Jan Zítko

Improving the convergence of iterative methods

Aplikace matematiky, Vol. 28 (1983), No. 3, 215–229

Persistent URL: http://dml.cz/dmlcz/104028

IMPROVING THE CONVERGENCE OF ITERATIVE METHODS

JAN ZÍTKO

(Received October 1, 1982)

1. INTRODUCTION. FORMULATION OF THE PROBLEM

In the paper [1] an acceleration of the convergence of a sequence $\{x_n\}_{n=0}^{\infty}$, which is obtained from some convergent iterative process $x_{n+1} = Gx_n$, is investigated. This process is used for solving an operator equation $Ax = b$ in Hilbert space with a linear operator A. For two iterations x_n and x_{n-1} the author constructs a new element $\bar{x}_n = x_n - c_n(x_n - x_{n-1})$ such that for the real constant c_n the norm $\|\bar{r}_n\| = \|b - A\bar{x}_n\|$ is minimized.

The aim of this paper it to present a general theoretical investigation of one class of methods for acceleration of convergence. This class extends essentially the extrapolation procedure presented in the paper [1]. The computational procedures which have been developed for iterative solution of operator equations $x = Tx + b$ on the basis of this theoretical investigation do not require the explicit knowledge of the spectrum or the spectral radius of the operator. T.

By C^n we denote the complex linear space of all column vectors $\mathbf{x} = (x_1, \ldots, x_n)^{\mathsf{T}}$ with complex components. The superscript T is used for transpose and H for conjugate and transpose. The vector $\mathbf{e}_i(n)$ is the i-th column of the $n \times n$ identity matrix $\mathbf{I}$ and

$$\mathbf{e}(n) = \sum_{i=1}^{n} \mathbf{e}_i(n) = (1, 1, \ldots, 1)^{\mathsf{T}}.$$

If $\mathbf{u}_i \in C^n$, $i = 1, 2, \ldots, s$, then $(\mathbf{u}_1, \ldots, \mathbf{u}_s)$ is the matrix with columns $\mathbf{u}_i$. By R^n we denote the real n-dimensional linear space. $\boldsymbol{\Theta}(n)$ denotes the null vector in R^n or C^n.

Let X be a linear space with inner product $(\cdot, \cdot)$. Let

$$T : X \to X, \quad H : X \to X$$

be linear operators. We will consider the operator equation

$$x = Tx + b \tag{1}$$

and an iterative process

(2) $$x_{n+1} = Tx_n + b\,,$$

where b is a fixed element of X. Starting from an $x_0 \in X$, form the sequence $\{x_n\}_{n=0}^{\infty}$ according to the formula (2) and let this sequence be convergent with a limit x^*. Let $l > 1$, k, m_0, m_1, ..., m_l be integers and let the inequalities

(3) $$m_l > m_{l-1} > \ldots > m_1 > m_0 = 0\,,$$

(4) $$k > m_l$$

be valid. The above presented notational conventions and relations are assumed to be valid throughout this paper.

The problem to be considered in this paper is that of finding complex numbers $\alpha_0^{(k)}, \ldots, \alpha_l^{(k)}$ to satisfy

(5) $$\sum_{i=0}^{l} \alpha_i^{(k)} = 1\,,$$

(6) $$\left\| H\left(x^* - \sum_{i=0}^{l} \alpha_i^{(k)} x_{k-m_i}\right)\right\| = \min_{\substack{\boldsymbol{\beta} \in \mathfrak{M} \\ \boldsymbol{\beta} = (\beta_0, \ldots, \beta_l)^{\mathrm{T}}}} \left\| H\left(x^* - \sum_{i=0}^{l} \beta_i x_{k-m_i}\right)\right\|\,,$$

where $\mathfrak{M} \subset C^{l+1}$ is a set of all vectors satisfying the relation $\sum_{i=0}^{l} \beta_i = 1$. (The norm is defined by the inner product in a usual way.)

The operator H should be chosen in such a form that it is possible to evaluate the expression in the norm. One special case is shown at the end of this paper.

We shall show that on the basis of certain assumptions there exists just one vector $\boldsymbol{\alpha}^{(k)} = (\alpha_0^{(k)}, \ldots, \alpha_l^{(k)})^{\mathrm{T}} \in C^{l+1}$ whose components are solutions to the problem (5), (6) and we shall give various formulas for the calculation of the $\alpha_i^{(k)}$. Moreover, we shall investigate some relations between the structure of the spectrum of T and the possibilities of constructing the $\boldsymbol{\alpha}^{(k)}$. The components of the vector $\boldsymbol{\alpha}^{(k)}$ will be called **coefficients of extrapolation.**

In forthcoming papers we shall investigate the convergence of the coefficients $\boldsymbol{\alpha}^{(k)}$ for $k \to \infty$. Moreover, we shall present special detailed extrapolation procedures and show that the sequence $\{y_k\}$ constructed by the formula $y_k = \sum_{i=0}^{k} \alpha_i^{(k)} x_{k-m_i}$ converges faster to x^* than the sequence $\{x_k\}$. Some remarks concerning the calculation are to be found at the end of this paper.

2. AUXILIARY THEOREMS

Let

$$\boldsymbol{M}_k = (\mu_0, \mu_1, \ldots, \mu_t)\,, \quad \mu_i \in X\,, \quad i = 0, \ldots, t\,,$$

$$\boldsymbol{N}_k = (\nu_0, \nu_1, \ldots, \nu_s)\,, \quad \nu_i \in X\,, \quad i = 0, \ldots, s$$

be two row vector with elements in X. By $N_k \otimes M_k$ we denote the complex $(s + 1) \times \times (t + 1)$ matrix of the form

$$\begin{pmatrix} (\mu_0, \nu_0), (\mu_1, \nu_0), \ldots, (\mu_t, \nu_0) \\ (\mu_0, \nu_1), (\mu_1, \nu_1), \ldots, (\mu_t, \nu_1) \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ (\mu_0, \nu_s), (\mu_1, \nu_s), \ldots, (\mu_t, \nu_s) \end{pmatrix}.$$

Further, we put

(7) $$\varepsilon_i = x^* - x_i, \ \eta_i = H\varepsilon_i ,$$

(7′) $$\boldsymbol{b} = (\beta_0, \beta_1, \ldots, \beta_l)^{\mathrm{T}} \in C^{l+1} ,$$

(7″) $$\boldsymbol{H}_k = (\eta_k, \eta_{k-m_1}, \ldots, \eta_{k-m_l}) ,$$

(7‴) $$\mathbf{Q}_k = \boldsymbol{H}_k \otimes \boldsymbol{H}_k .$$

Let us observe that the vector $\boldsymbol{\alpha}^{(k)}$ is a minimizing vector of the function f,

(8) $$f(\beta_0, \beta_1, \ldots, \beta_l) = \mathbf{b}^{\mathrm{H}} \mathbf{Q}_k \boldsymbol{b} ,$$

where the vector $\boldsymbol{b} = (\beta_0, \ldots, \beta_l)^{\mathrm{T}}$ is to be found on the set $\mathfrak{M}$.

If $\boldsymbol{A} = (a_{ij})_{i,j=1,n}$ is an $n \times n$ complex matrix then $\operatorname{Re} \boldsymbol{A}$ and $\operatorname{Im} \boldsymbol{A}$ are real $n \times n$ matrices $(\operatorname{Re} a_{ij})$ and $(\operatorname{Im} a_{ij})_{i,j=1,\ldots,n}$ and $\boldsymbol{A}^{(*)}$ is a $2n \times 2n$ real matrix

$$\begin{pmatrix} \operatorname{Re} \boldsymbol{A}, & -\operatorname{Im} \boldsymbol{A} \\ \operatorname{Im} \boldsymbol{A}, & \operatorname{Re} \boldsymbol{A} \end{pmatrix}.$$

Lemma 1. *Let $v_1, \ldots, v_n$ be n linearly independent elements from X and let us put*

$$N_1 = (v_1, v_2, \ldots, v_n) ,$$

$$N_2 = (v_1 - v_2, v_2 - v_3, \ldots, v_{n-1} - v_n) .$$

Then the following hold:

1) *The elements*

(9) $$v_1 - v_2, v_2 - v_3, \ldots, v_{n-1} - v_n$$

are linearly independent.

2) *The matrix $N_2 \otimes N_2$ is positive definite.*

3) *The matrix $(N_2 \otimes N_2)^{(*)}$ is symmetric and positive definite.*

4) *The $n \times n$ matrix* $\boldsymbol{M}$

$$\boldsymbol{M} = \begin{pmatrix} N_2 \otimes N_1 \\ \mathbf{e}(n)^{\mathrm{T}} \end{pmatrix}$$

is nonsingular.

5) *The matrix $\boldsymbol{M}^{(*)}$ is nonsingular.*

Remark. The matrix $N_2 \circledast N_1$ is an $(n - 1) \times n$ matrix.

Proof. Let $(\gamma_1, \ldots, \gamma_{n-1})^{\mathsf{T}}$ be an arbitrary vector from C^{n-1} and put

$$(10) \qquad (\delta_1, \ldots, \delta_n)^{\mathsf{T}} = \mathbf{K}^{\mathsf{T}}(\gamma_1, \ldots, \gamma_{n-1}, -\gamma_{n-1})^{\mathsf{T}},$$

where $\mathbf{K} = (\mathbf{e}_1(n), \mathbf{e}_2(n) - \mathbf{e}_1(n), \ldots, \mathbf{e}_{n-1}(n) - \mathbf{e}_{n-2}(n), \mathbf{e}_n(n))$ is a nonsingular matrix. From the identity

$$\sum_{i=1}^{n-1} \gamma_i(v_i - v_{i+1}) = \sum_{i=1}^{n} \delta_i v_i ,$$

from (10) and from the independence of $v_1, \ldots, v_n$ it follows easily that the elements (9) are linearly independent.

The statements of 2) and 3) are evident.

It is obvious that det $\boldsymbol{M}$ = det $(N_2 \circledast N_2)$, hence $\boldsymbol{M}$ is nonsingular.

Finally, let $\boldsymbol{M}^{(*)}\boldsymbol{u} = \Theta(2n)$, where

$$\boldsymbol{u} = \begin{pmatrix} \boldsymbol{u}_1 \\ \boldsymbol{u}_2 \end{pmatrix} \in R^{2n}$$

with $\boldsymbol{u}_1, \boldsymbol{u}_2 \in R^n$. Then the vector $\boldsymbol{v} = \boldsymbol{u}_1 + \mathrm{i}\boldsymbol{u}_2 \in C^n$ solves the equation $\boldsymbol{M}\boldsymbol{v} = \Theta(n)$ and conversely, if $\boldsymbol{v}$ solves the latter equation, then $\boldsymbol{u}$ solves the former one. But $\boldsymbol{M}$ is nonsingular and therefore $\boldsymbol{M}^*$ is nonsingular.

Theorem 1. *Let $v_1, \ldots, v_n$ be linearly independent elements in X. We put*

$$\boldsymbol{A} = (v_1, \ldots, v_n) \circledast (v_1, \ldots, v_n) .$$

Let us define for every vector $\mathbf{x} = (\vartheta_1, \ldots, \vartheta_n)^{\mathsf{T}} \in C^n$ the function f:

$$(11) \qquad f(\vartheta_1, \ldots, \vartheta_n) = \mathbf{x}^{\mathsf{H}}\boldsymbol{A}\mathbf{x}.$$

Finally, we define the set $\mathcal{M}$:

$$\mathcal{M} = \{(\vartheta_1, \ldots, \vartheta_n)^{\mathsf{T}} \in C^n \mid \vartheta_1 + \ldots + \vartheta_n = 1\} .$$

Then

1) *there exists one and only one vector $\mathbf{z} = (\zeta_1, \ldots, \zeta_n)^{\mathsf{T}} \in \mathcal{M}$ such that*

$$f(\zeta_1, \ldots, \zeta_n) = \min_{(\eta_1, \ldots, \eta_n)^{\mathsf{T}} \in \mathcal{M}} f(\eta_1, \ldots, \eta_n) .$$

2) *For the vector $\mathbf{z}$ from 1) we have the formula*

$$(12) \qquad \mathbf{z} = (\mathbf{e}(n)^{\mathsf{T}} \boldsymbol{A}^{-1} \mathbf{e}(n))^{-1} \boldsymbol{A}^{-1} \mathbf{e}(n) .$$

3) *If we denote*

$$\boldsymbol{A}_2 = (v_1 - v_2, \ldots, v_{n-1} - v_n) \circledast (v_1, \ldots, v_n)$$

and

$$\mathbf{A}_1 = \begin{pmatrix} \mathbf{A}_2 \\ \mathbf{e}(n)^\mathsf{T} \end{pmatrix},$$

then the matrix $\mathbf{A}_1$ *is nonsingular and the vector* $\mathbf{z}$ *from* 1) *which minimizes the function* f *is the solution of the equation*

$$\mathbf{A}_1 \mathbf{z} = \mathbf{e}_n(n)\,. \tag{12'}$$

Proof. Put $\mathbf{B} = \mathbf{A}^{(*)}$, $\xi_i = \operatorname{Re} \vartheta_i$, $\xi_{i+n} = \operatorname{Im} \vartheta_i$, $\mathbf{u} = (\xi_1, \ldots, \xi_n)^\mathsf{T}$, $\mathbf{v} = (\xi_{n+1}, \ldots \ldots, \xi_{2n})^\mathsf{T}$, where $(\vartheta_1, \ldots, \vartheta_n)^\mathsf{T} \in C^n$ is an arbitrary element. From (11) we obtain

$$f(\vartheta_1, \ldots, \vartheta_n) = (\mathbf{u}^\mathsf{T}, \mathbf{v}^\mathsf{T})\, \mathbf{B} \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix} \equiv F(\xi_1, \ldots, \xi_{2n})\,, \tag{13}$$

where $\begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix} \in R^{2n}$ and $(\mathbf{u}^\mathsf{T}, \mathbf{v}^\mathsf{T})$ is evidently its transpose.

Let g_1, g_2 be two real functions of $2n$ real variables defined by the relations

$$g_1(\xi_1, \ldots, \xi_{2n}) = \xi_1 + \ldots + \xi_n - 1$$

$$g_2(\xi_1, \ldots, \xi_{2n}) = \xi_{n+1} + \ldots + \xi_{2n}\,.$$

Let $\mathcal{N}$ be the set of all vectors $(\xi_1, \ldots, \xi_{2n})^\mathsf{T} \in R^{2n}$ which satisfy the conditions

$$g_1(\xi_1, \ldots, \xi_{2n}) = 0\,, \tag{14}$$

$$g_2(\xi_1, \ldots, \xi_{2n}) = 0\,. \tag{14'}$$

We have obtained an equivalent problem, i.e., to find a minimum of the function F on the set $\mathcal{N}$. Using the well known theorem from analysis, we first solve the system of equations

$$\frac{\partial F(\mathbf{z}_1)}{\partial \xi_j} + r_1 \frac{\partial g_1(\mathbf{z}_1)}{\partial \xi_j} + r_2 \frac{\partial g_2(\mathbf{z}_1)}{\partial \xi_j} = 0\,, \tag{15}$$

$$g_1(\mathbf{z}_1) = 0\,, \quad g_2(\mathbf{z}_1) = 0 \tag{16}$$

for $j = 1, \ldots, 2n$, where we have put $\mathbf{z}_1 = (\xi_1, \xi_2, \ldots, \xi_{2n})$, or equivalently

$$\mathbf{B}\mathbf{z}_1^\mathsf{T} + \frac{r_1}{2} \begin{pmatrix} \mathbf{e}(n) \\ \boldsymbol{\Theta}(n) \end{pmatrix} + \frac{r_2}{2} \begin{pmatrix} \boldsymbol{\Theta}(n) \\ \mathbf{e}(n) \end{pmatrix} = \boldsymbol{\Theta}(2n)\,, \tag{17}$$

$$\mathbf{e}(n)^\mathsf{T} (\xi_1, \ldots, \xi_n)^\mathsf{T} = 1\,, \quad \mathbf{e}(n)^\mathsf{T} (\xi_{n+1}, \ldots, \xi_{2n})^\mathsf{T} = 0\,,$$

where r_1 and r_2 are multiplicators. Substituting here $\lambda = \frac{1}{2} r_1 + \mathrm{i}\, \frac{1}{2} r_2$ and $\mathbf{z} = (\zeta_1, \ldots, \zeta_n)^\mathsf{T}$, where $\zeta_j = \xi_j + \mathrm{i}\xi_{j+n}$, we obtain from (17) an equivalent form of the equations (17), i.e.

$$\mathbf{A}\mathbf{z} + \lambda\, \mathbf{e}(n) = \Theta(n)\,, \tag{18}$$

$$\mathbf{e}(n)^{\mathsf{T}}\, \mathbf{z} = 1\,. \tag{18'}$$

The system (18), (18′) has a unique solution

$$\lambda = -(\mathbf{e}(n)^{\mathsf{T}}\, \mathbf{A}^{-1}\, \mathbf{e}(n))^{-1}\,; \quad \mathbf{z} = \frac{\mathbf{A}^{-1}\, \mathbf{e}(n)}{\mathbf{e}(n)^{\mathsf{T}}\, \mathbf{A}^{-1}\, \mathbf{e}(n)}\,, \tag{19}$$

which implies that the system (15) and (16) or (17) has a unique solution. Since the quadratic form

$$\sum_{j,m=1}^{2n} \left(\frac{\partial^2 F(\mathbf{z}_1)}{\partial \xi_j\, \partial \xi_m} + r_1 \frac{\partial^2 g_1(\mathbf{z}_1)}{\partial \xi_j\, \partial \xi_m} + r_2 \frac{\partial^2 g_2(\mathbf{z}_1)}{\partial \xi_j\, \partial \xi_m} \right) \tau_j \tau_m = \mathbf{y}^{\mathsf{T}}\, \mathbf{B}\mathbf{y}\,,$$

where $\mathbf{y} = (\tau_1, \ldots, \tau_{2n})^{\mathsf{T}} \in R^{2n}$, is positive definite, it follows that $\mathbf{z}_1$ is the minimum of the function F on $\mathcal{N}$ or equivalently $\mathbf{z}^{\mathsf{T}}$ is the unique minimum of the function f on $\mathcal{M}$.

By successive elimination of λ in the system (18) and by omitting the last equation we obtain together with (18′) directly the system (12′). The regularity of $\mathbf{A}_1$ follows immediately from Lemma 1. □

3. CONSTRUCTION OF THE NUMBERS $\alpha_i^{(k)}$

Let $\{u_k\}_{k=0}^{\infty} \subset X$. The notation $\delta_{ij} u_k$ and $\delta_i u_k$ will be used for differences

$$\delta_{ij} u_k = u_{k-m_{i-1}} - u_{k-m_j}\,, \tag{20}$$

$$\delta_i u_k = \delta_{ii} u_k = u_{k-m_{i-1}} - u_{k-m_i}\,, \tag{20'}$$

where the numbers m_i and l have been defined at the beginning of this paper. Further, putting

$$\boldsymbol{L}_k = (\delta_1 \eta_k, \delta_2 \eta_k, \ldots, \delta_l \eta_k)\,,$$

we can introduce the matrix

$$\mathbf{S}_k = \begin{pmatrix} \boldsymbol{L}_k \otimes \boldsymbol{H}_k \\ \mathbf{e}(l+1)^{\mathsf{T}} \end{pmatrix}.$$

Let us remark that we have defined η_k, $\boldsymbol{H}_k$ by formulae (7), (7′). A procedure for the construction of coefficients of extrapolation which we have denoted by $\alpha_i^{(k)}$ immediately follows from Theorem 1 and Lemma 1.

Theorem 2. *Let X be a linear space with inner product. Let $T : X \to X$, $H : X \to X$ be linear operators. Assume that for $x_0 \in X$ the sequence $\{x_n\}_{n=0}^{\infty}$ obtained from (2) tends to a limit x^*. Assume further that integers $l > 0$, $m_0, \ldots, m_l$ and k fulfil the inequalities (3) and (4).*

If the matrix $\mathbf{Q}_k$ *is positive definite, then there exists one and only one vector* $\boldsymbol{\alpha}^{(k)} = (\alpha_0^{(k)}, \alpha_1^{(k)}, \ldots, \alpha_l^{(k)})^{\mathsf{T}}$ *which solves the problem* (5), (6). *For the vector* $\boldsymbol{\alpha}^{(k)}$ *we have*

$$\boldsymbol{\alpha}^{(k)} = (\mathbf{e}(n)^{\mathsf{T}}\, \mathbf{Q}_k^{-1}\, \mathbf{e}(n))^{-1}\, \mathbf{Q}_k^{-1}\, \mathbf{e}(n) \tag{21}$$

or

$$\mathbf{S}_k \boldsymbol{\alpha}^{(k)} = \mathbf{e}_{l+1}(l+1)\,. \tag{21'}$$

Putting $\boldsymbol{\alpha}^{(k)} = \xi^{(k)} + \mathrm{i}\zeta^{(k)}$, *we can write* (21') *in the form*

$$\mathbf{S}_k^{(*)} \begin{pmatrix} \xi^{(k)} \\ \zeta^{(k)} \end{pmatrix} = \begin{pmatrix} \mathbf{e}_{l+1}(l+1) \\ \boldsymbol{\Theta}(l+1) \end{pmatrix}. \tag{22}$$

Proof. The statement of this theorem is an immediate consequence of Lemma 1 and Theorem 1. □

Example. Let $l = 1$, $m_1 = 1$. Then the system (21) or (21') has a solution

$$\alpha_0^{(k)} = -\frac{(\eta_{k-1}, \nabla\eta_k)}{(\nabla\eta_k, \nabla\eta_k)}\,; \quad \alpha_1^{(k)} = \frac{(\eta_k, \nabla\eta_k)}{(\nabla\eta_k, \nabla\eta_k)}\,,$$

where

$$\nabla\eta_k = \eta_k - \eta_{k-1}\,.$$

4. WHEN IS THE MATRIX $\mathbf{Q}_k$ POSITIVE DEFINITE?

Let us use the notation and assumptions of Sections 1–3; moreover, let X be Hilbert space, T a bounded linear operator and let H^{-1} exist. The spectrum of T has the following structure: There exist finite sequences $\{i_k\}_{k=1}^r$ of positive integers and $\{\lambda_k\}_{k=1}^r \subset C$ for some integer $r > 1$ such that each λ_k is a pole of the resolvent operator of order i_k,

$$|\lambda_1| \geqq |\lambda_2| \geqq \ldots \geqq |\lambda_r| > 0\,, \tag{23}$$

$$\lambda_i \neq \lambda_j \quad \text{for} \quad i \neq j$$

and

$$\{\lambda \in \sigma(T)\,, \quad \lambda \neq \lambda_i\,, \quad i = 1, \ldots, r\} \Rightarrow |\lambda| < |\lambda_r|\,.$$

For a fixed $j \in \langle 1, r\rangle$ let C_j be the circumference with center λ_j and radius $\varrho_j > 0$ such that

$$\{\lambda \in C \mid |\lambda - \lambda_j| \leqq \varrho_j\} \cap \sigma(T) = \{\lambda_j\}\,.$$

Moreover, let $K = \{\lambda \in C \mid |\lambda| = \tau\}$, where $\tau > r(T)$ and $K_C = \{\lambda \in C \mid |\lambda| = \varrho_C\}$, where ϱ_C is taken such that

$$\{\lambda \mid |\lambda| \leqq \varrho_C\} \cap \sigma(T) = \sigma(T) \dot{-} \{\lambda_1, \ldots, \lambda_r\}\,. \tag{23'}$$

Without any loss of generality we can assume that

(24) $$B_{ji_j}\varepsilon_0 \neq 0 \quad \text{for all} \quad j = 1, 2, \ldots, r\,,$$

and

(25) $$k > \max_{j=1,2,\ldots,r} (i_j)\,,$$

where

$$B_{ji} = \frac{1}{2\pi \mathrm{i}} \int_{C_j} (\lambda - \lambda_j)^{i-1} R(\lambda, T)\, \mathrm{d}\lambda\,.$$

From (1), (2) and (7) it follows that

(26) $$\varepsilon_k = T^k \varepsilon_0 = \frac{1}{2\pi \mathrm{i}} \int_K \lambda^k R(\lambda, T)\, \varepsilon_0\, \mathrm{d}\lambda = \frac{1}{2\pi \mathrm{i}} \sum_{j=1}^{r} \int_{C_j} \lambda^k R(\lambda, T)\, \varepsilon_0\, \mathrm{d}\lambda +$$

$$+ \frac{1}{2\pi \mathrm{i}} \int_{K_C} \lambda^k R(\lambda, T)\, \varepsilon_0\, \mathrm{d}\lambda = \sum_{j=1}^{r} \sum_{i=1}^{i_j} \binom{k}{i-1} \lambda_j^{k-i+1} B_{ji}\varepsilon_0 + \frac{1}{2\pi \mathrm{i}} \int_{K_C} \lambda^k R(\lambda, T)\, \varepsilon_0\, \mathrm{d}\lambda\,.$$

Lemma 2. *The vectors $B_{ji}\varepsilon_0$ for $j = 1, 2, \ldots, r$ and $i = 1, 2, \ldots, i_j$ for every j are linearly independent.*

Proof. The statement of this lemma immediately follows from the definition by using the relation

$$B_{j,k+1} = (T - \lambda_j I)\, B_{j,k}\,. \ \square$$

Put

(27) $$H B_{ji} \varepsilon_0 / \lambda_j^{i-1} = v_{ji}\,,$$

(28) $$H \left(\frac{1}{2\pi \mathrm{i}} \int_{K_C} \lambda^k R(\lambda, T)\, \varepsilon_0\, \mathrm{d}\lambda \right) = v(k)\,.$$

Lemma 3. *The vectors v_{ji} for $j = 1, 2, \ldots, r$ and $i = 1, 2, \ldots, i_j$ for every j* are *linearly independent.*

The proof is obvious. $\square$

From (24), (27), (28) we have

(29) $$\eta_k = H\varepsilon_k = \sum_{j=1}^{r} \sum_{i=1}^{i_j} \binom{k}{i-1} \lambda_j^k v_{ji} + v(k)\,.$$

Let $\mathscr{K}$ be the set of all pairs (j, i) for $j = 1, 2, \ldots, r$ and $i = 1, 2, \ldots, i_j$ for every j. Let

$$\mathscr{W}_{pq} = \mathscr{L}\{v_{ji}\}_{(j,i) \in \mathscr{K} - (p,q)}\,,$$

i.e., the linear space generated by all the vectors v_{ji}, $j = 1, \ldots, r$; $i = 1, \ldots, i_j$, except v_{pq}. Denote

$$\mathscr{W} = \mathscr{L}\{v_{ji}\}_{(j,i)\in\mathscr{K}}$$

and put for every (j, i)

$$\mathscr{W} = \mathscr{W}_{ji} \oplus \mathscr{W}_{ji}^{\perp}.$$

It is easy to see that $\dim \mathscr{W}_{ji}^{\perp} = 1$ and if

(30) $$\mathscr{W}_{ji}^{\perp} = \mathscr{L}\{w_{ji}\}, \quad (w_{ji} \in X),$$

then

(31) $$(v_{ji}, w_{ji}) \neq 0.$$

Denote

(32) $$t = \sum_{j=1}^{r} i_j \,(= \operatorname{card} \mathscr{K}).$$

Lemma 4. *Put* $y_k = \eta_k - v(k)$. *If* $m_l < t$ *and* $k - m_l > \max_{(j)} (i_j)$, *then the vectors* $y_k, y_{k-m_1}, \ldots, y_{k-m_l}$ *are linearly independent.*

Proof. For some $\beta_0, \ldots, \beta_l$ from C let $\sum_{i=0}^{l} \beta_i y_{k-m_i} = 0$. Then we have for the scalar products

(33) $$\left(\sum_{i=0}^{l} \beta_i y_{k-m_i}, w_{pq}\right) = 0$$

for every pair $(p, q) \in \mathscr{K}$. Using (29), (30) and (31) we obtain from (33)

(34) $$\sum_{i=0}^{l} \beta_i \binom{k - m_i}{q - 1} \lambda_p^{k-m_i-q+1} = 0.$$

If

(35) $$\sum_{i=0}^{l} |\beta_i|^2 > 0,$$

we can assume without any loss of generality that $\beta_0 \neq 0$. Then (34) implies that the polynomial P defined by

$$P(z) = \beta_0 z^k + \beta_1 z^{k-m_1} + \ldots + \beta_l z^{k-m_l}$$

has $k - m_l + \sum_{s=1}^{r} i_s = k - m_l + t$ roots. But $t > m_l$ and $k - m_l + t > k$ and therefore (35) does not hold. □

Order the set $\mathscr{K}$ in the following finite sequence:

(36) $$\begin{array}{l} (1, i_1), (1, i_1 - 1), \ldots, (1, 1), \\ (2, i_2), (2, i_2 - 1), \ldots, (2, 1), \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ (r, i_r), (r, i_r - 1), \ldots, (r, 1). \end{array}$$

This ordering of pairs will be observed throughout this paper.

Let the symbol $\mathbf{c}(k)$ denote a vector from C^t whose p-th component is $\binom{k}{i-1} \lambda_j^k$, where the pair (j, i) lies at the p-th place in the sequence (36). For a positive integer $v \leqq k$ the symbol $\mathscr{L}_{k,v}$ denotes the subspace in X generated by the vectors $v(k)$, $v(k-1), \ldots, v(k-v)$, where these vectors are defined by (28). For $\zeta_k \in \mathscr{L}_{k,v}$ the symbol $\mathbf{d}(\zeta_k)$ denotes a vector from C^t with components of the form

(37) $$(\zeta_k, w_{ji})/(v_{ji}, w_{ji}),$$

where w_{ji} are defined by (30).

Our aim is to prove that the vectors $\eta_k, \eta_{k-m_1}, \ldots, \eta_{k-m_l}$ are linearly independent for all $k \geqq k_0$ for some k_0.

Lemma 5. *Let $p(z)$ be a polynomial and $\mu > 0$ an integer. Let $j \in \langle 1, r \rangle$. Construct a sequence $\{\tau_k\} \subset C^t$ for $k > \max_{(j)} (i_j) + \mu$ in the following way:*

$$\tau_k = p(k)\, \lambda_j^{-k}\, \mathbf{d}(\zeta_k),$$

where $\zeta_k \in \mathscr{L}_{k,\mu}$ and the components of $\mathbf{d}(\zeta_k)$ have the form (37). *Then*

$$\lim_{k \to \infty} \tau_k = \Theta(t).$$

Proof. For an integer $n \in \langle 0, \mu \rangle$ let the symbol $f_s(k, n)$ denote the scalar product $(v(k-n), w_{pq})$, where (p, q) lies at the s-th place in the sequence (36). We have defined

$$v(k-n) = H\left(\frac{1}{2\pi i} \int_C \lambda^{k-n} R(\lambda, T)\, \varepsilon_0 \, d\lambda\right).$$

Let us estimate

(38) $$|p(k) \,.\, \lambda_j^{-k} f_s(k, n)| \leqq$$

$$\leqq \|w_{pq}\| \,.\, |p(k)| \,.\, |\lambda_j^{-k}| \,.\, \|H\| \,.\, \frac{1}{2\pi} \,.\, 2\pi\varrho_C \,.\, \varrho_c^{k-n} \,.\, \max_{|\lambda| = \varrho_C} \|R(\lambda, T\| \,.\, \|\varepsilon_0\| \,.$$

According to the assumption (23′), $\varrho_C < |\lambda_j|$. The estimate (38) yields

$$\lim_{k \to \infty} p(k)\, \lambda_j^{-k} f_s(k, n) = 0.$$

The rest is obvious. □

Theorem 3. *Let $m_l < t$. There exist an integer k_0 such that the vectors $\eta_k, \eta_{k-m_1}, \ldots$ $\ldots, \eta_{k-m_l}$ are linearly independent for all $k \geqq k_0$.*

Proof. The $t \times (l + 1)$ matrix

(39) $$\mathbf{F}_k = (\mathbf{c}(k), \mathbf{c}(k - m_1), \ldots, \mathbf{c}(k - m_l))$$

has the rank $l + 1$ for all $k \geqq \max_{(j)} (i_j) + m_l + 1 \equiv k_1$. This statement follows from Lemma 4.

The s-th row of the matrix $\mathbf{F}_k$ has the form

$$(40) \qquad \binom{k}{i-1} \lambda_j^k, \binom{k-m_1}{i-1} \lambda_j^{k-m_1}, \ldots, \binom{k-m_l}{i-1} \lambda_j^{k-m_l},$$

where the pair (j, i) lies at the s-th place in the sequence (36). Since $\mathbf{F}_{k_1}$ has a maximal rank, it follows that there exists a nonsingular matrix $\mathbf{G}_{k_1}^{(1)}$ with rows of the form (40) for $l + 1$ pairs

$$(41) \qquad (p_0, q_0), (p_1, q_1), \ldots, (p_l, q_l)$$

from the set $\mathscr{K}$. Let a pair (p_i, q_i) be the s_i-th term in the sequence (36). Let for $k \geqq k_1$,

$$(42) \qquad \sum_{i=0}^{l} \beta_i \eta_{k-m_i} = 0,$$

where $\beta_i \in C$. Then

$$(43) \qquad \left(\sum_{i=0}^{l} \beta_i \eta_{k-m_i}, w_{pq}\right) = 0$$

for any pair (p, q) from the set (41). Using (29), (30), (31) and (37) we obtain from (43)

$$(44) \qquad \sum_{i=0}^{l} \beta_i \left[\binom{k-m_i}{q_j - 1} \lambda_{pj}^{k-m_i} + \mathbf{e}_{s_j}(t)^{\mathrm{T}} \mathbf{d}(\zeta_k)\right] = 0,$$

where $\zeta_k \in \mathscr{L}_{k,m_l}$.

Let $\mathbf{G}_k$ be the matrix of the system (44), $\mathbf{G}_k^{(2)} = \mathbf{G}_k - \mathbf{G}_k^{(1)}$ and

$$(45) \qquad \mathbf{G}_k^{(3)} = \mathbf{G}_k^{(1)} \,.\, \mathrm{diag}\,(\lambda_{p0}^{-k}, \ldots, \lambda_{pl}^{-k}).$$

Since $\det \mathbf{G}_k^{(3)}$ is a polynomial in k and $\det \mathbf{G}_{k_1}^{(3)} \neq 0$, there exists an integer $k_2 \geqq k_1$ such that $\det \mathbf{G}_k^{(3)} \neq 0$ for all $k \geqq k_2$. Put

$$\det \mathbf{G}_k^{(3)} = \gamma_{n_0} k^{n_0} + \gamma_{n_0 - 1} k^{n_0 - 1} + \ldots + \gamma_0,$$

where $\gamma_{n_0} \neq 0$. Then (45) yields

$$(46) \qquad \det \mathbf{G}_k^{(1)} = \left(\prod_{j=0}^{l} \lambda_{pj}^k\right) k^{n_0} \left[\gamma_{n_0} + \sum_{j=0}^{n_0 - 1} \frac{\gamma_j}{k^{n_0 - j}}\right].$$

Developing $\det \mathbf{G}_k$, we have

$$(47) \qquad \det \mathbf{G}_k = \det \mathbf{G}_k^{(1)} + \sum_{\substack{\xi_0, \xi_1, \ldots, \xi_l = 1 \\ \mathrm{no}(\xi_0 = \xi_1 = \ldots = \xi_l = 1)}}^{2} \det\left(\mathbf{g}_0^{(\xi_0)}(k), \ldots, \mathbf{g}_l^{(\xi_l)}(k)\right),$$

where

$$\mathbf{g}_i^{(1)}(k) = \mathbf{G}_k^{(1)} \mathbf{e}_i(l+1) \quad \text{and} \quad \mathbf{g}_i^{(2)}(k) = \mathbf{G}_k^{(2)} \mathbf{e}_i(l+1) .$$

Lemma 5 immediately implies that

$$\lim_{k\to\infty} \left[k^{no} \prod_{j=0}^{l} \lambda_{p_j}^k\right)^{-1} \det \left(\mathbf{g}_0^{(\xi_0)}(k), \ldots, \mathbf{g}_l^{(\xi_l)}(k)\right] = 0$$

for every $(\xi_0, \ldots, \xi_l) \neq (1, 1, \ldots, 1)$. From (46) and (47) we have

$$\det \mathbf{G}_k = k^{n_0} \prod_{j=0}^{l} \lambda_{p_j}^k \left[\gamma_{n_0} + \sum_{n_0-1}^{j=0} \frac{\gamma_j}{k^{n_0-j}} + \right.$$

$$\left. + \left(k^{n_0} \prod_{j=0}^{l} \lambda_{p_j}^k\right)^{-1} \sum_{\substack{\xi_0, \xi_1, \ldots, \xi_l = 1 \\ \text{no}(\xi_1 = \xi_2 = \ldots = \xi_l = 1)}}^{2} \det \left(\mathbf{g}_0^{(\xi_0)}(k), \ldots, \mathbf{g}_l^{(\xi_l)}(k)\right) \right]$$

and therefore there exists $k_0 > k_2$ such that the for all $k \geqq k_0$, $\det \mathbf{G}_k \neq 0$. Hence the system (42) has only the trivial solution for all $k \geqq k_0$. □

Now, we can formulate the main theorem.

Theorem 4. *Let X be a Hilbert space, $T \in [X]$, $H \in [X]$ and let H^{-1} exist. We suppose that for $x_0 \in X$ the sequence $\{x_n\}_{n=0}^{\infty}$ obtained from (2) tends to a limit x^*. Let integers $l > 0$, $r > 1$ $m_0, m_1, \ldots, m_l$ fulfil the inequalities (3) and (4), $m_l < t$. Let $\lambda_1, \ldots, \lambda_r$ be poles of $R(\lambda, T)$ of order $i_1, \ldots, i_r$, respectively, and*

$$|\lambda_1| \geqq |\lambda_2| \geqq \ldots \geqq |\lambda_r| > |\lambda|$$

for any $\lambda \in \sigma(T)$, $\lambda \neq \lambda_j$, $j = 1, \ldots, r$. Let $\lambda_i \neq \lambda_j$ for $i \neq j$, $|\lambda_r| > 0$ and (24) hold.

Then there exists an integer $k_0 > \max_{(j)} (i_j) + m_l$ such that for all $k \geqq k_0$ the matrix $\mathbf{Q}_k$ is positive definite. Therefore, for every $k \geqq k_0$ there exists one and only one vector $\boldsymbol{\alpha}^{(k)} = (\alpha_0^{(k)}, \ldots, \alpha_l^{(k)})^{\mathsf{T}}$ such that (5) and (6) are valid. The numbers $\alpha_i^{(k)}$ are given by the formulae (21) or (21'). □

The statements of this main theorem follow from previous assertions.

5. REMARKS FOR CALCULATION

Detailed extrapolation procedures will be published in forthcoming papers. Here we shall make only a few remarks with a sample example. In practical calculation of coefficients of extrapolation we may proceed according to Theorem 2 (or Theorem 4). We usually take $l = 1, 2$ or 3. It is convenient to put $m_i = in \ (i = 1, \ldots, l)$ and $H = I - T^n$ for some integer n. In this case $\eta_k = x_{k+n} - x_k$ and the relation

$$\eta_{k+1} = T\eta_k \tag{48}$$

holds. Now we present one extrapolation procedure for acceleration of the iterative process (2).

1) Take an integer $k_1 > (l + 1) n$ and put $k := k_1$.

2) Evaluate $x_1, x_2, \ldots, x_{k-(l-1)n}$ according to (2) and a corresponding vector η_{k-ln}.

3) Evaluate $\eta_{k-ln+1}, \eta_{k-ln+2}, \ldots, \eta_k$ according to (48) (in what follows we need only $\eta_{k-ln}, \eta_{k-(l-1)n}, \ldots, \eta_{k-n}, \eta_k$) and

$$x_{k-(l-s-1)n} = x_{k-(l-s)n} + \eta_{k-(l-s)n} \quad \text{for} \quad s = 1, \ldots, l-1 .$$

4) Form a matrix $\mathbf{Q}_k$ or $\mathbf{S}_k$ and solve the system

$$\left.\begin{aligned} \mathbf{Q}_k \boldsymbol{\alpha}^{(k)} &= -\lambda \mathbf{e}(l+1) \\ \mathbf{e}(l+1)^{\mathrm{T}} \boldsymbol{\alpha}^{(k)} &= 1 \end{aligned}\right\} \text{ or } \mathbf{S}_k \boldsymbol{\alpha}^{(k)} = \mathbf{e}_{l+1}(l+1) .$$

5) Calculate

$$y_k = \alpha_0^{(k)} x_k + \alpha_1^{(k)} x_{k-n} + \ldots + \alpha_l^{(k)} x_{k-ln} ,$$

$$x_{k+n} = x_k + \eta_k$$

and the vectors $\eta_{k+1}, \ldots, \eta_{k+n}$.

6) Put $k + n \to k$ and repeat this procedure from step 4).

In another paper it will be proved that the sequence $\{y_{k_1+sn}\}_{s=0}^{\infty}$ converges faster to x^* than the sequence $\{x_{k_1+sn}\}_{s=0}^{\infty}$.

As an example we consider a model problem. Let a rectangle $\Omega = \mathsf{ABCD}$ in the plane be given. Suppose the cooidinates of the points A, B, C, D are $\mathsf{A} = (\mathsf{x}_0, \mathsf{y}_0)$, $\mathsf{B} = (\mathsf{x}_{N+1}, \mathsf{y}_0)$, $\mathsf{D} = (\mathsf{x}_0, \mathsf{y}_{M+1})$, $\mathsf{C} = (\mathsf{x}_{N+1}, \mathsf{y}_{M+1})$. Moreover, assume that a uniform mesh exists with the mesh size h such that

$$\mathsf{x}_{N+1} = \mathsf{x}_0 + (N+1) h , \quad \mathsf{y}_{M+1} = \mathsf{y}_0 + (M+1) h .$$

Consider now the equation

$$\begin{aligned} -\Delta u &= 0 \quad \text{on} \quad \Omega , \\ u(x, y) &= 0 \quad \text{on} \quad \partial(\Omega) . \end{aligned} \tag{49}$$

By the five point difference approximation of (49) we obtain the system of linear algebraic equations

$$\boldsymbol{A} \boldsymbol{x} = \boldsymbol{0} . \tag{50}$$

Let us expres

$$\boldsymbol{A} = \boldsymbol{D} - \boldsymbol{L} - \boldsymbol{U}$$

and rewrite the system (50) in the form

$$\boldsymbol{x} = \mathscr{L}_\omega \boldsymbol{x} ,$$

where

$$\mathscr{L}_\omega = (\mathbf{D} - \omega \mathbf{L})^{-1} (\omega \boldsymbol{U} + (1 - \omega) \mathbf{D}) .$$

For the initial approximation $\mathbf{x}_0$ we choose

$$\mathbf{x}_0 = (1, 1, \ldots, 1)^T .$$

We take $M = 20$, $N = 30$, $\omega = 1{\cdot}5$ and compare the sequence $\{\mathbf{x}_k\}$ obtained from the S.O.R. iteration $(\mathbf{x}_{k+1} = \mathscr{L}_\omega \mathbf{x}_k)$ with the sequence $\{\mathbf{y}_k\}$ obtained from $\{\mathbf{x}_k\}$ by using the extrapolation procedure described in this part. We take $l = 3$, $n = 6$. In both cases we compare norms of the error vectors.

Table

k	S.O.R. iterations	Extrapolated iterations
32	$0{\cdot}73632 \,.\, 10^{-2}$	$0{\cdot}50524 \,.\, 10^{-3}$
56	$0{\cdot}22313 \,.\, 10^{-2}$	$0{\cdot}26337 \,.\, 10^{-5}$
92	$0{\cdot}36119 \,.\, 10^{-3}$	$0{\cdot}20895 \,.\, 10^{-8}$

References

[1] *G. Maess:* Extrapolation bei Iterationsverfahren. ZAMM 56 (1976), 121—122.

[2] *G. Maess:* Iterative Lösung linearer Gleichungssysteme. Deutsche Akademie der Naturforscher Leopoldina Halle (Saale), 1979.

[3] *I. Marek, J. Zítko:* Ljusternik Acceleration and the Extrapolated S.O.R. Method. Apl. mat. 22 (1977), 116—133.

[4] *A. E. Taylor:* Introduction to Functional Analysis. J. Wiley Publ. New York 1958.

[5] *D. M. Young:* Iterative Solution of Large Linear Systems. Academic Press, New York—London 1971.

Souhrn

URYCHLENÍ KONVERGENCE ITERAČNÍCH METOD

JAN ZÍTKO

Nechť X je lineární prostor se skalárním součinem. Nechť $T : X \to X$ a $H : X \to X$ jsou lineární operátory. Uvažujme operátorovou rovnici

$$x = Tx + b \tag{1}$$

a iterační proces

$$x_{n+1} = Tx_n + b \tag{2}$$

Předpokládejme, že pro počáteční vektor x_0 posloupnost $\{x_n\}_{n=0}^{\infty}$ sestrojená podle předpisu (2) konverguje k vektoru x^*. Nechť $l > 1$, k, $m_0, m_1, \ldots, m_l$ jsou celá čísla splňující nerovnosti

$$(3) \qquad m_l > m_{l-1} > \ldots > m_1 > m_0 = 0\,,$$

$$(4) \qquad k > m_l\,.$$

V této práci je studován problém sestrojit čísla $\alpha_0^k, \ldots, \alpha_l^k$ splňující podmínky

$$(5) \qquad \sum_{i=0}^{l} \alpha_i^{(k)} = 1\,,$$

$$(6) \qquad \left\|H(x^* - \sum_{i=0}^{l} \alpha_i^{(k)} x_{k-m_i})\right\| = \min_{(\beta_0,\ldots,\beta_l)^\mathsf{T} \in \mathfrak{M}} \left\|H(x^* - \sum_{i=0}^{l} \beta_i x_{k-m_i})\right\|\,,$$

kde $\mathfrak{M} \subset C_{l+1}$ je množina všech vektorů $(\beta_0, \beta_1, \ldots, \beta_l)^\mathsf{T}$ splňujících relaci $\sum_{i=0}^{l} \beta_i = 1$. Norma je definována pomocí skalárního součinu obvyklým způsobem. Operátor H je volen tak, aby bylo možné vyčíslit výraz v normě ve vztahu (6).

V práci je ukázáno, že za jistých předpokladů existuje právě jeden vektor $(\alpha_0^{(k)}, \ldots$ $\ldots, \alpha_l^{(k)})^\mathsf{T} \in C^{l+1}$, jehož složky řeší problém (5) a (6) a jsou uvedeny různé postupy na jeho výpočet. Kromě toho je studována souvislost mezi strukturou spektra operátoru T a možností konstrukce čísel α_i^k, které nazýváme koeficienty extrapolace. Výsledek je formulován ve Větě 4 (Theorem 4). Práce je zakončena numerickým experimentem. Uvažujeme řešení Laplaceovy rovnice na obdélníku. Obvyklou pětibodovou diferenční aproximací obdržíme soustavu lineárních algebraických rovnic. Pro síť, kterou jsme zvolili v příkladu, má matice soustavy řád 600. Tuto soustavu řešíme iterační metodou S.O.R. K posloupnosti iterací, kterou obdržíme užitím metody S.O.R. pro $\omega = 1{,}5$ a počáteční aproximaci $\mathbf{x}_0 = (1, 1, \ldots, 1)^\mathsf{T}$ pak sestrojíme posloupnost $\{y_k\}$ pro vybraná k, kde

$$(7) \qquad y_k = \sum_{i=0}^{l} \alpha_i^{(k)} \mathbf{x}_{k-m_i}\,.$$

Volíme $H = I - T^6$, $l = 3$, $m_i = 6i$ a srovnáváme normy $\|\mathbf{x}_k - \mathbf{x}^*\|$ a $\|y_k - \mathbf{x}^*\|$. Tabulka na konci článku reprezentuje číselně tyto normy pro několik vybraných iterací.

V článku, který následuje vyšetříme konvergenci čísel $\alpha_i^{(k)}$ pro $k \to \infty$. Tato čísla jsme nazvali koeficienty extrapolace. Dále ukážeme obecně, že posloupnost $\{y_k\}$ definovaná vztahem (7) konverguje k $\mathbf{x}^*$ rychleji než posloupnost $\{\mathbf{x}_k\}$ sestrojená podle (2).

Author's address: RNDr. *Jan Zítko* CSc., Katedra numerické matematiky na MFF UK, Malostranské náměstí 25, 118 00 Praha 1.