BiCONJUGATE DIRECTION METHODS IN KRYLOV SUBSPACES

VALERY P. IL'IN*

Abstract. This paper analyzes orthogomal and variotional properties of the set of iterative algorithms in Krylov subspaces for solving the systems of linear algebraic equation (SLAEs) with sparse unsymmetric matrices. Biconjugate residual, conjugate residual squared and biconjugate residual stabilized method are proposed. The results of numerical experiments are discussed.

Key words. Biconjugate residual method, conjugate residul squared, biconjugate residual stabilized methods, iterative solver, preconditioning

AMS (MOS) subject classifications. 65F10

1 Introduction

Iterative solution of very large, sparse, non-symmetric SLAEs can be done by means of three main approaches in Krylov subspaces, see [1]–[5]. The first one is based on Gauss transform (left or right) of original SLAE and subsequent using the conjugate direction (conjugate gradient or conjugate residual) algorithms for solving the resulting symmetric algebraic system. The principal disadvantage here consists in growing of the condition numbers of the obtained symmetrized matrices AA^T or A^TA. The second approach implements generalized conjugate direction methods which provide semi-conjugate vectors computed with the help of long recurrent relations. The examples of such algorithms are popular GMRES [6] in various modifications and described in [7], [8], [2]-[5] different versions of Krylov iterative processes. These methods have a significant limitation because for large number of iterations too much volume of memory to save auxiliary vectors is necessary. In this case, the reduced variants of algorithms are used with periodical restarts or/and with truncated orthogonalization, but the dimension of Krylov subspaces and convergence rate are decreased for such simplifications.

The third strategy for solving nonsymmetric SLAEs implies the constructing biorthogonal sets of residual and correction vectors which can be computed by the short two-terms recurrent

*Institute of Computational Mathematics and Mathematical Geophysics, SBRAS, 6 Lavrentiev ave, Novosibirsk, 630090, Russia. The work was supported by RFBR grant N 08-01-00526a

relations. The biconjugate gradient method (BCG) was developed in [9], [10] and its transpose free versions, conjugate gradient squared (CGS) and biconjugate gradient stabilized method (BiCGSTAB), were proposed in [11], [12]. The stability and convergence behaviour of these algorithms were studied later experimentally by many authors, see [13]–[15] for example.

The aim of the present paper is to extend the set of iterative processes which are based on the biorthogonalization procedure. In the section 2 we describe the orthogonal and variational properties of the biconjugate direction methods which include the BCG algorithm and its analog, the biconjugate residual method (BCR). The section 3 includes the unified description of the conjugate residual squared (CRS) method and its prototype, the CGS algorithm. In the section 4, we describe, in a similar manner, the biconjugate residual stabilized algorithm (BiCRSTAB), together with BiCGTAB.

The last section is devoted to discussions of the results of numerical experiments for the considered iterative processes in application for solving model 3D diffusion-convection PDEs with different coefficients, which were considered in [16], [17] previously. The preconditioned Krylov methods are tested at the different grids with parametrized restart procedure.

2 Biconjugate direction methods

We consider the solution of the system

$$Au = f, \quad u, f \in R^N, \quad A \in R^{N,N}, \tag{2.1}$$

where the matrix A is supposed to be positive definite, i.e.

$$\begin{aligned} &(Au, u) \geq \delta ||u||^2, \quad \delta > 0, \\ &(u, v) = (v, u) = \sum_{i=1}^{N} u_i v_i, \quad ||u||^2 = (u, u). \end{aligned} \tag{2.2}$$

Let us consider the following iterative process for solving non-linear SLAE (2.1):

$$\begin{aligned} u^{n+1} &= u^n + \alpha_n p^n, \\ r^{n+1} &= r^n - \alpha_n A p^n, \\ \tilde{r}^{n+1} &= \tilde{r}^n - \tilde{\alpha}_n A^T \tilde{p}^n, \\ p^{n+1} &= r^{n+1} + \beta_n p^n, \\ \tilde{p}^{n+1} &= \tilde{r}^{n+1} + \tilde{\beta}_n \tilde{p}^n. \end{aligned} \tag{2.3}$$

Here $\alpha_n, \beta_n, \tilde{\alpha}_n, \tilde{\beta}_n$ are some real coefficients, u^n is the n-th iterative approximation of the sought solution u, r^n is corresponding residual vector and p^n is called the correction vector. We shall call also $\tilde{r}^n, \tilde{p}^n$ as dual residual and dual correction vectors. From the formulas

$$\begin{aligned} r^n &= r^0 - \alpha_0 A p^0 - \ldots - \alpha_{n-1} A p^{n-1}, \\ \tilde{r}^n &= \tilde{r}^0 - \tilde{\alpha}_0 A^T \tilde{p}^0 - \ldots - \tilde{\alpha}_{n-1} A^T \tilde{p}^{n-1}, \end{aligned} \tag{2.4}$$

we have that $r^n, \tilde{r}^n$ belong to Krylov subspaces

$$\begin{aligned} \mathcal{K}_n(A, r^0) &= span\{r^0, Ar^0, ..., A^{n-1}r^0\}, \\ \mathcal{K}_n(A^T, \tilde{r}^0) &= span\{\tilde{r}^0, A^T\tilde{r}^0, ..., (A^T)^{n-1}\tilde{r}^0\}. \end{aligned} \tag{2.5}$$

The vector u^0 in (2.3) is initial guess, and the corresponding initial residual being defined as $r^0 = f - Au^0$. The vectors $p^0, \tilde{r}^0, \tilde{p}^0$ can be choosen arbitrarily, in principal, but if we define $\tilde{r}^0 = \tilde{f} - A^T\tilde{u}^0$ for some vectors $\tilde{f}$ and $\tilde{u}^0$ then the vector sequence $\tilde{u}^{n+1} = u^n + \alpha_n \tilde{p}^n$, if it converges, have the limit vector $\tilde{u}$ which is the solution of dual equation $A^T\tilde{u} = \tilde{f}$.

The scalar iterative parameters $\alpha_n, \beta_n, \tilde{\alpha}_n, \tilde{\beta}_n$ are computed from the orthogonal vector properties. We suppose in the following that correction vectors are satisfied to the one of conditions:

$$(A^q p^n, A^T\tilde{p}^k) = \rho_n^{(q)}\delta_{k,n}, \quad \rho_n^{(q)} = (A^q p^n, A^T\tilde{p}^n), \tag{2.6}$$

where $q = 0$ or $q = 1$ and $\delta_{k,n}$ is Kronecker symbol. If the vectors $r^n, p^n, \tilde{r}^n, \tilde{p}^n$ are subjected to the conditions

$$(A^q r^n, \tilde{p}^k) = (A^q r^n, \tilde{p}^n)\delta_{k,n}, \quad (A^q p^k, \tilde{r}^n) = (A^q p^n, \tilde{r}^n)\delta_{k,n}, \tag{2.7}$$

then from (2.4) we have that the parameters $\alpha_n, \tilde{\alpha}_n$ are defined as

$$\alpha_n = (A^q r^0, \tilde{p}^n)/\rho_n^{(q)}, \quad \tilde{\alpha}_n = (\tilde{r}^0, A^q p^n)/\rho_n^{(q)}. \tag{2.8}$$

Now, from the relation

$$\begin{aligned} &(A^{q-1}r^n, \tilde{r}^n) = (A^{q-1}r^0, \tilde{r}^0) - \\ &- \sum_{k=0}^{n-1}[\alpha_k(A^q p^k, \tilde{r}^0) + \tilde{\alpha}_k(A^q r^0, \tilde{p}^k) - \alpha_k\tilde{\alpha}_k\rho_k^{(q)}], \end{aligned}$$

which is valid under the condition (2.6), it is easy to see that the values (2.8) provide the variational properties

$$\frac{\partial(A^{q-1}r^n, \tilde{r}^n)}{\partial\alpha_k} = \frac{\partial(A^{q-1}r^n, \tilde{r}^n)}{\partial\tilde{\alpha}_k} = 0, \quad k = 0, 1, ..., n-1. \tag{2.9}$$

So, the corresponding equation is true:

$$\begin{aligned} &(A^{q-1}r^n, \tilde{r}^n) = (A^{q-1}r^0, \tilde{r}^0) - \\ &- \sum_{k=0}^{n-1}(A^q r^0, \tilde{p}^k)(A^q p^k, \tilde{r}^0)/\rho_k^{(q)}, \quad q = 0, 1. \end{aligned} \tag{2.10}$$

Let us consider the other inner product:

$$\begin{aligned} &(A^q r^n, \tilde{r}^k) = \Big((A^q p^n - \beta_{n-1}A^q p^{n-1}), (\tilde{r}^0 - \sum_{i=0}^{k-1}\tilde{\alpha}_i A^T\tilde{p}^i)\Big) = \\ &\Big((A^q r^0 - \sum_{i=0}^{n-1}\alpha_i A^{q+1}p^i), (\tilde{p}^k - \tilde{\beta}_{k-1}\tilde{p}^{k-1})\Big). \end{aligned} \tag{2.11}$$

From here for $k = n$ we obtain the equalities

$$(A^q r^n, \tilde{r}^n) = (A^q p^n, \tilde{r}^0) = (A^q r^0, \tilde{p}^n)$$

and new formulas for the iterative parameters:

$$\alpha_n = \tilde{\alpha}_n = \sigma_n^{(q)}/\rho_n^{(q)}, \quad \sigma_n^{(q)} = (A^q r^n, \tilde{r}^n). \tag{2.12}$$

So, the relation (2.10) can be rewritten in the following form:

$$(A^{q-1} r^n, \tilde{r}^n) = (A^{q-1} r^0, \tilde{r}^0) - \sum_{k=0}^{n-1} (A^q r^k, \tilde{r}^k)^2 / \rho_k^{(q)}.$$

From (2.11), for $k > n$ we have also

$$(A^q r^n, \tilde{r}^n) = (A^q p^n, \tilde{r}^0) - \tilde{\alpha}_n (A^q p^n, A^T \tilde{p}^n) -$$
$$-\beta_{n-1} \Big[(A^q p^{n-1}, r^0) - \tilde{\alpha}_{n-1} (A^q p^{n-1}, A^T \tilde{p}^{n-1}) \Big] = 0,$$

and for $k < n$ the following equality is true:

$$(A^q r^n, \tilde{r}^k) = (A^q r^0, \tilde{p}^k) - \alpha_k (A^q p^k, A^T \tilde{p}^k) -$$
$$-\tilde{\beta}_{k-1} \Big[(A^q r^0, \tilde{p}^{k-1}) - \tilde{\alpha}_{k-1} (A^q p^{k-1}, A^T \tilde{p}^{k-1}) \Big] = 0.$$

So, an important orthogonal property is valid:

$$(A^q r^n, \tilde{r}^k) = \sigma_n^{(q)} \delta_{n,k}. \tag{2.13}$$

It should be remarked that we did not define $\beta_n, \tilde{\beta}_n$ yet. To do this, we just exploit the properties (2.6):

$$(A^q p^{n+1}, A^T \tilde{p}^n) = (A^q r^{n+1} + \beta_n A^q p^n, A^T \tilde{p}^n) = 0,$$
$$(A^q p^n, A^T \tilde{p}^{n+1}) = (A^q p^n, A^T \tilde{r}^{n+1} + \tilde{\beta}_n A^T \tilde{p}^n) = 0.$$

So, we have

$$\beta_n = -(A^q r^{n+1}, A^T \tilde{p}^n)/\rho_n^{(q)}, \quad \tilde{\beta}_n = -(A^q p^n, A^T \tilde{r}^{n+1})/\rho_n^{(q)}. \tag{2.14}$$

If we use now the relations

$$A^T \tilde{p}^n = \frac{1}{\tilde{\alpha}_n} (\tilde{r}^n - \tilde{r}^{n+1}), \quad A^q p^n = \frac{1}{\alpha_n} A^{q-1} (r^n - r^{n+1})$$

and substitude them into (2.14) then we obtain the resulting formulas for parameters $\beta_n, \tilde{\beta}_n$:

$$\beta_n = \tilde{\beta}_n = \sigma_{n+1}^{(q)}/\sigma_n^{(q)}. \tag{2.15}$$

Let us consider the variational property of inner product

$$(A^q p^n, A^T \tilde{p}^n) = \Big((A^q r^n + \beta_{n-1} A^q p^{n-1}), (A^T \tilde{r}^n + \tilde{\beta}_{n-1} A^T \tilde{p}^{n-1}) \Big) =$$
$$= (A^q r^n, A^T \tilde{r}^n) + \beta_{n-1}(A^q p^{n-1}, A^T \tilde{r}^n) + \tilde{\beta}_{n-1}(A^q r^n, A^T \tilde{p}^{n-1}) + \beta_{n-1}\tilde{\beta}_{n-1}\rho_{n-1}^{(q)}.$$

It is easy to show that the coefficients $\beta_n, \tilde{\beta}_n$ from (2.14), (2.15) provide the conditions

$$\frac{\partial (A^q p^n, A^T \tilde{p}^n)}{\partial \beta_{n-1}} = \frac{\partial (A^q p^n, A^T \tilde{p}^n)}{\partial \tilde{\beta}_{n-1}} = 0, \tag{2.16}$$

and corresponding values of $\rho_n^{(q)}$ for $q = 0, 1$ are

$$(A^q p^n, A^T \tilde{p}^n) = (A^q r^n, A^T \tilde{r}^n) - (A^q p^{n-1}, A^T \tilde{p}^{n-1}) \frac{(A^q r^n, \tilde{r}^n)^2}{(A^q r^{n-1}, \tilde{r}^{n-1})^2}. \tag{2.17}$$

From relations (2.12), (2.15) we can define the rules for choice of initial vectors $\tilde{r}^0, p^0, \tilde{p}^0$ in each case $q = 0$ or $q = 1$:

$$(A^q r^0, \tilde{r}^0) \neq 0, \quad (A^q p^0, A^T \tilde{p}^0) \neq 0. \tag{2.18}$$

In practice, the conventional guess is

$$p^0 = \tilde{p}^0 = \tilde{r}^0 = r^0. \tag{2.19}$$

Obviously, the formulas (2.3) for $q = 0$ and $q = 1$ define two different algorithms but we omit the index q in vectors and iterative parameters for brevity.

It is easy to see that for $q = 0$ from (2.3) and (2.15) we obtain the biconjugate gradient method which for symmetric matrix $A = A^T$, under conditions (2.19), provides the conjugate gradient algorithm.

In the case $q = 1$ and $A = A^T$ the formulas (2.3), (2.15) give the conjugate residual method, see [3]-[5]. For such reason, we shall call algorithm (2.3), (2.15) for unsymmetric matrix and $q = 1$ as biconjugate residual one.

It is known that for symmetric SLAEs conjugate gradient and conjugate residuals methods provide the minimization of the functionals $\Phi_n^{(q)} = (A^{q-1} r^n, r^n)$ in the Krylov subspace $\mathcal{K}_n(A, r^0)$, and this property give the estimation of iterative convergence rate via Chebyshev polynomials.

In unsymmetric case the relation (2.10) doesn't mean an optimization of the functional, and we can not have the convergence estimate. If algorithms (2.3) for $q = 0$ or $q = 1$ don't fail then the vectors $p^0, p^1, ..., p^n$ and $\tilde{p}^0, \tilde{p}^1, ..., \tilde{p}^n$ are linear independent respectively and these iterative processes converge to exact solution of (2.1) in not more N steps, under condition of exact arithmetics.

3 Biconjugate direction squared algorithms

The residual and correction vectors in biconjugate direction (BCD) methods (2.3) for $q = 0, 1$ can be expressed in terms of their initial values:

$$\begin{aligned} r^n &= \varphi_n^{(q)}(A)r^0, \quad p^n = \psi_n^{(q)}(A)p^0, \\ \tilde{r}^n &= \varphi_n^{(q)}(A^T)\tilde{r}^0, \quad \tilde{p}^n = \psi_n^q(A^T)\tilde{p}^0. \end{aligned} \tag{3.1}$$

Here $\varphi_n^{(q)}(t)$ and $\psi_n^{(q)}(t)$ are polynomials of order n with scaling of the following forms:

$$\psi_n^{(q)}(0) = n + 1, \quad \varphi_0^{(q)}(t) = \varphi_n^{(q)}(0) = \psi_0^{(q)}(t) = 1. \tag{3.2}$$

These polynomials are satisfied to the recursions

$$\begin{aligned} \varphi_{n+1}^{(q)}(t) &= \varphi_n^{(q)}(t) - \alpha_n^{(q)} t \psi_n^{(q)}(t), \\ \psi_{n+1}^{(q)}(t) &= \varphi_{n+1}^{(q)}(t) + \beta_n^{(q)} \psi_n^{(q)}(t). \end{aligned} \tag{3.3}$$

Also, note that the scalars α_n, β_n in BCD are given by formulaes

$$\begin{aligned} \alpha_n^{(q)} &= \frac{(A^q\varphi_n^{(q)}(A)r^0, \varphi_n^{(q)}(A^T)\tilde{r}^0)}{(A^q\psi_n^{(q)}(A)p^0, A^T\psi_n^{(q)}(A^T)\tilde{p}^0)} = \frac{(A^q(\varphi_n^{(q)})^2(A)r^0, \tilde{r}^0)}{(A^q(\psi_n^{(q)})^2(A)p^0, A^T\tilde{p}^0)}, \\ \beta_n^{(q)} &= (A^q(\varphi_{n+1}^{(q)})^2(A)r^0, \tilde{r}^0)/(A^q(\varphi_n^{(q)})^2(A)r^0, \tilde{r}^0), \end{aligned} \tag{3.4}$$

which indicate that if it is possible to get the recursions for the vectors

$$\begin{aligned} \bar{r}^n &= \Phi_n^{(q)}(A)r^0, \quad \bar{p}^n = \Psi_n^{(q)}(A)p^0, \\ \Phi_n^{(q)}(t) &= (\varphi_n^{(q)})^2(t), \quad \Psi_n^{(q)}(t) = (\psi_n^{(q)})^2(t), \end{aligned} \tag{3.5}$$

then computing $\alpha_n^{(q)}$ and, similarly, $\beta_n^{(q)}$ causes no problem.

For the new polynomials of order $2n$ the following recurrences can be derived, in which the argument t is omitted:

$$\begin{aligned} \Phi_{n+1}^{(q)} &= \Phi_n^{(q)} - \alpha_n^{(q)} t(Y_n^{(q)} + X_{n+1}^{(q)}), \\ \Psi_n^{(q)} &= Y_n^{(q)} + \beta_n^{(q)}(X_n^{(q)} + \beta_n^{(q)}\Psi_{n-1}^{(q)}), \\ X_{n+1}^{(q)} &= Y_n^{(q)} - \alpha_n^{(q)} t\Psi_n^{(q)}, \\ Y_n^{(q)} &= \Phi_n^{(q)} + \beta_n^{(q)} X_n^{(q)}. \end{aligned} \tag{3.6}$$

Here an auxiliary polinomial $X_n^{(q)} = \psi_n^{(q)}\varphi_{n-1}^{(q)}$ is introduced. If we define the vectors

$$v^n = X_n^{(q)}(A)r^0, \quad w^n = Y_n^{(q)}(A)r^0,$$

and take into account the relations

$$\rho_n^{(q)} \equiv (A^q\Phi_n^{(q)}(A)r^0, \tilde{r}^0) = (\bar{r}^n, (A^q)^T\tilde{r}^0),$$

$$\sigma_n^{(q)} \equiv (A^q\Psi_n^{(q)}(A)p^0, A^T\tilde{p}^0) = (A\bar{p}^n, (A^q)^T\tilde{p}^0),$$

then the following conjugate direction squared methods for $q = 0$ and $q = 1$ can be defined, under initial guess of type (2.19):

$$
\begin{aligned}
&r^0 = f - Au^0, \quad p^0 = \tilde{p}^0 = \tilde{r}^0 = w^0 = r^0, \quad r_q^0 = (A^T)^q r^0,\\
&\rho_n^{(q)} = (r^n, r_q^0), \quad \sigma_n^{(q)} = (Ap^n, r_q^0),\\
&\alpha_n^{(q)} = \rho_n^{(q)}/\sigma_n^{(q)}, \quad v^n = w^n - \alpha_n^{(q)} Ap^n,\\
&u^{n+1} = u^n + \alpha_n^{(q)}(w^n + v^n),\\
&r^{n+1} = r^n - \alpha_n^{(q)} A(w^n + v^n),\\
&\beta_n^{(q)} = \rho_{n+1}^{(q)}/\rho_n^{(q)}, \quad w^{n+1} = r^{n+1} + \beta_n^{(q)} v^n,\\
&p^{n+1} = w^{n+1} + \beta_n^{(q)}(v^n + \beta_n^{(q)} p^n).
\end{aligned}
\tag{3.7}
$$

In these formulas the index "q" and symbol "bar" for the vectors r^n, p^n are omitted where there is no ambiguity. Algorithm (3.7) for $q = 0$ provides conjugate gradient square (CGS) method and for $q = 1$ it will be called as conjugate residual squared (CRS).

Observe that there are no matrix-by-vector products with the transpose of A in CGS. The single difference between CGS and CRS in the formulas (3.7) consists in definition of r_q^0 which must be computed for $q = 1$ before iterations only. Also, we can change the definitions of the scalar parameters

$$\rho_n^{(q)} = (A^q r^n, r^0), \quad \sigma_n^{(q)} = (A^q Ap^n, r^0),$$

so CGS and CRS both present two transpose free Krylov's algorithms.

In general, one should expect obtained algorithms for $q = 0$ and $q = 1$ to converge twice as fast as BCG and BCR correspondingly. Really, if $\max_{t\in S}\{|\varphi(t)|\} = 1 - \delta, \delta \ll 1$, on the spectrum S of martix A, then $\max_{t\in S}\{|\Phi(t)|\} \approx 1 - 2\delta$, what should be cause the decreasing the number of iterations.

4 Biconjugate direction stabilized methods

We consider in unified form two algorithms which for $q = 0$ present the known biconjugate gradient stabilized (BiCGSTAB) and for $q = 1$ provide the new biconjugate residual stabilized (BiCRSTAB) method.

The motivation of BiCGSTAB by H.A. van der Vorst was to obtain more smoothly converging variant of BiCG, because of irregular convergence behavior of CGS in some situations.

So, let us find iterative processes in which the residual and correction vectors are defined by the formulas

$$\bar{r}_n^{(q)} = \eta_n(A)\varphi_n^{(q)}(A)r^0, \quad \bar{p}_n^{(q)} = \eta_n(A)\psi_n^{(q)}(A)p^0, \tag{4.1}$$

where $\varphi_n^{(q)}, \psi_n^{(q)}$ provide CGS and CRS methods for $q = 0$ and $q = 1$ respectively and the polynomial $\eta_n(t)$ is satisfied to the recurrent relation

$$\eta_{n+1}(t) = (1 - \omega_n t)\eta_n(t), \tag{4.2}$$

with some scalar parameter ω_n which will be refined later.

From the equations (3.3) we have

$$\begin{aligned} \eta_{n+1}\varphi_{n+1}^{(q)} &= (1 - \omega_n t)(\eta_n \varphi_n^{(q)} - \alpha_n t \eta_n \psi_n^{(q)}), \\ \eta_n \psi_n^{(q)} &= \eta_n \varphi_n^{(q)} + \beta_{n-1}^{(q)}(1 - \omega_{n-1} t)\eta_{n-1}\psi_{n-1}^{(q)}. \end{aligned} \tag{4.3}$$

Thus, we can write the following recurrencies for the vectors (4.1):

$$\begin{aligned} \bar{r}_{n+1}^{(q)} &= (I - \omega_n A)(\bar{r}_n^{(q)} - \alpha_n^{(q)} A \bar{p}_n^{(q)}), \\ \bar{p}_{n+1}^{(q)} &= \bar{r}_{n+1}^{(q)} + \beta_n^{(q)}(I - \omega_n A)\bar{p}_n^{(q)}. \end{aligned} \tag{4.4}$$

Due to orthogonal properties (2.7) of the residuals and of the correction vectors, the iterative parameters $\alpha_n^{(q)}$ and $\beta_n^{(q)}$ can be rewritten in the unified form, see [4], [12] for details in the case $q = 0$:

$$\begin{aligned} \alpha_n^{(q)} &= \rho_n^{(q)}/\sigma_n^{(q)}, \quad \beta_n^{(q)} = \alpha_n^{(q)} \rho_{n+1}^{(q)}/(\omega_n \rho_n^{(q)}), \\ \rho_n^{(q)} &= (\bar{r}_n^{(q)}, (A^T)^q r^0), \quad \sigma_n^{(q)} = (A\tilde{p}_n^{(q)}, (A^T)^q p^0). \end{aligned} \tag{4.5}$$

Next, an additional free parameter ω_n in polinomial $\eta_{n+1}(t)$ must be defined. One of the simplest choices is to select ω_n to achieve the steepest step in the residual direction obtained before multiplying the corresponding vector by $(I - \omega_n A)$.

The first of the equations (4.4) can be rewritten as

$$\begin{aligned} \bar{r}_{n+1}^{(q)} &= \bar{r}_n^{(q)} - \alpha_n^{(q)} A \bar{p}_n^{(q)} - \omega_n A s^n = (I - \omega_n A) s^n, \\ s^n &= \bar{r}_n^{(q)} - \alpha_n^{(q)} A \bar{p}_n^{(q)}, \quad \bar{r}_0^{(q)} = r^0. \end{aligned} \tag{4.6}$$

Then the minimization condition $\partial ||\bar{r}_{n+1}^{(q)}||^2/\partial \omega_n = 0$ provides the optimal value for ω_n as

$$\omega_n = (As^n, s^n)/(As^n, As^n). \tag{4.7}$$

Finally, a formula is needed to update the approximate solution $u_{n+1}^{(q)}$ from $u_n^{(q)}$. Equation (4.6) for residual yields

$$u_{n+1}^{(q)} = u_n^{(q)} + \alpha_n^{(q)} \bar{p}_n^{(q)} + \omega_n s^n, \quad \bar{r}_n^{(q)} = f - A u_n^{(q)}. \tag{4.8}$$

After putting the above relations together, we obtain the following unified form of the BiCGSTAB and BiCRSTAB, for $q = 0$ and $q = 1$ respectively:

$$
\begin{aligned}
&r^0 = f - Au^0, \quad r_0^{(q)} = p_0^{(q)} = r^0, \quad n = 0, 1, \ldots : \\
&\alpha_n^{(q)} = (r^n, (A^T)^q r^0)/(Ap^n, (A^T)^q r^0), \quad u_0^{(q)} = u^0, \\
&s^n = r_n^{(q)} - \alpha_n^{(q)} Ap_n^{(q)}, \quad \omega_n = (As^n, s^n)/(As^n, As^n), \\
&u_{n+1}^{(q)} = u_n^{(q)} + \alpha_n^{(q)} p_n^{(q)} + \omega_n s^n, \quad r_{n+1}^{(q)} = s^n - \omega_n As^n, \\
&\beta_n^{(q)} = [\alpha_n^{(q)} (r^{n+1}, (A^T)^q r^0)]/[\omega_n (r^n, (A^T)^q r^0], \\
&p_{n+1}^{(q)} = r_{n+1}^{(q)} + \beta_n^{(q)} (p_n^{(q)} - \omega_n Ap_n^{(q)}).
\end{aligned}
\tag{4.9}
$$

5 Numerical results

We consider the comparative convergence efficiency of the above six iterative solvers, in application to the representative set of SLAEs which is provided by the Dirichlet three-dimensional boundary value problem (BVP) for diffusion-convection equation

$$
\begin{aligned}
&\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} + p\frac{\partial u}{\partial x} + q\frac{\partial u}{\partial y} + r\frac{\partial u}{\partial z} = f(x, y, z), \\
&(x, y, z) \in \Omega = [0, 1]^3, \quad u|_\Gamma = g(x, y, z).
\end{aligned}
\tag{5.1}
$$

The approximation of this BVP was made by means of exponential fitting seven-diagonal finite volume scheme on the cubic grids with the meshsteps $h = 1/(N+1)$, see [13]. The obtained matrices are monotone for any values of convection coefficients p, q, r, which were taken constant and variable, positive, negative and of different signs. In total, 10 various combinations of p, q, r were used, which are presented in the Table 1-6. The dimensions of tested systems are $(N-1)^3$, for $N = 32, 64, 128, 256$.

The stopping criteria

$$(r^n, r^n) \leq (f, f)\varepsilon^2, \quad \varepsilon = 10^{-7},$$

was used in all experiments. The simplest functions f, g were choosen in (5.1) to provide unit exact solution $u(x, y, z) = 1$. The initial guess u^0 for iterative processes was

$$u^0(x, y, z) = x^2 + y^2 + z^2. \tag{5.2}$$

Iterative processes in Krylov subspace were realized with preconditioning matrix

$$
\begin{aligned}
&B = (G - L)G^{-1}(G - U), \quad G = \frac{1}{\omega}D - \theta S, \\
&Se = \Big(\frac{1-\omega}{\omega}D + LG^{-1}U\Big)e,
\end{aligned}
\tag{5.3}
$$

where D, L and U are diagonal, low triangular and upper triangular parts of the original matrix, G and S are diagonal matrices, e is the vector with unit entries, ω and θ are relaxation and compensation parameters respectively. Here, for $\theta = 1$ we have row sum condition $Be = Ae$.

In fact, the solvers were applied for preconditioned SLAE

$$\begin{aligned}
&\bar{A}\bar{u} = \bar{f} = (I - \bar{L})^{-1}G^{-1/2}f, \quad \bar{u} = (I - \bar{U})G^{-1/2}u, \\
&\bar{A} = (I - \bar{L})^{-1} - (I - \bar{U})^{-1} - (I - \bar{L})^{-1}(2I - \bar{D})(I - \bar{U})^{-1}, \\
&\bar{L} = G^{-1/2}LG^{-1/2}, \quad \bar{U} = G^{-1/2}UG^{-1/2}, \quad \bar{D} = G^{-1/2}DG^{-1/2}.
\end{aligned} \tag{5.4}$$

An implementation of iterations was done by Eisenstat modification, of preconditioning, see [5], with cheap multiplication of the vector by matrix $\bar{A}$ in the following form which demands almost the same number of arithmetic operations as the vector multiplication by the original matrix $A = D - L - U$:

$$\bar{A}v = (I - \bar{L})^{-1}[v - (2I - \bar{D})w] + w, \quad w = (I - \bar{U})^{-1}v. \tag{5.5}$$

In the presented results, we use the periodic restarted variants of Krylov's iterations with integer parameter m: for each step with number $n_l = lm, l = 0, 1, 2, ...$ the residual vector is computed not by recursion formula but from equation $r^{n_l} = f - Au^{n_l}$, and orthogonalization process starts again.

In each cell of the Tables 1-6 the numbers of iterations are given, from the top to down, for three values $m = 100, 20, 10$ respectively. We use for all cases the parameters $\omega = \theta = 1$ for simplicity.

The results for preconditioned BiCG and BiCR methods are presented in the Tables 1, 2.

As we can see here, the numbers of iterations in BiCR are smaller, compare to BiCG methods. For $p = q = r = 0$ we have symmetric matrices, and BiCG, BiCR correspond to the "classic" conjugate gradient and conjugate resigual CG and CR methods.

Table 1. Preconditioned BiCG method, $\omega = \theta = 1$

p	-64	-16	-4	0	4	16	64	64	64	1-2x
q	-64	-16	-4	0	4	16	64	64	-64	0
r	-64	-16	-4	0	4	16	64	-64	-64	0
N										
	7	13	22	23	20	14	7	31	30	26
32	7	13	20	24	22	14	7	31	28	26
	7	14	21	26	23	14	7	31	28	27
	10	20	30	35	32	21	11	45	49	38
64	10	20	32	37	29	20	11	45	41	39
	10	25	31	42	30	21	10	80	78	44
	20	34	42	51	43	35	21	78	75	55
128	20	33	44	55	43	32	20	85	98	59
	21	38	42	65	44	31	20	138	177	67
	38	52	58	74	58	52	41	108	108	79
256	41	48	64	86	61	49	43	142	148	90
	31	47	61	104	70	62	32	244	265	108

Table 2. Preconditioned BiCR method, $\omega = \theta = 1$

p	-64	-16	-4	0	4	16	64	64	64	1-2x
q	-64	-16	-4	0	4	16	64	64	-64	0
r	-64	-16	-4	0	4	16	64	-64	-64	0
N										
	7	13	21	23	20	13	6	30	29	25
32	7	13	20	23	20	13	6	28	28	25
	7	14	20	25	27	14	6	32	31	27
	10	20	30	33	29	19	11	44	48	35
64	10	20	28	35	29	19	11	42	41	37
	10	22	28	39	29	20	10	68	80	42
	19	31	41	48	43	32	19	108	78	51
128	19	29	39	53	41	30	19	82	102	56
	18	33	40	61	41	34	18	178	233	64
	37	50	58	70	58	50	39	109	105	75
256	33	43	60	78	56	46	41	173	185	87
	27	41	58	101	66	45	31	189	282	106

The Tables 3, 4 present the similar results for the preconditioned conjugate direction squared methods (CGS and CRS). The symbol "∞" means here the divergence of iterations which is explained by the numerical non-stability of algorithms in some cases.

Table 3. Preconditioned CGS method, $\omega = \theta = 1$

p	-64	-16	-4	0	4	16	64	64	64	1-2x
q	-64	-16	-4	0	4	16	64	64	-64	0
r	-64	-16	-4	0	4	16	64	-64	-64	0
N										
	3	9	13	14	14	8	3	16	14	17
32	3	9	13	14	14	8	3	16	14	17
	3	9	14	16	14	8	3	17	15	19
	6	14	18	23	18	13	6	24	22	27
64	6	14	18	25	18	13	6	24	24	30
	6	16	19	25	19	12	6	52	41	28
	16	20	25	∞	26	19	16	38	38	43
128	16	22	33	∞	28	19	16	41	54	39
	21	39	39	38	34	23	16	367	∞	54
	31	30	37	∞	38	29	34	61	70	68
256	37	44	43	∞	43	37	38	224	200	72
	48	44	94	66	53	55	31	144	∞	72

Table 4. Preconditioned CRS method, $\omega = \theta = 1$

p	-64	-16	-4	0	4	16	64	64	64	1-2x
q	-64	-16	-4	0	4	16	64	64	-64	0
r	-64	-16	-4	0	4	16	64	-64	-64	0
N										
	3	9	12	14	11	8	3	15	14	17
32	3	9	12	14	11	8	3	15	14	17
	3	9	11	14	18	8	3	19	17	15
	6	13	18	21	18	12	6	23	23	24
64	6	13	18	21	18	12	6	22	23	20
	6	13	16	22	16	11	6	38	31	25
	14	20	25	∞	26	19	15	38	37	39
128	14	20	21	28	24	19	15	37	44	30
	11	19	30	30	24	18	11	68	58	30
	31	30	37	∞	30	28	33	58	69	58
256	23	24	33	40	34	24	23	78	87	44
	15	24	36	42	40	27	15	92	87	46

As we can see, CRS method is more stable, and it provides a good convergence for the moderate values of restart parameter ($m \approx 20$).

In the Tables 5, 6 the numbers of iterations are given for the stabilized algorithms. Really, the preconditioned BiCGSTAB and BiCRSTAB demonstrate a good stability and convergence rate for all considered values of convection coefficients p, q, r, restart parameter m and meshstep number N.

Table 5. Preconditioned BiCGSTAB method, $\omega = \theta = 1$

p	-64	-16	-4	0	4	16	64	64	64	1-2x
q	-64	-16	-4	0	4	16	64	64	-64	0
r	-64	-16	-4	0	4	16	64	-64	-64	0
N										
	4	9	12	16	12	8	4	18	16	17
32	4	9	12	16	12	8	4	18	16	17
	4	9	13	16	13	8	4	18	17	16
	6	12	16	24	18	12	6	28	27	22
64	6	12	16	24	18	12	6	28	28	23
	6	13	17	24	20	13	6	35	33	24
	11	17	24	38	25	16	9	45	43	33
128	11	17	23	31	24	16	9	51	57	32
	11	20	25	33	27	18	9	71	87	35
	16	23	34	55	33	25	16	66	74	52
256	16	24	35	46	36	25	16	86	131	52
	15	27	35	46	37	25	16	109	139	49

Table 6. Preconditioned BiCRSTAB method, $\omega = \theta = 1$

p	-64	-16	-4	0	4	16	64	64	64	1-2x
q	-64	-16	-4	0	4	16	64	64	-64	0
r	-64	-16	-4	0	4	16	64	-64	-64	0
N										
	4	9	12	17	12	8	4	18	16	16
32	4	9	12	17	12	8	4	18	16	16
	4	9	14	16	14	8	4	18	17	16
	6	12	16	26	17	12	6	29	27	25
64	6	12	16	25	17	12	6	28	29	26
	6	13	18	22	18	13	6	33	32	24
	9	17	25	40	26	17	9	44	43	37
128	9	17	24	32	25	17	9	48	54	32
	9	19	29	35	25	19	9	65	66	36
	16	24	34	53	36	23	16	69	77	53
256	16	24	32	45	34	25	16	82	110	48
	15	25	32	48	49	27	17	118	99	49

Because the convergence rate of the considered algorithms depends on the quality of the preconditioning matrix B in (5.3) and preconditioned system (5.4), (5.5), we compare the above results with the new ones which are obtained for empiric definition of the relaxation parameter ω from the condition $(Be, e) = (Ae, e)$, which provides the value

$$\omega = \frac{(e,e) - \sqrt{(e,e)^2 - 4(\bar{L}\bar{U}e,e)(e,e)}}{2(\bar{L}\bar{U}e,e)} \,. \tag{5.6}$$

Into the right hand site of (5.6) the values $\omega = \omega_0 = 1$ and $\theta = 0.975$ were used. The corresponding data for BiCR and CRS methods are presented in the Tables 7, 8. Also, for CRS we used the following trick in this case Namely, at the first iteration of each restart the simple minimal residual step was realized by the formulas

$$\begin{aligned} &u^{n+1} = u^n + \alpha_n r^n, \quad r^n = f - Au^n, \quad n = n_l = lm, \quad l = 0, 1, ..., \\ &r^{n+1} = r^n - \alpha_n A r^n, \quad \alpha_n = (Ar^n, r^n)/(Ar^n, Ar^n). \end{aligned} \tag{5.7}$$

Table 7. Preconditioned BiCR method, $\omega_0 = 1, \theta = 0.975$

p	-64	-16	-4	0	4	16	64	64	64	1-2x
q	-64	-16	-4	0	4	16	64	64	-64	0
r	-64	-16	-4	0	4	16	64	-64	-64	0
N										
	6	14	21	22	21	14	6	29	30	23
32	6	14	22	22	21	14	6	29	27	24
	6	13	21	24	23	13	6	28	27	24
	10	22	30	30	28	21	10	44	45	31
64	10	21	26	31	26	21	10	43	41	32
	10	21	27	33	30	24	10	64	84	34
	22	37	40	40	38	33	19	73	75	42
128	20	30	36	42	36	34	19	74	81	43
	20	35	36	45	37	41	22	124	147	46
	-	59	54	54	60	59	42	106	109	56
256	-	69	53	58	64	50	44	188	213	60
	-	50	56	68	62	52	35	235	236	69

Table 8. Preconditioned CRS method, $\omega_0 = 1, \theta = 0.975$

p	-64	-16	-4	0	4	16	64	64	64	1-2x
q	-64	-16	-4	0	4	16	64	64	-64	0
r	-64	-16	-4	0	4	16	64	-64	-64	0
N										
	4	11	12	13	10	15	4	21	15	14
32	4	11	12	13	10	15	4	21	15	14
	4	11	12	13	10	15	4	21	15	14
	6	19	14	17	17	13	6	54	43	19
64	6	19	14	17	17	13	6	54	43	19
	6	19	15	21	17	13	6	54	43	19
	10	19	29	22	23	18	11	55	88	27
128	10	19	29	23	23	18	11	55	88	24
	10	19	29	25	21	18	11	47	75	25
	31	49	62	31	42	54	20	97	149	32
256	31	49	62	32	32	54	20	97	149	33
	31	49	62	34	37	54	20	97	141	37

Table 9. Preconditioned CSR method, $\omega_0 = 1, \theta = 0.975$

p	-64	-16	-4	0	4	16	64	64	64	1-2x
q	-64	-16	-4	0	4	16	64	64	-64	0
r	-64	-16	-4	0	4	16	64	-64	-64	0
N										
	6	13	18	22	18	12	5	25	24	23
32	6	13	18	22	18	12	5	29	27	24
	6	13	19	24	18	12	5	38	36	24
	9	19	26	30	25	19	9	35	35	31
64	9	19	26	31	26	19	9	55	56	32
	9	20	26	33	26	19	9	68	79	34
	16	29	35	40	36	29	16	53	53	42

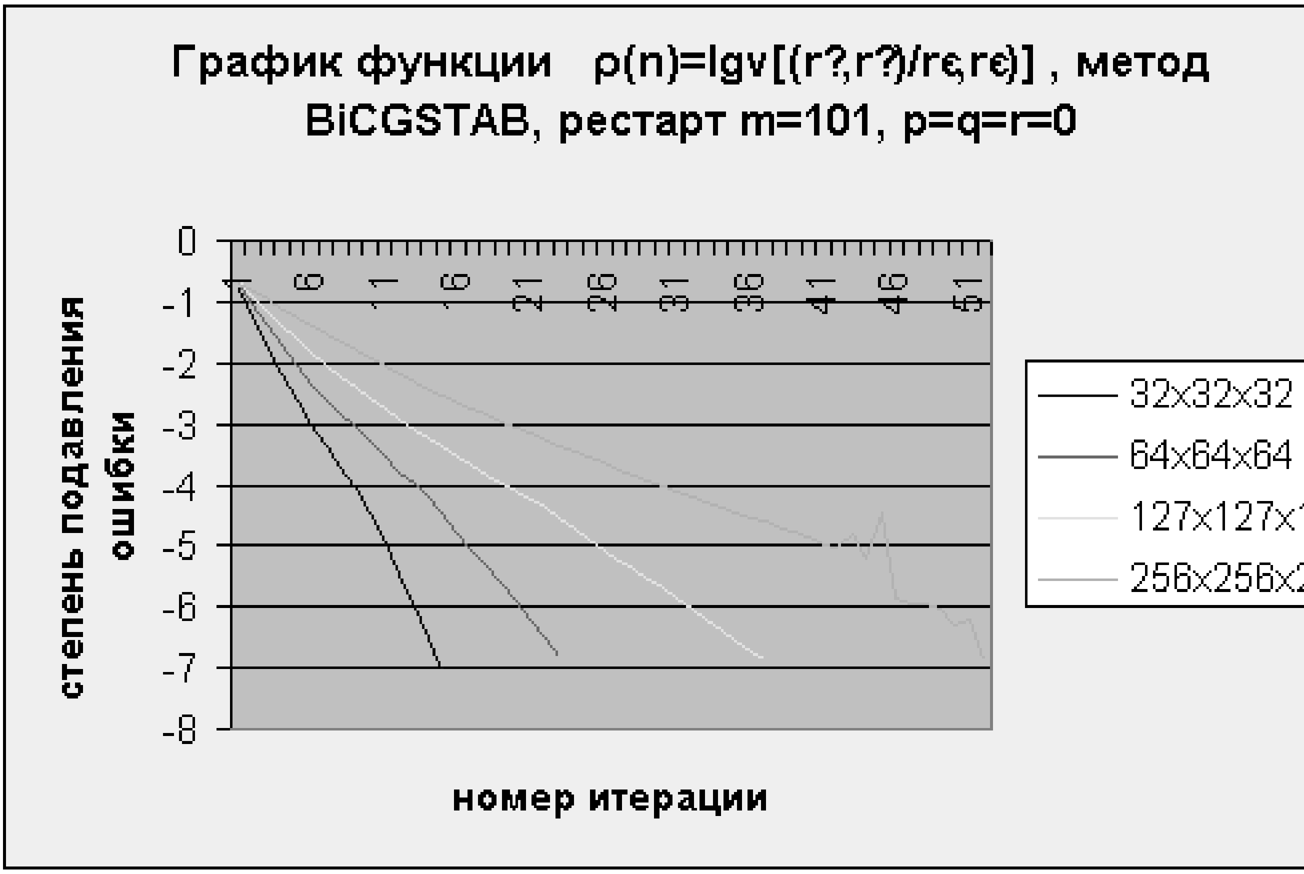

Figure 1: jkahgj'ahg

BiCO.

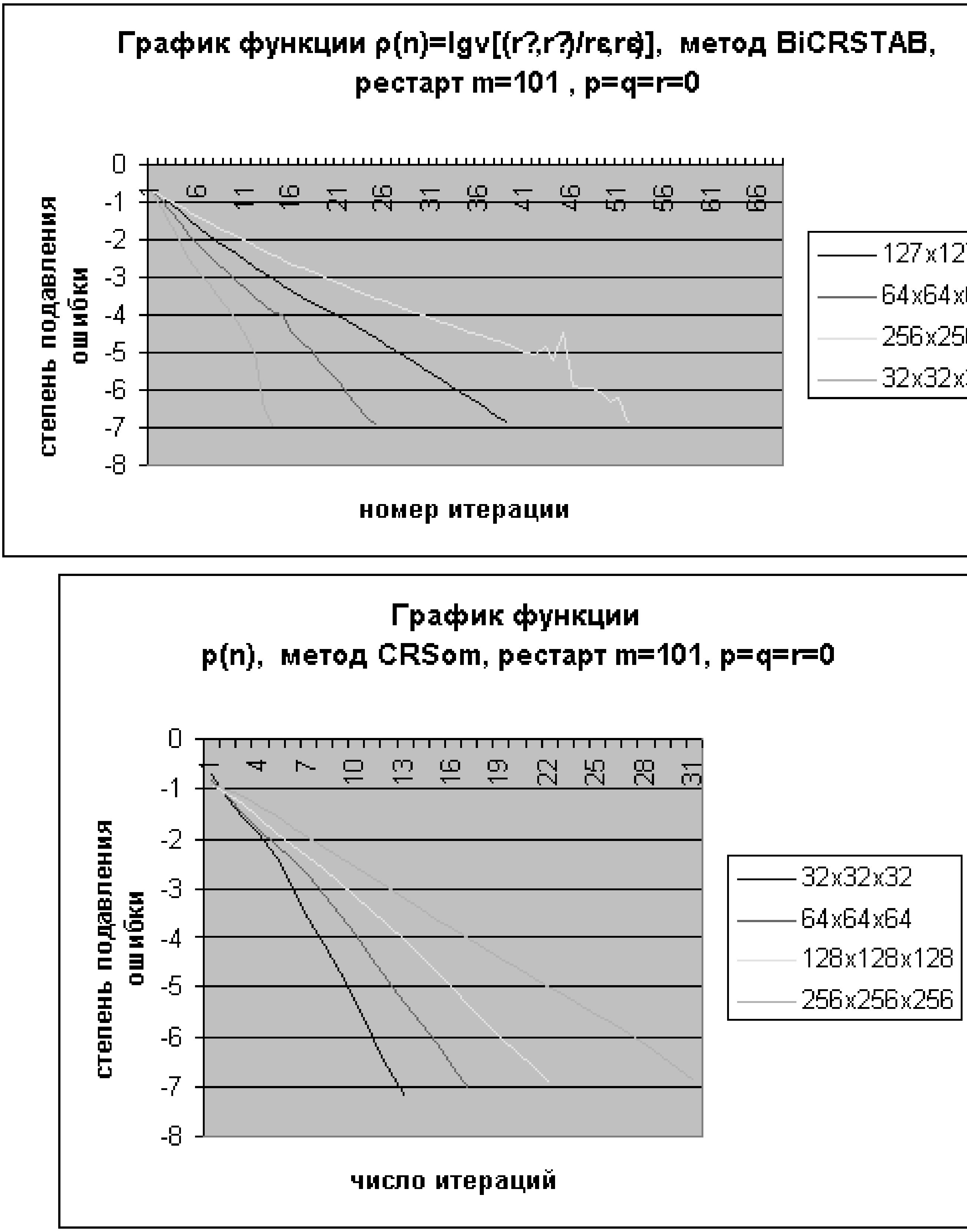

Figure 3: jkahgj'ahg

In conslusion, we make the following remarks on the given numerical results.

References

[1] D.K.Faddeev and V.N.Faddeeva, *Numerical Methods of Linear Algebra (in Russian),* Phys-Mathlit, Moscow, 1963.

[2] O.Axelsson, *Iterative Solution Methods*, Cambridge University Press, New York, 1994.

[3] G.H.Golub and C.Van Loan, *Matrix Computations*, The John Hopkins University Press, Baltimore,1996.

[4] Y.Saad, *Iterative methods for sparse linear systems*, PWS Publishing: New York, 1996.

[5] V.P.Il'in, *Finite Element Methods and Technologies (in Russian)*, Publ: Institute of Computational Math. and Mathematical Geophysics, SBRAS, Novosibirsk, 2007.

[6] J.Saad, M.Schultz, *GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems*, SIAM J. Sci. Stat. Comput., v. 7, (1986), 856-869.

[7] G.W.Stewart, *Conjugate direction methods for solving systems of linear equations*, Numer. Math., v.21, (1973), 285-297.

[8] J.Y.Juan, G.H.Golub, R.J.Plemmons, *Cecilio. Semi-conjugate direction methods for real positive definite systems*, Techn. Rep. SCCM Pe 02-02, Stanford Univ., 2003.

[9] C.Lanczos, *Solution of Systems of Linear Equations by Minimized Iterations*, Journal of Research of the National Bureau of Standards, 49, (1952), pp. 33-53.

[10] R.Fletcher, *Conjugate Gradient Methods for Indefinite Systems*, In G.A.Watson, editor, Proceedings of the Dundee Biennal Conference on Numerical Analysis 1974, pages 73-89. Springer Verlag, New York, 1975.

[11] P.Sonneveld, *CGS, a Fast Lanczos-type Solver for Non-symmetric Linear Systems*, SIAM Journal on Scientific and Statistical Computing, 10 (1), (1989), pp. 36-52.

[12] H.A. van der Vorst, *Bi-CGSTAB: A fast amd s,oothly converging variant of Bi-CG for the solution of non-symmetric linear systems*, SIAM Journal on Scientific and Statistical Computing, 12, (1992), pp. 631-644.

[13] M.Yu.Andreeva, V.P.Il'in, E.A.Itskovich, *Two solvers for nonsymmetric SLAE*, Bull. Novosibirsk Computing Center, ser. Numer. Anal, N 12, (2004), pp. 1-16.