_{k}$ it easily follows that $B^{*}=F^{\prime}\left(x^{*}\right)^{-1}$ and that

$$
\left\|F^{\prime}\left(x^{*}\right)^{-1}-B_{k}\right\| \leq \frac{1656 a d^{3^{k}}}{2401\left(1-d^{2 \cdot 3^{k}}\right)}, \quad k=0,1, \ldots
$$

The proof is completed.

## 3 Numerical examples

We shall consider two test matrices ${ }^{1}$ in order to study the behavior of the considered methods. The programs were written in Matlab ${ }^{2}$ and were run on a PC.

Pores1 matrix. This matrix arise from oil reservoir simulation. It is real, unsymmetric, of dimension 30 and has 20 real eigenvalues. We have chosen to study the largest eigenvalue $\lambda^{*}=-1.8363 e+1$. The initial approximation was taken $\lambda_{0}=\lambda^{*}+0.5$; for the initial vector $v_{0}$ we perturbed the solution $v^{*}$ (computed by Matlab and then properly scaled to fulfill the norming equation) with random vectors having the components uniformly distributed on $(-\varepsilon, \varepsilon), \varepsilon=0.2$. The initial matrices $\Gamma_{0}$ and $B_{0}$ were taken in each case as the inverse of $F^{\prime}\left(x_{0}\right)$ computed by Matlab. The following results are typical for the runs made (we have considered here the same vector $\varepsilon$ for the four initial approximations).

|  | Choice I |  | Choice II |  |
| :---: | :---: | :--- | :--- | :--- |
| $k$ | $\left\\|x^{*}-x_{k}\right\\|$ | $\left\\|F\left(x_{k}\right)\right\\|$ | $\left\\|x^{*}-x_{k}\right\\|$ | $\left\\|F\left(x_{k}\right)\right\\|$ |
| 0 | $7.8042 \mathrm{e}-01$ | $5.5828 \mathrm{e}+06$ | $7.8042 \mathrm{e}-01$ | $5.5828 \mathrm{e}+06$ |
| 1 | $1.4111 \mathrm{e}-01$ | $3.9355 \mathrm{e}-01$ | $2.3565 \mathrm{e}-02$ | $3.0227 \mathrm{e}-01$ |
| 2 | $1.8788 \mathrm{e}-02$ | $5.1300 \mathrm{e}-02$ | $4.6685 \mathrm{e}-05$ | $9.6691 \mathrm{e}-04$ |
| 3 | $3.7663 \mathrm{e}-04$ | $6.3248 \mathrm{e}-04$ | $5.7799 \mathrm{e}-10$ | $9.0156 \mathrm{e}-09$ |
| 4 | $1.4161 \mathrm{e}-07$ | $2.1373 \mathrm{e}-07$ |  |  |
| 5 | $4.5991 \mathrm{e}-10$ | $5.0482 \mathrm{e}-10$ |  |  |

[^1]Table 1. Newton-type method for Pores1.

|  | Choice I |  | Choice II |  |
| :---: | :---: | :--- | :--- | :--- |
| $k$ | $\left\\|x^{*}-x_{k}\right\\|$ | $\left\\|F\left(x_{k}\right)\right\\|$ | $\left\\|x^{*}-x_{k}\right\\|$ | $\left\\|F\left(x_{k}\right)\right\\|$ |
| 0 | $7.8042 \mathrm{e}-01$ | $5.5828 \mathrm{e}+06$ | $7.8042 \mathrm{e}-01$ | $5.5828 \mathrm{e}+06$ |
| 1 | $5.6679 \mathrm{e}-02$ | $1.0406 \mathrm{e}-01$ | $1.5461 \mathrm{e}-03$ | $1.2236 \mathrm{e}-02$ |
| 2 | $2.8973 \mathrm{e}-06$ | $4.0907 \mathrm{e}-06$ | $5.6407 \mathrm{e}-10$ | $5.5530 \mathrm{e}-09$ |
| 3 | $4.5959 \mathrm{e}-10$ | $6.5014 \mathrm{e}-10$ |  |  |

Table 2. Chebyshev-type method for Pores1.

FidAP002 matrix. This real symmetric matrix of dimension $n=$ 441 arise from finite element modeling. Its eigenvalues are all simple and range from $-7 \cdot 10^{8}$ to $3 \cdot 10^{6}$. We have chosen to study the smallest eigenvalue, which is well separated. The initial approximation was taken $\lambda_{0}=\lambda^{*}+10^{3}=-6.9996 \cdot 10^{8}+1000$; for the initial vector $v_{0}$ we perturbed the solution $v^{*}$ with random vectors having the components uniformly distributed on $(-\varepsilon, \varepsilon), \varepsilon=0.1$. The following results are typical for the runs made (we have considered a common vector $\varepsilon$ ).

|  | Choice I |  | Choice II |  |
| :--- | :--- | :--- | :--- | :--- |
| $k$ | $\left\\|x^{*}-x_{k}\right\\|$ | $\left\\|F\left(x_{k}\right)\right\\|$ | $\left\\|x^{*}-x_{k}\right\\|$ | $\left\\|F\left(x_{k}\right)\right\\|$ |
| 0 | $1.0000 \mathrm{e}+3$ | $8.5068 \mathrm{e}+8$ | $1.0000 \mathrm{e}+3$ | $8.5068 \mathrm{e}+8$ |
| 1 | $2.3611 \mathrm{e}+2$ | $1.5730 \mathrm{e}+3$ | $3.3415 \mathrm{e}+0$ | $1.2195 \mathrm{e}+3$ |
| 2 | $2.5528 \mathrm{e}+2$ | $8.8353 \mathrm{e}+2$ | $8.4714 \mathrm{e}-3$ | $1.0600 \mathrm{e}+0$ |
| 3 | $4.3758 \mathrm{e}+1$ | $8.2481 \mathrm{e}+1$ | $5.9605 \mathrm{e}-7$ | $4.1725 \mathrm{e}-6$ |
| 4 | $1.3141 \mathrm{e}+0$ | $2.0458 \mathrm{e}+0$ |  |  |
| 5 | $9.9051 \mathrm{e}-4$ | $1.4631 \mathrm{e}-3$ |  |  |
| 6 | $5.9605 \mathrm{e}-7$ | $1.0662 \mathrm{e}-7$ |  |  |

Table 3. Newton-type method for Fidap002.

Remarks on Some Newton and Chebyshev-type Methods for . . .

|  | Choice I |  | Choice II |  |
| :--- | :--- | :--- | :--- | :--- |
| $k$ | $\left\\|x^{*}-x_{k}\right\\|$ | $\left\\|F\left(x_{k}\right)\right\\|$ | $\left\\|x^{*}-x_{k}\right\\|$ | $\left\\|F\left(x_{k}\right)\right\\|$ |
| 0 | $1.0000 \mathrm{e}+3$ | $8.5068 \mathrm{e}+8$ | $1.0000 \mathrm{e}+3$ | $8.5068 \mathrm{e}+8$ |
| 1 | $3.9756 \mathrm{e}+2$ | $9.4690 \mathrm{e}+2$ | $8.5282 \mathrm{e}-1$ | $2.5509 \mathrm{e}+1$ |
| 2 | $4.6275 \mathrm{e}-1$ | $6.5442 \mathrm{e}-1$ | $7.1526 \mathrm{e}-7$ | $5.0390 \mathrm{e}-6$ |
| 3 | $4.7684 \mathrm{e}-7$ | $3.1597 \mathrm{e}-7$ |  |  |

Table 4. Chebyshev-type method for Fidap002.

If $\varepsilon$ was increased to 0.15 then the Newton-type method with choice I has not converged for some of the initial approximations, even if we took $\lambda^{*}$ for $\lambda_{0}$. The Newton-type method with choice II and the Chebyshev-type method converged as above. The explanation seem to reside in the fact that the eigenvector $v^{*}$ has a larger norm with choice II than with choice I, the relative error of $v^{*}+\varepsilon$ with the second choice being much smaller than of $v^{*}+\varepsilon$ with the first choice.

## References

[1] M.P. Anselone and L.B. Rall, The Solution of Characteristic Value-Vector Problems by Newton Method, Numer. Math., 11 (1968), pp. 38-45.
[2] E. Cǎtinaş and I. Pǎvǎloiu, On the Chebyshev Method for Approximating the Eigenvalues of Linear Operators, Rev. Anal. Numér. Théorie Approximation 25 (1996) 1-2, pp. 43-56.
[3] E. Cătinaş and I. Păvăloiu, On the Chebyshev Method for Approximating the Eigenvalues of Linear Operators, Proceedings of International Conference on Approximation and Optimization, ClujNapoca, July 29 - August 1, 1996, Vol. 1, pp. 219-226.
[4] E. Cǎtinaş and I. Pǎvǎloiu, On Approximating the Eigenvalues and Eigenvectors of Linear Continuous Operators, Rev. Anal. Numér. Théorie Approximation 25 (1996) 1-2, pp. 43-56.
[5] F. Chatelin, Valeurs propres de matrices, Mason, Paris, 1988.
[6] L. Collatz, Functionalanalysis und Numerische Mathematik, Springer-Verlag, Berlin, 1964.
[7] A. Diaconu and I. Pǎvǎloiu, Sur quelque méthodes itératives pour la résolution des équations opérationelles, Rev. Anal. Numér. Théorie Approximation 1, 1 (1972), pp. 45-61.
[8] A. Diaconu, On the Convergence of an Iterative Method of Chebyshev Type, Rev. Anal. Numér. Théorie Approximation, 24 (1995) 1-2, pp. 91-102.
[9] J.J. Dongarra, C.B. Moler and J.H. Wilkinson, Improving the Accuracy of the Computed Eigenvalues and Eigenvectors, SIAM J. Numer. Anal., 20 (1983) no. 1, pp. 23-45.
[10] S.M. Grzegórski, On the Scaled Newton Method for the Symmetric Eigenvalue Problem, Computing, 45 (1990), pp. 277-282.
[11] V. S. Kartîşov and F. L. Iuhno, O nekotorîh Modifikaţiah Metoda Niutona dlea Resenia Nelineinoi Spektralnoi Zadaci, J. Vîcisl. matem. i matem. fiz. 33 (1973) 9, pp. 1403-1409.
[12] J.M. Ortega and W.C. Rheinboldt, Iterative Solution of Nonlinear Equations in Several Variables, Academic Press, New York, 1970.
[13] I. Pǎvǎloiu, Sur les procédés itératifs à un ordre élevé de convergence, Mathematica (Cluj) 12 (35) (1970) 2, pp. 309-324.
[14] I. Pǎvǎloiu, Introduction to the Approximation Theory for the Solutions of Equations, Ed. Dacia, Cluj-Napoca, 1986 (in romanian).
[15] I. Pǎvǎloiu, Observations concerning some Approximation Methods for the Solutions of Operator Equations, Rev. Anal. Numér. Théorie Approximation, 23 (1994) 2, pp. 185-196.
[16] G. Peters and J.H. Wilkinson, Inverse Iteration, Ill-Conditioned Equations and Newton's Method, SIAM Review, 21 (1979) no. 3, pp. 339-360.
[17] M.C. Santos, A Note on the Newton Iteration for the Algebraic Eigenvalue Problem, SIAM J. Matrix Anal. Appl., 9 (1988) no. 4, pp. 561-569.
[18] R.A. Tapia and L.D. Whitley, The Projected Newton Method has Order $1+\sqrt{2}$ for the Symmetric Eigenvalue Problem, SIAM J. Numer. Anal., 25 (1988) 6, pp. 1376-1382.
[19] S. Ul'm, On the Iterative Method with Simultaneous Approximation of the Inverse of the Operator, Izv. Acad. Nauk. Estonskoi S.S.R., 16 (1967) 4, pp. 403-411.
[20] K. Wu, Y. Saad and A. Stathopoulos, Inexact Newton Preconditioning Techniques for Eigenvalue Problems, Lawrence Berkeley National Laboratory report number 41382 and Minnesota Supercomputing Institute report number UMSI 98-10, 1998.
[21] T. Yamamoto, Error Bounds for Computed Eigenvalues and Eigenvectors, Numer. Math. 34 (1980), pp. 189-199.
I.Pǎvǎloiu, E.Cǎtinaş,
"T. Popoviciu" Institute of
Numerical Analysis
P.O. Box 68

3400 Cluj-Napoca 1, Romania
e-mail: pavaloiu@ictp.math.ubbcluj.ro ecatinas@ictp.math.ubbcluj.ro


[^0]:    © 1999 by I.Păvăloiu, E.Cătinaş

[^1]:    ${ }^{1}$ These matrices are available from MatrixMarket at the following address: http://math.nist.gov/MatrixMarket/.
    ${ }^{2}$ MATLAB is a registered trademark of the MathWorks, Inc.

