DOI 10.1007/s10958-018-3916-8
Journal of Mathematical Sciences, Vol. 232, No. 6, August, 2018

TWO-LEVEL LEAST SQUARES METHODS IN KRYLOV SUBSPACES

V. P. Il'in* UDC 519.6

Two-level least squares acceleration approaches are applied to the Chebyshev acceleration method and the restarted conjugate residual method in solving systems of linear algebraic equations with sparse unsymmetric coefficient matrices arising from finite volume or finite element approximations of boundary-value problems on irregular grids. Application of the proposed idea to other iterative restarted processes also is considered. The efficiency of the algorithms suggested is investigated numerically on a set of model Dirichlet problems for the convection-diffusion equation. Bibliography: 6 *titles.*

1. Introduction

The paper considers the classical problem of numerical linear algebra, i.e., the solution of a system of linear algebraic equations (SLAE) of the form

$$Au = f, \quad A = \{a_{k,l}\} \in \mathcal{R}^{N,N}; \quad u, f \in \mathcal{R}^N, \tag{1}$$

where the matrix A and the right-hand side vector f are given. It is assumed that the order N of the system is extremely large, and the coefficient matrix is sparse and ill-conditioned. For these reasons, SLAE (1) is numerically solved by parallel preconditioned iterative methods in Krylov subspaces on a Multiprocessor Computer System (MCS). In the case where the coefficient matrix is unsymmetric, the iterative methods optimal in the order of the convergence rate of an iterative process (such as methods of semi-conjugate residuals (SCR) or generalized minimum residuals (GMRES) [1,2]) use long vector recursions. But in practice such recursions are necessarily shortened by using restarts, which are repeated every m iterations. In this case, the current residual vector is determined from the original equation, and the Krylov iterations are restarted. This results in a significant deceleration of convergence, which is a due fee for limited resources of a computer. Another approach to saving computer resources consists in allowing a limited orthogonality of direction vectors. Obviously, both approaches can also be combined.

The aim of the present paper is to study methods for accelerating the convergence of restarted iterations based on the Least Squares Method (LSM) for decreasing the Euclidean norm of the residual. The approach suggested is described in application to Multi-Preconditioned Semi-Conjugate Residual (MP-SCR) methods [3], which, in some special cases, are equivalent to the widely-used Flexible Generalized Minimum Residual (FGMRES) method [1]. Also we consider application of less expensive Conjugate Residual (CR) algorithms and Chebyshev acceleration, see [2]. Since, in the general unsymmetric case, the latter methods do not possess variational properties and may even diverge in some cases, residual minimization at restarts for the current period is carried out using LCM. Moreover, we additionally use the second optimization level, consisting in using linear combinations of the restart approximations computed by the LSM.

The paper is organized as follows. Section 2 describes the suggested two-level methods for accelerating iterations in Krylov subspaces. The next Sec. 3 is devoted to presenting results of numerical investigation of the algorithms under consideration. In the Conclusion, we consider

*Institute of Computational Mathematics and Mathematical Geophysics SO RAS, Novosibirsk, Russia; Novosibirsk State University, Novosibirsk, Russia, e-mail: ilin@sscc.ru.

Translated from *Zapiski Nauchnykh Seminarov POMI*, Vol. 463, 2017, pp. 224–239. Original article submitted November 1, 2017.

prospects of parallelizing the approaches in question and of applying them in solving practical problems.

2. Two-level iteration methods in Krylov subspaces

Consider a multi-preconditioned semi-conjugate residual method, with a given initial approximation u^0, described by the following formulas:

$$\begin{aligned} r^0 = f - Au^0, \quad p_l^0 = (B_0^{(l)})^{-1} r^0, \quad l = 1, \ldots, M_0, \quad P_0 = [p_1^0 \cdots p_{M_0}^0] \in \mathcal{R}^{N,M_0}; \\ n = 0, 1, \ldots : u^{n+1} = u^n + P_n \bar{\alpha}_n = u^0 + P_0 \bar{\alpha}_0 + \cdots + P_n \bar{\alpha}_n, \\ r^{n+1} = r^n - AP_n \bar{\alpha}_n = r^0 - AP_0 \bar{\alpha}_0 - \ldots - AP_n \bar{\alpha}_n, \\ \bar{\alpha}_n = (\alpha_n^1, \ldots, \alpha_n^{M_n})^T, \quad P_n = [p_1^n \cdots p_{M_n}^n] \in \mathcal{R}^{N,M_n}. \end{aligned} \tag{2}$$

Here, $p_l^0 = (B_0^{(l)-1}) r^n,\ l = 1, \ldots, M_n$, are the direction vectors, which are determined from the initial residual vector r^0 and some nonsingular preconditioning matrices $B_0^{(l)}$; $\bar{\alpha}_n \in \mathcal{R}^{M_n}$ are vectors of iteration parameters, and M_n is the number of preconditioners used at the nth iteration. Considering the norm $||r^{n+1}||^2 = (r^{n+1}, r^{n+1})$, we see that it attains its minimum in the block Krylov subspace

$$K_{\bar{M}_n} = \mathrm{Span}\{P_0, \ldots, A^{n-1} P_{n-1}\}, \quad \bar{M}_n = \sum_{k=0}^{n-1} M_k, \tag{3}$$

provided that the following orthogonality conditions are fulfilled:

$$\begin{aligned} P_n^T A^T A P_k = D_{n,k} = 0 \text{ for } k \neq n, \\ D_{n,n} = \mathrm{diag}\{\rho_{n,l} = (p_n^l)^T A^T A p_n^l\} \in \mathcal{R}^{M_n, M_n}, \end{aligned} \tag{4}$$

and the corresponding coefficients are determined by the formulas

$$\bar{\alpha}_n = \{\alpha_{n,l}\} = D_{n,n}^{-1} P_n^T A^T r^0. \tag{5}$$

In this case, the following relations hold:

$$\begin{aligned} \| r^{n+1} \|^2 = (r^n, r^n) - (C_n r^0, r^0) = (r^0, r^0) - (C_0 r^0, r^0) - \cdots - (C_n r^0, r^0), \\ C_n = P_n A D_{n,n}^{-1} A^T P_n^T. \end{aligned} \tag{6}$$

A direct verification of the orthogonality relations (4) demonstrates that for them to be valid, it is sufficient that the "direction matrices" P_{n+1} be determined from the recursion relations

$$\begin{aligned} P_{n+1} = Q_{n+1} - \sum_{k=0}^{n} P_k \bar{\beta}_{k,n}, Q_{n+1} \left[q_1^{n+1} \ldots\ q_{m_s}^{n+1} \right] \in \mathcal{R}^{N,m_s}, \\ q_l^{n+1} = \left(B_{n+1}^{(l)} \right)^{-1} r^{n+1}, \quad l = 1, \ldots, M_n, \quad \bar{\beta}_{k,n} = D_{k,k}^{-1} P_k^T A^T A Q_{n+1}. \end{aligned} \tag{7}$$

Here, $\bar{\beta}_{k,n} = (\beta_n^1, \ldots, \beta_n^{M_k})^T \in \mathcal{R}^{M_k}$ are vector coefficients, and $B_{n+1}^{(l)}$ are easily invertible nonsingular preconditioning matrices, chosen in a certain way, whose number M_n can vary from iteration to iteration. Formulas (7) can be realized using a stable modification of the Gram–Schmidt orthogonalization [4].

Theorem 1. *Let the matrices A and $B_n^{(l)}$, $n = 0, 1, \ldots$; $l = 1, \ldots, M_n$, be nonsingular and let the matrices P_n be of full rank. Then the iterative process* (2)–(7) *minimizes $\|r^{n+1}\|$ in*

the block Krylov subspace (3) *of dimension* $\bar{M}_n$. *Moreover, the following semi-orthogonality conditions for the residual vectors are fulfilled:*

$$\left(A(B_n^{(l)})^{-1}r^n, r^k\right) = \begin{cases} 0, & k < n, \\ \sigma_n = \left(A(B_n^{(l)})^{-1}r^n, r^n\right), & k = n, \end{cases}$$

and the coefficients of the recursion relations can be computed by the formula $\alpha_{n,l} = \sigma_n/\rho_{n,l}$.

The above-considered version of MP-SCR generalizes FGMRES in the sense that, at different iterations, the preconditioners are allowed to differ not only in their form but also in their number. A specific feature of such algorithms is the necessity to store, in the course of iterations, $\bar{M}_{n+1}$ vectors, which is too expensive for n large. In order to alleviate this drawback, periodic restarts are repeated every m_r iterations and a limited orthogonalization is carried out, where the summation in relations (7) is only performed for $k = n, n-1, \ldots, \max\{0, n-m_0+1\}$, i.e., for $n > m_0$, only the last m_0 direction matrices are stored.

By a restart we mean an iterative procedure of computing the current residual vector r^{n+1} directly from the original equation, as the initial residual r^0, rather than from the recursion relation (2). The subsequent m_r approximations are again determined from the recursion relations, then a new restart occurs, etc.

This iterative process can be generalized if the period length at the sth restart, which is denoted by $m_r^{(s)}$, is allowed to vary, i.e., $m_r^{(s)} = n_s - n_{s-1}$, where $n_0 = 0$ and n_s is the number of the iteration at which the sth restart occurs. In addition, the number of matrices P_n to be orthogonalized, which is denoted by $m_0^{(n)}$ in what follows, can change not only at different restarts but at each iteration as well. In the resulting reduced R-SCR method the recursion relations (2) are written in the standard way, but at restarts the residual is computed from the original equation. In this case, the first equality in (7) takes the form

$$P_{n+1} = Q_{n+1} - \sum_{k=\bar{m}_0^{(n)}}^{n} P_k \bar{\beta}_{k,n}, \quad n = n_s, n_s+1, \ldots, n_{s+1}-1, \tag{8}$$

where $\bar{m}_0^{(n)} = \min\{0, n-m_0^{(n)}\}$, $s = 1, 2, \ldots$. In the R-SCR method, described by formulas (2) (but only for $n = n_s,\ n_s+1, \ldots,\ n^{s+1}-1;\ s = 1, 2, \ldots$) and relations (8), the dimensions of the Krylov subspaces are obviously reduced, which leads to slowing down the convergence rate of the iterative process. In order to eliminate or, at least, alleviate this drawback, we additionally accelerate the reduced algorithm under consideration by forming linear combinations of vectors at restarts and optimizing them by minimizing the residuals. The suggested accelerated reduced method AR-SCR is a two-level iterative process in Krylov subspaces and can be represented in terms of the auxiliary vectors

$$v^s = u^{n_s} - u^{n_{s-1}}, \quad w^s = Av^s = r^{n_{s-1}} - r^{n_s}, \quad s = 1, \ldots, M_s, \tag{9}$$

where M_s is the number of the restarts performed. We will search for the corrected values of the restart approximations u^{n_s} in the form

$$\begin{aligned} \widehat{u}^{n_s} &= u^{n_s} + c_1 v^1 + \cdots + c_s v^s = u^{n_s} + V_s \bar{c}_s, \\ V_s &= (v^1 \ldots v^s) \in \mathcal{R}^{N,s}, \quad \bar{c}_s = (c_1, \ldots, c_s)^T \in \mathcal{R}^s. \end{aligned} \tag{10}$$

The corresponding residual vectors are written as

$$\widehat{r}^{n_s} = f - A\widehat{u}^{n_s} = r^{n_s} - W_s \bar{c}_s, \quad W_s = (w^1 \ldots w^s) \in \mathcal{R}^{N,s}, \quad s = 1, \ldots, M_s. \tag{11}$$

The unknown coefficient vectors from (11) satisfy the overdetermined SLAE

$$W_s \bar{c}_s = r^{n_s}, \quad s = 1, \ldots, M_s, \tag{12}$$

which characterizes, in a certain sense, the smallness of the residual $\widehat{r}^{n_s}$ and has, in general, no classical solution. In order to find $\bar{c}_s$, one can use the generalized inverse matrix W_s^+ or apply the left Gaussian transformation to the equation. This results in the least squares generalized solution [5],

$$B_s\bar{c}_s = W_s^T r^{n_s}, \quad B_s = W_s^T W_s \in \mathcal{R}^{s,s}, \tag{13}$$

which provides for minimization of the residual

$$\widehat{r}^{n_s} = \min_{c_1,\ldots,c_s}\{f - A\widehat{u}^{n_s}\} \tag{14}$$

in the Krylov subspace $\mathcal{K}_{\bar{M}_s}(r^0, A)$, whose dimension equals $\bar{M}_s = M_s + \min\{m_0^{(n)}\}$. Note that the system of equations (13) with symmetric positive semidefinite matrix B_s always is compatible, but it is nonsingular only whenever the rectangular matrix W_s has full rank s. The numerical solutions of systems (12) and (13) obtained, for example, by using the singular value decomposition (SVD) [4] are identical, provided that the exact arithmetic is used. However, if round-off errors are taken into consideration one should bear in mind that the condition number of SLAE (13) is an order of magnitude larger than that of (12).

Remark 1. The norms of the residual vectors $\widehat{r}^{n_s}$ attain their minima in multi-preconditioned block Krylov subspaces with rather complicated structures, whose dimensions are formally determined by the number of parameters.

Now we comment on potential generalizations or, on the contrary, on special cases of the approach suggested. First, instead of using the MP-SCR methods, intermediate iterations in between restarts can be based on the Semi-Conjugate Residual algorithm [2], as well as on the full orthogonalization methods (FOM) and GMRES [1], which rely on Arnoldi's orthogonalization and are asymptotically equivalent in the order of the convergence rate of iterations. Second, in the recursion relations (8), one can take a single term with the scalar multiplier $\beta_n = \beta_{n,n}$ determined in (7) rather than a sum. This situation can be described as a special case of formula (8), in which one sets $\bar{m}_0^{(n)} = n$:

$$\begin{gathered}
s = 0, 1\ldots, \quad r^{n_s} = f - Au^{n_s}, \quad P_l^{n_s} = \left(B_{n_s}^{(l)}\right)^{-1} r^{n_s}, \\
n = n_s, n_s+1, \ldots, n_{s+1}-1 : u^{n+1} = u^n - P_n\bar{\alpha}_n, \\
P_{n+1} = Q_{n+1} - P_n\bar{\beta}_n, \quad r^{n+1} = r^n - Ap_n\bar{\alpha}_n, \\
P_n = [p_1^n \ldots p_{M_n}^n] \in \mathcal{R}^{N,M_n}, \quad \bar{\alpha}_n = (\alpha_n^1, \ldots, \alpha_n^{M_n})^T.
\end{gathered} \tag{15}$$

Formally, in between restarts, we have the multi-preconditioned conjugate residual method (MP-CR) with short recursions. For unsymmetric SLAEs, this method only ensures local residual minimization in one-dimensional Krylov subspaces. However, it can be regarded as a method for constructing certain auxiliary vectors, defined similarly to (9), but at intermediate iterations in between restarts only:

$$\widetilde{v}^n = u^n - u^{n-1}, \quad \widetilde{w}^n = A\widetilde{v}^n, \quad n = n_s+1, n_s+2, \ldots, n_{s+1} = n_s + m_r^{(s)}. \tag{16}$$

Given the latter vectors and using LSM, one can minimize the residuals at restarts. To this end, for $s = 0, 1, \ldots$ the following formulas, similar to (10)–(14), can be used:

$$\begin{gathered}
\widetilde{u}^{n_{s+1}} = u^{n_{s+1}} + \widetilde{V}_s\widetilde{c}_s, \quad \widetilde{V}_s = (v^{n_s}+1, v^{n_s}+2, \ldots, v^{n_{s+1}}) \in \mathcal{R}^{N,m_r^{(s)}}, \quad \widetilde{c}_s \in \mathcal{R}^{m_r^{(s)}}, \\
\widetilde{r}^{n_{s+1}} = r^{n_{s+1}} - \widetilde{W}_s\widetilde{c}_s, \quad \widetilde{W}_s = AV_s, \\
\widetilde{W}_s\widetilde{c}_s = r^{n_{s+1}}, \quad \widetilde{B}_s\widetilde{c}_s = \widetilde{W}_s^T W_s^T\widetilde{c}_s = W_s^T r^{n_{s+1}}.
\end{gathered} \tag{17}$$

Subsequent intermediate iterations are performed with account for the corrections, i.e., for $n = n_s$, the vector u^{n_s} in the recursion relations (15) should be replaced by the vector $\widetilde{u}_s^n$ from (17). The resulting method, which is referred to as LSM-CR, can be improved in its turn if the approximations obtained at the restart iterations are corrected once again by the same least squares method using formulas (9)–(14). Such an algorithm with two-level minimization of residuals is denoted by ALSM-CR.

In the approach considered, the subscript s can formally be regarded as the number of a certain iterative process in which the operations between restarts, performed in accordance with the CR method by formulas (15), realize a special polynomial preconditioning [6].

Note that at intermediate iterations, one can use the spectral Chebyshev acceleration method rather than the conjugate residual method. This method can be represented in various forms, one of which is similar to (15). We present it for the case of a single constant preconditioner $B_n^{(l)} = B$, to which, at every iteration, a single direction vector p^n corresponds, and the coefficients $\bar{\alpha}_n = \alpha_n$ and $\bar{\beta}_n = \beta_n$ are scalars [2]:

$$\begin{aligned}
s = 0, 1 \ldots : r^{n_s} = f - Au^{n_s}, \quad p^{n_s} &= B^{-1} r^{n_s}, \\
n = n_s, n_s + 1, \ldots, n_{s+1} - 1 : u^{n+1} &= u^n + \alpha_n p^n, \\
p^{n+1} &= B^{-1} r^{n+1} + \beta_n p^n, \quad r^{n+1} = r^n - A\alpha_n p^n, \\
\alpha_0 = \tau, \quad \alpha_n &= \tau_n \tau, \\
\beta_n &= (\tau_n - 1)\alpha_{n-1}\alpha_n, \quad \tau = 2/(\lambda_1 + \lambda_N), \\
\tau_n &= 2/\Big[\lambda_N + \lambda_1 - (\lambda_N - \lambda_1)\cos\frac{(2n-1)\pi}{2m}\Big], \\
n = 1, \ldots, m_1 &= n_{s+1} - n_s,
\end{aligned} \tag{18}$$

In (18), the values λ_1 and λ_N, which are assumed to be real, positive, and available, are the smallest and largest eigenvalues of the matrix $B^{-1}A$. In a number of applications, they can be estimated which accuracy sufficient for practical purposes. The algorithm resulting from combination of the Chebyshev iterations (18) with the correction of restart approximations by the least squares formulas (9)–(14) will be denoted by LSM-CH. This iterative process, by analogy with ALSM-CR, can additionally be accelerated by using linear combinations of restarts in accordance with (16)–(17). This yields the ALSM-CH method, which formally determines consecutive approximations in Krylov subspaces as well.

Theorem 2. *Let the ALSM-CR and ALSM-CH algorithms be defined for the intermediate iterations with a constant period length $m_r^{(s)} = m_r$ from the recursion relations* (15) *and* (18)*, respectively. If, in addition, the restart approximations u^{n_s} are twice corrected in accordance with formulas* (9)–(14) *and* (16)–(17) *of the least squares method, then the resulting iterative processes minimize the residuals in the Krylov subspaces*

$$\mathcal{K}_{M_s}(r^0, A, B^{-1}) = \mathrm{Span}(r^0, AB^{-1}r^0, \ldots, (AB^{-1})_s^M r^0), \quad M_s = sm_r. \tag{19}$$

Remark 2. The least squares acceleration methods under consideration can be interpreted as implicit algorithms in Krylov subspaces. An important feature of such algorithms is a high degree of parallelizability, which results from the fact that the entries of the matrices B_s and $\bar{B}_s$ from (13) and (17) can be computed concurrently on different processing units. Furthermore, the simultaneous computation of the vectors $\bar{c}_s$ and $\check{c}_s$ from the corresponding SLAEs of small order $s \ll N$ by different arithmetic units allows one to avoid extra communications.

3. Experimental investigation of the algorithms

The efficiency of the algorithms suggested will be considered on the Dirichlet problem for the convection-diffusion equation

$$-\frac{\partial^2 u}{\partial x^2}-\frac{\partial^2 u}{\partial y^2}+p\frac{\partial u}{\partial x}+q\frac{\partial u}{\partial y}=f(x,y),\quad (x,y)\in\Omega,\quad u|_\Gamma=g(x,y), \tag{20}$$

in the square computational domain $\Omega=(0,1)^2$ with boundary Γ and convective coefficients p and q, which are considered constant for simplicity. This boundary-value problem is approximated on a square grid with step size $h=1/(L+1)$ and total number of interior nodes $N=L^2$,

$$x_i=ih,\quad y_j=jh,\quad i,j=0,1,\ldots,L+1,$$

by using five-points finite volume monotone approximations of the exponential type, which have the second-order accuracy. The SLAE to be solved is preliminarily scaled in such a way that in our numerical experiments we actually solve the normalized equations resulting from the following transformations with the diagonal matrix $D=\mathrm{diag}\{a_{l,l}\}$:

$$\begin{aligned} &D^{-1/2}AD^{-1/2}D^{1/2}u=D^{-1/2}f,\\ &\bar{A}\bar{u}=\bar{f},\quad \bar{A}=D^{-1/2}AD^{-1/2},\quad \bar{u}=D^{1/2}u,\quad \bar{f}=D^{-1/2}f. \end{aligned} \tag{21}$$

Numerical experiments have been conducted in the standard double precision arithmetic for values of the functions $f(x,y)=0$ and $g(x,y)=1$, which correspond to the exact solution $u(x,y)=1$ of problem (20). Since the convergence rate depends on the initial error $u-u^0$, its influence has been analyzed by comparing results for the different initial guesses $u^0=0$ and $u^0=P_2(x,y)=x^2+y^2$. We have used the stopping criterion $(r^n,r^n)\leqslant\varepsilon^2(f,f)$ with $\varepsilon=10^{-7}$, the square grids with $N=7^2,15^2,31^2,63^2$, and 127^2 nodes, and the period lengths $m=8,16,32,64$, and 128. The computations have been performed on the computer resources of the Siberian Supercomputer Center (SSKC IVMMG SO RAN).

In the tables below, we present the results of numerical solution of problem (20) with the convective coefficients $p=q=0$ and $p=q=4$. Tables 1–4 provide data for the unpreconditioned conjugate residual method (8) (all $B_n^{(l)}$ are the identity matrices I) with two-level correction by the least squares algorithm. This method is denoted by ALSM-CR. Every cell of these tables contains three numbers, which are the number of iterations n, the ultimate residual norm $||r^n||^2$, and the absolute error $\rho=\max\{|1-u^n|\}$.

As is readily seen from the data presented, the numbers of iterations for symmetric and unsymmetric SLAEs are sufficiently close and weakly depend on the initial guess. The resulting values of the residual norm and absolute error are acceptable and show good agreement with the stopping criterion used, which testifies that the algorithm is numerically stable. In dependence of the lengths of restart periods, the convergence rate is different in different experiments. It either has local extrema or decreases as m grows. It is of importance to emphasize that the introduction of the second level of least squares acceleration results in a significant (up to ten times) reduction of the number of iterations. This is observed if one compares our results with those obtained in [5], which are not reproduced here for shortness.

Tables 5–8 provide similar results for the method ALSM-CH, in which the intermediate iterations (performed in between restarts) realize the Chebyshev acceleration in accordance with formulas (18).

For problems with zero and nonzero convection we have used the exact values λ_1 and λ_N of the spectral bounds for the model grid Laplace equation. However, as supplementary computations have demonstrated, the number of iterations only weakly depends on the accuracy of bounding λ_1, provided that λ_N is chosen from the natural estimate of the norms of the

matrices A. For the algorithm in question, as can be seen from comparison with the results in [5], the additional level of the LSM acceleration also provides for a considerable increase of the convergence rate. For instance, in Table 8 for $N = 127^2$ and $m = 8$ we have 216 iteration for the ALSM-CH method, whereas without the second LSM acceleration (see Table 5 in [5]) the number of iterations is $n = 3987$.

Table 1. Results of numerical experiments with ALSM-CR for $p = q = 4$ and $u^0 = x^2 + y^2$.

$N \setminus m$	8	16	32	64	128
7^2	35 $2.1 \cdot 10^{-7}$ $8.0 \cdot 10^{-14}$	31 $6.7 \cdot 10^{-8}$ $3.4 \cdot 10^{-14}$	63 $1.4 \cdot 10^{-9}$ $1.6 \cdot 10^{-17}$	127 $5.2 \cdot 10^{-12}$ $6.7 \cdot 10^{-22}$	255 $4.1 \cdot 10^{-3}$ $5.0 \cdot 10^{-24}$
15^2	50 $1.4 \cdot 10^{-7}$ $5.1 \cdot 10^{-14}$	65 $6.9 \cdot 10^{-7}$ $1.4 \cdot 10^{-13}$	94 $1.5 \cdot 10^{-7}$ $8.2 \cdot 10^{-14}$	127 $2.9 \cdot 10^{-9}$ $9.6 \cdot 10^{-18}$	255 $3.4 \cdot 10^{-12}$ $7.2 \cdot 10^{-23}$
31^2	85 $2.1 \cdot 10^{-7}$ $1.2 \cdot 10^{-13}$	101 $1.7 \cdot 10^{-6}$ $2.8 \cdot 10^{-13}$	135 $8.1 \cdot 10^{-7}$ $2.6 \cdot 10^{-13}$	190 $4.9 \cdot 10^{-8}$ $2.1 \cdot 10^{-15}$	255 $7.28 \cdot 10^{-7}$ $4.2 \cdot 10^{-14}$
63^2	158 $1.6 \cdot 10^{-16}$ $6.1 \cdot 10^{-13}$	176 $2.6 \cdot 10^{-6}$ $5.9 \cdot 10^{-13}$	201 $3.4 \cdot 10^{-6}$ $6.3 \cdot 10^{-13}$	294 $5.8 \cdot 10^{-6}$ $6.4 \cdot 10^{-13}$	382 $1.6 \cdot 10^{-7}$ $2.4 \cdot 10^{-15}$
127^2	302 $1.9 \cdot 10^{-6}$ $6.8 \cdot 10^{-13}$	330 $6.1 \cdot 10^{-6}$ $1.2 \cdot 10^{-12}$	349 $5.6 \cdot 10^{-6}$ $1.2 \cdot 10^{-12}$	386 $1.1 \cdot 10^{-5}$ $1.3 \cdot 10^{-12}$	582 $1.0 \cdot 10^{-5}$ $1.3 \cdot 10^{-12}$

Table 2. Results of numerical experiments with ALSM-CR for $p = q = 0$ and $u^0 = x^2 + y^2$.

$N \setminus m$	8	16	32	64	128
7^2	35 $3.0 \cdot 10^{-7}$ $6.9 \cdot 10^{-14}$	31 $4.2 \cdot 10^{-8}$ $1.7 \cdot 10^{-14}$	63 $5.4 \cdot 10^{-9}$ $6.2 \cdot 10^{-16}$	127 $5.4 \cdot 10^{-9}$ $6.2 \cdot 10^{-16}$	255 $5.4 \cdot 10^{-9}$ $6.2 \cdot 10^{-16}$
15^2	50 $3.9 \cdot 10^{-8}$ $1.9 \cdot 10^{-14}$	65 $1.1 \cdot 10^{-6}$ $1.5 \cdot 10^{-13}$	94 $3.1 \cdot 10^{-7}$ $1.2 \cdot 10^{-13}$	127 $8.8 \cdot 10^{-8}$ $1.4 \cdot 10^{-13}$	255 $8.8 \cdot 10^{-8}$ $1.4 \cdot 10^{-13}$
31^2	76 $7.2 \cdot 10^{-7}$ $2.5 \cdot 10^{-13}$	98 $3.0 \cdot 10^{-6}$ $3.0 \cdot 10^{-13}$	129 $2.8 \cdot 10^{-6}$ $2.8 \cdot 10^{-13}$	83 $1.2 \cdot 10^{-6}$ $2.7 \cdot 10^{-13}$	80 $2.4 \cdot 10^{-7}$ $3.3 \cdot 10^{-13}$
63^2	141 $4.9 \cdot 10^{-7}$ $3.3 \cdot 10^{-13}$	151 $5.5 \cdot 10^{-7}$ $4.6 \cdot 10^{-13}$	190 $2.7 \cdot 10^{-6}$ $6.5 \cdot 10^{-13}$	259 $6.5 \cdot 10^{-6}$ $6.3 \cdot 10^{-13}$	160 $2.6 \cdot 10^{-6}$ $5.5 \cdot 10^{-13}$
127^2	267 $2.2 \cdot 10^{-6}$ $9.0 \cdot 10^{-13}$	286 $2.2 \cdot 10^{-6}$ $9.9 \cdot 10^{-13}$	301 $4.0 \cdot 10^{-6}$ $1.3 \cdot 10^{-12}$	378 $2.1 \cdot 10^{-5}$ $1.2 \cdot 10^{-12}$	495 $2.3 \cdot 10^{-5}$ $1.2 \cdot 10^{-12}$

Table 3. Results of numerical experiments with ALSM-CR for $p = q = 0$ and $u^0 = 0$.

$N \setminus m$	8	16	32	64	128
7^2	15 $3.3 \cdot 10^{-16}$ $1.4 \cdot 10^{-30}$	10 $3.3 \cdot 10^{-16}$ $2.0 \cdot 10^{-33}$	10 $3.3 \cdot 10^{-16}$ $2.0 \cdot 10^{-33}$	10 $3.3 \cdot 10^{-16}$ $2.0 \cdot 10^{-33}$	10 $3.3 \cdot 10^{-16}$ $2.0 \cdot 10^{-33}$
15^2	43 $5.9 \cdot 10^{-8}$ $4.4 \cdot 10^{-14}$	48 $3.0 \cdot 10^{-7}$ $1.1 \cdot 10^{-13}$	28 $2.4 \cdot 10^{-8}$ $5.1 \cdot 10^{-14}$	28 $2.4 \cdot 10^{-8}$ $5.1 \cdot 10^{-14}$	28 $2.4 \cdot 10^{-8}$ $5.1 \cdot 10^{-14}$
31^2	64 $4.6 \cdot 10^{-7}$ $2.8 \cdot 10^{-13}$	93 $1.1 \cdot 10^{-6}$ $3.0 \cdot 10^{-13}$	95 $6.5 \cdot 10^{-7}$ $2.5 \cdot 10^{-13}$	58 $6.6 \cdot 10^{-8}$ $1.2 \cdot 10^{-13}$	58 $6.6 \cdot 10^{-8}$ $1.2 \cdot 10^{-13}$
63^2	102 $9.9 \cdot 10^{-7}$ $6.1 \cdot 10^{-13}$	123 $1.7 \cdot 10^{-6}$ $5.6 \cdot 10^{-13}$	200 $7.4 \cdot 10^{-6}$ $5.9 \cdot 10^{-13}$	190 $1.1 \cdot 10^{-6}$ $3.7 \cdot 10^{-13}$	110 $4.5 \cdot 10^{-7}$ $3.9 \cdot 10^{-13}$
127^2	190 $1.1 \cdot 10^{-6}$ $8.9 \cdot 10^{-13}$	208 $1.4 \cdot 10^{-6}$ $1.2 \cdot 10^{-12}$	231 $4.8 \cdot 10^{-6}$ $1.2 \cdot 10^{-12}$	382 $4.9 \cdot 10^{-6}$ $1.2 \cdot 10^{-12}$	377 $1.5 \cdot 10^{-5}$ $1.3 \cdot 10^{-12}$

Table 4. Results of numerical experiments with ALSM-CR for $p = q = 4$ and $u^0 = 0$.

$N \setminus m$	8	16	32	64	128
7^2	36 $1.3 \cdot 10^{-8}$ $6.3 \cdot 10^{-15}$	33 $7.2 \cdot 10^{-8}$ $3.9 \cdot 10^{-14}$	63 $2.3 \cdot 10^{-9}$ $4.5 \cdot 10^{-17}$	127 $3.3 \cdot 10^{-11}$ $2.1 \cdot 10^{-20}$	130 $9.0 \cdot 10^{-8}$ $4.9 \cdot 10^{-14}$
15^2	51 $2.3 \cdot 10^{-7}$ $1.2 \cdot 10^{-13}$	76 $7.3 \cdot 10^{-8}$ $1.7 \cdot 10^{-14}$	81 $7.8 \cdot 10^{-7}$ $1.6 \cdot 10^{-13}$	127 $7.6 \cdot 10^{-9}$ $3.9 \cdot 10^{-16}$	255 $7.3 \cdot 10^{-11}$ $2.7 \cdot 10^{-20}$
31^2	85 $2.4 \cdot 10^{-7}$ $1.7 \cdot 10^{-13}$	104 $1.5 \cdot 10^{-6}$ $2.9 \cdot 10^{-13}$	128 $2.4 \cdot 10^{-6}$ $3.1 \cdot 10^{-13}$	190 $4.2 \cdot 10^{-7}$ $1.2 \cdot 10^{-14}$	255 $2.2 \cdot 10^{-8}$ $5.0 \cdot 10^{-16}$
63^2	161 $1.9 \cdot 10^{-6}$ $5.7 \cdot 10^{-13}$	180 $3.3 \cdot 10^{-6}$ $5.0 \cdot 10^{-13}$	205 $3.6 \cdot 10^{-6}$ $6.3 \cdot 10^{-13}$	312 $5.8 \cdot 10^{-6}$ $6.2 \cdot 10^{-13}$	382 $5.6 \cdot 10^{-7}$ $6.3 \cdot 10^{-15}$
127^2	302 $2.4 \cdot 10^{-6}$ $8.8 \cdot 10^{-13}$	331 $2.3 \cdot 10^{-6}$ $6.3 \cdot 10^{-13}$	350 $7.8 \cdot 10^{-6}$ $1.2 \cdot 10^{-12}$	403 $8.7 \cdot 10^{-6}$ $1.2 \cdot 10^{-12}$	509 $4.2 \cdot 10^{-6}$ $4.9 \cdot 10^{-13}$

Table 5. Results of numerical experiments with ALSM-CH for $p = q = 0$ and $u^0 = x^2 + y^2$.

$N \setminus m$	8	16	32	64	128
7^2	37 $2.2 \cdot 10^{-7}$ $7.9 \cdot 10^{-14}$	29 $4.5 \cdot 10^{-8}$ $4.2 \cdot 10^{-14}$	32 $6.7 \cdot 10^{-15}$ $1.3 \cdot 10^{-27}$	44 $1.0 \cdot 10^{-7}$ $3.1 \cdot 10^{-14}$	44 $1.2 \cdot 10^{-7}$ $4.3 \cdot 10^{-14}$
15^2	56 $7.3 \cdot 10^{-8}$ $1.9 \cdot 10^{-14}$	72 $4.5 \cdot 10^{-7}$ $9.6 \cdot 10^{-14}$	64 $3.5 \cdot 10^{-8}$ $1.1 \cdot 10^{-14}$	64 $4.1 \cdot 10^{-8}$ $4.0 \cdot 10^{-14}$	88 $1.6 \cdot 10^{-7}$ $1.1 \cdot 10^{-13}$
31^2	88 $8.1 \cdot 10^{-8}$ $3.4 \cdot 10^{-14}$	112 $1.3 \cdot 10^{-6}$ $2.7 \cdot 10^{-13}$	144 $8.0 \cdot 10^{-7}$ $2.5 \cdot 10^{-13}$	127 $9.8 \cdot 10^{-8}$ $2.7 \cdot 10^{-13}$	160 $1.1 \cdot 10^{-7}$ $1.6 \cdot 10^{-13}$
63^2	160 $5.9 \cdot 10^{-7}$ $3.3 \cdot 10^{-13}$	160 $7.5 \cdot 10^{-7}$ $4.6 \cdot 10^{-13}$	224 $1.1 \cdot 10^{-6}$ $1.1 \cdot 10^{-13}$	288 $2.5 \cdot 10^{-6}$ $4.4 \cdot 10^{-13}$	249 $5.5 \cdot 10^{-7}$ $5.1 \cdot 10^{-13}$
127^2	304 $2.2 \cdot 10^{-6}$ $9.0 \cdot 10^{-13}$	304 $2.3 \cdot 10^{-6}$ $9.9 \cdot 10^{-13}$	320 $6.9 \cdot 10^{-7}$ $1.0 \cdot 10^{-13}$	384 $2.9 \cdot 10^{-6}$ $3.7 \cdot 10^{-13}$	512 $1.2 \cdot 10^{-5}$ $7.5 \cdot 10^{-13}$

Table 6. Results of numerical experiments with ALSM-CH for $p = q = 4$ and $u^0 = x^2 + y^2$.

$N \setminus m$	8	16	32	64	128
7^2	38 $9.0 \cdot 10^{-8}$ $7.0 \cdot 10^{-14}$	32 $1.9 \cdot 10^{-9}$ $2.2 \cdot 10^{-17}$	32 $3.1 \cdot 10^{-12}$ $3.3 \cdot 10^{-22}$	48 $2.8 \cdot 10^{-8}$ $5.0 \cdot 10^{-15}$	48 $5.2 \cdot 10^{-8}$ $8.4 \cdot 10^{-15}$
15^2	56 $1.4 \cdot 10^{-7}$ $5.1 \cdot 10^{-14}$	79 $1.2 \cdot 10^{-7}$ $2.5 \cdot 10^{-14}$	64 $1.0 \cdot 10^{-8}$ $2.6 \cdot 10^{-16}$	80 $8.8 \cdot 10^{-8}$ $7.2 \cdot 10^{-14}$	96 $6.0 \cdot 10^{-8}$ $2.7 \cdot 10^{-14}$
31^2	96 $2.1 \cdot 10^{-7}$ $1.2 \cdot 10^{-13}$	112 $1.7 \cdot 10^{-7}$ $3.3 \cdot 10^{-14}$	158 $3.1 \cdot 10^{-7}$ $2.9 \cdot 10^{-13}$	128 $7.9 \cdot 10^{-9}$ $8.6 \cdot 10^{-17}$	160 $3.1 \cdot 10^{-7}$ $1.7 \cdot 10^{-13}$
63^2	184 $5.0 \cdot 10^{-7}$ $1.4 \cdot 10^{-13}$	192 $3.2 \cdot 10^{-7}$ $6.9 \cdot 10^{-14}$	224 $8.5 \cdot 10^{-8}$ $6.2 \cdot 10^{-15}$	256 $2.2 \cdot 10^{-6}$ $6.4 \cdot 10^{-13}$	256 $2.0 \cdot 10^{-8}$ $5.2 \cdot 10^{-17}$
127^2	344 $1.9 \cdot 10^{-6}$ $6.8 \cdot 10^{-13}$	352 $1.8 \cdot 10^{-6}$ $4.6 \cdot 10^{-13}$	384 $5.1 \cdot 10^{-7}$ $5.1 \cdot 10^{-14}$	384 $6.4 \cdot 10^{-6}$ $1.1 \cdot 10^{-12}$	512 $2.1 \cdot 10^{-6}$ $9.8 \cdot 10^{-14}$

Table 7. Results of numerical experiments with ALSM-CH for $p = q = 4$ and $u^0 = 0$.

$N \setminus m$	8	16	32	64	128
7^2	40 $1.3 \cdot 10^{-8}$ $6.3 \cdot 10^{-15}$	32 $8.7 \cdot 10^{-10}$ $9.8 \cdot 10^{-18}$	32 $1.5 \cdot 10^{-12}$ $7.2 \cdot 10^{-23}$	48 $3.5 \cdot 10^{-8}$ $4.8 \cdot 10^{-15}$	48 $4.7 \cdot 10^{-8}$ $7.9 \cdot 10^{-15}$
15^2	57 $2.4 \cdot 10^{-7}$ $1.4 \cdot 10^{-13}$	80 $7.2 \cdot 10^{-8}$ $1.6 \cdot 10^{-14}$	64 $1.5 \cdot 10^{-8}$ $8.3 \cdot 10^{-16}$	80 $7.0 \cdot 10^{-8}$ $5.8 \cdot 10^{-14}$	96 $5.9 \cdot 10^{-8}$ $2.7 \cdot 10^{-14}$
31^2	96 $2.4 \cdot 10^{-7}$ $1.7 \cdot 10^{-13}$	112 $2.6 \cdot 10^{-7}$ $5.9 \cdot 10^{-14}$	160 $5.9 \cdot 10^{-7}$ $1.5 \cdot 10^{-13}$	128 $2.2 \cdot 10^{-8}$ $7.1 \cdot 10^{-16}$	161 $3.6 \cdot 10^{-7}$ $1.9 \cdot 10^{-13}$
63^2	184 $7.0 \cdot 10^{-7}$ $1.9 \cdot 10^{-13}$	192 $3.1 \cdot 10^{-7}$ $1.0 \cdot 10^{-13}$	224 $4.3 \cdot 10^{-7}$ $2.5 \cdot 10^{-14}$	256 $2.5 \cdot 10^{-6}$ $3.9 \cdot 10^{-13}$	256 $5.0 \cdot 10^{-8}$ $5.8 \cdot 10^{-16}$
127^2	344 $2.4 \cdot 10^{-6}$ $8.8 \cdot 10^{-13}$	352 $2.3 \cdot 10^{-6}$ $6.3 \cdot 10^{-13}$	384 $3.5 \cdot 10^{-6}$ $1.1 \cdot 10^{-12}$	448 $5.4 \cdot 10^{-6}$ $6.0 \cdot 10^{-13}$	512 $7.1 \cdot 10^{-7}$ $1.4 \cdot 10^{-14}$

Table 8. Results of numerical experiments with ALSM-CH for $p = q = 0$ and $u^0 = 0$.

$N \setminus m$	8	16	32	64	128
7^2	24 $2.4 \cdot 10^{-8}$ $2.8 \cdot 10^{-14}$	16 $8.9 \cdot 10^{-15}$ $3.7 \cdot 10^{-27}$	32 $9.0 \cdot 10^{-15}$ $6.1 \cdot 10^{-27}$	44 $3.2 \cdot 10^{-8}$ $6.1 \cdot 10^{-15}$	44 $1.2 \cdot 10^{-8}$ $4.1 \cdot 10^{-15}$
15^2	48 $1.0 \cdot 10^{-7}$ $4.4 \cdot 10^{-14}$	61 $2.1 \cdot 10^{-7}$ $1.7 \cdot 10^{-13}$	32 $8.3 \cdot 10^{-9}$ $7.8 \cdot 10^{-13}$	64 $2.0 \cdot 10^{-15}$ $4.3 \cdot 10^{-28}$	88 $8.4 \cdot 10^{-8}$ $1.5 \cdot 10^{-13}$
31^2	80 $1.7 \cdot 10^{-7}$ $1.5 \cdot 10^{-13}$	112 $4.5 \cdot 10^{-7}$ $1.2 \cdot 10^{-13}$	112 $2.0 \cdot 10^{-7}$ $2.2 \cdot 10^{-13}$	89 $3.2 \cdot 10^{-8}$ $4.2 \cdot 10^{-14}$	160 $6.1 \cdot 10^{-8}$ $1.6 \cdot 10^{-13}$
63^2	120 $1.4 \cdot 10^{-7}$ $5.1 \cdot 10^{-14}$	144 $6.4 \cdot 10^{-7}$ $3.6 \cdot 10^{-13}$	224 $1.6 \cdot 10^{-7}$ $2.5 \cdot 10^{-14}$	192 $1.4 \cdot 10^{-6}$ $3.7 \cdot 10^{-13}$	224 $1.8 \cdot 10^{-7}$ $1.0 \cdot 10^{-13}$
127^2	216 $1.2 \cdot 10^{-6}$ $9.0 \cdot 10^{-13}$	224 $3.9 \cdot 10^{-7}$ $1.4 \cdot 10^{-13}$	256 $4.6 \cdot 10^{-7}$ $5.2 \cdot 10^{-14}$	448 $1.9 \cdot 10^{-6}$ $2.0 \cdot 10^{-13}$	384 $1.0 \cdot 10^{-6}$ $9.3 \cdot 10^{-14}$

4. Conclusion

As is seen from the presented preliminary numerical results, the suggested two-level minimization of residuals is a promising approach to accelerating the convergence of restarted iterative methods when solving SLAEs with unsymmetric matrices. In addition to reducing the number of iterations, this allows one to save computer resources considerably. As concerns efficiency, the most promising approach is combination of the spectral Chebyshev acceleration algorithm with two-level corrections by the least squares method in Krylov subspaces. Also

there is a huge resource for increasing efficiency in concurrent realization of the algorithms considered based on using hybrid programming for effecting scalable parallelization on an MCS with distributed and hierarchical shared memory. In the present paper, no preconditioning has been used, which provides yet another resource for accelerating solution of large algebraic systems arising in practice. Of course, the above conclusions should be considered preliminary, and further theoretical investigation of the approaches considered is needed.

This work was supported by the Russian Science Foundation (project No. 14-11-00485P) and Russian Foundation for Fundamental Research (project No. 16-29-1522/17 ofi-m.)

Translated by L. Yu. Kolotilina.

REFERENCES

1. Y. Saad, *Iterative Methods for Sparse Linear System*, PWS Publ. (2002).
2. V. P. Il'in, *Methods and Technologies of Finite Element Methods* [in Russian], IVMMG Publ., Novosibirsk (2007).
3. V. P. Il'in, "Problems of parallel solution of large systems of linear algebraic equations," *Zap. Nauchn. Semin. POMI*, **439**, 112–127 (2015).
4. C. L. Lawson and R. Z. Hanson, *Solving Least Squares Problems* [Russian translation], Nauka, Moscow (1986).
5. V. P. Il'in, "Least squares methods in Krylov subspaces," *Zap. Nauchn. Semin. POMI*, **453**, 131–147 (2016).
6. D. P. O'Leary, "Yet another polynomial preconditioner for the conjugate gradient algorithm," *Linear Algebra Appl.*, **154-156**, 377–388 (1991).