FILOMAT (Niš) **11** (1997), 89–102

ON THE CONVERGENCE OF AN ITERATIVE VECTOR ALTERNATING DIRECTIONS DIFFERENCE SCHEME

Boško S. Jovanović

Abstract. In this paper we consider a vector iterative alternating directions difference scheme for solving multidimensional Poisson equation. The scheme reduces to a modified block Jacobi overrelaxation (JOR) algorithm. We investigate the stability and the convergence of the scheme, determine the optimal iterative parameters, and estimate the error.

1. Introduction

By discretisation of boundary value problems for linear partial differential equations one obtains large linear systems. Their matrices are sparse and have distinctive structure. Iterative alternating direction methods are often used to solve such systems (see [5], [2], [6] and [4]).

A new class of so called multicomponent difference schemes was proposed recently for solving multidimensional evolutive problems (see [1], [8] and [3]). The main idea lies in vectorization of the problem, i.e. the unknown solution is approximated by vector mesh–function. Convergence of these schemes was proved and their numerical stability was checked on numerous test and real problems.

It is well known that every difference scheme for solving initial boundary value problems of parabolic type can be interpreted as an iterative method for solving the corresponding stationary problem. The aim of this paper is the investigation of one class of iterative multicomponent methods.

As a model problem we consider the Dirichlet boundary value problem for the Poisson equation in the region $\Omega = (0,\,1)^n$

$$\begin{aligned} -\Delta u &= f\,, \qquad x \in \Omega\,, \\ u(x) &= 0\,, \qquad x \in \Gamma = \partial\Omega\,. \end{aligned} \tag{1}$$

Received June 14, 1997

1991 *Mathematics Subject Classification*: 65N22.

Key words and phrases. Finite differences, linear system, Jacobi overrelaxation.

Slightly supported by MST of Serbia, grant number 04M03/C

Let $\overline{\omega}$ be an uniform mesh in $\overline{\Omega}$, with the step size $h = 1/(N+1)$. Let us use the notation $\omega = \overline{\omega} \cap \Omega$ and $\gamma = \overline{\omega} \setminus \omega$. For a function v defined on the mesh $\overline{\Omega}$ we introduce the finite differences $v_{x_i} = \big(v(x+h\,r_i) - v(x)\big)/h$ and $v_{\bar{x}_i} = \big(v(x) - v(x-h\,r_i)\big)/h$, where r_i is the unit vector of the x_i axis [7].

Let H_h denote the set of discrete functions defined on the mesh $\overline{\omega}$, which vanish on γ. We introduce the finite difference operators

$$\Lambda_i\, v = \begin{cases} -v_{x_i\bar{x}_i}\,, & x \in \omega \\ 0\,, & x \in \gamma \end{cases} \qquad \text{and} \qquad \Lambda\, v = \sum_{i=1}^{n} \Lambda_i\, v\ .$$

Let I denote the unit operator on H_h. We also define the discrete inner product

$$(v,\, w) = h^n \sum_{x\in\omega} v(x)\, w(x)$$

and the norm

$$\|v\| = (v,\, v)^{1/2} = \Big(h^n \sum_{x\in\omega} v^2(x)\Big)^{1/2}.$$

We approximate the boundary value problem (1) with the standard $(2\,n+1)$–point difference scheme

$$\Lambda\, v = \widetilde{f}\,, \qquad v,\, \widetilde{f} \in H_h\,, \tag{2}$$

where $\widetilde{f}$ is some approximation of f.

For solving (2) we use the following multicomponent alternating directions scheme [1]

$$(I+\sigma\tau\Lambda_i)\frac{v_k^i - v_{k-1}^i}{\tau} + \sum_{j=1}^{n} \Lambda_j v_{k-1}^j = \widetilde{f}, \quad i = 1, 2, \dots, n; \quad k = 1, 2, \dots \tag{3}$$

where k is the iteration number, while σ and τ are free parameters. The scheme (3) is a system with n unknown mesh functions v_k^i. To determine v_k^i we must solve a linear system whose matrix can be represented in a tridiagonal form. These systems (for $i = 1, 2, \dots, n$) can be solved simultaneously (paralelly), contrary to the other variants of the alternating direction method, such as factorized scheme [6]

$$(I + \sigma\tau\, \Lambda_1) \cdots (I + \sigma\tau\, \Lambda_n)\, \frac{v_k - v_{k-1}}{\tau} + \Lambda\, v_{k-1} = \widetilde{f}\,.$$

In the paper, we shall often use the same notation for linear operators in H_h and their matrix representation.

2. Convergence Result

To investigate the convergence of the method (3) let us represent the equation (3) in the matrix form (see [8] and [3])

$$(\mathbf{I} + \sigma\,\tau\,\Lambda)\,\frac{\mathbf{v}_k - \mathbf{v}_{k-1}}{\tau} + \mathbf{E}\,\Lambda\mathbf{v}_{k-1} = \mathbf{f}\,, \tag{4}$$

where

$$\begin{aligned} \mathbf{v}_k &= (v_k^1,\, v_k^2,\, \dots\,,\, v_k^n)^T\,, \\ \mathbf{f} &= (\widetilde{f},\, \widetilde{f},\, \dots\,,\, \widetilde{f})^T\,, \\ \mathbf{I} &= \operatorname{diag}\,(I,\, I,\, \dots\,,\, I)\,, \\ \Lambda &= \operatorname{diag}\,(\varLambda_1,\, \varLambda_2,\, \dots\,,\, \varLambda_n) \end{aligned} \qquad \text{and} \qquad \mathbf{E} = \begin{pmatrix} I & I & \dots & I \\ I & I & \dots & I \\ \vdots & \vdots & \ddots & \vdots \\ I & I & \dots & I \end{pmatrix}.$$

In such a manner, the method (3)–(4) reduces to a modified block JOR algorithm (see [9]).

Let us also define the inner product and the norm of vector–functions

$$(\mathbf{v},\, \mathbf{w}) = \sum_{i=1}^{n} (v^i,\, w^i) \qquad \text{and} \qquad \|\mathbf{v}\| = (\mathbf{v},\, \mathbf{v})^{1/2}\,.$$

Expressing $\mathbf{v}_k$ from (4) we get the canonical form of the method

$$\mathbf{v}_k = \mathbf{B}\,\mathbf{v}_{k-1} + \mathbf{c}\,, \tag{5}$$

where $\mathbf{B} = \mathbf{I} - \tau\,(\mathbf{I} + \sigma\,\tau\,\Lambda)^{-1}\,\mathbf{E}\,\Lambda$ and $\mathbf{c} = \tau\,(\mathbf{I} + \sigma\,\tau\,\Lambda)^{-1}\,\mathbf{f}$. It is well known (see [9]), that the method (5) converges for arbitrary vectors $\mathbf{v}_0$ and $\mathbf{c}$ from H_h^n if and only if the moduli of all eigenvalues of the matrix $\mathbf{B}$ are less than 1. Such eigenvalues are the roots of the equation

$$\det\,(\mathbf{B} - \lambda\,\mathbf{I}) = 0\,. \tag{6}$$

Lemma 1. *Eigenvalues of the matrix* $\mathbf{B}$ *are* $\lambda = 1$ *and*

$$\lambda = \lambda_{k_1,\,k_2,\,\dots\,,\,k_n} = 1 - \tau \sum_{j=1}^{n} \frac{\lambda_{k_j}}{1 + \sigma\,\tau\,\lambda_{k_j}}\,, \qquad k_1,\, k_2,\, \dots\,,\, k_n = 1,\, 2,\, \dots\,,\, N$$

where $\lambda_{k_j} = \frac{4}{h^2}\,\sin^2\,\frac{k_j\,\pi\,h}{2}$.

Proof. From (6), after simple transformations, we obtain the equation

$$D = \det\,(\mathbf{E} - \mathbf{P}) = 0\,,$$

where $\mathbf{P} = \mathbf{P}(\lambda) = \frac{1-\lambda}{\tau}\,(\sigma\,\tau\,\mathbf{I} + \Lambda^{-1}) = \frac{1-\lambda}{\tau}\,\mathbf{Q} = \operatorname{diag}(P_1, \ldots, P_n)$ and $P_i = \frac{1-\lambda}{\tau}\,Q_i = \frac{1-\lambda}{\tau}\,(\sigma\,\tau\,I + \Lambda_i^{-1})$. Further

$$0 = D = \begin{vmatrix} I-P_1 & I & I & \ldots & I \\ I & I-P_2 & I & \ldots & I \\ I & I & I-P_3 & \ldots & I \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ I & I & I & \ldots & I-P_n \end{vmatrix}$$

$$= \begin{vmatrix} I-P_1 & I & I & \ldots & I \\ P_1 & -P_2 & 0 & \ldots & 0 \\ P_1 & 0 & -P_3 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ P_1 & 0 & 0 & \ldots & -P_n \end{vmatrix} = \begin{vmatrix} 0 & 0 & 0 & \ldots & I \\ P_1 & -P_2 & 0 & \ldots & 0 \\ P_1 & 0 & -P_3 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ D_1 & P_n & P_n & \ldots & -P_n \end{vmatrix},$$

where $D_1 = P_n + P_1 - P_n\,P_1$. From here we obtain

$$0 = \begin{vmatrix} P_1 & -P_2 & 0 & \ldots & 0 & 0 \\ P_1 & 0 & -P_3 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ P_1 & 0 & 0 & \ldots & 0 & -P_{n-1} \\ D_1 & P_n & P_n & \ldots & P_n & P_n \end{vmatrix}$$

$$= \begin{vmatrix} P_1 & -P_2 & 0 & \ldots & 0 & 0 \\ 0 & P_2 & -P_3 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & P_2 & 0 & \ldots & 0 & -P_{n-1} \\ 0 & D_2 & P_n & \ldots & P_n & P_n \end{vmatrix}$$

$$= \det\,P_1 \cdot \begin{vmatrix} P_2 & -P_3 & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ P_2 & 0 & \ldots & 0 & -P_{n-1} \\ D_2 & P_n & \ldots & P_n & P_n \end{vmatrix},$$

where $D_2 = P_n + D_1\,P_1^{-1}\,P_2$. Continuing in the same manner, we get

$$\det\,P_1 \cdot \det\,P_2 \cdots \det\,P_{n-2} \cdot \det\,D_{n-1} = 0\,, \tag{7}$$

where $D_i = P_n + D_{i-1}\,P_{i-1}^{-1}\,P_i\,,\quad i = 2, 3, \ldots, n-1\,.$

Using the commutativity of operators Λ_i, by recursion we obtain

$$\begin{aligned} D_{n-1} &= P_n\,\left(P_1^{-1} + \cdots + P_n^{-1} - I\right)\,P_{n-1} \\ &= \frac{1-\lambda}{\tau}\,Q_n\,\left(Q_1^{-1} + \cdots + Q_n^{-1} - \frac{1-\lambda}{\tau}\,I\right)\,Q_{n-1}\,, \end{aligned}$$

which, together with (7), yields

$$\left(\frac{1-\lambda}{\tau}\right)^{(n-1)\,N^n} \det Q_1 \cdot \det Q_2 \cdots \det Q_n \times \times \det\left(Q_1^{-1}+\cdots+Q_n^{-1}-\frac{1-\lambda}{\tau}\,I\right)=0\,. \tag{8}$$

From (8) it follows that $\lambda = 1$ is the eigenvalue of the matrix $\mathbf{B}$ of the multiplicity $(n-1)\,N^n$. The other eigenvalues can be expressed using eigenvalues of the matrix $Q_1^{-1}+\cdots+Q_n^{-1}$. Because $Q_j = \sigma\,\tau\,I + \Lambda_j^{-1} = \Lambda_j^{-1}\left(I+\sigma\,\tau\,\Lambda_j\right)$ and the eigenvalues of Λ_j equal to λ_{k_j}, $k_j = 1, 2, \ldots, N$, we get

$$\lambda = \lambda_{k_1,\,k_2,\,\ldots,\,k_n} = 1-\tau\sum_{j=1}^{n}\frac{\lambda_{k_j}}{1+\sigma\tau\,\lambda_{k_j}} \qquad k_1,\,k_2,\,\ldots,\,k_n = 1,\,2,\,\ldots,\,N\,. \quad \blacksquare$$

One can directly verify the following result.

Lemma 2. *The eigenvectors of the matrix* $\mathbf{B}$ *corresponding to the eigenvalue* $\lambda = 1$ *are*

$$\phi^{1;\,k_1,\,k_2,\,\ldots,\,k_n} = \left(\frac{\varphi^{k_1,\,k_2,\,\ldots,\,k_n}}{\lambda_{k_1}},\,-\frac{\varphi^{k_1,\,k_2,\,\ldots,\,k_n}}{\lambda_{k_2}},\,0,\,\ldots,\,0\right)^T,$$
$$\phi^{2;\,k_1,\,k_2,\,\ldots,\,k_n} = \left(\frac{\varphi^{k_1,\,k_2,\,\ldots,\,k_n}}{\lambda_{k_1}},\,0,\,-\frac{\varphi^{k_1,\,k_2,\,\ldots,\,k_n}}{\lambda_{k_3}},\,0,\,\ldots,\,0\right)^T,$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$\phi^{n-1;\,k_1,\,k_2,\,\ldots,\,k_n} = \left(\frac{\varphi^{k_1,\,k_2,\,\ldots,\,k_n}}{\lambda_{k_1}},\,0,\,\ldots,\,0,\,-\frac{\varphi^{k_1,\,k_2,\,\ldots,\,k_n}}{\lambda_{k_n}}\right)^T,$$

where

$$\varphi^{k_1,\,k_2,\,\ldots,\,k_n} = \sin\,x_1\cdot\sin\,x_2\cdots\sin\,x_n\,, \qquad (x_1,\,x_2,\,\ldots,\,x_n)\in\omega\,.$$

For the eigenvalue $\lambda_{k_1,\,k_2,\,\ldots,\,k_n}$ *corresponding eigenvector is*

$$\phi^{k_1,\,k_2,\,\ldots,\,k_n} = \phi^{n;\,k_1,\,k_2,\,\ldots,\,k_n}$$
$$= \left(\frac{\varphi^{k_1,\,k_2,\,\ldots,\,k_n}}{1+\sigma\,\tau\,\lambda_{k_1}},\,\frac{\varphi^{k_1,\,k_2,\,\ldots,\,k_n}}{1+\sigma\,\tau\,\lambda_{k_2}},\,\ldots,\,\frac{\varphi^{k_1,\,k_2,\,\ldots,\,k_n}}{1+\sigma\,\tau\,\lambda_{k_n}}\right)^T.$$

The vectors $\varphi^{k_1,\,k_2,\,\dots,\,k_n}$ represent an orthogonal basis of H_h. From the representation

$$\widetilde{f} = \sum_{k_1,\,k_2,\,\dots,\,k_n} \alpha_{k_1,\,k_2,\,\dots,\,k_n}\, \varphi^{k_1,\,k_2,\,\dots,\,k_n}$$

we immediately obtain

$$\mathbf{c} = \tau \sum_{k_1,\,k_2,\,\dots,\,k_n} \alpha_{k_1,\,k_2,\,\dots,\,k_n}\, \phi^{k_1,\,k_2,\,\dots,\,k_n}\,.$$

Choosing $\mathbf{v}_0 = \mathbf{0}$, from (5) follows

$$\mathbf{v}_k = (\mathbf{B}^{k-1} + \mathbf{B}^{k-2} + \dots + \mathbf{I})\,\mathbf{c} \tag{9}$$

$$= \tau \sum_{k_1,\,k_2,\,\dots,\,k_n} \alpha_{k_1,\,k_2,\,\dots,\,k_n}\, (\lambda^{k-1}_{k_1,\,k_2,\,\dots,\,k_n} + \lambda^{k-2}_{k_1,\,k_2,\,\dots,\,k_n} + \dots + 1)\, \phi^{k_1,\,k_2,\,\dots,\,k_n}\,,$$

and this iterative process converges if

$$\max_{k_1,\,k_2,\,\dots,\,k_n} |\lambda_{k_1,\,k_2,\,\dots,\,k_n}| < 1\,. \tag{10}$$

If (10) is satisfied, for $k \to \infty$ from (9) follows

$$\mathbf{v}_k \to \mathbf{v} = \tau \sum_{k_1,\,k_2,\,\dots,\,k_n} \alpha_{k_1,\,k_2,\,\dots,\,k_n}\, \frac{1}{1 - \lambda_{k_1,\,k_2,\,\dots,\,k_n}}\, \phi^{k_1,\,k_2,\,\dots,\,k_n}\,. \tag{11}$$

Using the orthogonality of vectors $\varphi^{k_1,\,k_2,\,\dots,\,k_n}$ and the Parseval equation, from (9) and (11) follows

$$\|\mathbf{v}_k - \mathbf{v}\| \le \big(\max_{k_1,\,k_2,\,\dots,\,k_n} |\lambda_{k_1,\,k_2,\,\dots,\,k_n}|\big)^k\, \|\mathbf{v}_0 - \mathbf{v}\|\,, \qquad \mathbf{v}_0 \in \widetilde{\mathbf{H}}_h\,. \tag{12}$$

The same conclusion holds if $\mathbf{v}_0$ is an arbitrary linear combination of vectors $\phi^{k_1,\,k_2,\,\dots,\,k_n}$. The set of such vectors is a subspace of H_h^n, which will be denoted by $\widetilde{\mathbf{H}}_h$.

Let us estimate $\max_{k_1,\,k_2,\,\dots,\,k_n} |\lambda_{k_1,\,k_2,\,\dots,\,k_n}|$. The values λ_{k_j} belong to the interval $[m,\, M]$, where

$$m = \frac{4}{h^2}\sin^2\frac{\pi h}{2} \ge 8\,, \qquad M = \frac{4}{h^2}\cos^2\frac{\pi h}{2} < \frac{4}{h^2}\,.$$

Further

$$\max_{k_1,\,k_2,\,\dots,\,k_n} |\lambda_{k_1,\,k_2,\,\dots,\,k_n}| \le \max_{\lambda_{k_j} \in [m,\,M]} \Big|1 - \tau \sum_{j=1}^{n} \frac{\lambda_{k_j}}{1 + \sigma\tau\,\lambda_{k_j}}\Big|$$

$$= \max_{\lambda \in [m,\,M]} \Big|1 - \frac{n\tau\,\lambda}{1 + \sigma\tau\,\lambda}\Big|\,.$$

From here, one can directly check that the condition (10) is satisfied for $\tau > 0$ and $n/2 \le \sigma < \infty$. For $0 < \sigma < n/2$ the condition (10) is satisfied for $0 < \tau < \frac{2}{M\,(n-2\,\sigma)}$. Thus the following result holds true.

Lemma 3. *If* $\mathbf{v}_0 \in \widetilde{\mathbf{H}}_h$ *and* $\tau > 0$, $n/2 \le \sigma < \infty$ *or* $0 < \tau < \frac{2}{M\,(n-2\,\sigma)}$, $0 < \sigma < n/2$ *then the iterative process* (5) *converges and the error estimate* (12) *holds.*

If the condition (10) is satisfied, the convergence rate of the sequence (9) is optimal when $\max_{k_1,\,k_2,\,\dots,\,k_n} |\lambda_{k_1,\,k_2,\,\dots,\,k_n}|$ is minimal. Supposing that σ and τ are nonnegative, in a natural way we obtain the following inf–sup problem: find

$$q = q(\sigma) = \inf_{\tau \ge 0} \sup_{\lambda \in [m,\,M]} \left| 1 - \frac{n\tau\,\lambda}{1 + \sigma\tau\,\lambda} \right| .$$

Let us first consider the case $0 \le \sigma < n$. Denote

$$\psi(t) = \left| 1 - \frac{nt}{1 + \sigma t} \right|$$

Let $\tau > 0$ be fixed. The function

$$\psi_0(\lambda) = \psi(\tau\,\lambda) = \left| 1 - \frac{n\tau\lambda}{1 + \sigma\tau\lambda} \right|$$

is decreasing on the interval $[0, \frac{1}{(n-\sigma)\,\tau}]$, increasing on the interval $[\frac{1}{(n-\sigma)\tau}, +\infty]$, and $\psi_0(0) = 1$, $\psi_0(\frac{1}{(n-\sigma)\,\tau}) = 0$, $\lim_{\lambda\to+\infty} \psi_0(\lambda) = (n-\sigma)/\sigma$ (fig. 1). It follows that

$$\sup_{\lambda \in [m,\,M]} \left| 1 - \frac{n\tau\lambda}{1 + \sigma\tau\lambda} \right| = \sup_{\lambda \in [m,\,M]} \psi_0(\lambda) = \max\{\psi_0(m),\, \psi_0(M)\} .$$

Treated as functions of τ, $\psi_1(\tau) = \psi_0(m) = \psi(m\,\tau)$ and $\psi_2(\tau) = \psi_0(M) = \psi(M\,\tau)$ have an analogous behaviour as $\psi_0(\lambda)$ (fig. 2). Because $0 < \frac{1}{(n-\sigma)\,M} < \frac{1}{(n-\sigma)\,m}$, there exists a point $\tau_0 \in (\frac{1}{(n-\sigma)\,M}, \frac{1}{(n-\sigma)\,m})$ such that $\psi_1(\tau_0) = \psi_2(\tau_0)$. Furthermore, $\psi_1(\tau) > \psi_2(\tau)$ for $\tau \in (0, \tau_0)$ and $\psi_1(\tau) < \psi_2(\tau)$ for $\tau \in (\tau_0, +\infty)$. In such a way

$$\max\{\psi_0(m),\, \psi_0(M)\} = \max\{\psi_1(\tau),\, \psi_2(\tau)\} = \begin{cases} \psi_1(\tau)\,, & \tau \in (0,\, \tau_0) \\ \psi_2(\tau)\,, & \tau \in (\tau_0,\, +\infty) \end{cases}$$

and

$$\inf_{\tau \ge 0} \sup_{\lambda \in [m,\,M]} \left| 1 - \frac{n\tau\lambda}{1 + \sigma\tau\lambda} \right| = \psi_1(\tau_0) = \psi_2(\tau_0) < 1 .$$

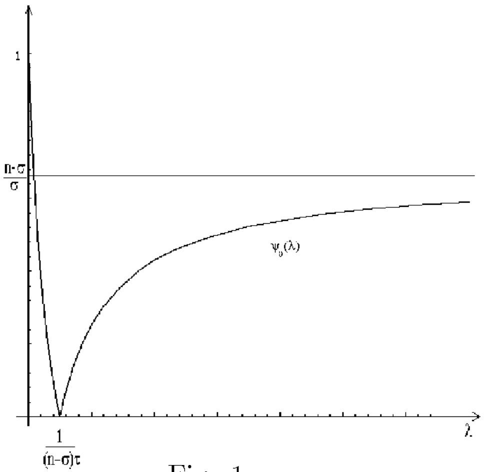

Fig. 1

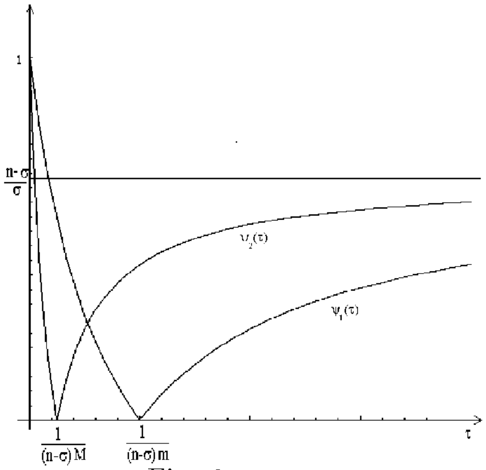

Fig. 2

The value τ_0 can be easily obtained as

(13)

$$\tau_0 = \tau_0(\sigma) = \frac{2}{\sqrt{(\sigma - n/2)^2\,(M-m)^2 + n^2\,Mm} - (\sigma - n/2)\,(M+m)}\,,$$

and from here

(14)

$$q = q(\sigma) = \inf_{\tau \geq 0} \sup_{\lambda \in [m,\,M]} \left| 1 - \frac{n\tau\lambda}{1+\sigma\tau\lambda} \right|$$

$$= \frac{n-\sigma}{\sigma}\,\frac{(\sigma + n/2)\,M - (\sigma - n/2)\,m - \sqrt{(\sigma - n/2)^2\,(M-m)^2 + n^2\,Mm}}{(3n/2 - \sigma)\,M + (\sigma - n/2)\,m + \sqrt{(\sigma - n/2)^2\,(M-m)^2 + n^2\,Mm}}\,.$$

Notice that the function $\tau_0(\sigma)$ is increasing on the interval $[0,\,n)$ (fig. 3). In particular

$$\tau_0(0) = \frac{2}{n\,(M+m)} = \frac{h^2}{2n}; \quad \tau_0(n/2) = \frac{2}{n\,\sqrt{Mm}} = \frac{h^2}{n\,\sin\,\pi h} \asymp \frac{h}{n\,\pi}, \quad h \to 0;$$
$$\lim_{\sigma \to n-0} \tau_0(\sigma) = +\infty\,.$$

The function $q(\sigma)$ is decreasing on $[0,\,n]$ (fig. 4), and

$$q(0) = \frac{M-m}{M+m} = \cos\,\pi h \asymp 1 - \frac{\pi^2\,h^2}{4}\,, \quad h \to 0\,;$$
$$q(n/2) = \frac{\sqrt{M} - \sqrt{m}}{\sqrt{M} + \sqrt{m}} = \frac{1 - \sin\,\pi h}{\cos\,\pi h} \asymp 1 - \pi\,h\,, \quad h \to 0\,; \qquad q(n) = 0\,.$$

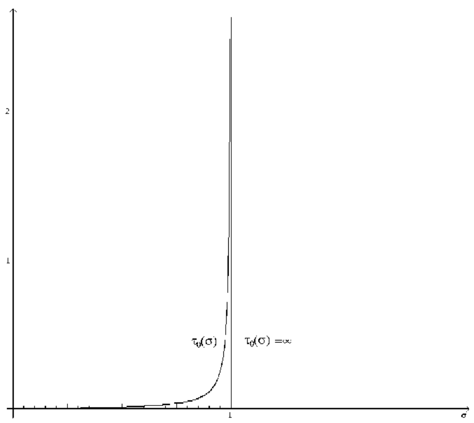

Fig. 3

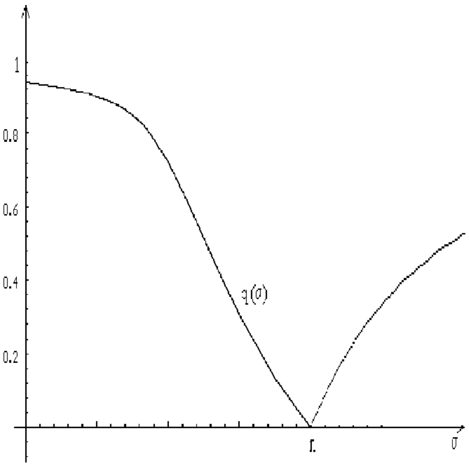

Fig. 4

For $\sigma \geq n$ the function $\psi_0(\lambda)$ is positive and monotonically decreasing for $\lambda > 0$, so

$$\sup_{\lambda \in [m,\, M]} \psi_0(\lambda) = \psi_0(m) = \psi(m\,\tau) = 1 - \frac{n\,\tau\, m}{1 + \sigma\,\tau\, m} = \psi_1(\tau)\,.$$

The function $\psi_1(\tau) = \psi(m\,\tau)$ is also decreasing for $\tau \geq 0$, so

$$1 = \psi_1(0) \geq \psi_1(\tau) > \lim_{\tau \to +\infty} \psi_1(\tau) = \frac{\sigma - n}{\sigma}\,.$$

Consequently,

$$\begin{aligned} q = q(\sigma) &= \inf_{\tau \geq 0} \sup_{\lambda \in [m,\, M]} \left| 1 - \frac{n\,\tau\,\lambda}{1 + \sigma\,\tau\,\lambda} \right| \\ &= \inf_{\tau \geq 0} \psi_1(\tau) = \lim_{\tau \to +\infty} \psi_1(\tau) = \frac{\sigma - n}{\sigma}\,. \end{aligned} \tag{15}$$

In such a manner we have proved the following assertion:

Theorem 1. *If the initial vector* $\mathbf{v}_0$ *of the sequence* (4) *belongs to the subspace* $\widetilde{\mathbf{H}}_h$ *then the optimal iterative parameter* τ*, by the maximal convergence rate criterion is* $\tau = \tau_0(\sigma)$, *for* $\sigma \in (0,\, n)$, *or* $\tau = +\infty$, *for* $\sigma \geq n$ *. In this case, the following convergence rate estimate holds*

$$\|\mathbf{v}_k - \mathbf{v}\| \leq q^k\, \|\mathbf{v}_0 - \mathbf{v}\|\,,$$

where $q = q(\sigma)$ *is defined by* (14) *or* (15).

Remark 1. In the case when $\sigma = n$, $\tau = \infty$ the method (4) becomes

$$n\,\Lambda\,(\mathbf{v}_k - \mathbf{v}_{k-1}) + \mathbf{E}\,\Lambda \mathbf{v}_{k-1} = \mathbf{f}$$

and converges in a single step if $\mathbf{v}_0 \in \widetilde{\mathbf{H}}_h$. For example, if $\mathbf{v}_0 = \mathbf{0}$ then

$$\mathbf{v}_1 = \mathbf{v}_2 = \ldots = \mathbf{v} = \frac{1}{n}\,\Lambda^{-1}\,\mathbf{f} = \frac{1}{n}\,(\Lambda_1^{-1}\,\widetilde{f},\, \Lambda_2^{-1}\,\widetilde{f},\, \ldots,\, \Lambda_n^{-1}\,\widetilde{f})^T .$$

Note that in this case the limit vector does not determine the solution of the starting problem (2).

3. Internal Error of the Method

In the previous paragraph we have proved that for a suitable choice of iterative parameter τ the iterative process (4) converges to the limit vector (11). On the other hand, the exact solution of the problem (2) is

(16)
$$v^* = (\Lambda_1 + \Lambda_2 + \ldots + \Lambda_n)^{-1}\,\widetilde{f} = \sum_{k_1,\,k_2,\,\ldots,\,k_n} \frac{\alpha_{k_1,\,k_2,\,\ldots,\,k_n}}{\lambda_{k_1} + \cdots + \lambda_{k_n}}\,\varphi^{k_1,\,k_2,\,\ldots,\,k_n} .$$

Let us estimate the distance from the exact solution v^* to the arithmetical mean $\overline{v} = (v^1 + v^2 + \cdots + v^n)/n$ of components of the limit vector $\mathbf{v}$. From (16) and (11) after some algebraic manipulation we obtain

(17)
$$\overline{v} - v^* = \sum_{k_1,\,k_2,\,\ldots,\,k_n} \alpha_{k_1,\,k_2,\,\ldots,\,k_n} \left(\frac{1}{\sum_{j=1}^n \frac{\lambda_{k_j}}{1+\sigma\,\tau\,\lambda_{k_j}}} - \frac{\sigma\,\tau}{n} - \frac{1}{\sum_{j=1}^n \lambda_{k_j}} \right) \varphi^{k_1,\,k_2,\,\ldots,\,k_n} .$$

In such a manner, estimation of the error $\overline{v} - v^*$ reduces to estimation of the maximum of modulus of the function

$$\chi(\lambda_1,\,\lambda_2,\,\ldots,\,\lambda_n) = \frac{1}{\sum_{j=1}^n \lambda_j\,(1+\sigma\tau\,\lambda_j)^{-1}} - \frac{\sigma\tau}{n} - \frac{1}{\sum_{j=1}^n \lambda_j}$$

in the domain $[m,\,M]^n$.

The function $\chi(\lambda_1,\,\lambda_2,\,\ldots,\,\lambda_n)$ is a symmetric function of variables $\lambda_1,\,\lambda_2,\,\ldots,\,\lambda_n$. Consequently, $|\chi|$ reaches its maximum in the domain $[m,\,M]^n$ either on the diagonal or on the boundary of the domain. Because

$\chi(\lambda, \lambda, \dots, \lambda) = 0$, it only remains to consider the second possibility. It is sufficient to examine the function χ on hyperplanes:

a) $\lambda_{k+1} = \dots = \lambda_n = m$ (or $\lambda_{k+1} = \dots = \lambda_n = M$), $1 \le k \le n-1$,

and

b) $\lambda_{k+1} = \dots = \lambda_{k+j} = m$, $\lambda_{k+j+1} = \dots = \lambda_n = M$, $k \ge 1$, $j \ge 1$,

$k + j \le n - 1$,

as well as in vertices

c) $\lambda_1 = \dots = \lambda_i = M$, $\lambda_{i+1} = \dots = \lambda_n = m$, $1 \le i \le n-1$.

On hyperplanes of the type a) the only stationary point is at the vertix $\lambda_1 = \dots = \lambda_k = m$ (respectively $\lambda_1 = \dots = \lambda_k = M$).

On hyperplanes of the type b) there is a stationary point

$$\lambda_1 = \dots = \lambda_k = \frac{C_{kj} - \widehat{C}_{kj}}{\sigma\tau\, \widehat{C}_{kj}},$$

where

$$C_{kj} = j\, m + (n-k-j)\, M, \qquad \widehat{C}_{kj} = \frac{j\, m}{1 + \sigma\tau\, m} + \frac{(n-k-j)\, M}{1 + \sigma\tau\, M}.$$

Then

$$\chi = \frac{\sigma\tau}{n} \cdot \frac{(n-k)\,(C_{kj} - \widehat{C}_{kj}) - \sigma\tau\, C_{kj}\, \widehat{C}_{kj}}{k\,(C_{kj} - \widehat{C}_{kj}) + \sigma\tau\, C_{kj}\, \widehat{C}_{kj}},$$

and

$$|\chi| \le \sigma\tau. \tag{18}$$

At the vertix c)

$$\chi = \frac{1}{i\, M/(1+\sigma\tau\, M) + (n-i)\, m/(1+\sigma\tau\, m)} - \frac{\sigma\tau}{n} - \frac{1}{i\, M + (n-i)\, m}.$$

If we denote $t_i = \big[i\, M + (n-i)\, m\big]/n$ the last equality can be rewritten as

$$\chi = \kappa(t_i) = \frac{1}{n}\Big[\frac{1 + \sigma\tau\,(M+m) + \sigma^2\,\tau^2\, M\, m}{t_i + \sigma\tau\, M\, m} - \sigma\,\tau - \frac{1}{t_i}\Big],$$

and the problem reduces to the determination of $\max_{1\le i\le n-1} |\kappa(t_i)|$.

The function $\kappa(t)$ (fig. 5) is nonnegative on the interval $[m,\, M]$, vanishes on its boundaries and has one maximum in the point

$$t' = \frac{\sigma\,\tau\,M\,m}{\sqrt{1+\sigma\,\tau\,(M+m)+\sigma^2\,\tau^2\,M\,m}-1}\,.$$

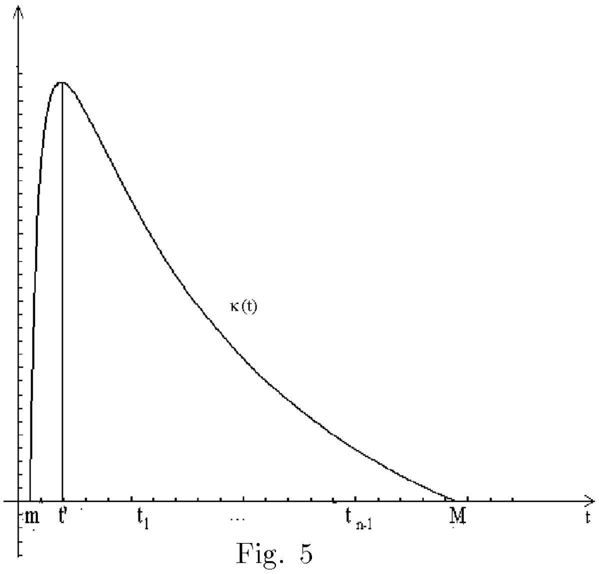

Fig. 5

For sufficiently small h, the inequalities $3\,h^{-2} < M < 4\,h^{-2}$ and $8 < m < 10$ hold. After a few simple estimation we obtain

$$m < t' < 10 + 11/h = O(h^{-1})\,.$$

On the other hand

$$t_i = m + \frac{i}{n}\,(M-m) \geq \frac{3}{n\,h^2} = O(h^{-2})\,,$$

so

$$m < t' < t_1 < t_2 < \cdots < t_{n-1} < M\,.$$

From here we conclude that the desired maximum is reached in the point t_1:

$$\max_{1\leq i\leq n-1}\,|\kappa(t_i)| = \kappa(t_1)\,.$$

Further

$$\kappa(t_1) = \frac{(n-1)\,\sigma\tau\,M + \sigma\tau\,m + n}{n\,[M+(n-1)\,m + n\,\sigma\tau\,Mm]} - \frac{1}{M+(n-1)\,m} \leq \sigma\tau\,. \tag{19}$$

In such a manner, from (18) and (19) follows

$$\max_{\lambda_j\in[m,M]} \left| \frac{1}{\sum_{j=1}^{n} \lambda_j\,(1+\sigma\tau\,\lambda_j)^{-1}} - \frac{\sigma\tau}{n} - \frac{1}{\sum_{j=1}^{n}\lambda_j} \right| \le \sigma\,\tau\ . \tag{20}$$

Notice that for $\sigma\,\tau = O(1)$ estimates (19) and (20) are of optimal order. Really, when $h\to 0$, we get

$$\kappa(t_1) \asymp \frac{n-1}{n}\,\frac{\sigma\,\tau}{1+n\,\sigma\,\tau\,\pi^2} \ge \frac{n-1}{n}\,\frac{\sigma\,\tau}{1+C\,n\,\pi^2} = C_1\,\sigma\,\tau\,.$$

From (17) and (20) we immediately obtain the following proposition:

Theorem 2. *For the iterative process* (4)*, in the case of convergence, the error estimate*

$$\|\overline{v}-v^*\| \le \sigma\,\tau\,\|\widetilde{f}\|$$

holds.

Remark 2. From the previous it follows that the free parameter σ must be determined by two criteria: to maximize the convergence rate (minimize $q(\sigma)$) and to minimize the error $\overline{v}-v^*$ (in practice it is sufficient to set $\sigma\,\tau = O(h^2)$). Unfortunately, these conditions are contradictory. Setting

$$\sigma = \frac{n}{2} - C\,n\,h^{\alpha}\,,$$

and using assimptotical expansion when $h\to 0$, we obtain from (13) and (14)

$$q = 1 - O(h^{2-\alpha})\,, \qquad \sigma\,\tau_0 = O(h^{2-\alpha}) \qquad \text{for} \qquad 0\le\alpha\le 1$$

and

$$q = 1 - O(h)\,, \qquad \sigma\,\tau_0 = O(h) \qquad \text{for} \qquad \alpha\ge 1\,.$$

In the case $\sigma \ge n/2$, we have

$$\sigma\,\tau_0 \ge 0.5\,n\,\tau_0(0.5\,n) = O(h)\,.$$

In such a way, in the case of "fast" convergence – internal error is "large" ($\alpha = 1$), while in the case of "small" internal error – the convergence is "slow" ($\alpha = 0$).

Due to its paralellelism, the method can be used for the fast calculation of a rough approximation of the solution, which can be improved with other iterative methods.

References

[1] Abrashin, V.N., Mukha, V.A., *A class of efficient difference schemes for solving multidimensional problems in mathematical physics*, Differentsial'nye Uravneniya **28** (1992), 1786–1799 (Russian).

[2] Douglas, J., *On numerical integration of* $\partial^2 u/\partial x^2 + \partial^2 u/\partial y^2 = \partial u/\partial t$ *by implicit methods*, J. Soc. Industr. Appl. Math. **3** (1955), 42–65.

[3] Jovanović, B.S., *On the convergence of a multicomponent alternating direction difference scheme*, Publ. Inst. Math. **56 (70)** (1994), 129–134.

[4] Marchuk, G.I., *Splitting methods*, Moscow: Nauka, 1988 (Russian).

[5] Peaceman, D.W.; Rachford, H.H., *The numerical solution of parabolic and elliptic differential equations*, J. Soc. Industr. Appl. Math. **3** (1955 28–42).

[6] Samarskiĭ, A.A., *Additive difference schemes*, International congress of mathematics, Moscow, Section 14, 1966, pp. 46–47 (Russian).

[7] Samarskiĭ, A.A., *Theory of difference schemes*, Moscow: Nauka, 1983 (Russian).

[8] Vabishchevich, P.N., *Vector additive difference schemes for first order evolution equations*, Zh. Vychisl. Mat. Mat. Fiz. **36** (1996), 44–51 (Russian).

[9] Young, D.M., *Iterative solution of large linear systems*, New York – London: Academic Press, 1971.

University of Belgrade, Faculty of Mathematics, Studentski trg 16, P.O. Box 550, 11000 Beograd, Yugoslavia,
E-mail: bosko@matf.bg.ac.yu