DOI 10.1007/s10958-021-05365-8
Journal of Mathematical Sciences, Vol. 255, No. 3, June, 2021

CONJUGATE DIRECTION METHODS FOR MULTIPLE SOLUTION OF SLAES

Y. L. Gurieva* **and V. P. Il'in†** UDC 519.6

Conjugate gradient and conjugate residual methods for multiple solution of systems of linear algebraic equations (SLAEs) with the same matrices but different successively determined right-hand sides are considered. In order to speed up the iterative solution of the second and subsequent SLAEs, deflation algorithms are applied. These algorithms use the direction vectors obtained in the course of solving the first system as the basis ones. Results of numerical experiments for model examples, illustrating the efficiency of the approaches under consideration, are provided. Bibliography: 27 *titles.*

1. Introduction

The problem of solving large systems of linear algebraic equations (SLAEs) with the same coefficient matrix but different right-hand sides,

$$Au^{(l)}u^{(l)} = f^{(l)}, \quad f^{(l)u^{(l)}} \in \mathcal{R}^N, \quad A \in \mathcal{R}^{N,N}, \quad l = 1, \ldots, L, \tag{1}$$

in the case where the vectors $f^{(l)}$ are not available simultaneously but are determined consecutively, arises in many topical applications. For example, such situations occur in solving nonstationary and/or nonlinear initial-boundary-value multidimensional problems with complex geometry of computational domains approximated by implicit methods of finite differences, finite elements, finite volumes, or discontinuous Galerkin algorithms of various orders of accuracy [1], and a SLAE must be solved at every time step or at different nonlinear iterations. Other cases arise when solving resource-intensive large-block algebraic systems using two-level iterative processes, in which at every outer iteration one must compute an approximate solution of an auxiliary SLAE.

In such problems, it is natural to expect that in the course of iterative solution of the first SLAE, one can generate some useful information on properties of the matrix A, which can be used to speed up solution of the subsequent systems. A similar approach, called deflation, was proposed for the conjugate gradient method by Nicolaides [2] and Dostal [3], and later investigated by many authors, see [4–18] and the references therein.

Naturally, in this case, iterative processes in Krylov subspaces are used. This makes it possible to store the direction vectors or an approximate solution of the spectral problem for the original matrix, obtained when solving the first SLAE. Then, for the subsequent systems, initial approximations can be computed in a special way, and/or a preconditioner can be constructed, or projection and other approaches can be applied.

In the approaches under consideration, from a practical point of view, it is not so much an increase in the convergence rate of the resulting iterative procedures that matters but the ultimate performance of the algorithms implemented on modern multiprocessor computer systems with distributed and hierarchical shared memory. Here, two issues are of importance, namely, the scalability of parallelization on a supercomputer of heterogeneous architecture and

*Institute of Computational Mathematics and Mathematical Geophysics SB RAS, Novosibirsk, Russia, e-mail: yana@apasrv.sscc.ru.

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

Translated from *Zapiski Nauchnykh Seminarov POMI*, Vol. 496, 2020, pp. 26–42. Original article submitted October 23, 2020.

the efficiency of working with big data when using compressed sparse matrix formats, see [19] and the references therein.

The paper is organized as follows. In Sec. 2, we consider some features of preconditioned conjugate direction iterative methods in Krylov subspaces, including those with approximate iterative inversion of the preconditioning matrix. For simplicity, we consider linear algebraic systems with symmetric positive definite (spd) matrices. Section 3 discusses various algorithms of deflation and projection types, as well as their application to multiple solution of SLAEs with the same matrices. The final section presents results of numerical experiments for typical model problems, which demonstrate the efficiency of the algorithms in question, and also discusses some issues of further investigation of the problem under consideration.

2. Preconditioned methods in Krylov subspaces

First we consider iterative algorithms for solving a SLAE resulting from two-sided preconditioning of an original system of the form (1),

$$\bar{A}\bar{u} = \bar{f}, \quad \bar{A} = L_B^{-1} A U_B^{-1}, \quad \bar{u} = U_B u, \quad \bar{f} = L_B^{-1} f. \tag{2}$$

Here, L_B and U_B are the factors in a decomposition of a nonsingular preconditioning matrix

$$B = L_B U_B, \quad B^{-1} = U_B^{-1} L_B^{-1}. \tag{3}$$

In order to solve the preconditioned SLAE $\bar{A}\bar{u} = \bar{f}$ with the symmetric positive definite matrix $\bar{A} = \bar{A}^t$, we consider the following iterative processes:

$$\begin{aligned} \bar{r}^0 &= \bar{f} - \bar{A}\bar{u}^0, \quad \bar{p}^0 = \bar{r}^0, \quad n = 0, 1, \dots : \\ \bar{u}^{n+1} &= \bar{u}^n + \alpha_n \bar{p}^n = \bar{u}^0 + \alpha_0 \bar{p}^n + \dots + \alpha_n \bar{p}^n, \\ \bar{r}^{n+1} &= \bar{f} - \bar{A}\bar{u}^{n+1} = \bar{r}^n - \alpha_n \bar{A}\bar{p}_n = \bar{r}^0 - \alpha_0 \bar{A}\bar{p}^0 - \dots - \alpha_n \bar{A}\bar{p}^n. \end{aligned} \tag{4}$$

Here, $\bar{p}^n$ are some direction vectors; α_n are iteration parameters; $\bar{u}^0 = U_B u^0$ and $\bar{r}^0$ are the preconditioned vectors of the initial guess and residual, and u^0 is an arbitrary vector.

In formulas (4), it is assumed that the vectors p^n satisfy the orthogonality relations

$$\left(\bar{A}^\gamma p^n, p^k\right) = (p^n, p^k)_\gamma = \rho_n^{(\gamma)} \delta_{k,n}, \quad \rho_n^{(\gamma)} = (p^n, p^n)_\gamma, \tag{5}$$

where $\gamma = 1, 2$, and $\delta_{k,n}$ is the Kronecker symbol. Under these assumptions, we have the following relation for the residual:

$$\left(\bar{r}^{n+1}, \bar{r}^{n+1}\right)_{\gamma-2} = \left(\bar{r}^0, \bar{r}^0\right)_{\gamma-2} - \sum_{k=0}^{n} \left[2\alpha_k \left(\bar{r}^0, p^k\right)_{\gamma-1} - \alpha_k^2 \left(p^k, p^k\right)_\gamma\right].$$

From this relation we obtain that if the coefficients α_n are determined via

$$\alpha_n = \sigma_n / \rho_n, \quad \sigma_n = \left(\bar{r}^0, p^n\right)_{\gamma-1}, \tag{6}$$

then the residual functional can be written as

$$\Phi_\gamma\left(\bar{r}^{n+1}\right) = \Phi_\gamma\left(\bar{r}^0\right) - \sum_{k=0}^{n} \sigma_k^2 / \rho_k, \tag{7}$$

where in the values α_k, σ_k, and ρ_k the symbol "γ" is omitted for simplicity. In this case, the functional attains its minimum in the Krylov subspaces

$$\mathcal{K}_{n+1}\left(\bar{r}^0, \bar{A}\right) = \mathrm{Span}\left\{\bar{r}^0, \bar{A}\bar{r}^0, \dots, \bar{A}^n \bar{r}^0\right\}. \tag{8}$$

For the orthogonality conditions (5) to be fulfilled, the direction vectors are determined from the recurrence relation

$$p^0 = \bar{r}^0, \quad p^{n+1} = \bar{r}^{n+1} + \beta_n p_n, \quad \beta_n = -\left(\bar{r}^{n+1}, p^n\right)_\gamma / \rho_n, \tag{9}$$

where the conventional method for determining the initial vector p^0 is generally accepted but not mandatory (strictly speaking, it can be chosen arbitrarily). In this case, the vectors posses the following additional orthogonality properties:

$$\left(r^k, r^n\right)_{\gamma-1} = \|r^n\|_{\gamma-1}\delta_{k,n}, \quad \left(r^n, p^k\right)_{\gamma-1} = 0 \quad \text{for} \quad k < n. \tag{10}$$

Also they satisfy the relations $(\bar{r}^0, \rho^n)_{\gamma-1} = (\bar{r}^n, \bar{r}^n)_{\gamma-1}$, and for the coefficients σ_n, β_n we obtain the new formulas

$$\sigma_n = (\bar{r}^n, \bar{r}^n)_{\gamma-1}, \quad \beta_n = \sigma_{n+1}/\sigma_n. \tag{11}$$

The above conjugate direction (CD) algorithms for solving the SLAE (2) with $\gamma = 1, 2$ are called the conjugate gradient and conjugate residual methods, respectively (CG and CR, see [12, 20, 21] and the references therein). From the above relations for $\gamma = 1, 2$ we obtain the following formulas in terms of the matrices A and B:

for the conjugate gradient method,

$$\begin{gathered} r^0 = f - Au^0, \quad p^0 = B^{-1}r^0, \quad \alpha_n = \sigma_n/\rho_n, \\ u^{n+1} = u^n + \alpha_n p^n, \quad r^{n+1} = r^n - \alpha_n A p^n, \quad p^{n+1} = B^{-1}r^{n+1} + \beta_n p^n, \\ \sigma_n = (r^n, p^n) = \left(B^{-1}r^n, r^n\right), \quad \rho_n = (Ap^n, p^n), \quad \beta_n = \sigma_{n+1}/\sigma_n; \end{gathered} \tag{12}$$

for the conjugate residual method,

$$\begin{gathered} r^0 = f - Au^0, \quad \widehat{r}^0 = \widehat{p}^0 = B^{-1}r^0, \quad \alpha_n = \sigma_n/\rho_n, \\ u^{n+1} = u^n + \alpha_n \widehat{p}^n, \quad \widehat{r}^{n+1} = \widehat{r}^n - \alpha_n B^{-1}A\widehat{p}^n, \quad \widehat{p}^{n+1} = \widehat{r}^{n+1} + \beta_n \widehat{p}_n, \\ \sigma_n = \left(B^{-1}\widehat{r}^n, A\widehat{p}^n\right) = (A\widehat{r}^n, \widehat{r}^n), \quad \rho_n = \left(B^{-1}A\widehat{p}^n, A\widehat{p}^n\right) \quad \beta_n = \sigma_{n+1}/\sigma_n. \end{gathered} \tag{13}$$

Note that at every iteration of each of these methods, one multiplication by each of the matrices A and B^{-1} is performed, and, in (13), $\widehat{r}^n$ is the preconditioned residual vector rather than the "true" one, i.e., in the case of exact computations, we have $\widehat{r}^n = B^{-1}r^n = B^{-1}(f - Au^n)$, and at every iteration the value $(B^{-1}r^n, r^n)$ is minimized. Substituting the vectors $\widehat{r}^n = B^{-1}r^n$ and $\widehat{p}^n = B^{-1}p^n$, we can write relations (13) in terms of the ordinary residuals r^n as follows:

$$\begin{aligned} p^0 &= r^0 = f - Au^0, & & n = 0, 1, \ldots : \\ u^{n+1} &= u^n + \alpha_n B^{-1}p^n, & & \alpha_n = \sigma_n/\rho_n, \\ r^{n+1} &= r^n - \alpha_n AB^{-1}p^n, & & \rho_n = \left(B^{-1}AB^{-1}p^n, AB^{-1}p^n\right), \\ p^{n+1} &= r^{n+1} + \beta_n p_n, & & \beta_n = \sigma_{n+1}/\sigma_n, \quad \sigma_n = \left(B^{-1}AB^{-1}r^n, r^n\right). \end{aligned}$$

However, in this case, an additional multiplication by the matrix B^{-1} must be performed at every iteration. The preconditioned algorithms (12) and (13) are known under the standard abbreviations PCG and PCR, respectively.

In implementing methods (12), (13), an important issue arises if the inversion of the preconditioning matrix B actually reduces to iterative solution of the corresponding auxiliary SLAE, allowing for a finite residual and an error. In this case, the iterative process turns into a two-level one. Although the problem of choosing the stopping criteria for iterations is

rather subtle (see [22, 23]), for simplicity we assume that in the preconditioned CG and CR algorithms the outer iterations are continued until

$$\|r^n\| \le \varepsilon_e \|f\|, \quad \varepsilon_e \ll 1. \tag{14}$$

Note that in the case of exact computations in (12), (13) (including the case of a positive semidefinite matrix A), the number of iterations n_ε sufficient for condition (14) to be fulfilled with $\varepsilon = \varepsilon_e$ is bounded as follows:

$$n(\varepsilon) \le \frac{1}{2}\Big|\ln\frac{\varepsilon}{2}\Big|(\mathrm{cond}_r(B^{-1}A)^{1/2} + 1).$$

Here, $\mathrm{cond}_r(B^{-1}A) = \lambda_{\max}/\lambda_{\mathrm{mnz}}$ is the so-called reduced, or effective condition number, and λ_{mnz} is the minimal nonzero eigenvalue of the matrix $B^{-1}A$.

At inner iterations, we proceed in a similar way, choosing a possibly different accuracy parameter $\varepsilon_i \ll 1$. For example, the approximate initial direction vector $\overline{p}^0$ is determined as the solution of the system $B\overline{p}^0 = r^0$ as follows:

$$\delta^0 = r^0 - Bp^0, \quad \overline{p}^0 = B^{-1}(r^0 - \delta^0), \quad \|\delta^0\| \le \varepsilon_i \|r^0\|. \tag{15}$$

Relations (15) can be interpreted in such a way that the inversion of the preconditioner B is performed in an approximate iterative manner, δ^0 being the corresponding residual.

For the CG method, we will analyze the perturbations of the computed vectors $\overline{u}^n$, $\overline{r}^n$, $\overline{p}^n$, assuming that at the nth step they are computed with some errors, which result, among other things, from approximate inversion of the matrix B,

$$\overline{u}^n = u^n + z^n, \quad \overline{r}^n = r^n + \varphi^n, \quad \overline{p}^n = p^n + \psi^n. \tag{16}$$

At the same time, we assume that the coefficients α_n, β_n are computed exactly and the arithmetic operations are performed in exact arithmetic. From (12), (16) for the approximate solutions and residuals we obtain

$$\begin{aligned} \overline{u}^{n+1} &= \overline{u}^n + \alpha_n \overline{p}^n = u^{n+1} + z^{n+1}, \quad z^{n+1} = z^n + \alpha_n \psi^n, \\ \overline{r}^{n+1} &= \overline{r}^n + \alpha_n A\overline{p}^n = r^{n+1} + \varphi^{n+1}, \quad \varphi^{n+1} = \varphi^n + \alpha_n A\psi^n. \end{aligned} \tag{17}$$

Then, by analogy with (14), we have

$$\begin{aligned} B\overline{q}^{n+1} &= \overline{r}^{n+1}, \quad \delta^{n+1} = \overline{r}^{n+1} - B\widetilde{q}^{n+1}, \\ \widetilde{q}^{n+1} &= B^{-1}(\overline{r}^{n+1} - \delta^{n+1}) = q^{n+1} + B^{-1}(\varphi^{n+1} - \delta^{n+1}) \\ &= q^{n+1} + B^{-1}(\varphi^{n+1} - \delta^{n+1}), \quad \|\delta^{n+1}\| \le \varepsilon_i \|r^{n+1}\|. \end{aligned} \tag{18}$$

Here, we assume that perturbations are small and $\overline{r}^{n+1} \approx r^{n+1}$; $\overline{q}^{n+1}$ and $\widetilde{q}^{n+1}$ are the solutions of the SLAE corresponding to the exact and iterative inversion of the matrix B, respectively. Now the direction vectors are determined by the formulas

$$\overline{p}^{n+1} = \widetilde{q}^{n+1} + \beta_n \overline{p}^n = p^{n+1} + \psi^{n+1}, \quad \psi^{n+1} = B^{-1}(\varphi^{n+1} - \delta^{n+1}) + \beta_n \psi^n. \tag{19}$$

As a result, from (16)–(18) we arrive at the inequalities

$$\begin{aligned} &\|z^{n+1}\| \le \|z^n\| + \alpha_n \|\psi^n\|, \quad \|\psi^{n+1}\| \le \|\psi^n\| + \alpha_n \|A\| \|\psi^n\|, \\ &\|\varphi^{n+1}\| \le \|B^{-1}\| \|\varphi^n\| + (\alpha_n \|B^{-1}A\| + \beta_n)\|\psi^n\| + \varepsilon_i \|B^{-1}\| \|r^{n+1}\|. \end{aligned} \tag{20}$$

Since, in (17), $z^0 = \varphi^0 = 0$ and $\|\psi^0\| \le \varepsilon_i \|r^0\|$, from the recurrence formulas (20) we obtain that all errors are of order $O(\varepsilon_i)$, provided that the values $\|B^{-1}\|$, $\|B^{-1}A\|$, α_n, and β_n are bounded and the number of iterations n is small, which is assumed. Note that although above the same values ε_i have been used for all outer iterations, even in the case of approximate inversion of the matrix B with the same number of inner iterations of any of the conjugate

direction methods, the resulting preconditioners must in general be regarded as variable matrices B_n. In its turn, this leads us to the necessity of switching to the flexible conjugate gradient method (FCG) [12] with long recursions, which makes the algorithm and its analysis essentially more complicated. An alternative approach here is to apply, at the inner iterations, Chebyshev acceleration with the same bounds for the eigenvalues of the matrix B and a fixed degree of the polynomial at all outer iterations.

3. Deflated approaches in Krylov subspaces

Let $V = (v_1 \dots v_m) \in \mathcal{R}^{N,m}$, $m < N$, be a rectangular matrix of full rank m with linearly independent column vectors v_k. An approximate solution of the SLAE $Au = f$ can be sought for as a linear combination of the vectors v_k, i.e.,

$$\begin{aligned} &\widetilde{u} = u^{-1} + c_1 v_1 + \dots + c_m v_m = u^{-1} + Vc, \qquad c \in \mathcal{R}^m, \\ &\widetilde{r} = f - A\widetilde{u} = r^{-1} - AVc = r^{-1} - Wc, \qquad W = AV, \end{aligned} \tag{21}$$

where u^{-1} is an arbitrary vector. The coefficients c_k will be found from the condition of orthogonality of the residual $\widetilde{r}$ either to the vectors v_k, i.e.,

$$V^T\widetilde{r} = 0, \quad \widehat{A}\widehat{c} \equiv V^T AV\widehat{c} = V^T r^{-1}, \quad \widehat{c} = \widehat{A}^{-1}V^T r^{-1}, \tag{22}$$

or to the vectors $w_k = Av_k$, i.e.,

$$W^T\widetilde{r} = 0, \quad \check{A}\check{c} \equiv W^T W\check{c} = W^T r^{-1}, \quad \check{c} = \check{A}^{-1}W^T r^{-1}. \tag{23}$$

In the latter case, $\check{c}$ is the normal solution of the SLAE $Wc = r^0$, i.e., it minimizes the residual $\widetilde{r}$ and has the smallest length $\|c\| = (c, c)^{1/2}$. The approximate solution of the original equation and its residual are given by

$$\begin{aligned} &\widehat{u}^0 = u^{-1} + V\big(V^T AV\big)^{-1}V^T r^{-1} \equiv u^{-1} + \widehat{H}r^{-1}, \\ &\widehat{H} = V\widehat{A}^{-1}V^T, \quad \widehat{r}^0 = (I - A\widehat{H})r^{-1} \end{aligned} \tag{24}$$

if formulas (21) are used and by

$$\begin{aligned} &\check{u}^0 = u^{-1} + V\big(V^T AAV\big)^{-1}V^T Ar^{-1} \equiv u^{-1} + \check{H}r^{-1}, \\ &\check{H} = W\check{A}^{-1}W^T, \quad \check{r}^0 = (I - A\check{H})r^{-1} \end{aligned} \tag{25}$$

in the other case, which actually is the method of moments, see the monographs [24,25], or the least squares method, see [26,27]. The matrices $\widehat{H}$ and $\check{H}$ are low-rank approximations to the inverse matrix A^{-1}, and, in the limit, we have $\widehat{H} = A^{-1}$ or $\check{H} = A^{-1}$, i.e., $\widehat{u}^0 = u$ or $\check{u}^0 = u$. In what follows, we will also bear in mind the geometric interpretation of relations (22) and (23), which mean that $\widetilde{r} \perp \mathcal{V}$ or $\widetilde{r} \perp \mathcal{W}$, respectively, where $\mathcal{V} = \mathrm{Span}\{v_s\}$ and $\mathcal{W} = \mathrm{Span}\{w_s\}$ are m-dimensional subspaces.

Note that in (24) and (25), the vectors $\widehat{u}$ and $\check{u}$ can formally be considered as the initial guess u^0 in the stationary iterative processes with the preconditioning matrices $B_1^{-1} = \widehat{H}$ and $B_2^{-1} = \check{H}$, which are singular in view of the orthogonality conditions (22) and (23). Also observe that the matrices $H_1 = \widehat{H}$ and $H_2 = \check{H}$ are symmetric, whereas the matrices $P_i = I - AH_i,\ i = 1, 2$, are projectors, i.e.,

$$P_i^2 = P, \quad (I - P_i)^2 = I - P_i, \quad i = 1, 2.$$

Moreover, the matrices P_i are singular because of the equalities $V^T P_1 = 0$ and $W^T P_2 = 0$, which mean that the vectors v_s and w_s are eigenvectors of the matrices P_1 and P_2, respectively, corresponding to the zero eigenvalue.

If the vectors v_s or w_s possess some orthogonality properties, then the computation of the entries of the matrices $\widehat{H}$ and $\check{H}$ significantly simplifies. For example, let m of the first direction vectors p^n from the CG method ($\gamma = 1$, formula (12)), which are A-orthogonal, be taken as v_s. Then in (22) and (24) we have

$$\widehat{A} = V^T AV = R_1 = \operatorname{diag}\{\rho_s^{(1)} = (v_s, Av_s)\}. \tag{26}$$

In the other case, where v_s are the direction vectors from the CR method (13), in (23) and (25) we have

$$\check{A} = W^T W = R_2 = \operatorname{diag}\{\rho_s^{(2)} = (Av_s, Av_s)\}. \tag{27}$$

This implies, in particular, that if in (21) one sets $u^{-1} = u^0$, where u^0 is the initial guess from (12) or (13), then the components $\widehat{c}_s$ and $\check{c}_s$ of the vectors $\widehat{c}$ and $\check{c}$ from (22) and (23) coincide with the coefficients α_s, $s = 0, 1, \ldots, m$, from (12) or (13), respectively.

Based on the extension of the Krylov subspaces (8) by the vectors v_s or w_s, we construct preconditioned deflated conjugate direction algorithms (PDCD), in which the direction and residual vectors possess additional orthogonality properties. Following [17], we describe the PDCG conjugate gradient method corresponding to $\gamma = 1$ using the projection of the A-orthogonal complement $\mathcal{V}^{\perp A}$ onto $\mathcal{V}$ along $\mathcal{V}$:

$$Q_1 = I - V\widehat{A}^{-1}V^T A, \quad \widehat{A} = V^T AV. \tag{28}$$

The formulas of this algorithm are as follows:

$$\begin{aligned}
u^0 &= \widehat{u}^0, \quad r^0 = f - Au^0, \\
p^0 &= q^0 = Q_1 B^{-1} r^0; \quad n = 0, 1, \ldots \ : \\
u^{n+1} &= u^n + \alpha_n p^n, \ r^{n+1} = r^n - \alpha_n A p_n, \\
\alpha_n &= (p^n, r^n)/(Ap^n, p^n) = (B^{-1}r^n, r^n)/(Ap^n, p^n), \\
q^{n+1} &= Q_1 B^{-1} r^{n+1}, \quad p^{n+1} = q^{n+1} + \beta_n p^n, \\
\beta_n &= -(q^{n+1}, Ap^n)/(Ap^n, p^n) = (B^{-1}r^{n+1}, r^{n+1})/(B^{-1}r^n, r^n).
\end{aligned} \tag{29}$$

The vectors defined here satisfy the following orthogonality relations:

$$\begin{aligned}
&V^T r^k = 0, \quad V^T A p^k = 0, \quad k = 0, 1, \ldots, n; \\
&(r^n, p^k) = 0, \quad (r^n, B^{-1} r^k) = 0, \quad (p^n, Ap^k) = 0, \quad k < n.
\end{aligned} \tag{30}$$

The resulting error of the approximate solution $z^n = u - u^n$ is orthogonal to the subspace $\mathcal{U} = \operatorname{Span}\left\{\mathcal{V}, \widehat{Q}\mathcal{K}_n(B^{-1}r^0, B^{-1}A\widehat{Q})\right\}$, and its norm $\|u - u^n\|_A = (A^{-1}r^n, r^n)$ attains its minimum in this subspace.

The preconditioned deflated conjugate residual method can be considered in a similar way. In this case, the A^2-orthogonal projection operator is constructed,

$$Q_2 = I - QA_2^{-1}W^T A, \quad A_2 = W^T AW, \tag{31}$$

and the iterative process itself is described by formulas (13), in which u^0 is replaced by $\check{u}^0$, and the matrix B^{-1} is replaced by the product $Q_2 B^{-1}$. In this case, the residual and direction vectors satisfy, instead of conditions (30), the following orthogonality relations:

$$\begin{aligned}
&W^T r^k = 0, \quad W^T A p^k = 0, \quad k = 0, 1, \ldots, n; \\
&(r^n, Ap^k) = 0, \quad (r^n, AB^{-1} r^k) = 0, \quad (Ap^n, Ap^k) = 0, \quad k < n.
\end{aligned} \tag{32}$$

Note that the projectors considered satisfy the relations

$$AQ_i = Q_i^T A = Q_i^T AQ_i = \overline{A}_i, \quad i = 1, 2, \tag{33}$$

and the preconditioned deflated conjugate direction methods generated by them can be interpreted as a result of two-sided preconditioning with the matrices Q_i^T and Q_i in accordance with (33), whereas the matrices $\overline{A}_i$ are singular. In particular, if in (21) we take the eigenvectors of the matrices A_i corresponding to their m minimal eigenvalues $0 < \lambda_1 \le \cdots \le \lambda_m$ as $v_1, \ldots, v_m$, then the effective condition numbers of the matrices $\overline{A}_i$ will be bounded as follows (see [15]):

$$\mathrm{cond}_r(\overline{A}_i) \le \lambda_{\max}(\overline{A}_i)/\lambda_{m+1}(\overline{A}_i).$$

One of the main approaches to choosing the deflation vectors v_s in (21) is to use approximate eigenvectors of the coefficient matrix, which can be computed when performing iterations of the conjugate direction methods.

In the CD methods, for system (1) with $\gamma = 1$ or $\gamma = 2$ from formulas (4), (9) (in which the bar over $\overline{A}$ and $\overline{r^n}$ is dropped) we obtain the following three-term recurrence relation for the residual vectors (here and below, the subscript "γ " of the vectors and coefficients is omitted for simplicity):

$$\begin{aligned} &Ar^1 = \alpha_1^{-1} r^1 - \alpha_2^{-1} r^2, \quad n = 2, 3, \ldots : \\ &Ar^n = -\frac{\beta_{n-1}}{\alpha_{n-1}} r^{n-1} + \Big(\frac{1}{\alpha_n} + \frac{\beta_{n-1}}{\alpha_{n-1}}\Big) r^n - \frac{1}{\alpha_n} r^{n+1}. \end{aligned} \tag{34}$$

Hence, as is readily seen, the normalized residual vectors

$$\widetilde{r}^n = r^n / \|r^n\|_{\gamma-1}, \quad \|r^n\|^2_{\gamma-1} = (r^n, r^n)_{\gamma-1} = \sigma_n,$$

satisfy the relation

$$AR_n = R_n T_n - \nu_n \widetilde{r}^{n+1} e_{n+1}^T, \tag{35}$$

where $e_{n+1} = (0 \ldots 01)$ is a coordinate row vector in $\mathcal{R}^{n+1}$; $R_n = [\widetilde{r}^0 \ldots \widetilde{r}^n] \in \mathcal{R}^{N,n+1}$, and $T_n \in \mathcal{R}^{n+1,n+1}$ is the symmetric tridiagonal matrix

$$T_n = \begin{bmatrix} \mu_0 & -\nu_0 & 0 & \ldots & 0 \\ -\nu_0 & \mu_1 & -\nu_1 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & -\nu_{n-2} & \mu_{n-1} & -\nu_{n-1} \\ 0 & \ldots & 0 & -\nu_{n-1} & \mu_n \end{bmatrix} \tag{36}$$

with its entries

$$\mu_0 = \alpha_0^{-1}, \quad \mu_n = \alpha_n^{-1} + \beta_{n-1}/\alpha_{n-1}, \quad \nu_n = \sqrt{\beta_n}/\alpha_n. \tag{37}$$

By virtue of the orthogonality conditions (10) for the residuals, upon multiplying Eq. (35) by the matrix $R_n^T A^\gamma$ on the left, we obtain the matrix equality

$$T_n = R_n^T A^\gamma A R_n \in \mathcal{R}^{n+1,n+1}.$$

Thus, the matrix T_n is a low-rank approximation of the matrix $A^{\gamma+1}$.

Let λ_k and z_k, $k = 1, \ldots, n+1$, be the eigenvalues and eigenvectors of the matrix T_n, i.e.,

$$T_n z_k = \lambda_k z_k, \quad z_k \in \mathcal{R}^{n+1}.$$

Then the vectors $v_k = R_n z_k \in \mathcal{R}^N$ and values $\nu_k = (Aw_k, w_k)/(w_k, w_k)$ are called the Ritz vectors and values of the matrix A.

The vectors v_k corresponding to $m \le n$ minimal Ritz values of the matrix A are frequently used as the deflation vectors, see (21).

Note that in constructing a tridiagonal matrix T_n of the type (36), which is, in a certain sense, an approximation of the matrix $A^{\gamma+1}$, one can use other sequences of orthogonal vectors, for example, p^n or Ap^n, which also satisfy three-term recurrence relations.

We dwell on the application of this variant of the deflation algorithm to solving SLAEs (1) with two different right-hand sides f^1, f^2. Let the first algebraic system be solved by the CG or CR method in n_1 iterations. For simplicity, we consider the unpreconditioned algorithms, i.e., we set $B = I$ in (12), (13). In this case, the approximate solution obtained is denoted by u^{n_1}; the value $\gamma = 1$ or $\gamma = 2$, corresponding to the CG or CR method, respectively, will be determined below.

Let, in solving the second SLAE (by the CG or CR method), u^{-1} be an arbitrary initial guess. Before starting iteration, we carry out certain corrections, or deflations of the initial guess and initial residual. This is carried out in accordance with the following formulas of the form (21):

$$\begin{aligned} u^0 &= u^{-1} + c_0 p^0 + \cdots + c_{m-1} p^{m-1} = u^{-1} + Pc, \quad c \in \mathcal{R}^m, \\ r^0 &= r^{-1} - c_0 A p^0 - \cdots - c_{m-1} A p^{m-1} = r^{-1} - APc, \quad P \in \mathcal{R}^{N,m}. \end{aligned} \tag{38}$$

Here, $r^{-1} = f^2 - Au^{-1}$, and $p^0, \dots p^{m-1}$, $m \le n_1$, are the direction vectors computed in the course of solving the first SLAE.

In order to find the unknown coefficient vector c, two different orthogonality conditions for the vector r^0 can be used:

$$\begin{aligned} P^T r^0 = 0: \quad & c = c^{(1)} = A_1^{-1} P^T r^{-1}, \quad A_1 = P^T A P, \\ & u^0 = u^{-1} + P c^{(1)}, \quad r^0 = r^{-1} - Q c^{(1)}, \quad Q = AP; \end{aligned} \tag{39}$$

$$\begin{aligned} Q^T r^0 = 0: \quad & c = c^{(2)} = A_2^{-1} Q^T r^{-1}, \quad A_2 = Q^T Q, \\ & u^0 = u^{-1} + P c^{(2)}, \quad r^0 = r^{-1} - Q c^{(2)} = (I - Q A_2^{-1} Q^T) r^{-1}. \end{aligned} \tag{40}$$

Note that in (39) we have $A_1 = \widehat{A}$ and $c^{(1)} = \widehat{c}$ (see (22)) if the direction vectors $p^s = v^s$ are computed by the CG method, and in (40) we have $A_2 = \check{A}$ and $c^{(2)} = \check{c}$ (see (23)) if the direction vectors $p^s = v^s$ are determined by the CR method. Moreover, in both cases, in accordance with (26) and (27), the above matrices are diagonal. The orthogonality conditions for r^0 in (39) and (40) are similar to the properties of the residual vectors r^n they have in the CG and CR methods, respectively. For the conjugate direction methods, there is a certain arbitrariness in the choice of the initial direction vector p^0. Here, we consider three possible different orthogonality conditions, which can readily be ensured prior to iterating the second SLAE:

$$P^T A p^0 = 0: \quad p^0 = r^0 - P\overline{c}^{(1)}, \quad \overline{c}^{(1)} = A_1^{-1} P^T A r^0; \tag{41}$$

$$P^T A A p^0 = 0: \quad p^0 = r^0 - P A_2^{-1} P^T A A r^0; \tag{42}$$

$$Q^T A p^0 = 0: \quad p^0 = r^0 - Q c^{(3)}, \quad c^{(3)} = A_3^{-1} Q^T r^0, \quad A_3 = Q^T A Q. \tag{43}$$

It can be shown that the matrix $A_3 = P^T A^3 P$ defined in (43) is tridiagonal if the direction vectors p^k are determined from the solution of the first SLAE using formulas (13) of the CR method. Moreover, A_3 can be transformed into the symmetric matrix $\overline{A}_3 = \overline{P}^T A^3 \overline{P} = \overline{Q}^T A \overline{Q}$ if the vectors p^k and $q^k = Ap^k$ are normalized as follows: $\overline{q}^k = q^k/\|q^k\|$, $\overline{p}^k = p^k/\|p^k\|$. Indeed, if in (13) with $B = I$ we set $r^n = \widehat{r}^n$ and $q^n = Ap^n$, then we come to the relations

$$\begin{aligned} A r^{n+1} &= A r^n - \alpha_n A q^n, \\ q^{n+1} &= A r^{n+1} = \beta_n q^n, \end{aligned}$$

from which the following three-term recursion is obtained:

$$A q^n = -(\beta_{n-1} q^{n-1} + (1 + \beta_n) q^n - q^{n+1})/\alpha_n. \tag{44}$$

On substituting the expressions $\alpha_n = \sigma_n/\rho_n$ and $\beta_n = \sigma_{n+1}/\sigma_n$ from (13) into (44), introducing the orthonormal vectors $\overline{q}^n = q^n/\|q^n\| = q^n/\rho_n^{1/2}$, and using the relation $\overline{Q}^T Q = I$, we obtain

$$A\overline{q}^n = -\frac{(\rho_{n-1}\rho_n)^{1/2}}{\sigma_{n-1}}\overline{q}^{n-1} + \frac{\rho_n^{1/2}}{\sigma_n}(1+\beta_n)\overline{q}^n - \frac{(\rho_n\rho_{n+1})^{1/2}}{\sigma_n}\overline{q}^{n+1}, \tag{45}$$

$$A\overline{Q} = \overline{Q}T, \quad T = 3\text{-diag}\left\{-\frac{(\rho_{n-1}\rho_n)^{1/2}}{\sigma_{n-1}}\frac{\rho_n^{1/2}}{\sigma_n}(1+\beta_n) - \frac{(\rho_n\rho_{n+1})^{1/2}}{\sigma_n}\right\} = T^T = \overline{A}_3. \tag{46}$$

Note that in (41) and (42), the orthogonality conditions are similar to those satisfied by the direction vectors in formulas (12) and (13) for the CG and CR methods, respectively. Note that in (42) and (43) the conditions for p^0 coincide. However, in the first case, the correction formula for p^0 is simpler. An analysis of formulas (41), (42) demonstrates that the deflation methods of conjugate directions that provide for the orthogonality of the direction vectors $P^T A p^n$ for DCG ($\gamma = 1$) and $P^T A A p^n$ for DCR ($\gamma = 2$) at all iterations can uniformly be written in the following form:

$$\begin{aligned} r^0 &= r^{-1} - APA_\gamma^{-1}P^T A^{\gamma-1}r^{-1}, \\ p^0 &= r^0 - PA_\gamma^{-1}P^T A^\gamma r_0, \quad A_\gamma = P^T A^\gamma P, \\ u^{n+1} &= u^n + \alpha_n p^n, \quad r^{n+1} = r^n - \alpha_n A p_n, \quad \alpha_n = \sigma_n/\rho_n, \\ p^{n+1} &= r^{n+1} + \beta_n p^n - PA_\gamma^{-1}P^T A^\gamma r^{n+1}, \quad \beta_n = \sigma_{n+1}/\sigma_n, \\ \sigma_n &= (A^{\gamma-1}r^n, r^n), \quad \rho_n = (A^\gamma p^n, p^n). \end{aligned} \tag{47}$$

4. Results of numerical experiments

We present and discuss results of numerical experiments on applying some of the deflation approaches considered to the conjugate gradient and conjugate residual methods. The computations were carried out on model grid SLAEs resulting from the standard five-point approximations of the second order on square grids with $N \times N$ cells for the two-dimensional Poisson equation with the Dirichlet boundary conditions in the square domain $\Omega = [0,1]^2$, see [1]. In each of the experiments, computations were carried out for two algebraic systems with the same coefficient matrices and different right-hand sides, corresponding to different boundary conditions in the differential problems. The right-hand side of the first SLAE corresponds to the exact solution of the Poisson equation $u(x,y) = 1$, and the initial guess $u^0 = \{u^0_{i,j} = x_i^2 + y_j^2\}$ for the iterations at the grid nodes ($x_i = ih,\ y_j = jh,\ h = 1/N$) is used. For the second SLAE, the right-hand side was determined from the exact solution of the Dirichlet problem $u(x,y) = x^2 + y^2$, and the zero vector was used as the initial guess. (In formulas (24), (25) these initial vectors are denoted by u^{-1}.) All computations were carried out in the standard double-precision arithmetic on consecutively refined grids with the node numbers $N^2 = 8^2, 16^2, \ldots, 256^2, 512^2$. In both cases, the stopping criterion (14) with $\varepsilon = 10^{-7}$ was used.

In order to compare the efficiency of different algorithms, we only present the numbers of iterations $n(\varepsilon)$. An analysis of the performance of different variants of deflation methods should be carried out with account for the execution times of the algorithms with a scalable parallelization on different configurations of multiprocessor computer systems, which is outside the scope of this paper.

Table 1 presents the results of a series of computations for the CG method in solving the second SLAE using as the deflation vectors v^s from (21) the A-orthogonal direction vectors p^n of the CG method from (12) with $B = I$, obtained in the course of solving the first SLAE (the number of iterations for the first SLAE is denoted by n_1). For the second SLAE, the

DCG approach was implemented either only for the initial guess and initial residual vectors, corrected in accordance with formulas (24), or in accordance with formulas (29), with the complete set of the direction vectors p^n used as the columns of the matrix V from (28) and obtained in the course of solving the first SLAE by the CG method (the corresponding numbers of iterations in solving the second SLAE are denoted by n_2 and n_3). Additionally, the last two rows of this table provide the values of the squared norms of the initial and corrected residuals $\|r^{-1}\|$, $\|r^0\|$ before solving the second SLAE.

N	8	16	32	64	128	256	512
n_1	20	42	83	161	314	610	1185
n_2	10	26	53	96	190	351	745
n_3	1	17	36	73	144	271	538
$\|r^{-1}\|^2$	2.4	4.6	8.8	17.1	33.7	66.8	132.9
$\|r^0\|^2$	$3.5{\cdot}10^{-2}$	1.7	3.8	7.2	13.5	25.5	49.3

Table 1. Numbers of iterations for the CG method with $\varepsilon = 10^{-7}$:
n_1 corresponds to solution of the first SLAE without deflation;
n_2 corresponds to solution of the second SLAE with deflation of the initial guess and initial residual;
n_3 corresponds to solution of the second SLAE with deflation of all the direction vectors.

In conclusion, we can say that the approaches proposed in this work for the deflated conjugate gradient and conjugate residual methods, in which the direction vectors, which are orthogonal in the corresponding metrics, are used as the bases for extending the Krylov spaces, are promising, because of their cost-effectiveness and natural parallelizability, when solving the urgent problem of multiple solution of SLAEs with the same matrices and different consecutively determined right-hand sides. In the literature, many different techniques are available, and their qualitative comparative analysis requires additional investigations, both theoretical and experimental, which are an immediate goal of the authors. Issues of parallelizing deflated algorithms require a special consideration, because introduction of a large number of additional vectors significantly increases the complexity of every iteration.

This work was supported by the Russian Foundation for Basic Research (grant No. 18-01-00295).

Translated by the authors.

REFERENCES

1. V. P. Il'in, *Mathematical Modeling. Part I. Continuous and Discrete Models* [in Russian], SBRAS Publ., Novosibirsk (2017).
2. R. Nicolaides, "Deflation of conjugate gradients with applications to boundary value problems," *SIAM J. Numer. Anal.*, **24**, 355–365 (1987).
3. Z. Dostal, "Conjugate gradient method with preconditioning by projector," *Int. J. Computer Math.*, **23**, 315–323 (1988).
4. L. Mansfield, "On the use of deflation to improve the convergence of conjugate gradient iteration, " *Communs. Appl. Numer. Math.*, **4**, 151–156 (1998).
5. L. Mansfield, "Damped Jacobi preconditioning and coarse grid deflation for conjugate gradient iteration on parallel computers," *SIAM J. Sci. Stat. Comput.*, **12**, 1314–1323 (1991).

6. A. Gaul, M. H. Gutknecht, J. Liesen, and R. A. Nabben, "A framework for deflated and augmented Krylov subspace methods," *SIAM J. Anal. Appl.*, **34**, 495–518 (2013).
7. M. H. Gutknecht, "Deflated and augmented Krylov subspace methods: A framework for the deflated BiCG and related solvers," *SIAM J. Matrix Appl.*, **35**, 1444–1466 (2014).
8. K. Ahuja, E. Sturler, and L. Feng, "Recycling BiCG STAB with an application to parametric model order reduction," *SIAM J. Sci. Comput.*, **37**, 429–446 (2015).
9. M. P. Neuenhofen and C. Greif, "MSTAB: Stabilized inducted dimension reduction for Krylov subspace recycling," *SIAM J. Sci. Comput.*, **40**, No. 2, 554–571 (2018).
10. R. Nabben and C. Vuik, "A comparison of deflation and coarse grid correction applied to porous media flow," *SIAM J. Numer. Anal.*, **42**, 631– 647 (2004).
11. J. M. Tang, R. Nabben, C. Vuik, and Y. A. Erlangga, "Comparison of two-level preconditioners derived from deflation, domain decomposition and multigrid methods," *J. Sci. Comput.*, **39**, 340–370 (2009).
12. Y. Saad, *Iterative Methods for Sparse Linear Systems*, 2nd edn., SIAM (2003).
13. Y. Saad, M. Yeung, J. Erhel, and F. Guyomarch, "A deflated version of conjugate gradient algorithm," *SIAM J. Sci. Comput.*, **24**, 1909–1926 (2000).
14. V. Simoncini and D. B. Szyld, "Recent computational developments in Krylov subspace methods for linear systems," *Numer. Linear Algebra Appl.*, **14**, 1–59 (2007).
15. J. I. Aliaga, D. L. Boley, R. W. Freund, and V. Hernandez, "A Lanczos-type method for multiple starting vectors," *Math. Comp.*, **69**, 1577–1601 (2000).
16. J. Erhel and F. Guyomarch, "An augmented conjugate gradient method for solving consecutive symmetric positive definite linear systems," *SIAM J. Matrix Anal. Appl.*, **21**, 1279–1299 (2000).
17. L. Yu. Kolotilina, "Twofold deflation preconditioning of linear algebraic systems. I. Theory," *Zap. Nauchn. Semin. POMI*, **229**, 95–152 (1995).
18. N. Venkovic, P. Mycek, L. Giraud, and O. Le Maitre, "Comparative study of harmonic and Rayleigh–Ritz procedures with applications for deflated conjugate gradients," Pressed Report Cerfacs (2020), hal-02434043.
19. V. P. Il'in, "Problems of parallel solution of large systems of linear algebraic equations," *Zap. Nauchn. Semin. POMI*, **439**, 112–127 (2015).
20. V. P. Il'in, *Finite Element Methods and Technologies* [in Russian], ICM&MG SB RAS Publ., Novosibirsk (2007).
21. M. A. Olshanskii and E. E. Tyrtyshnikov, *Iterative Methods for Linear Systems. Theory and Applications*, SIAM (2014).
22. M. Arioli, "Generalized Golub–Kahan diagonalization and stopping criteria," *SIAM J. Matrix Anal. Appl.*, **34**, No. 2, 571–592 (2013).
23. N. J. Higham, *Accuracy and Stability of Numerical Algorithms*, SIAM (2002).
24. Yu. V. Vorobyev, *Method of Moments in Applied Mathematics* [in Russian], Fizmatgiz, Moscow (1958).
25. J. Liesen and Z. Strakos, *Krylov Subspace Methods, Principles and Analysis*, Oxford University Press (2013).
26. Y. L. Gurieva and V. P. Il'in, "On coarse grid correction methods in Krylov subspaces," *Zap. Nauchn. Semin. POMI*, **463**, 44–57 (2017).
27. V. P. Il'in, "Two-level least squares methods in Krylov subspaces," *Zap. Nauchn. Semin. POMI*, **463**, 224–239 (2017).