REVUE D'ANALYSE NUMÉRIQUE ET DE THÉORIE DE L'APPROXIMATION
Tome 25, N°s 1–2, 1996, pp. 43–56

ON THE CHEBYSHEV METHOD FOR APPROXIMATING THE EIGENVALUES OF LINEAR OPERATORS

EMIL CĂTINAŞ and ION PĂVĂLOIU
(Cluj-Napoca)

1. INTRODUCTION

Approaches to the problem of approximating the eigenvalues of linear operators by the Newton method have been done in a series of papers ([1], [3], [4], [5]). There is a special interest in using the Newton method because the operatorial equation to be solved has a special form, as we shall see. We shall study in the following the convergence of the Chebyshev method attached to this problem and we shall apply the results obtained for the approximation of an eigenvalue and of a corresponding eigenvector for a matrix of real or complex numbers. It is known that the convergence order of Newton method is 2 and the convergence order of Chebyshev method is 3. Taking into account as well the number of operations made at each step, we obtain that the Chebyshev method is more efficient than the Newton method.

Let E be a Banach space over K, where $K = \mathbf{R}$ or $K = \mathbf{C}$, and $T : E \to E$ a linear operator. It is well known that the scalar λ is an eigenvalue of T if the equation

$$Tx - \lambda x = \theta \tag{1.1}$$

has at least one solution $\overline{x} \neq \theta$, where θ is the null element of the space E. The elements $x \neq \theta$ that satisfy equation (1.1) are called eigenvectors of the operator T, corresponding to the eigenvalue λ.

For the simultaneous determination of the eigenvalues and eigenvectors of T we can proceed in the following way.

We attach to equation (1.1) an equation of the form

$$Gx = 1 \tag{1.2}$$

where G is a linear functional $G : E \to K$.

Consider the real Banach space $F = E \times K$, with the norm given by

$$(1.3)\quad \|u\| = \max\{\|x\|, |\lambda|\}\ ,\quad u \in F\ ,\quad u = \begin{pmatrix} x \\ \lambda \end{pmatrix} \quad \text{with}\quad x \in E \quad \text{and} \quad \lambda \in K.$$

In this space we consider the operator $f : F \to F$ given by

$$(1.4)\qquad f\begin{pmatrix} x \\ \lambda \end{pmatrix} = \begin{pmatrix} Tx - \lambda x \\ Gx - 1 \end{pmatrix}.$$

If we denote by $\theta_1 = \begin{pmatrix} \theta \\ 0 \end{pmatrix}$ the null element of the space F, then the eigenvalues and the corresponding eigenvectors of the operator T are solutions of the equation

$$(1.5)\qquad f(u) = \theta_1.$$

Obviously, f is not a linear operator.

It can be easily seen that the first order Fréchet derivative of f has the following form [4]:

$$(1.6)\qquad f'(u_0)h = \begin{pmatrix} Th_1 - \lambda_0 h_1 - \lambda_1 x_0 \\ Gh_1 \end{pmatrix},$$

where $u_0 = \begin{pmatrix} x_0 \\ \lambda_0 \end{pmatrix}$ and $h = \begin{pmatrix} h_1 \\ \lambda_1 \end{pmatrix}$. For the second order derivative of f we obtain the expression

$$(1.7)\qquad f''(u_0)hk = \begin{pmatrix} -\lambda_2 h_1 - \lambda_1 h_2 \\ 0 \end{pmatrix},$$

where $k = \begin{pmatrix} h_2 \\ \lambda_2 \end{pmatrix}$.

The Fréchet derivatives of order higher than 2 of f are null.

Considering the above forms of the Fréchet derivatives of f, we shall study in the following the convergence of the Chebyshev method for the operators having the third order Fréchet derivative the null operator.

2. THE CONVERGENCE OF CHEBYSHEV METHOD

The iterative Chebyshev method for solving equation (1.5) consists in the successive construction of the elements of the sequence $(u_n)_{n\ge 0}$ given by

$$(2.1)\qquad u_{n+1} = u_n - \Gamma_n f(u_n) - \frac{1}{2}\Gamma_n f''(u_n)\left(\Gamma_n f(u_n)\right)^2,\quad n = 0,1,\ldots,\ u_0 \in F,$$

where $\Gamma_n = [f'(u_n)]^{-1}$.

Let $u_0 \in F$ and $\delta > 0$, $b > 0$ be two real numbers. Write $S = \left\{ u \in F \,\middle|\, \|u - u_0\| \le \delta \right\}$.

If $m_2 = \sup_{u\in S}\|f''(u)\|$, then $\sup_{u\in S}\|f'(u)\| \le \|f'(u_0)\| + m_2\delta$ and $\sup_{u\in S}\|f(u)\| \le \|f(u_0)\| + +\delta\|f'(u_0)\| + m_2\delta^2 = m_0$. Consider the numbers

$$(2.2)\qquad \mu = \frac{1}{2}m_2^2 b^4\left(1 + \frac{1}{4}m_2 m_0 b^2\right),\qquad \nu = b\left(1 + \frac{1}{2}m_2 m_0 b^2\right)$$

With the above notation, the following theorem holds:

THEOREM 2.1 *If the operator f is three times differentiable with $f'''(u) \equiv \theta_3$ for all $u \in S$ (θ_3 being the 3-linear null operator) and if, moreover, there exist $u_0 \in F$, $\delta > 0$, $b > 0$ such that the following relations hold*

i. *the operator $f'(u)$ has a bounded inverse for all $u \in S$, and*

$$\left\| f'(u)^{-1} \right\| \le b;$$

ii. *the numbers μ and ν given by (2.2) satisfy the relations*

$$\rho_0 = \sqrt{\mu}\,\|f(u_0)\| < 1$$

and

$$\frac{\nu\rho_0}{\sqrt{\mu}(1 - \rho_0)} \le \delta,$$

then the following properties hold:

j. $(u_n)_{n\ge 0}$ *given by (2.1) is convergent;*

jj. *if* $\bar{u} = \lim_{n\to\infty} u_n$, *then* $\bar{u} \in S$ *and* $f(\bar{u}) = \theta_1$;

jjj. $\|u_{n+1} - u_n\| \le \dfrac{\nu\rho_0^{3^n}}{\sqrt{\mu}},\quad n = 0,1,\ldots;$

jv. $\|\bar{u} - u_n\| \le \dfrac{\nu\rho_0^{3^n}}{\sqrt{\mu}\left(1 - \rho_0^{3^n}\right)},\quad n = 0,1,\ldots.$

Proof. Denote by $g : S \to F$ the following mapping:

$$(2.3)\qquad g(u) = -\Gamma(u)f(u) - \frac{1}{2}\Gamma(u)f''(u)\left[\Gamma(u)f(u)\right]^2,$$

where $\Gamma(u) = [f'(u)]^{-1}$.

It can be easily seen that for all $u \in S$ the following identity holds

$$f(u)+f'(u)g(u)+\frac{1}{2}f''(u)g^2(u)=$$

$$=\frac{1}{2}f''(u)\left([f'(u)]^{-1}f(u),[f'(u)]^{-1}f''(u)\left\{[f'(u)]^{-1}f(u)\right\}^2\right)+$$

$$+\frac{1}{8}f''(u)\left\{[f'(u)]^{-1}f''(u)\left\{[f'(u)]^{-1}f(u)\right\}^2\right\}^2$$

whence we obtain

$$(2.4)\qquad \left\|f(u)+f'(u)g(u)+\frac{1}{2}f''(u)g^2(u)\right\|\le\frac{1}{2}m_2^2b^4\left(1+\frac{1}{4}m_0m_2b^2\right)\|f(u)\|^3,$$

or

$$(2.5)\qquad \left\|f(u)+f'(u)g(u)+\frac{1}{2}f''(u)g^2(u)\right\|\le\mu\|f(u)\|^3,\quad \text{for all } u\in S.$$

Similarly, by (2.3) and taking into acount the notation we made, we get

$$(2.6)\qquad \|g(u)\|\le\nu\|f(u)\|,\ \text{for all } u\in S.$$

Using the hypotheses of the theorem, inequality (2.5) and the fact that $f'''(u)=\theta_3$, we obtain the following inequality:

$$\|f(u_1)\|\le\left\|f(u_1)-f(u_0)-f'(u_0)g(u_0)-\frac{1}{2}f''(u_0)g^2(u_0)\right\|+$$

$$+\left\|f(u_0)+f'(u_0)g(u_0)+\frac{1}{2}f''(u_0)g^2(u_0)\right\|\le\mu\|f(u_0)\|^3.$$

Since $u_1-u_0=g(u_0)$, by (2.5) we have

$$\|u_1-u_0\|\le\nu\|f(u_0)\|=\frac{\nu\sqrt{\mu}\|f(u_0)\|}{\sqrt{\mu}}<\frac{\nu\rho_0}{\sqrt{\mu}(1-\rho_0)}\le\delta,$$

whence it follows that $u_1\in S$.

Suppose now that the following properties hold:

a) $u_i\in S,\ i=\overline{0,k}$;

b) $\|f(u_i)\|\le\mu\|f(u_{i-1})\|^3,\ i=\overline{1,k}$.

By the fact that $u_k\in S$, using (2.5) it follows

$$(2.7)\qquad \|f(u_{k+1})\|\le\mu\|f(u_k)\|^3,$$

and from relation $u_{k+1}-u_k=g(u_k)$

$$(2.8)\qquad \|u_{k+1}-u_k\|<\nu\|f(u_k)\|.$$

The inequalities **b)** and (2.7) lead us to

$$(2.9)\qquad \|f(u_i)\|\le\frac{1}{\sqrt{\mu}}\left(\sqrt{\mu}\|f(u_0)\|\right)^{3^i},\ i=\overline{1,k+1}.$$

We have that $u_{k+1}\in S$:

$$\|u_{k+1}-u_0\|\le\sum_{i=1}^{k+1}\|u_i-u_{i-1}\|\le\sum_{i=1}^{k+1}\nu\|f(u_{i-1})\|\le\frac{\nu}{\sqrt{\mu}}\sum_{i=1}^{k+1}\rho_0^{3^{i-1}}\le\frac{\nu\rho_0}{(1-\rho_0)\sqrt{\mu}}.$$

Now we shall prove that the sequence $(u_n)_{n\ge0}$ is Cauchy. Indeed, for all $m, n\in\mathbf{N}$ we have

$$(2.10)\qquad \begin{aligned}\|u_{n+m}-u_n\|&\le\sum_{i=0}^{m-1}\|u_{n+i+1}-u_{n+i}\|\le\nu\sum_{i=0}^{m-1}\|f(u_{n+i})\|\le\\&\le\frac{\nu}{\sqrt{\mu}}\sum_{i=0}^{m-1}\rho_0^{3^{n+i}}=\frac{\nu}{\sqrt{\mu}}\rho_0^{3^n}\sum_{i=0}^{m-1}\rho_0^{3^{n+i}-3^n}\le\frac{\nu\rho_0^{3^n}}{\sqrt{\mu}\left(1-\rho_0^{3^n}\right)},\end{aligned}$$

whence, taking into account that $\rho_0<1$, it follows that $(u_n)_{n\ge0}$ converges. Let $\bar{u}=\lim_{n\to\infty}u_n$. Then, for $m\to\infty$ in (2.10) it follows **jv**. The consequence **jjj** follows from (2.8) and (2.9). □

3. THE APPROXIMATION OF THE EIGENVALUES AND EIGENVECTORS OF THE MATRICES

In the following we shall apply the previously studied method to the approximation of eigenvalues and eigenvectors of matrices with elements real or complex numbers.

Let $p\in\mathbf{N}$ and the matrix $A=(a_{ij})_{i,j=\overline{1,p}}$, where $a_{ij}\in K,\ i,j=\overline{1,p}$.

Using the above notation, we shall consider $E=K^p$ and $F=K^p\times K$. Any solution of equation

$$(3.1)\qquad f\begin{pmatrix}x\\\lambda\end{pmatrix}=\begin{pmatrix}Ax-\lambda x\\x_{i_0}-1\end{pmatrix}=\begin{pmatrix}\theta\\0\end{pmatrix},\quad i_0\in\{1,\dots,p\}\ \text{ being fixed,}$$

where $x=(x_1,\dots,x_p)\in K^p$ and $\theta=(0,\dots,0)\in K^p$, will lead us to an eigenvalue of A and to a corresponding eigenvector. For the sake of simplicity write $\lambda=x_{p+1}$, so that equation (3.1) is represented by the system

$$f_i(x_1,\dots,x_p,x_{p+1})=0,\quad i=\overline{1,p+1} \tag{3.2}$$

where

$$f_i(x_1,\dots,x_{p+1})=a_{i1}x_1+\dots+(a_{ii}-x_{p+1})x_i+\dots+a_{ip}x_p,\quad i=\overline{1,p} \tag{3.3}$$

and

$$f_{p+1}(x)=x_{i_0}-1. \tag{3.4}$$

Denote by P the mapping $P: K^{p+1}\to K^{p+1}$ defined by relations (3.3) and (3.4).

Let $\overline{x}_n=(x_1^n,\dots,x_{p+1}^n)\in K^{p+1}$. Then the first order Fréchet derivative of the operator P at $\overline{x}_n$ has the following form

$$P'(\overline{x}_n)h=\begin{pmatrix} a_{11}-x_{p+1}^n & a_{12} & \cdots & a_{1i_0} & \cdots & a_{1p} & -x_1^n \\ a_{21} & a_{22}-x_{p+1}^n & \cdots & a_{2i_0} & \cdots & a_{2p} & -x_2^n \\ \vdots & \vdots & & & & \vdots & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pi_0} & \cdots & a_{pp}-x_{p+1}^n & -x_p^n \\ 0 & 0 & \cdots & 1 & \cdots & 0 & 0 \end{pmatrix}\begin{pmatrix} h_1 \\ h_2 \\ \vdots \\ h_p \\ h_{p+1}\end{pmatrix}, \tag{3.5}$$

where $\overline{h}=(h_1,\dots,h_{p+1})\in K^{p+1}$. If we write $\overline{k}=(k_1,\dots,k_{p+1})\in K^{p+1}$ then for the second order Fréchet derivative we get

$$P''(\overline{x}_n)\overline{k}\,\overline{h}=\begin{pmatrix} -k_{p+1} & 0 & \cdots & 0 & -k_1 \\ 0 & -k_{p+1} & \cdots & 0 & -k_2 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & -k_{p+1} & -k_p \\ 0 & 0 & \cdots & 0 & 0 \end{pmatrix}\begin{pmatrix} h_1 \\ h_2 \\ \vdots \\ h_p \\ h_{p+1}\end{pmatrix}. \tag{3.6}$$

Denote by $\Gamma(\overline{x}_n)$ the inverse of the matrix attached to the operator $P'(x_n)$ and $\overline{u}_n=\Gamma(\overline{x}_n)P(\overline{x}_n)=(u_1^n,u_2^n,\dots,u_{p+1}^n)$. Let $\overline{v}_n=P''(\overline{x}_n)(\Gamma(\overline{x}_n)P(\overline{x}_n))^2=$ $=P''(\overline{x}_n)\overline{u}_n^2$. We obtain the following representation

$$\overline{v}_n=P''(\overline{x}_n)\overline{u}_n^2=\begin{pmatrix} -u_{p+1}^n & 0 & \cdots & 0 & -u_1^n \\ 0 & -u_{p+1}^n & \cdots & 0 & -u_2^n \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & -u_{p+1}^n & -u_p^n \\ 0 & 0 & \cdots & 0 & 0 \end{pmatrix}\begin{pmatrix} u_1^n \\ u_2^n \\ \vdots \\ u_p^n \\ u_{p+1}^n\end{pmatrix} \tag{3.7}$$

From this relation we can easily deduce the equalities

$$\begin{aligned} v_i^n &= -2u_{p+1}^n u_i^n,\quad i=\overline{1,p}; \\ v_{p+1}^n &= 0. \end{aligned} \tag{3.8}$$

Writing $\overline{w}_n=(w_1^n,w_2^n,\dots,w_{p+1}^n)=\Gamma(x_n)\overline{v}_n$ and supposing that $\overline{x}_n$ is an approximation of the solution of system (3.2), then the next approximation $\overline{x}_{n+1}$ given by method (2.1) is obtained by

$$\overline{x}_{n+1}=\overline{x}_n-\overline{u}_n-\frac{1}{2}\overline{w}_n,\quad i=0,1,\dots. \tag{3.9}$$

Consider K^p with the norm of an element $x=(x_1,\dots,x_p)$ given by the equality

$$\|x\|=\max_{1\le i\le p}\{|x_i|\}, \tag{3.10}$$

and consequently

$$\|A\|=\max_{1\le i\le p}\sum_{j=1}^{p}|a_{ij}|. \tag{3.11}$$

It can be easily seen that $\|P''(\overline{x}_n)\|=2$, for all $\overline{x}_n\in K^{p+1}$. Let $\overline{x}_0\in K^{p+1}$ be an initial approximation of the solution of system (3.2). Consider a real number $r>0$ and the set $\overline{S}=\{x\in K^{p+1}\,|\,\|x-\overline{x}_0\|\le r\}$. Write $m_0=\|P(\overline{x}_0)\|+r\|P'(\overline{x}_0)\|+2r^2$, $\overline{\mu}=2\overline{b}^4\left(1+\frac{1}{2}\overline{m}_0\overline{b}^2\right)$ and $\overline{\nu}=\overline{b}(1+\overline{m}_0\overline{b}^2)$, where $\overline{b}=\sup_{x\in\overline{S}}\|\Gamma(x)\|$, $\Gamma(x)$ being the inverse of the matrix attached to the operator $P'(x)$.

Taking into account the results obtained in Section 2 the following consequence holds:

COROLLARY 3.1 *If $\overline{x}_0\in K^{p+1}$ and $r\in\mathbf{R}$, $r>0$, are chosen such that the matrix attached to the operator $P'(x)$ is nonsingular for all $x\in\overline{S}$, and the following inequalities hold:*

$$\rho_0<\sqrt{\overline{\mu}}\,\|P(\overline{x}_0)\|<1$$

$$\frac{\overline{\nu}\rho_0}{\sqrt{\overline{\mu}}(1-\overline{\rho}_0)}\le r$$

then the following properties are true

j_1 *the sequence* $(\bar{x}_n)_{n\geq 0}$ *generated by* (3.9) *is convergent;*

jj_1 *if* $\bar{x} = \lim\limits_{n\to\infty} \bar{x}_n$, *then* $P(\bar{x}) = \theta_1 = (0,\dots,0) \in K^{p+1}$;

jjj_1 $\|\bar{x}_{n+1} - \bar{x}_n\| \leq \dfrac{\bar{\nu}\rho_0^{3^n}}{\sqrt{\bar{\mu}}}$, $n = 0, 1, \dots$;

jv_1 $\|\bar{x} - \bar{x}_n\| \leq \dfrac{\bar{\nu}\rho_0^{3^n}}{\sqrt{\bar{\mu}}\left(1 - \rho_0^{3^n}\right)}$, $n = 0, 1, \dots$.

Remark. If the radius r of the ball $\bar{S}$ is given and there exists $[P'(x_0)]^{-1}$ and $2r\left\|[P'(x_0)]^{-1}\right\| < 1$, then

$$\left\|[P'(x)]^{-1}\right\| \leq \frac{\left\|[P'(x_0)]^{-1}\right\|}{1 - 2r\left\|[P'(x_0)]^{-1}\right\|},$$

for all $x \in \bar{S}$ and in the above Corollary, taking into account the proof of Theorem 2.1, we can take

$$\bar{b} = \frac{\left\|[P'(x_0)]^{-1}\right\|}{1 - 2r\left\|[P'(x_0)]^{-1}\right\|}.$$

4. A COMPARISON WITH THE NEWTON METHOD

Note that if in (3.9) we neglect the term $-\frac{1}{2}\bar{w}_n$, then we get the Newton method:

$$\bar{x}_{n+1} = \bar{x}_n - \bar{u}_n = \bar{x}_n - \Gamma(\bar{x}_n)F(\bar{x}_n), \quad n = 0,1,\dots .$$

In order to compare the Newton method and the Chebyshev method, we shall consider that the linear systems which appear in both methods are solved by the Gauss method.

While the Newton method requires at each step the solving of a system $Ax = b$, the Chebyshev method requires the solving of two linear systems $Ax = b$, $Ay = c$ with the same matrix A and the vector c depending on the solution x of the first

system. So, we shall adapt the Gauss method in order to perform as few as possible multiplications and divisions. When comparing the two methods we shall neglect the number of addition and subtraction operations.

The solving of a given linear system $Ax = b$, $A \in M_m(K)$, $b, x \in K^m$, (where we have written $m = p + 1$) using the Gauss method consists in two stages. In the first stage, the given system is transformed into an equivalent one but with the matrix of the coefficients being upper triangular. In the second stage the unknowns $(x_i)_{i=\overline{1,m}}$ are determined by backward substitution.

The first stage. There are performed $m-1$ steps, at each one vanishing the elements on the same column below the main diagonal.

We write the initial system in the form

$$\begin{pmatrix} a_{11}^1 & \cdots & a_{1m}^1 \\ \vdots & & \vdots \\ a_{m1}^1 & \cdots & a_{mm}^1 \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} = \begin{pmatrix} b_1^1 \\ \vdots \\ b_m^1 \end{pmatrix}.$$

Suppose that $a_{11}^1 \neq 0$, a_{11}^1 being called the first pivote. The first line in the system is multiplied by $\alpha = -\dfrac{a_{21}^1}{a_{11}^1}$ and is added to the second one, which becomes $0, a_{22}^2, a_{23}^2, \dots, a_{2m}^2, b_2^2$, after performing $m + 1$ multiplication or division (M/D) operations. After $m-1$ such transformations, the system becomes

$$\begin{pmatrix} a_{11}^1 & a_{12}^1 & \cdots & a_{1m}^1 \\ 0 & a_{22}^2 & \cdots & a_{2m}^2 \\ \vdots & \vdots & & \vdots \\ 0 & a_{m2}^2 & \cdots & a_{mm}^2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix} = \begin{pmatrix} b_1^1 \\ b_2^2 \\ \vdots \\ b_m^2 \end{pmatrix}.$$

Hence at the first step there were performed $(m-1)(m+1)M/D$ operations. In the same manner, at the k-th step we have the system

$$\begin{pmatrix} a_{11}^1 & a_{12}^1 & \cdots & a_{1k}^1 & a_{1,k+1}^1 & \cdots & a_{1m}^1 \\ 0 & a_{22}^2 & \cdots & a_{2k}^2 & a_{2,k+1}^2 & \cdots & a_{2m}^2 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{kk}^k & a_{k,k+1}^k & \cdots & a_{km}^k \\ 0 & 0 & \cdots & a_{k+1,k}^k & a_{k+1,k+1}^k & \cdots & a_{k+1,m}^k \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{mk}^k & a_{m,k+1}^k & \cdots & a_{mm}^k \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \\ x_{k+1} \\ \vdots \\ x_m \end{pmatrix} = \begin{pmatrix} b_1^1 \\ b_2^2 \\ \vdots \\ b_k^k \\ b_{k+1}^k \\ \vdots \\ b_m^k \end{pmatrix}.$$

Supposing the k-th pivote $a_{kk}^k \neq 0$ and performing $(m-k)(m-k+2)\ M/D$ operations we get

$$\begin{pmatrix} a_{11}^1 & a_{12}^1 & \cdots & a_{1k}^1 & a_{1,k+1}^1 & \cdots & a_{1m}^1 \\ 0 & a_{22}^2 & \cdots & a_{2k}^2 & a_{2,k+1}^2 & \cdots & a_{2m}^2 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{kk}^k & a_{k,k+1}^k & \cdots & a_{k,m}^k \\ 0 & 0 & \cdots & 0 & a_{k+1,k+1}^{k+1} & \cdots & a_{k+1,m}^{k+1} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & a_{m,k+1}^{k+1} & \cdots & a_{mm}^{k+1} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \\ x_{k+1} \\ \vdots \\ x_m \end{pmatrix} = \begin{pmatrix} b_1^1 \\ b_2^2 \\ \vdots \\ b_k^k \\ b_{k+1}^{k+1} \\ \vdots \\ b_m^{k+1} \end{pmatrix}.$$

At each step k, the elements below the k-th pivote vanishing, they are not needed any more in the solving of the system.

The corresponding memory in the computer is used keeping in it the coefficients

$$-\frac{a_{k+1,k}^k}{a_{kk}^k}, \ldots, -\frac{a_{mk}^k}{a_{kk}^k},$$

which, of course, will be needed only for solving another system $Ay = c$, with c depending on x, the solution of $Ax = b$.

At the first stage there are performed

$$(m-1)(m+1) + \ldots + 1 \cdot 3 = \frac{2m^3 + 3m^2 - 5m}{6}\ M/D \text{ operations.}$$

The second stage. Given the system

$$\begin{pmatrix} a_{11}^1 & a_{12}^1 & \cdots & a_{1m}^1 \\ 0 & a_{22}^2 & \cdots & a_{2m}^2 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{mm}^m \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix} = \begin{pmatrix} b_1^1 \\ b_2^2 \\ \vdots \\ b_m^m \end{pmatrix}$$

the solution x is computed in the following way:

$$x_m = b_m^m / a_{mm}^m,$$
$$\vdots$$
$$x_k = \left(b_k^k - \left(a_{k,k+1}^k x_{k+1} + \ldots + a_{km}^k x_m\right)\right) / a_{kk}^k,$$
$$\vdots$$
$$x_1 = \left(b_1^1 - \left(a_{12}^1 x_2 + \ldots + a_{1m}^1 x_m\right)\right) / a_{11}^1.$$

At this stage there are performed $1 + 2 + \ldots + m = \dfrac{m(m+1)}{2}\ M/D$ operations. In both stages, there are totally performed

$$\frac{m^3}{3} + m^2 - \frac{m}{3}\ M/D \text{ operations.}$$

In the case when we solve the systems $Ax = b$, $Ay = c$, where the vector c depends on the solution x, we first apply the Gauss method for the system $Ax = b$ and at the first stage we keep below the main diagonal the coefficients by which the pivotes were multiplied.

Then we apply to the vector c the transformations performed to the vector b when solving $Ax = b$.

Write $c = \left(c_i^1\right)_{i=1,m}$.

At the first step

$$c_2^2 := a_{21} c_1^1 + c_2^1$$
$$\vdots$$
$$c_m^2 := a_{m1} c_1^1 + c_m^1.$$

At the k-th step

$$c_{k+1}^{k+1} := a_{k+1,k} c_k^k + c_{k+1}^k$$
$$\vdots$$
$$c_m^{k+1} := a_{mk} c_k^k + c_m^k.$$

At the m-th step

$$c_m^m := a_{m,m-1} c_{m-1}^{m-1} + c_m^{m-1}.$$

There were performed $m - 1 + m - 2 + \ldots + 1 = \dfrac{m(m-1)}{2}\ M/D$ operations.

Now the second stage of the Gauss method is applied to

$$\begin{pmatrix} a_{11}^1 & \cdots & a_{1m}^1 \\ \vdots & & \vdots \\ 0 & \cdots & a_{mm}^m \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} c_1^1 \\ \vdots \\ c_m^m \end{pmatrix}.$$

In addition to the case of a single linear system, in this case were performed $\dfrac{m(m-1)}{2}\ M/D$ operations, getting

$$\frac{m^3}{3} + \frac{3}{2}m^2 - \frac{5}{6}m \quad M/D \text{ operations},$$

and taking into account (3.8) we add $(m-1)$ more M/D operations, obtaining

$$\frac{m^3}{3} + \frac{3}{2}m^2 + \frac{m}{6} - 1 \quad M/D \text{ operations}.$$

Remark. At the first stage, if for some k we have $a_{kk}^k \approx 0$, then an element $a_{i_0,k}^k \neq 0$, $i_0 \in \{k+1, \ldots, m\}$ must be found, and the lines i_0 and k in A and b be swapped.

In order to avoid the error accumulations, a partial or total pivote strategy is recommended even if $a_{kk}^k \neq 0$ for $k = \overline{1, m-1}$.

For partial pivoting, the pivote is chosen such that $a_{i_0,k}^k = \max_{i=\overline{k,m}} |a_{ik}^k|$.

The effective interchange of lines can be bypassed by using a permutation vector $p = (p_i)_{i=\overline{1,m}}$, which is first initialized so that $p_i = i$. The elements in A and b are then referred to as $a_{ij} := a_{p(i),j}$ and $b_i := b_{p(i)}$, and swapping the lines k and i_0 is done by swapping the k-th and i_0-th elements in p.

For the Chebyshev method, the use of the vector p can't be avoided by the effective interchanging of the lines, because we must keep track for the permutations made, in order to apply them in the same order to the vector c.

Moreover we need two extra vectors t and q, in t storing the transpositions applied to the lines in $Ax = b$, and which are successively applied to q. At the first stage of the Gauss method, when the k-th pivote is $a_{i_0,k}^k$ and $i_0 \neq k$, the k-th and i_0-th elements in p are swapped, and we assign $t_k := i_0$ to indicate that at the k-th step we applied to p the transposition (k, i_0).

After computing the solution of $Ax = b$, we initialize the vector c by (3.7), the permutation vector q by $q_i := i$, $i := \overline{1, m}$, and then we successively apply the transforms operated to b, taking into account the eventual transpositions.

The algorithm is as follows:

```
for k:= 1 to m – 1 do
  begin
    if t[k] < > k
      then    {at the k-th step the transposition}
      begin   {(k, t[k]) has been applied to p}
        auxi : = q[k];
        q[k] := q[t[k]];
        q[t[k]] : = auxi;
      end;
    for i := k + 1 to m do c[q[i]] := c[q[i]] + A[q[i], k] * c[q[k]]
  end;
{the solution y is now computed}
for i : = m down to 1 do
  begin
    sum := 0;
    for j := i + 1 to m do sum := sum + A[p[i], j] * y[j]; {now p = q}
    y[i] := (c[p[i]] – sum) / A[p[i], i];
  end.
```


We adopt as the efficiency measure of an iterative method M the number

$$E(M) = \frac{\ln q}{s},$$

where q is the convergence order and s is the number of M/D operations needed at each step.

We obtain

$$E(N) = \frac{3\ln 2}{m^3 + 3m^2 - m}$$

for the Newton method and

$$E(C) = \frac{6\ln 3}{3m^3 + 9m^2 + m - 6}$$

for the Chebyshev method.

It can be easily seen that we have $E(C) > E(N)$ for $n \geq 2$, i.e. the Chebyshev method is more efficient than the Newton method.

5. NUMERICAL EXAMPLE

Consider the real matrix

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}$$

which has the following eigenvalues and eigenvectors

$$\lambda_{1,2,3} = 2, \quad x_1 = (1,1,0,0), \quad x_2 = (1,0,1,0), \quad x_3 = (1,0,0,1) \text{ and}$$

$$\lambda_4 = -2, \quad x_4 = (1,-1,-1,-1).$$

Taking the initial value $x_0 = (1, -1.5, -2, -1.5, -1)$, and applying the two methods we obtain the following results:

Newton method

n	x_1	x_2	x_3	x_4	$x_5 = \lambda$
0	1.0	−1.5000000000	−2.0000000000	−1.5000000000	−1.0000000000
1	1.0	−0.9000000000	−0.80000000000	−0.90000000000	−1.6000000000
2	1.0	−1.0125000000	−1.0250000000	−1.0125000000	−2.0500000000
3	1.0	−1.0001524390	−1.0003048780	−1.0001524390	−2.0006097561
4	1.0	−1.0000000232	−1.0000000465	−1.0000000232	−2.0000000929
5	1.0	−1.0000000000	−1.0000000000	−1.0000000000	−2.0000000000

Chebyshev method

n	x_1	x_2	x_3	x_4	$x_5 = \lambda$
0	1.0	−1.5000000000	−2.0000000000	−1.5000000000	−1.0000000000
1	1.0	−0.97200000000	−0.94400000000	−0.97200000000	−1.8880000000
2	1.0	−0.99995000189	−0.99990000377	−0.99995000189	−1.9998000075
3	1.0	−1.0000000000	−1.0000000000	−1.0000000000	−2.0000000000

REFERENCES

1. Anselone, M. P., Rall, B. L., *The Solution of Characteristic Value-Vector Problems by Newton Method*, Numer. Math. 11 (1968), 38-45.
2. Ciarlet, P.G., *Introduction à l'analyse numérique matricielle et à l'optimisation*, Mason, Paris Milan Barcelone Mexico,1990.
3. Chatelin, F., *Valeurs propres de matrices*, Mason, Paris Milan Barcelone Mexico,1988.
4. Collatz, L., *Functionalanalysis und Numerische Mathematik*, Berlin-Göttingen-Heidelberg, Springer-Verlag,1964.
5. Kartîşov, V. S., Iuhno, F. L., *O nekotoryh h Modifikatsiah Metoda Niutona dlea Resenia Nelineinoi Spektralnoi Zadaci*, J. Vîcisl. matem. i matem. fiz. (33) 9 (1973),1403-1409.
6. Păvăloiu, I., *Sur des procédés itératifs à un ordre élevé de convergence*, Mathematica (Cluj) 12 (35) 2 (1970), 309-324.
7. Traub, F. J., *Iterative Methods for the Solution of Equations*, Prentice-Hall Inc., Englewood Cliffs, N. J.,1964.

Received 07.03.1996

Romanian Academy
"T. Popoviciu" Institute of Numerical Analysis
P.O. Box 68 Cluj-Napoca 1
RO-3400
Romania