Applications of Mathematics

Jan Zítko

Combining the preconditioned conjugate gradient method and a matrix iterative method

Applications of Mathematics, Vol. 41 (1996), No. 1, 19–39

Persistent URL: http://dml.cz/dmlcz/134311

COMBINING THE PRECONDITIONED CONJUGATE GRADIENT METHOD AND A MATRIX ITERATIVE METHOD

JAN ZÍTKO, Praha

(Received July 1, 1994)

Summary. The preconditioned conjugate gradient method for solving the system of linear algebraic equations with a positive definite matrix is investigated. The initial approximation for conjugate gradient is constructed as a result of a matrix iteration method after m steps. The behaviour of the error vector for such a combined method is studied and special numerical tests and conclusions are made.

Keywords: conjugate gradients, preconditioning, iterative method, numerical experiments

AMS classification: 65F10

1. INTRODUCTION

Let us consider the system of n linear algebraic equations

$$Ax = b \tag{1.1}$$

supposing that the matrix A is real, symmetric and positive definite. Let x^* denote the solution of (1.1). Before presenting an overview of the results of our work, in this introduction, we will make the following consideration, which will contribute towards better understanding of the whole problem. Let us mention here the notation used in this article. By $\mathbb{R}^n$ we denote the real linear space of all column vectors $x = (x_1, \ldots, x_n)^T$ with real components. The symbol $L(\mathbb{R}^n)$ denotes the set of all real $n \times n$ matrices. The vector $e_i(n)$ is the ith column of the identity matrix

This paper was supported by the Grant Agency of the Czech Republic under Grant No 201/93/0429.

$I \in L(\mathbb{R}^n)$. We will write only e_i, if the dimension n is prescribed. Let

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

If $u_i \in \mathbb{R}^n$, $i = 1, 2, \ldots, s$, then $(u_1, \ldots, u_s)$ is the matrix with columns u_i. The symbol Θ denotes the null vector in $\mathbb{R}^n$.

Let C be a nonsingular matrix $n \times n$. If we multiply the system (1.1) from the left by the matrix C^T, we obtain

$$C^T ACC^{-1}x = C^T b\,, \tag{1.2}$$

and if we substitute

$$\tilde{A} = C^T AC\,; \quad \tilde{x} = C^{-1}x\,; \quad \tilde{b} = C^T b\,, \tag{1.3}$$

we obtain from (1.2) the system

$$\tilde{A}\tilde{x} = \tilde{b} \tag{1.4}$$

with a symmetric and positive definite matrix $\tilde{A}$, having the solution $\tilde{x}^* = C^{-1}x^*$.

Let $\tilde{x}_0 \in \mathbb{R}^n$ and let us put $\tilde{r}_0 = \tilde{b} - \tilde{A}\tilde{x}_0$. Let us suppose that the Krylov subspace $\mathrm{span}\{\tilde{r}_0, \tilde{A}\tilde{r}_0, \ldots, \tilde{A}^{k-1}\tilde{r}_0\}$ has the dimension k. The vector $\tilde{x}_k$ that we obtain in the kth step when applying the conjugate gradient method is just the vector which lies in the linear variety

$$\tilde{x}_0 + \mathrm{span}\{\tilde{r}_0, \tilde{A}\tilde{r}_0, \ldots, \tilde{A}^{k-1}\tilde{r}_0\} \tag{1.5}$$

and fulfils the condition

$$\tilde{r}_k = \tilde{b} - \tilde{A}\tilde{x}_k \perp \mathrm{span}\{\tilde{r}_0, \tilde{A}\tilde{r}_0, \ldots, \tilde{A}^{k-1}\tilde{r}_0\}. \tag{1.6}$$

Let us define

$$M = CC^T\,, \quad p_k = C\tilde{p}_k\,, \quad x_k = C\tilde{x}_k\,, \quad b = C^{-T}\tilde{b}. \tag{1.7}$$

Thus

$$\tilde{r}_k = \tilde{b} - \tilde{A}\tilde{x}_k = C^T b - C^T ACC^{-1}x_k = C^T(b - Ax_k)$$

and if we write in the sequel

$$r(x) = b - Ax\,, \tag{1.8}$$

then

$$\tilde{r}_k = C^T r(x_k). \tag{1.9}$$

By analogy with the relations (1.5) and (1.6) the vector x_k that we obtain in the kth step when applying the well known conjugate gradient method with preconditioning given by the matrix M is such a vector in the linear variety

$$x_0 + \operatorname{span}\{Mr(x_0), (MA)Mr(x_0), \dots, (MA)^{k-1}Mr(x_0)\} \tag{1.10}$$

that fulfils the projection condition

$$r(x_k) \perp \operatorname{span}\{Mr(x_0), MAMr(x_0), \dots, (MA)^{k-1}Mr(x_0)\}. \tag{1.11}$$

(For details see [D.O'L].)

There are more possibilities how to practically realise the calculation of x_k. For programming purposes the preconditioned conjugate gradient method for solving the system (1.1) is usually formulated by the following sequence of recurrences.

Algorithm 1.1

1) Choose $x_0 \in \mathbb{R}^n$ and put $r(x_0) = b - Ax_0$, $p_0 = Mr(x_0)$.
2) For $k = 1, \dots, n$ do
   if $r(x_{k-1}) = 0$
   then
      set $x^* = x_{k-1}$ and quit
   else

$$\begin{aligned} \lambda_{k-1} &= r(x_{k-1})^T Mr(x_{k-1}) / p_{k-1}^T A p_{k-1} \\ x_k &= x_{k-1} + \lambda_{k-1} p_{k-1} \\ r(x_k) &= r(x_{k-1}) - \lambda_{k-1} A p_{k-1} \quad (\text{or } r(x_k) = b - Ax_k) \\ \beta_k &= r(x_k)^T Mr(x_k) / r(x_{k-1})^T Mr(x_{k-1}) \\ p_k &= Mr(x_k) + \beta_k p_{k-1} \end{aligned}$$

   endif

   endfor
3) $x^* = x_n$

The other possibility is to construct the sequence $\{x_k\}_{k=0}^n$ by the three-term recurrence algorithm given by the formula

$$x_{k+1} = x_{k-1} + \omega_{k+1}(\gamma_{k+1} z_k + x_k - x_{k-1}) \tag{1.12}$$

where the coefficients ω_{k+1} and γ_{k+1} are calculated according to well known formulas (see [G-L]) and $z_k = Mr(x_k)$. Putting $\gamma_k = \omega_k = 1$ for all k, we obtain the iterative process

$$x_{k+1} = (I - MA)x_k + Mb. \tag{1.13}$$

Therefore the conjugate gradient method represents an acceleration of convergence of the successive iteration (1.13). To give some quantitative connection between the iterative process (1.13) and the preconditioned conjugate gradient algorithm we will consider the following algorithm.

Algorithm 1.2

1) Choose $y_0 \in \mathbb{R}^n$ and an integer m.
2) Calculate the mth iteration y_m by the following iterative process:

$$y_l = (I - MA)y_{l-1} + Mb, \quad l = 1, 2, \ldots, m. \tag{1.14}$$

3) Put $x_0 = y_m$ and carry out k steps of the preconditioned conjugate gradient method (i.e., Algorithm 1.1).

In the sequel we set

$$S = MA, \quad Q = I - S. \tag{1.15}$$

Let us note that in the notation (1.15), the iterations (1.14) have the form

$$y_l = Qy_{l-1} + Mb. \tag{1.14'}$$

From the equalities (1.14)′ and $x^* = Qx^* + Mb$ it follows that

$$y_l - x^* = Q(y_{l-1} - x^*). \tag{1.16}$$

If the matrix Q is convergent then the sequence $\{y_l\}_{l=0}^\infty$ constructed by an iterative process (1.14) converges to the solution of the equation (1.1). From (1.16) it follows that

$$y_l - x^* = Q^l(y_0 - x^*). \tag{1.17}$$

The aim of this paper is to investigate theoretically the behaviour of the sequence $\{x_k - x^*\}_{k=0}^n$, where the sequence $\{x_k\}_{k=0}^n$ is obtained by using Algorithm 1.2, and on the basis of the results obtained to draw conclusions for practical calculations. In view of the connection between extrapolation and projection given in the paper [Si 88], it could be expected that we obtain an analogous formula as in [Si 86] or [Zi 84]. The formula (3.16) in Theorem 3.1 inspired us to use some small number of successive iterations (1.14) before starting conjugate gradients. This approach could be of advantage if one iteration (1.14) costs substantially less work than one step of Algorithm 1.1. For the demonstration of this idea two examples are presented. The numerical experiments show that if we do not calculate with a high accuracy then the use of Algorithm 1.2 could be more advantageous.

In Section 2, preparatory considerations are made. In Section 3, the formula for the difference $x_k - x^*$ is derived and formulated in Theorem 1.3. Some numerical experiments are demonstrated in Section 4.

2. Preparatory considerations

For the vector $x \in \mathbb{R}^n$ we have defined $r(x) = b - Ax$. Let us further set

$$\hat{r}(x) = Mb - (I - Q)x = Mr(x). \tag{2.1}$$

First, we shall prove the following lemma.

Lemma 2.1. *For any positive integer k and any nonzero vector v define spaces*

$$\begin{aligned} W_1 &= \operatorname{span}\{Mv, SMv, S^2Mv, \ldots, S^{k-1}Mv\}, \\ W_2 &= \operatorname{span}\{Mv, QMv, Q^2Mv, \ldots, Q^{k-1}Mv\}. \end{aligned} \tag{2.2}$$

Then

$$W_1 = W_2. \tag{2.3}$$

Proof. Let us consider a matrix polynomial G in the form

$$G = G(S) = \mu_0 I + \mu_1 S + \ldots + \mu_{k-1} S^{k-1}. \tag{2.4}$$

Then let us modify the equality (2.4):

$$G = \sum_{j=1}^{k} \mu_{k-j}(-1)^{k-j}(I - S - I)^{k-j}$$

$$
\begin{aligned}
&= \mu_{k-1}(-1)^{k-1}\sum_{i=0}^{k-1}\binom{k-1}{i}(-1)^i(I-S)^{k-1-i} \\
&\quad + \mu_{k-2}(-1)^{k-2}\sum_{i=0}^{k-2}\binom{k-2}{i}(-1)^i(I-S)^{k-2-i} + \dots \\
&\quad + \mu_1(-1)\sum_{i=0}^{1}(-1)^i(I-S)^{1-i} + \mu_0 I \\
&= \nu_{k-1}(I-S)^{k-1} + \nu_{k-2}(I-S)^{k-2} + \dots + \nu_1(I-S) + \nu_0 I,
\end{aligned}
$$

where numbers ν_{k-l} are defined by the last equality according to which

$$
\nu_{k-l} = (-1)^{k-l}\sum_{i=1}^{l}\mu_{k-i}\binom{k-i}{l-i} \quad \text{for } l = 1, \dots, k. \tag{2.5}
$$

Let N_{k-1} be a lower triangular matrix whose lth row is the row vector

$$
(-1)^{k-l}\left[\binom{k-1}{l-1}, \binom{k-2}{l-2}, \dots, \binom{k-l}{0}, 0, \dots, 0\right]. \tag{2.6}
$$

If we set

$$
\begin{aligned}
\mu &= (\mu_{k-1}, \mu_{k-2}, \dots, \mu_0)^T, \\
\nu &= (\nu_{k-1}, \nu_{k-2}, \dots, \nu_0)^T,
\end{aligned}
$$

then the above mentioned procedure reveals that

$$
\nu = N_{k-1}\mu. \tag{2.7}
$$

At the same time N_{k-1} is a lower triangular matrix with $+1$ and -1 alternating on the diagonal. Thus, it is nonsingular.

If $q \in W_1$ then $q = \sum_{i=0}^{k-1} \mu_i S^i(Mv)$. However,

$$
\sum_{i=0}^{k-1}\mu_i S^i(Mv) = \sum_{i=0}^{k-1}\nu_i(I-S)^i(Mv),
$$

where the numbers ν_i are defined by the relation (2.7) and thus $q \in W_2$. As N_{k-1} is nonsingular, it is clear that the inverse implication, i.e., $q \in W_2 \Rightarrow q \in W_1$, holds. $\square$

Before examining Algorithm 1.2 let us present one of the characteristics of Krylov subspaces. The matrix S is similar to the symmetric matrix C^TAC and thus the eigenvectors of the matrix S form a base in $\mathbb{R}^n$. Let (ξ_1, w_1), (ξ_2, w_2), $\dots (\xi_n, w_n)$ be all eigenpairs of the matrix S. The pairs are written as follows (eigenvalue, eigenvector). Let us also define the sets $\mathcal{N} = \{1, 2, \dots, n\}$ and $\mathcal{N}_i = \mathcal{N} - \{i\}$ for all i. Let $\mathcal{W} = \operatorname{span}\{w_i\}_{i\in\mathcal{N}}$, $\mathcal{W}_i = \operatorname{span}\{w_j\}_{j\in\mathcal{N}_i}$. Since $\dim \mathcal{W}_i = n - 1$ there exists a nonzero vector $v_i \perp \mathcal{W}_i$ such that $v_i^T w_i \neq 0$. (The equality relation is met only for the null vector v_i.) Let us also denote by P_i $(i = 1, \dots, n)$ the projection $\mathbb{R}^n$ to $\operatorname{span}\{w_i\}$.

Theorem 2.1. *Let ξ_i $(i = 1, \dots, n)$ be mutually different numbers and let at least k vectors $(k \leqslant n)$ from the set $P_1Mr(x_0), P_2Mr(x_0), \dots, P_nMr(x_0)$ be nonzero. Then the Krylov subspace*

$$\operatorname{span}\{Mr(x_0), SMr(x_0), \dots, S^{k-1}Mr(x_0)\} \tag{2.8}$$

has the dimension k.

Proof. Without any loss of generality, let us suppose that the first k vectors

$$P_1Mr(x_0),\ P_2Mr(x_0),\ \dots,\ P_kMr(x_0)$$

are nonzero. Let us write the spectral decomposition

$$Mr(x_0) = \sum_{i=0}^{n} P_iMr(x_0). \tag{2.9}$$

This implies that

$$S^jMr(x_0) = \sum_{i=1}^{n} \xi_i^j P_iMr(x_0), \quad j = 0, 1, \dots, k-1. \tag{2.10}$$

Let us suppose that the vectors on the left hand side of (2.10) for $j = 0, 1, \dots, k-1$ are linearly dependent. Then there exist real numbers $\gamma_0, \dots, \gamma_{k-1}$ of which at least one is nonzero, such that

$$\sum_{j=0}^{k-1} \gamma_j S^jMr(x_0) = \Theta. \tag{2.11}$$

Let us multiply the equality (2.11) from the left successively by the vectors v_s^T $(s = 1, 2, \dots, k)$ defined above (i.e., $v_s \perp \mathcal{W}_s$ and $v_s^T w_s \neq 0$) and let us substitute from the relation (2.10), then we obtain

$$\sum_{j=0}^{k-1} \gamma_j \xi_s^j v_s^T P_sMr(x_0) = 0 \quad \text{for } s = 1, 2, \dots, k. \tag{2.12}$$

Since the vector $P_s Mr(x_0)$ is, due to the difference of the eigenvalues ξ_i, a multiple of the vector w_s, then $v_s^T P_s Mr(x_0) \neq 0$ and after reducing in (2.12), we obtain

$$\gamma_{k-1}\xi_s^{k-1} + \gamma_{k-2}\xi_s^{k-2} + \ldots + \gamma_0 = 0 \quad \text{for } s = 1, 2, \ldots, k, \tag{2.13}$$

which means that the polynomial $\gamma_{k-1}z^{k-1} + \gamma_{k-2}z^{k-2} + \ldots + \gamma_0$ of the $(k-1)$th degree has k different roots, which is a contradiction. □

From the proof of Theorem 2.1 it is easy to see how to reformulate this theorem in the case that the eigenvalues are not mutually different. But such investigations are not the purpose of our paper and therefore we will omit them. Let us mention that according to Lemma 2.1 also the space

$$\operatorname{span}\{Mr(x_0), QMr(x_0), Q^2Mr(x_0), \ldots, Q^{k-1}Mr(x_0)\} \tag{2.14}$$

has the dimension k.

Let us suppose the following:

Supposition 1. The space (2.14) has the dimension k.

And now we will go back to the conjugate gradient method.

With respect to Lemma 2.1

$$x_k = x_0 + \nu_0 Mr(x_0) + \nu_1 QMr(x_0) + \ldots + \nu_{k-1}Q^{k-1}Mr(x_0), \tag{2.15}$$

where the numbers $\nu_0, \nu_1, \ldots, \nu_{k-1}$ are constructed so that the conditions of verticality (11.1) are fulfilled, i.e.,

$$r(x_k) \perp Q^s Mr(x_0) \quad \text{for } j = 0, 1, \ldots, k-1. \tag{2.16}$$

Let us define the numbers $\alpha_0^{(k)}, \alpha_1^{(k)}, \ldots, \alpha_k^{(k)}$ as solutions of the following system of linear algebraic equations:

$$\begin{array}{llllllll}
\alpha_k^{(k)} & +\alpha_{k-1}^{(k)} & +\ldots & & +\alpha_1^{(k)} & +\alpha_0^{(k)} & = 1 \\
\alpha_k^{(k)} & +\alpha_{k-1}^{(k)} & +\ldots & & +\alpha_1^{(k)} & & = \nu_0 \\
\alpha_k^{(k)} & +\alpha_{k-1}^{(k)} & +\ldots & +\alpha_2^{(k)} & & & = \nu_1 \\
\multicolumn{7}{c}{\dotfill} \\
\alpha_k^{(k)} & +\alpha_{k-1}^{(k)} & & & & & = \nu_{k-2} \\
\alpha_k^{(k)} & & & & & & = \nu_{k-1}
\end{array} \tag{2.17}$$

According to (2.1), $\hat{r}(x) = Qx + Mb - x$. If $\{y_l\}_{l=0}^{\infty}$ is a sequence obtained by the iterative process (1.14′) then

$$\hat{r}(x_0) = \hat{r}(y_m) = y_{m+1} - y_m \tag{2.18}$$

and it follows that

$$Q^t \hat{r}(x_0) = y_{m+t+1} - y_{m+t}. \tag{2.19}$$

If we put $\hat{r}(x_0) = Mr(x_0)$ in the relation (2.15) and substitute for $\nu_0, \ldots, \nu_{k-1}$ from the equalities (2.17), we obtain according to (2.19)

$$\begin{aligned} x_k &= \left(\sum_{t=0}^{k} \alpha_t^{(k)}\right) y_m + \left(\sum_{t=1}^{k} \alpha_t^{(k)}\right)(y_{m+1} - y_m) \\ &+ \left(\sum_{t=2}^{k} \alpha_t^{(k)}\right)(y_{m+2} - y_{m+1}) + \ldots + \left(\sum_{t=k}^{k} \alpha_t^{(k)}\right)(y_{m+k} - y_{m+k-1}) \\ &= \alpha_0^{(k)} y_m + \alpha_1^{(k)} y_{m+1} + \ldots + \alpha_k^{(k)} y_{m+k}. \end{aligned} \tag{2.20}$$

Let us calculate $r(x_k)$ by substituting from the first equality (2.17) and (2.20):

$$\begin{aligned} r(x_k) &= b - Ax_k = b - \sum_{t=0}^{k} \alpha_t^{(k)} A y_{m+t} \\ &= \sum_{t=0}^{k} \alpha_t^{(k)} (b - Ay_{m+t}) = \sum_{t=0}^{k} \alpha_t^{(k)} r(y_{m+t}). \end{aligned} \tag{2.21}$$

Now we will formulate the conditions of verticality (2.16) by substituting from (2.20), (2.21) and (2.1) thus obtaining the following system of linear algebraic equations:

$$\begin{aligned} &\sum_{t=0}^{k} (Q^s \hat{r}(x_0))^T r(y_{m+t}) \alpha_t^{(k)} = 0, \quad s = 0, \ldots, k-1, \\ &\sum_{t=0}^{k} \alpha_t^{(k)} = 1. \end{aligned} \tag{2.22}$$

Lemma 2.2. *If we denote by* $B = (b_{st})_{\substack{s=0,\ldots,k \text{ (rows)} \\ t=0,\ldots,k \text{ (columns)}}}$ *the matrix of the system (2.22), then*

$$b_{st} = (C^{-1} Q^s \hat{r}(x_0))^T (C^{-1} Q^t \hat{r}(x_0)) \tag{2.23}$$

for $t = 0, \ldots, k$, $s = 0, \ldots, k-1$, *and*

$$b_{k,t} = 1 \tag{2.24}$$

for $t = 0, \ldots, k$.

P r o o f. Let us modify the coefficients of the system (2.22) according to (2.1) and (1.7)

$$\begin{aligned}(Q^s\hat{r}(x_0))^T r(y_{m+t}) &= \hat{r}(x_0)^T (Q^s)^T M^{-1}\hat{r}(y_{m+t}) \\ &= \hat{r}(x_0)^T (Q^s)^T C^{-1} C^{-1} Q^t \hat{r}(x_0) \\ &= (C^{-1}Q^s\hat{r}(x_0))^T \cdot (C^{-1}Q^t\hat{r}(x_0)).\end{aligned}$$

□

It follows from the relations (2.20) and (2.17) that

$$x_k - x^* = \sum_{t=0}^{k} \alpha_t^{(k)} (y_{m+t} - x^*). \tag{2.25}$$

The conditions (2.16) yield a system of linear algebraic equations for the unknowns $\nu_0, \dots, \nu_{k-1}$. The matrix of this system is positive definite because of Supposition 1 and the vector $\alpha^{(k)} = (\alpha_0^{(k)}, \alpha_1^{(k)}, \dots, \alpha_k^{(k)})^T$ fulfilling (2.17) solves the system (2.22). If the matrix B of the system (2.22) is singular then the kernel $N(B) \supsetneqq \{0\}$ and every vector $\alpha^{(k)} + w$, with $w \in N(B)$ solves (2.22) which is in contradiction to the existence of only one solution of (2.16) in view of (2.17).

Therefore, the matrix B is nonsingular.

Note. We do not indicate the dependence of the matrix B on m and write only B.

3. Calculating the difference $x_k - x^*$

Let (λ_i, u_i) be eigenpairs of the matrix $Q = I - MA$. Since the matrix Q is similar to the symmetric matrix $I - C^T AC$, all the numbers λ_i are real. Let the eigenvalues be numbered so that

$$|\lambda_1| \geqslant |\lambda_2| \geqslant \dots \geqslant |\lambda_n|.$$

In the following it will be seen that without any loss of generality we can suppose:

Supposition 2. Let

$$|\lambda_{k+1}| > |\lambda_{k+2}| \quad \text{and} \quad \lambda_i \neq 1 \ \forall i. \tag{3.1}$$

Let

$$y_0 - x^* = \sum_{i=1}^{n} \beta_i u_i \tag{3.2}$$

be the spectral decomposition of the vector $y_0 - x^*$. According to (1.17) and (3.2),

$$y_{m+t} - x^* = \sum_{i=1}^{n} \beta_i \lambda_i^{m+t} u_i, \tag{3.3}$$

$$y_{m+t+1} - y_{m+t} = \sum_{i=1}^{n} \beta_i (\lambda_i - 1) \lambda_i^{m+t} u_i = Q^t \hat{r}(x_0). \tag{3.4}$$

The last equality follows from (2.19).

If we set

$$v_i = C^{-1} \beta_i (\lambda_i - 1) u_i, \quad i = 1 \ldots, n, \tag{3.5}$$

then v_i is the eigenvector of the symmetric matrix $I - C^T AC$ and thus

$$v_i^T v_j = 0 \quad \text{for } i \neq j. \tag{3.6}$$

Lemma 3.1. *If $\gamma_i = v_i^T v_i$ then for the elements of the matrix B, defined by the relation* (2.23),

$$b_{st} = \sum_{i=1}^{n} \gamma_i \lambda_i^{2m+s+t} \tag{3.7}$$

holds for $s = 0, \ldots, k-1$, $t = 0, \ldots, k$.

P r o o f. According to (3.4), (3.5) and (3.6),

$$\begin{aligned} b_{st} &= \left(C^{-1} Q^s \hat{r}(x_0) \right)^T (C^{-1} Q^t \hat{r}(x_0)) \\ &= \left(\sum_{i=1}^{n} \lambda_i^{m+s} v_i \right)^T \left(\sum_{j=1}^{n} \lambda_j^{m+t} v_j \right) = \sum_{i=1}^{n} \gamma_i \lambda_i^{2m+s+t} \end{aligned}$$

holds. □

If

$$\alpha^{(k)} = (\alpha_0^{(k)}, \ldots, \alpha_k^{(k)})^T,$$

we can rewrite the system (2.22) in the form

$$B\alpha^{(k)} = e_{k+1}. \tag{3.8}$$

Now we will calculate the coefficients $\alpha_i^{(k)}$ according to Crammer's rule, substitute them into the sum (2.25) and express this sum in the form of a quotient of determinants. This procedure is well-known (see [Si 86]) and so we can write directly the expression for $x_k - x^*$ using the following simplifying notation.

For any positive integer t and an integer $s \in \langle 1, \dots, n \rangle$ let us define the following vectors $c_t \in \mathbb{R}^k$, $g_t \in \mathbb{R}^n$ and $d_{ts} \in \mathbb{R}^k$:

$$
\begin{aligned}
c_t &= \left(\sum_{i=1}^{n} \gamma_i \lambda_i^t, \sum_{i=1}^{n} \gamma_i \lambda_i^{t+1}, \dots, \sum_{i=1}^{n} \gamma_i \lambda_i^{t+k-1} \right)^T, \\
d_{ts}^{(k)} &= (\lambda_s^t, \lambda_s^{t+1}, \dots, \lambda_s^{t+k-1})^T, \\
g_t &= \sum_{i=1}^{n} \beta_i \lambda_i^t u_i.
\end{aligned}
\tag{3.9}
$$

According to (2.25), (3.7), (3.8) and (3.9),

$$
x_k - x^* = \frac{\det \begin{pmatrix} c_{2m}, & c_{2m+1}, & \dots, & c_{2m+k} \\ g_m, & g_{m+1}, & \dots, & g_{m+k} \end{pmatrix}}{\det \begin{pmatrix} c_{2m}, & c_{2m+1}, & \dots, & c_{2m+k} \\ 1, & 1, & \dots, & 1 \end{pmatrix}}.
\tag{3.10}
$$

In view of the notation (3.9) we have

$$
\begin{pmatrix} c_{2m}, & c_{2m+1}, & \dots, & c_{2m+k} \\ g_m, & g_{m+1}, & \dots, & g_{m+k} \end{pmatrix}
$$

$$
= \begin{pmatrix}
\sum_{i=1}^{n} \gamma_i \lambda_i^{2m}, & \sum_{i=1}^{n} \gamma_i \lambda_i^{2m+1}, & \dots, & \sum_{i=1}^{n} \gamma_i \lambda_i^{2m+k} \\
\sum_{i=1}^{n} \gamma_i \lambda_i^{2m+1}, & \sum_{i=1}^{n} \gamma_i \lambda_i^{2m+2}, & \dots, & \sum_{i=1}^{n} \gamma_i \lambda_i^{2m+k+1} \\
\cdots & \cdots & \cdots & \cdots \\
\sum_{i=1}^{n} \gamma_i \lambda_i^{2m+k-1}, & \sum_{i=1}^{n} \gamma_i \lambda_i^{2m+k}, & \dots, & \sum_{i=1}^{n} \gamma_i \lambda_i^{2m+2k-1} \\
\sum_{i=1}^{n} \beta_i \lambda_i^{m} u_i, & \sum_{i=1}^{n} \beta_i \lambda_i^{m+1} u_i, & \dots, & \sum_{i=1}^{n} \beta_i \lambda_i^{m+k} u_i
\end{pmatrix}
$$

and analogously in the denominator of (3.10). The determinant of this matrix is taken exactly according to definition. Thus it is the sum of $(k+1)!$ vectors. Each term of addition is a multiple of one of the vectors $u_1, \dots, u_m$, which will become evident from further modifications. A modification of the determinant of a similar type is described in the works [Si 86], [Zi 84] or [He]. For our case the process presented in [He] is the most suitable and we will now briefly outline it to elucidate the whole situation.

By $\mathcal{K}$ we denote the set of all $(k+1)$-tuples $(l_1, l_2, \dots, l_{k+1})$ of integers such that $l_i \in \{1, 2, \dots, n\}$ and $l_i \neq l_j$ for $i \neq j$. Further, let $\mathcal{L}$ be the system of all permutations of the numbers $1, 2, \ldots, k$. According to this notation,

$$\det\begin{pmatrix} c_{2m}, & c_{2m+1}, & \ldots, & c_{2m+k} \\ g_m, & g_{m+1}, & \ldots, & g_{m+k} \end{pmatrix}$$
$$= \sum_{(l_1,\ldots,l_{k+1})\in\mathcal{K}} \det\begin{pmatrix} \gamma_{l_1} d^{(k)}_{2m,l_1}, & \gamma_{l_2} d^{(k)}_{2m+1,l_2}, & \ldots, & \gamma_{l_{k+1}} d^{(k)}_{2m+k,l_k+1} \\ \beta_{l_1}\lambda_{l_1}^m u_{l_1}, & \beta_{l_2}\lambda_{l_2}^{m+1} u_{l_2}, & \ldots, & \beta_{l_{k+1}}\lambda_{l_{k+1}}^{m+k} u_{l_{k+1}} \end{pmatrix}$$
$$= \sum_{(l_1,\ldots,l_{k+1})\in\mathcal{K}} \lambda_{l_1}^m \lambda_{l_2}^m \ldots \lambda_{l_{k+1}}^m \det\begin{pmatrix} \gamma_{l_1} d^{(k)}_{m,l_1}, & \gamma_{l_2} d^{(k)}_{m+1,l_2}, & \ldots, & \gamma_{l_{k+1}} d^{(k)}_{m+k,l_k+1} \\ \beta_{l_1} u_{l_1}, & \beta_{l_2}\lambda_{l_2} u_{l_2}, & \ldots, & \beta_{l_{k+1}}\lambda_{l_{k+1}}^{k} u_{l_{k+1}} \end{pmatrix}$$

(we expand the determinant with respect to the last line)

$$= \sum_{(l_1,\ldots,l_{k+1})\in\mathcal{K}} \Bigg\{ \beta_{l_1} u_{l_1} \lambda_{l_1}^m \Bigg(\prod_{\substack{j=1\\ j\neq 1}}^{k+1} \lambda_{l_j}\Bigg)^{2m} \Bigg(\prod_{\substack{j=1\\ j\neq 1}}^{k+1} \gamma_{l_j}\Bigg) \lambda_{l_2}\lambda_{l_3}^2 \ldots, \lambda_{l_{k+1}}^k$$
$$\times \det(d^{(k)}_{0l_2}, d^{(k)}_{0l_3}, \ldots, d^{(k)}_{0l_{k+1}})(-1)^k$$
$$+ \beta_{l_2}\lambda_{l_2} u_{l_2} \lambda_{l_2}^m \Bigg(\prod_{\substack{j=1\\ j\neq 2}}^{k+1} \lambda_{l_j}\Bigg)^{2m} \Bigg(\prod_{\substack{j=1\\ j\neq 2}}^{k+1} \gamma_{l_j}\Bigg) \lambda_{l_1}^0 \lambda_{l_3}^2 \ldots \lambda_{l_{k+1}}^k$$
$$\times \det(d^{(k)}_{0l_1}, d^{(k)}_{0l_3}, \ldots, d^{(k)}_{0l_{k+1}})(-1)^{k-1} + \ldots$$
$$+ \beta_{l_{k+1}}\lambda_{l_{k+1}}^k u_{l_{k+1}} \lambda_{l_{k+1}}^m \Bigg(\prod_{\substack{j=1\\ j\neq k+1}}^{k+1} \lambda_{l_j}\Bigg)^{2m} \Bigg(\prod_{\substack{j=1\\ j\neq k+1}}^{k+1} \gamma_{l_j}\Bigg) \lambda_{l_1}^0 \lambda_{l_2}^1 \ldots \lambda_{l_k}^{k-1}$$
$$\times \det(d^{(k)}_{0l_1}, \ldots, d^{(k)}_{0l_k})(-1)^0 \Bigg\} = (\,*\,).$$

In what follows, let the inequality

$$(3.1') \qquad |\lambda_k| > |\lambda_{k+1}|$$

be valid. Then due to Supposition 2 we can find a dominant member in each term of summation. E.g. for the first term of summation we obtain it if we put $l_1 = k+1$ and $\prod_{\substack{j=1\\ j\neq 1}}^{k+1} \lambda_{l_j} = \lambda_1\lambda_2\ldots\lambda_k$ and analogously for the other terms of summation. Let us proceed further in the modification from $(\,*\,)$ and to make it shorter let us also set for $i = 1, 2, \ldots, k+1$

$$(3.11) \qquad D_i^{(k)} = \det(d^{(k)}_{0l_1}, \ldots, d^{(k)}_{0l_{i-1}}, d^{(k)}_{0l_{i+1}}, \ldots, d^{(k)}_{0l_{k+1}}).$$

We have

$$(*) = \beta_{k+1} u_{k+1} \lambda_{k+1}^{m} \left(\prod_{j=1}^{k} \lambda_j \right)^{2m}$$
$$\times \left[\sum_{p=1}^{k+1} \lambda_{k+1}^{p-1} \sum_{(l_1,\dots,l_{p-1},l_{p+1},\dots,l_{k+1}) \in \mathcal{L}} \left\{ \left(\prod_{\substack{j=1 \\ j\neq p}}^{k+1} \gamma_{l_j} \right) \left(\prod_{\substack{j=1 \\ j\neq p}}^{k+1} \lambda_{l_j}^{j-1} \right) D_p^{(k)} (-1)^{k+1-p} \right\} \right]$$
$$+ \tilde{w}(m) = (**),$$

where $\tilde{w}(m)$ is a vector for which

$$\lim_{m\to\infty} \tilde{w}(m) \Big/ \left(\lambda_{k+1}^{m} \left(\prod_{j=1}^{k} \lambda_j \right)^{2m} \right) = \Theta \tag{3.12}$$

holds.

To make things clear we shall expand $\left[\sum_{p=1}^{k+1} \cdot\right]$ and denote the k-tuple from the set $\mathcal{L}$ by $(s_1, \dots, s_k)$.

$$(**) = \beta_{k+1} u_{k+1} \lambda_{k+1}^{m} \left(\prod_{j=1}^{k} \lambda_j \right)^{2m} \left(\prod_{j=1}^{k} \gamma_j \right) \det(d_{01}^{(k)}, \dots, d_{0k}^{(k)})$$
$$\times \left\{ \left[(-1)^k \sum_{(s_1,\dots,s_k)\in\mathcal{L}} \operatorname{sign} \begin{pmatrix} 1, & 2, & \dots, & k \\ s_1, & s_2, & \dots, & s_k \end{pmatrix} \lambda_{s_1} \lambda_{s_2}^2 \lambda_{s_3}^3 \dots \lambda_{s_k}^k \right. \right.$$
$$+ (-1)^{k-1} \lambda_{k+1} \sum_{(s_1,\dots,s_k)\in\mathcal{L}} \operatorname{sign} \begin{pmatrix} 1, & 2, & \dots, & k \\ s_1, & s_2, & \dots, & s_k \end{pmatrix} \lambda_{s_1}^0 \lambda_{s_2}^2 \lambda_{s_3}^3 \dots \lambda_{s_k}^k$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$\left. + (-1)^0 \lambda_{k+1}^k \sum_{(s_1,\dots,s_k)\in\mathcal{L}} \operatorname{sign} \begin{pmatrix} 1, & 2, & \dots, & k \\ s_1, & s_2, & \dots, & s_k \end{pmatrix} \lambda_{s_1}^0 \lambda_{s_2}^1 \lambda_{s_3}^2 \dots \lambda_{s_k}^{k-1} \right]$$
$$\left. + w_1(m) \right\}$$

where

$$\lim_{m\to\infty} w_1(m) = \Theta. \tag{3.13}$$

The expression in the brackets of (3.11) is equal to

$$\det \begin{pmatrix} 1 & \lambda_1 & \lambda_1^2 & \dots & \lambda_1^k \\ 1 & \lambda_2 & \lambda_2^2 & \dots & \lambda_2^k \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \lambda_{k+1} & \lambda_{k+1}^2 & \dots & \lambda_{k+1}^k \end{pmatrix},$$

which equals, according to (3.9), to

$$\det\left(d_{01}^{(k+1)}, d_{02}^{(k+1)}, \dots, d_{0k+1}^{(k+1)}\right)^T.$$

Let us mention that if l is an integer then $(d_{01}^{(l)}, d_{02}^{(l)}, \dots, d_{0l}^{(l)})$ is the Vandermond matrix.

Thus the numerator in (3.10) equals

$$\beta_{k+1}\lambda_{k+1}^m\left(\prod_{j=1}^k \lambda_j\right)^{2m}\left(\prod_{j=1}^k \gamma_j\right) \times \det(d_{01}^{(k)}, \dots, d_{0k}^{(k)})\det(d_{01}^{(k+1)}, \dots, d_{0k}^{(k+1)})^T(u_{k+1} + w(m)) \tag{3.14}$$

and the sequence $\{w(m)\}$ satisfies (3.13).

The denominat in (3.10) is handled in exactly the same way. Let us write only the result. The denominator in (10.3) is equal to the expression

$$\left(\prod_{j=1}^k \lambda_j\right)^{2m}\left(\prod_{j=1}^k \gamma_j\right)\det(d_{01}^{(k)}, \dots, d_{0k}^{(k)})\det(d_{01}^{(k+1)}, \dots, d_{0k}^{(k+1)}, e(k+1))^T(1+\varphi(m)), \tag{3.15}$$

with $\lim_{m\to\infty}\varphi(m) = 0$.

And now we can summarize the results of the whole study in the following theorem.

Theorem 3.1. *Let the matrix A in the system* (1.1) *be symmetric and positive definite. Let (λ_i, u_i) be eigenpairs of the matrix Q. Let x^* be the solution of the system* (1.1) *found by applying Algorithm* 1.2. *Let Supposition* 1, *Supposition* 2 *and* (3.1′) *be fulfilled. Then*

$$x_k - x^* = \delta_{k+1}\lambda_{k+1}^m(\beta_{k+1}u_{k+1} + v(m)) \tag{3.16}$$

holds, where

$$\delta_{k+1} = \frac{\det(d_{01}^{(k+1)}, d_{02}^{(k+1)}, \dots, d_{0k}^{(k+1)}, d_{0k+1}^{(k+1)})}{\det(d_{01}^{(k+1)}, d_{02}^{(k+1)}, \dots, d_{0k}^{(k+1)}, e)} \tag{3.17}$$

does not depend on m and

$$\lim_{m\to\infty} v(m) = \Theta. \tag{3.18}$$

Proof of (3.16) follows directly from the relations (3.14) and (3.15), the relation (3.18) follows from (3.13). □

N o t e 1. The sequence $\{v(m)\}$ converges towards zero asymptotically as the quotient $(\frac{\lambda_{k+1}}{\lambda_k})^m$.

N o t e 2. From (3.16) we can see directly that after n steps we obtain the exact solution x^*.

Let only Supposition 2 be fulfilled. For example, let the following equality be valid:

$$|\lambda_{k_0}| = |\lambda_{k_0+1}| = \ldots = |\lambda_k| = |\lambda_{k+1}|.$$

It is easy to see from the calculations before Theorem 3.1 that the relation

$$x_k - x^* = \tilde{\delta}_{k+1}\lambda_{k+1}^m(\tilde{u}^{(k+1)} + \tilde{v}(m)) \tag{3.16'}$$

holds, where

$$\tilde{u}^{(k+1)} \in \operatorname{span}(u_{k_0}, u_{k_0+1}, \ldots, u_{k+1}) \tag{3.17'}$$

and $\tilde{\delta}_{k+1}$ is a constant depending only on k. We do not know any elegant formula for $\tilde{\delta}_{k+1}$ analogous to (3.17).

N o t e 3. It is well known (see [D.O'L]) that the vector x_k minimizes the functional $(x - x^*)^T A(x - x^*)$ over all x such that

$$x \in x_0 + \operatorname{span}\{Mr(x_0), QMr(x_0), \ldots, Q^{k-1}Mr(x_0)\}.$$

Let $\|x\|_A = \sqrt{x^T Ax}$. From (2.25) we have

$$\|x_k - x^*\|_A = \left\| \sum_{t=0}^{k} \alpha_t^k (y_{m+t} - x^*)^T \right\|_A . \tag{3.19}$$

If we substitute in the formula (3.19) instead of $\alpha_k^{(k)}, \alpha_{k-1}^{(k)}, \ldots, \alpha_0^{(k)}$ successively the coefficients of the polynom $q(t)/q(1)$, where

$$q(t) = (t - \lambda_1)(t - \lambda_2)\ldots(t - \lambda_k),$$

and instead of $y_{m+t} - x^*$ the right-hand side of (3.3) we obtain the estimate

$$\|x_k - x^*\|_A \leqslant \left\| \sum_{i=k+1}^{n} \lambda_i^m \beta_i q(\lambda_i) u_i \right\|_A = \lambda_{k+1}^m \chi(m) \tag{3.20}$$

where $\limsup_{m\to\infty} \chi(m) < \infty$. We have that $q(1) \neq 0$ according to Supposition 2.

N o t e 4. It follows from the relations (3.16), (3.17), (3.18) or (3.16′) or (3.17') that if we set $\eta(m) = \|x_k - x^*\|$ for a given m in Algorithm 1.2, then there exists $\hat{m} > m$ such that $\eta(\hat{m}) < \eta(m)$. Numerical experiments show that $\eta(m+1) < \eta(m)$ for all m.

4. Numerical results

In the paper [E-G] the authors consider the convection-diffusion equation

$$-\Delta\mathbf{x}(s,t)+\sigma(s,t)\mathbf{x}_s+\tau(s,t)\mathbf{x}_t=\mathbf{f}(s,t) \tag{4.1}$$

on the unit square $\Omega=(0,1)\times(0,1)$ with the homogeneous Dirichlet boundary condition on $\partial\Omega$. We suppose that $\tau(s,t)>0$ and $\sigma(s,t)>0$ on Ω. For discretization we consider a uniform mesh with the mesh size $h=1/(N+1)$, where N is a number of inner mesh points in both directions s and t.

For the five-point finite difference discretization and *red*-**black** ordering of mesh points the authors obtain a system of linear algebraic equations

$$\begin{pmatrix} D & C \\ E & F \end{pmatrix}\begin{pmatrix} x^{(r)} \\ x^{(b)} \end{pmatrix}=\begin{pmatrix} f^{(r)} \\ f^{(b)} \end{pmatrix} \tag{4.2}$$

for the approximate solution at mesh points represented by the vector $(x^{(r)},x^{(b)})^T$. The components denote successively the approximate values of $\mathbf{x}$ at *red* and **black** points. The matrix of the system (2.4) is σ_1-ordered. With one step of cyclic reduction the *red* points are eliminated. If the reduced **black** points are ordered by diagonal lines in the NW-SE direction, then the matrix of the reduced system

$$\Big(F-ED^{-1}C\Big)x^{(b)}=f^{(b)}-ED^{-1}f^{(r)} \tag{4.3}$$

has the block tridiagonal form and the diagonal blocks are tridiagonal matrices. Let us assume that $\sigma(s,t)\equiv\sigma$ and $\tau(s,t)\equiv\tau$ are constant. If $\max(\sigma,\tau)<2/h$ then the matrix in (4.3) can be symmetrized with a real diagonal matrix. (See [E-G] or [Zi 92].)

Acording to this, if we leave this NW-SE direction alone and number the **black** points again by *red*-**black** then after easy transformations by a diagonal and permutation matrix the system (4.3) can be transformed into the form

$$\begin{pmatrix} \tilde{D}_1 & \tilde{C}^T \\ \tilde{C} & \tilde{D}_2 \end{pmatrix}\begin{pmatrix} \tilde{x}_1^{(b^{(red)})} \\ \tilde{x}_2^{(b)} \end{pmatrix}=\begin{pmatrix} \tilde{f}_1^{(b^{(red)})} \\ \tilde{f}_2^{(b)} \end{pmatrix}, \tag{4.4}$$

where in the case of constant coefficients all nonzero elements of the matrix $\tilde{C}$ equal -1, $\tilde{D}_1$ and $\tilde{D}_2$ are Stieltjes matrices and the matrix of the system (4.4) is positive definite. The dimension of the system (4.4) equals the number of the original black points.

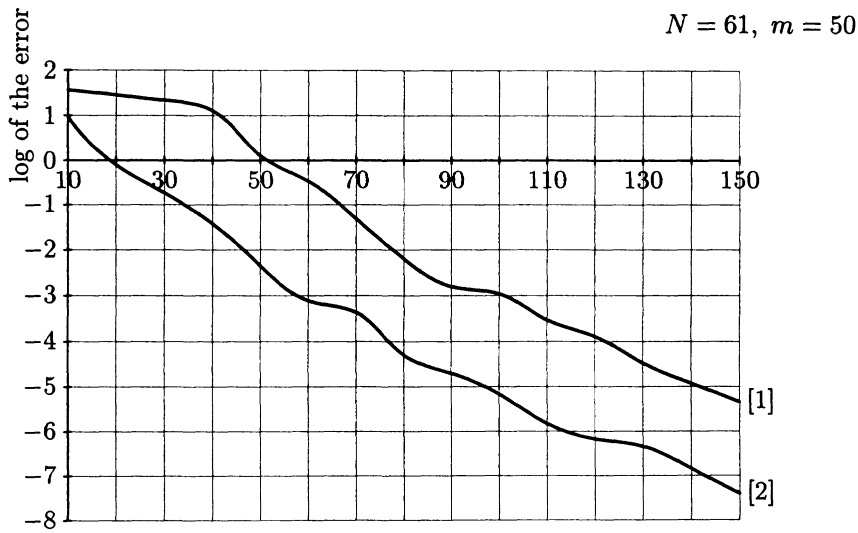
Now, we compare the numerical results which we obtain by applying Algorithms 1.1 and 1.2, respectively. We have used Algorithms 1.1 and 1.2 with

$$f=0,\quad \sigma=1,\quad \tau=2\quad\text{and}\quad M^{-1}=\operatorname{diag}(\tilde{D}_1,\tilde{D}_2).$$

The numbers $\log_{10}\|x_k - x^*\|$ are compared for both the algorithms. On the following Graph 1 the lines [1] and [2] correspond to the behaviour of $\log_{10}\|x_k - x^*\|$ for Algorithms 1.1 and 1.2, respectively. We have taken $m = 50$. The system was tested for a sequence for various N and the behaviour was similar. In this case we see that if $k_{11}(\varepsilon)$ and $k_{12}(\varepsilon)$ denote the first integer for which

$$\log_{10}\|x_k - x^*\| < \varepsilon,$$

obtained by using Algorithm 1.1 and 1.2, respectively, then $k_{11}(\varepsilon) - k_{12}(\varepsilon) \doteq m$ for sufficiently small ε. However, one iteration of (1.14) needs only the inverse of $\operatorname{diag}(\tilde{D}_1, \tilde{D}_2)$ if we have in view that the other elements out of the block diagonal equal -1. One iteration in Algorithm 1.1 needs, moreover, $10n$ multiplications where $n = N^2/2$ or $(N^2+1)/2$ if N is even or odd, respectively. The numerical tests show that using Algorithm 1.2 could be more advantageous for $m \in \langle 5, 50\rangle$.



Graph 1

The second example is analogous to the first which is in Appendix B in [V]. The matrix equation arises from discrete approximation of the second-order selfadjoint elliptic partial differential equation

$$(4.5)\qquad \begin{aligned} -(D(s,t)x_s)_s - (D(s,t)x_t)_t + \sigma(s,t)u = S(s,t) \quad \text{on } R,\\ \partial x/\partial n = 0 \quad \text{on } \Gamma(R), \end{aligned}$$

where R is the square $0 < s,t < 2.1$ divided into nine regions and the functions $D(s,t) > 0$, $\sigma(s,t) > 0$, and $S(s,t)$ are piecewise constant in every region. They are defined by the table given in [V].

We have used a finer mesh so that we have solved the system of linear algebraic equations having 8100 unknowns. In the following table we compare the norms of the errors $\|x_k - x^*\|$ for chosen iterations k by using Algorithms 1.1 and 1.2, respectively. We take the iteration (1.14) as preparatory work and compare only conjugate gradients in both cases. The following table is for $m = 5$.

k	ALGORITHM 1.1	ALGORITHM 1.2
25	$0.438_{10}+1$	$0.131_{10}+0$
30	$0.817_{10}-1$	$0.303_{10}-2$
35	$0.214_{10}-2$	$0.527_{10}-4$
40	$0.403_{10}-4$	$0.879_{10}-7$
45	$0.440_{10}-7$	$0.197_{10}-9$
50	$0.150_{10}-9$	$0.839_{10}-12$

Table 1

The following table is for the same linear system for $m = 15$.

k	ALGORITHM 1.1	ALGORITHM 1.2
15	$0.234_{10}+4$	$0.533_{10}+1$
20	$0.267_{10}+3$	$0.364_{10}-1$
25	$0.438_{10}+1$	$0.113_{10}-2$
30	$0.817_{10}-1$	$0.585_{10}-5$
35	$0.335_{10}-2$	$0.335_{10}-8$
40	$0.403_{10}-4$	$0.305_{10}-10$
45	$0.440_{10}-7$	0.000
50	$0.150_{10}-9$	0.000
55	$0.530_{10}-12$	0.000

Table 2

Remark. Let us put the following question: "What happens if we apply Algorithm 1.2 to nonsymmetric linear systems?" For the demonstration we have formed the system

$$(I - \mathcal{L}_\omega)x = c_\omega,$$

where

$$\mathcal{L}_\omega = (D - \omega C_L)^{-1}(\omega C_U + (1-\omega)D),$$
$$c_\omega = (D - \omega C_L)^{-1}\omega b.$$

We have considered the system (1.1) with a block tridiagonal Stieltjes matrix A arising from the five-point difference approximation of the second-order selfadjoint elliptic partial differential equation on a triangle. As usual $A = D - C_L - C_U$ where D is a diagonal, C_L a strictly lower and C_U a strictly upper triangular matrix, respectively. We have substituted part 3) in Algorithm 1.2 by the well known GMRES process and the SOR iterations introduced in part 2). We have used restarted GMRES after ten steps and the following table compares the errors for various m. We have taken $n = 210$ and $\omega = 1.55$.

k	$m = 0$	$m = 20$	$m = 50$
10	$0.511_{10} + 1$	$0.863_{10} - 2$	$0.747_{10} - 9$
15	$0.118_{10} + 0$	$0.181_{10} - 2$	$0.327_{10} - 10$
20	$0.996_{10} - 3$	$0.464_{10} - 4$	$0.424_{10} - 11$
25	$0.155_{10} - 3$	$0.212_{10} - 6$	0.0
30	$0.352_{10} - 5$	$0.130_{10} - 7$	0.0
35	$0.106_{10} - 6$	$0.174_{10} - 8$	0.0
40	$0.711_{10} - 11$	0.0	0.0

Table 3.

Let us add that k denotes the number of restarts of GMRES.
Table 3 is very interesting in view of the fact that one restart of GMRES needs approximately the same number of multiplications as twenty iterations SOR. The program for GMRES in FORTRAN 77 was prepared by my student Miroslav Folprecht.

Acknowledgement. I wish to thank Dr. Michal Křížek, DrSc. for reading and commenting the manuscript of this paper.

References

[G-L] *G.H. Golub, C.F. Van Loan*: Matrix Computation. The John Hopkins University Press, Baltimore, 1984.

[E-G] *H.C. Elman, G.H. Golub*: Block Iterative Methods for Cyclically Reduced Non-Self-Adjoint Elliptic Problems. Chapter 6 in the book "Iterative Methods for Large Linear Systems" edited by David R. Kincaid and Linda J. Hayes, Center for Numerical Analysis The University of Texas at Austin. Academic Press, 1989.

[He] *P. Henrici*: The Quotient-Difference Algorithm. Further Contribution to the Solution of Simultaneous Linear Equations and the Determination of Eigenvalues, Vol. 49, National Bureau of Standards Applied Mathematics Series. 1958.

[D.O'L] *D.P. O'Leary*: The Block Conjugate Gradient Algorithm and Related Methods. Linear Algebra Appl. *29* (1980), 293–322.

[Si 88] *A. Sidi*: Extrapolation vs. Projection Methods for Linear Systems of Equations. J. Comput. Appl. Math. *22* (1988), 71–88.

[Si-F-Sm] *A. Sidi, W.F. Ford, D.A. Smith*: Acceleration of Convergence of Vector Sequences. SIAM J. Numer. Anal. *23* (1986), no. 1, 178–196.

[Si 86] *A. Sidi*: Convergence and Stability Properties of Minimal Polynomial and Reduced Rank Extrapolation Algorithms. SIAM J. Numer. Anal. *23* (1986), no. 1, 197–209.

[S-S 86] *Y. Saad, M.H. Schultz*: GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems. SIAM J. Sci. Stat. Comput. *7* (1986), no. 3, 856–869.

[V] *R.L. Varga*: Matrix iterative analysis. Prentice-Hall Englewood Clifs, New Jersey, 1962.

[V-V 93] *H.A. Van der Vorst, C. Vuik*: The superlinear convergence behaviour of GMRES. J. Comput. Appl. Math. *48* (1993), 327–341.

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

[Zi 83] *J. Zítko*: Improving the Convergence of Iterative Methods. Apl. Mat. *28* (1983), 215–229.

[Zi 84] *J. Zítko*: Convergence of Extrapolation Coefficients. Apl. Mat. *29* (1984), 114–133.

[Zi 92] *J. Zítko*: Numerical experiments with extrapolated procedures. Programy a algoritmy numerické matematiky 6, Sborník kursu, Bratříkov 1992. pp. 178–187. (In Czech.)

[Zi 93] *J. Zítko*: Combining the preconditioned conjugate gradient method and the norm-reducing matrix iterative method. Technical report No 106/93, Prague 1993. pp. 1–17.

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