ITERATIVE PROCESSES IN THE KRYLOV–SONNEVELD SUBSPACES

V.P.Il'in*

In this paper we present the generalized block version of Induced Dimension Reduction (IDR) methods, in compare with Multi–Preconditioned Semi-Conjugate Direction (MPSCD) algorithms in the Krylov subspaces, with deflation and low rank matrix approximation approaches. The common and individual orthogonal and variational properties of these two methodologies are analysed. It is demonstrated, in particular, that for any set of Krylov subspaces with dimension extension exists the set of corresponding shrinking subspaces with dimension reduction. The main conclusion consists in the statement that IDR procedures, proposed by P.Sonneveld and other authors, do not present the alternative but the further development of the general principles of the iterative processes in the Krylov subspaces.

1. INTRODUCTION

In 1980, a new iterative method [1] for solving nonsymmetric systems of linear algebraic equations (SLAEs) was published in the Proceedings of a symposium on the numerical solution of Navier–Stokes problems. This Induced Dimension Reduction (IDR) method received hardly any attention, and it was devoted to search the solution in the embedded subspaces of decreasing dimension. The author, P.Sonneveld published, almost 30 years later, the joint with van Gijzen paper [2] on the set of algorithms which follows from his old idea. In the following, many papers were published on this topic, see [3]–[9] for example. Some researches proposed the name "methods in Sonneveld subspaces", and these algorithms considered as alternative to the methods in Krylov subspaces.

Let us remark, that P.Sonneveld proposed the conjugate gradient square method CGS for solving nonsymmetric SLAEs. It was transposed free version of bi-conjugate gradient algorithm BiCG [10], [11]. This approah initiated the appearance of the stabilized bi-conjugate gradient method BiCGStab [10], after which the new generalizations revealed, BiCGStab(l), [11] in particular. In contrast to widspead used generalized minimal residual method GMRES [10], the proposed algorithms (with different types of preconditioning including) are based on the using the short recursions for computing residuals and other vectors, which became possible due to their bi-orthogonalization. Let us remark also, that in [12], [13], the similar set of algorithms with various preconditioners are constructed, under changing the bi-orthogonalization to A-biorthogonalization of the vectors. The corresponding methods A-BiCG (or BiCR, from Biconjugate Residual), A-CGS and A-BiCGStab,or CGR and BiCRStab respectivaly, describe in experiments a good and even better stability, in compare with their prototypes with classical bi-orthogonalization.

In the consequent papers ([14]–[17] and many others), the different authors investigated the versions of IDR(s) methods. Their connections with BiCGStab(l) were established, and

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

the generalization IDRStab(s,l) was proposed as well as various versions in particular cases. These algorithms became use not only for solving SLAEs, but for computing the eigenvalues and corresponding eigenvectors. Formally, the parameter "s" in the last algorithm means the dimension of "shadow residual" space S, to which the Sonneveld subspaces are orthogonalized, and the second parameter "l' defines the order of auxiliar matrix polinomials, or the corresponding subspaces, by which the residual norm is minimized at each iterative step.

The goal of this paper consists, at first, in block representation of the IDRStab(s,l) methods, and to give different known variants, in compare with the modern modifications of the clasical iterative processes in the Krylov subspaces. Secondly, we want to remove mis-understanding with opposition of the algorithms in Krylov and Sonneveld subspaces, on the example of the proposed Multi-Preconditioned Semi-Conjugate Direction (MPSCD, see [18], [19]) methods, which use three following approaches. The first one consists in the possibility of A-biorthogonalization of the vectors as well as classical biorthogonalization, and the second – in employing the multipreconditioned algorithms of semi-conjugate direction (in particular – Semi-Conjugate Residual, SCR) methods, which present the further development of considered in [20], [21] methods for realization of the matrix polynomials, which minimize the residual norm. Another principial moment of the considered MPSCD method consists in the orthogonalization of the residual vectors, at each iteration, to originaly predescribed "deflation" subspace, which in fact presents the analogy to "shadow residual" subspace S in IDR methods. The idea of deflation in Krylov subspaces was proposed in 1987 by R.Nicoloides in the paper [22], and was developed later in different directions (quite independetly of the induction dimension reduction methods), under the names aggregation, augmentation, coarse grid correction and low rank matrix approximation, see reviews in [23] – [27].

The third aspect of the considered algorithms consists in the application at every iteration of the several preconditioned matrices. It was proposed firstly by R.Bridson and C.Greif in [28] for conjugate direction method. Let us remark, that in our consideration, the deflation, or aggregation, procedure is interpretated as using the additional prconditioner of special type.

This paper is organized as follows. In Section 2 we give generalized block presentation and short review of IDR methods, with consideration of them as special algorithms in Krylov subspaces. In Section 3, we describe the MPSCD methods which possess some properties in the shrink Sonneveld subspaces, but on the base of classical modified Krylov subspaces. In conclusion we discuss the obtained results.

2. BLOCK PRESENTATION OF IDRStab(s,l) METHODS

Let we have to solve by some iterative method the real nonsingular SLAEs

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

under some initial guess u^0 and residual vector $r^0 = f - Au^0$. In this Section we make accent on the property of induced dimension reduction methods only, without stopping on the possibility of the different type preconditioning the algebraic system. In particular, the matrix A in (1) can be considered as preconditioned already.

Let we have two subspaces $\mathcal{G}_0$ and $\mathcal{S}$, which have no the common nontrivial invariant subspaces of matrix A, and are defined by the following way:

$$\begin{aligned} &\mathcal{G}_0 = \mathcal{K}_N(A, g_1) = \text{Span } (g_1, Ag_1, ..., A^{N-1}g_1), \;\; g_1 \in \mathcal{R}^N, \;\; g_1 \neq 0, \\ &\mathcal{S} = \mathcal{N}(\tilde{R}^T), \;\; \tilde{R} = (\tilde{r}_1\, \tilde{r}_2 ... \tilde{r}_s) \in \mathcal{R}^{N,s}, \;\; \tilde{r}_k \in \mathcal{R}^N, \;\; k = 1, ..., s. \end{aligned} \tag{2}$$

Here g_1 is some nonzero vector, $\mathcal{K}_N$ – full Krylov subspace, $\mathcal{S}$ is the left null-subspace of the given rectangular matrix $\tilde{R}$ of full rank, and $s \ll N$ columns $\tilde{r}_k$ of which present the basis of $\mathcal{S}$ and called as initial shadow residual vectors, and idex "T" means the transposition. In other words, if some vector g belongs to $\mathcal{S}$, then it is orthogonalized to all $\tilde{r}_k$ ($g \perp \tilde{r}_k$, or $(g, \tilde{r}_k) = 0$, $k = 1, ..., s$), which can be written also, as $\mathcal{S} = \tilde{R}^{\perp}$).

The assumtion, made on the subspace from (2), means for example that no eigenvector of matrix A can belong simultaneously to $\mathcal{G}_0$ and $\mathcal{S}$, i.e. any egeinvector from $\mathcal{G}_0$ is orthoganal to all columns of matrix $\tilde{R}$.

Now, we consider the matrix polynomials

$$P_{l_n}(A) = \prod_{k=1}^{l_n} (\mu_k I - A), \;\; l_n \ll N, \tag{3}$$

where I is identity matrix, l_n are the orders of the corresponding polynomials, and μ_k – some real values, or shift parameters of the matrix A, which will discussed later. We construct the sequence of subspaces, named as Sonneveld ones:

$$\mathcal{G}_n = P_{l_n}(A)(\mathcal{G}_{n-1} \bigcap \mathcal{S}), \tag{4}$$

which dimensions will be denoted by α_n. The properties of these embedded subspaces are placed in the based of the considered in the following iterative methods, and they follows from the so called IDR-theorem, which is proved for $l_n = 1$ in [2], and for fixed $\ell_n = \ell > 1$ – in the paper [6].

Theorem 1. *Let $A \in \mathcal{R}^{N,N}$ and $\tilde{R} \in \mathcal{R}^{N,s}$ be non-singular matrix and full rank matrix, $\mathcal{G}_0$ and $\mathcal{S}$ be defined in (2) subspaces, which do not include the eigenvalues of matrix A, and $P_{l_n}(A)$ be the polynomial of order $l_n < s$ from (3). The Sonneveld subspaces (4) are satisfied to embedding condition $\mathcal{G}_n \subseteq \mathcal{G}_{n-1}$, if $\mathcal{G}_{n-1} \neq \{0\}$, and their dimensions are satisfied to the condition of monotone non-icreasing the differences $d_n - d_{n-1}$:*

$$0 \leq d_n - d_{n+1} \leq d_{n+1} - d_{n+2} \leq s. \tag{5}$$

Remark 1. The sequence of subspaces are embedded and finite, e.i. $\mathcal{G}_n \subset \mathcal{G}_{n-1}$ and $d_n < d_{n-1}$, if $\mathcal{G}_n = \mathcal{G}_{n-1} \neq \{0\}$. If for $0 < n < N$ the relation $\mathcal{G}_n = \mathcal{G}_{n-1} \neq \{0\}$ is valid, then $\mathcal{G}_{n-1} \cap \mathcal{S}$ contains the eigenvector of matrix A. In fact, istead $\mathcal{G}_0$ from (2) the any linear subspace, invariant to multiplication of the matrix A ($A\mathcal{G}_0 \subset \mathcal{G}_0$), can be taken.

Remark 2. The Sonneveld subspaces $\mathcal{G}_n$ (it is shown in [4]) can be presented by means of special block Krylov subspaces, which are naturally extended with the grows of n:

$$\mathcal{K}_n(A^T, \tilde{R}) = \{\sum_{k=0}^{n-1} (A^T)^k \tilde{R} \bar{c}_k |\; \bar{c}_k \in \mathcal{R}^s\}, \tag{6}$$

where c_k are coefficient vectors, which form the linear combinations of the columns of matrix $\tilde{R}$, are defined in the following by the orthogonalization conditions. Really, in accordsnce to (2)–(4) for $n = 1$ we have

$$\mathcal{G}_n = \{P_{l_n}(A)g |\; P_{l_n}(A)g \perp \tilde{R} \in \mathcal{K}_n(A^T, \tilde{R})\}. \tag{7}$$

But the orthogonality condition $P_{l_n}(A)g \perp \tilde{R}$ can be satisfied, if and only if $A^k g \perp \tilde{R}$ for all $k \leq l_n$. This is equivalent to relation $g \perp (A^T)^k \tilde{R}$, from which we have by induction

$$\mathcal{G}_n = \{P_{l_n}(A)g |\; g \perp \mathcal{K}_n(A^T, \tilde{R})\}. \tag{8}$$

The last means that for each extending sequence of block Krylov subspaces it is possible to associate the sequence of compressible Sonneveld subspaces.

The considered preliminary geometric representations about Sonneveld subspaces help to construct the general scheme of iterative process, in which the residual vectors $r^n = f - Au^n$ will belong to $\mathcal{G}_n$. If the vectors u^{n-1} and $r^{n-1} \in \mathcal{G}_{n-1}$ are known, then one iteration of the IDRStab(s,l) method can be presented in two stages, at which we define the vectors

$$r^{n-1/2} \in \mathcal{G}_{n-1} \cap \mathcal{S}, \;\; r^n = P_{l_n}(A) r^{n-1/2}. \tag{9}$$

Here, the first stage includes the orthogonalization of residual r^{n-1} to the shadow residual vectors $\tilde{r}_k$, and the second is responsible for the minimization of r^n. Let us remark that the convential algorithm BiCGStab is corresponded to the case $s = \ell = 1$. The orthoganalization step can be done by means of modified Gram–Smidt method. In accordance with (9), for computing the vector $r^{n-1/2} \in \mathcal{G}_{n-1} \cap \mathcal{S}$ we define firstly $s+1$ vectors $g_{-s}, ..., g_0 \in \mathcal{G}_0 = \mathcal{K}_N(A, g_1)$ for $n = 0$. When this is done, we use the presentation

$$r^{-1/2} = g_0 - \sum_{k=-s}^{-1} g_k c_k,$$

where unknown coefficients c_k are defined from orthogonalization condition

$$\tilde{R}^T r^{-1/2} = 0, \;\; \tilde{R} \in \mathcal{R}^{N,s}, \tag{10}$$

which gives the system of ℓ linear equations

$$Cc = R^T G\, c = R^T g_0, \;\; G = (g_{-1} \cdots g_{-s}) \in \mathcal{R}^{N,s}, \;\; c = \{c_k\} \in \mathcal{R}^s. \tag{11}$$

It has unique solution, if the matrix $C = R^T G \in \mathcal{R}^{s,s}$ is non-singular.

Remark 3. In fact, the computation of vector c from (10) presents the solving least square problem [29]. It's normal solution (with minimal norm) can be computing directly from overdetermined system $G\, c = g_0$ with rectangular matrix G. If the matrices G or C are singular, then it is reasonable to use generalized inverse matrices or, for example, the singular value decomposition SVD.

To solve SLAEs (11), it is necessary preliminary to find vectors $g_k \in \mathcal{K}_N(A, g_1)$. It can be implemented by some convential algorithm in Krylov subspaces. One of the way here consists in choosing $g_0 = r^0$ and computing the next g_k by Arnoldi orthogonalization process for the vectors $A^k g_0$.

The following stage of algorithm includes computing the vector $r^1 = P_{l_1}(A) r^{-1/2} \in \mathcal{G}_1$. Multiplication by the matrix polynomial of A is realized after defining the shift parameters μ_k in (3), and it will be discussed later. After completing the first step, the second stage begin with the realization of the orthogonalization condition (11). Further, the computational process continues in a similar way. In fact, two level iterative process is fulfilled here. The high level corresponds to the sequence of Sonneveld subspaces, which are numbered by n, and the low level presents the implementation of internal iterations in Krylov subspaces.

In total, for each $n = 1, 2, ...$ two kind of presented in (9) operations are implemented:

a. Multimplication of the residual vector r^{n-1} from the previous Sonneveld subspace by the polynomial $P_{l_n}(A)$:
$$\hat{r}^n = P_{l_n}(A) r^{n-1}, \;\; r^{n-1} \in \mathcal{G}_{n-1} \cap \mathcal{S}.$$

b. Solving the SLAEs for the coefficients c_k, which provide the orthogonalization:

$$R^T r^n = 0, \;\; r^n = \hat{r}^n - \sum_{k=-s}^{-1} g_k c_k = \hat{r} - G\,c.$$

It is possible to take the number of basis vectors $g_k \in \mathcal{G}_n$ not equal to s, and consider it as the least square problem, but we will not stay on these details. Let us remark, that we use here and in the following the block representation of the iterative process, when the orthogonalization of residual vector to all columns of matrix $\tilde{R}$ is included in one iteration.

The other interesting questions on the IDR methods concern to definition of the matrix $\tilde{R}$, initialization of the iterative process, and choosing the space $\mathcal{G}_0$ and vector $r^0 \in \mathcal{G}_0$, as well as the algorithm of computing the consequent residual vectors $r^n \in \mathcal{G}_n$. In all considered algorithms, we suppose that stopping criteria for iterations is

$$||r^{n(\varepsilon)}|| \le \varepsilon ||f||, \;\; \varepsilon \ll 1, \;\; ||f||^2 = (f,f). \tag{12}$$

Remark 4. As it was mentioned in [6], the IDR(s,l) methods relate to Petrov–Galerkin algorithms in the following sense. Let the vectors $u^n - u^0$ and r^n belong to the Krylov subspaces $\mathcal{K}_n$ and $\mathcal{K}_{n+1}(A, r^0)$ respectively (if the order of minimal polynomial $P_n(A) = 0$ equals to $n = N$, we have $\mathcal{K}_N = \mathcal{R}^N$). Then Petrov–Galerkin method with respect to the set of embedded subspaces $\{\mathcal{L}_m\}$ is defined by orthogonalization of residual vector r^n to subspace $\mathcal{L}_m$ of order m, i.e. $r^n \perp \mathcal{L}_m^{\perp}$, or $r^n \in \mathcal{L}_m^{\perp}$. In particular, if $\mathcal{L}_m = A\mathcal{K}_n$, then this scheme includes GMRES. Because $\mathcal{R}^N = \mathcal{K}_{n+1} \oplus \mathcal{K}_{n+1}^{\perp}$, the dimensions of $\mathcal{K}_{n+1}$ and $\mathcal{K}_{n+1}^{\perp}$ increase and decrease simultaneously, with growing n, then orthogonal comlement $\mathcal{K}_{n+1}^{\perp}$ to Krylov subspace play a role of Sonneveld subspace, in some sense. The papers from the Reference and citied therein works contain the wide reviews and existing interpretations of the various IDR methods, which witness about notstable methodology of these iterative processes yet.

3. MULTI-PRECONDITIONED SEMI-CONJUGATE DIRECTION METHODS

In this Section, we consider the application of some of any block semi-conjugate direction methods [18], [19] which are the development of generalized conjugate residual algorithm [20]. If we put in (2) $g_1 = r^0$, the proposed set of methods MPSCD in Krylov subspaces $\mathcal{K}_m(A, r^0)$ can be written as

$$u^{m+1} = u^m + P_m \bar{\alpha}_m, \;\; r^{m+1} = r^m - AP_m\bar{\alpha}_m, \;\; m = 0, 1, ..., \tag{13}$$

where $P_m = (p_1^m ... p_{M_m}^m) \in \mathcal{R}^{N,M_m}$ is the matrix consisted of direction vectors p_k^m, and $\bar{\alpha}_m = (\alpha_{m,1} ... \alpha_{m,M_m})^T$ is the iterative parameter vector, which are defined by orthogonal properties

$$\begin{aligned} &(Ap_k^m, A^\gamma p_{k'}^{m'}) = \rho_{m,k}^{(\gamma)} \delta_{m,m'}^{k,k'}, \;\; \rho_{m,k}^{(\gamma)} = (Ap_k^m, A^\gamma p_k^m), \\ &\gamma = 0, 1, ; m' = 0, 1, ..., m-1; \;\; k, k' = 1, 2, ..., M_m. \end{aligned} \tag{14}$$

Here $\delta_{m,m'}^{k,k'}$ is Kronecker symbol which equals to unit for $m = m'$, $k = k'$ and zero in other cases, and the values $\gamma = 0, 1$ define semi-conjugate gradient or semi-conjugate residual method respectivaly. In contrast to convential semi-conjugate direction algorithms, in formulas (13) at

every m-th iteration we have not one but M_n direction vector, and its can change, in general, at different iterations.

The coefficient vectors $\bar{\alpha}_m = \{\alpha_{m,l}^{(\gamma)}\}$ in (13), (14) for $\gamma = 0, 1$, under extremum condition

$$\partial \Phi_m^{(\gamma)} / \partial \alpha_{m,l} = 0, \;\; \Phi_m^{(\gamma)}(r^{m+1}) \equiv (r^{m+1}, A^{\gamma-1} r^{m+1}), \tag{15}$$

are defined by formulae

$$\alpha_{m,l}^{(\gamma)} = (A^{\gamma} B_{m,l}^{-1} r^m, r^m) / \rho_{m,l}^{(\gamma)}. \tag{16}$$

The direction vectors p_ℓ^m are defined by (14) for $\gamma = 0, 1$ in the following block form:

$$P_0 = \{p_l^0 = B_{0,l}^{-1} r^0\}, \;\; P_{m+1} = P_{m+1,0} - \sum_{k=0}^{m} P_k \bar{\beta}_{m,k}^{(\gamma)} = \{p_l^{m+1} = B_{m+1,l}^{-1} r^{m+1} - \sum_{k=o}^{m} \sum_{l=1}^{M_k} \beta_{m,k,l}^{(\gamma)} p_l^k\},$$
$$m = 0, 1, ...; \;\; B_{m,l} \in \mathcal{R}^{N,N}, \;\; l = 1, ..., M_m, \;\; \gamma = 0, 1. \tag{17}$$

Here $\bar{\beta}_{m,k}^{(\gamma)} = \{\beta_{m,k,l}^{(\gamma)}\} = (\beta_{m,k,1}^{(\gamma)} ... \beta_{m,k,M_m}^{(\gamma)})^T \in \mathcal{R}^{M_m}$ are coefficient vectors, and $B_{m,l} \in \mathcal{R}^{N,N}$ are preconditioning matrices, which are choosen by the reasons of non-singularity, easy invertability, and efficient acceleration of the constructed iterative process. The considered preconditioners $B_{m,\ell}$ are dynamic one, or flexible, as in FGMRES [10], and they depend on the iteration number m, for any ℓ.

After substituting (17) in orthogonality condition (14), we obtain the formulae for coefficients:

$$\begin{array}{l} \beta_{m,k,l}^{(\gamma)} = (A^{\gamma} p_l^k, A B_{m+1,l}^{-1} r^{m+1}) / \rho_{m,l}^{(\gamma)} = (A^{\gamma} p_l^k, A p_l^{m,k}) / \rho_{m,l}^{(\gamma)}, \;\; m = 0, 1, ...; \\ k = 0, ..., m; \;\; l = 1, ..., M_m, \end{array} \tag{18}$$

in which the coefficients $p_\ell^{m,k}$ are defined by the relations

$$\begin{array}{l} p_l^{m,k} = p_l^{m,k-1} - \beta_{m,k-1,l}^{(\gamma)} p_l^{k-1} = B_{m+1,l}^{-1} r^{m+1} - \sum_{i=0}^{k-1} \beta_{m,i,l}^{(\gamma)} p_l^m, \;\; l = 1, ..., M_m, \\ k = 0, 1, ..., m+1; \;\; p_l^{m,0} = B_{m+1,l}^{-1} r^{m+1}, \;\; p_l^{m,m+1} = p_l^{m+1}, \end{array} \tag{19}$$

on the base of modified Gram–Smidt method.

The following statement is true (see proof in (19)).

Theorem 2. *For the iterative process MPSCD, which is defined by the relations (13), (16)–(19) for $\gamma = 0, 1$ and non-singular matrices $A, B_{m,\ell}$, the following assertions are valid:*

- *direction vectors p_k^m are satisfied to orthogonality conditions (14);*
- *residual vectors r^m are generalized semi-conjugate, i.e.*

$$(A^{\gamma} B_{k,l}^{-1} r^m, r^k) = \begin{cases} 0, & k < m, \\ \sigma_m^{(\gamma)} = (A^{\gamma} B_{m,l}^{-1} r^m, r^m), & k = m, \end{cases} \tag{20}$$

- *the functionals $\Phi_m^{(\gamma)}(r^{m+1})$ satisfy to extremality conditions (15) and for any $q = 0, 1, ..., m$ the following relations are valid:*

$$\Phi_m^{(\gamma)}(r^{m+1}) = \Phi_q^{(\gamma)}(r^q) - \sum_{k=q}^{m} \sum_{l=1}^{Mm} (A^{\gamma} B_{q,l}^{-1} r^q, r^q)^2 / \rho_{k,l}^{(\gamma)}. \tag{21}$$

Remark 5. If $\gamma = 1$ or if matrix A is symmetric, the method MPSCD provides minimizing norm $||r^{m+1}|| = (r^{m+1}, r^{m+1})^{1/2}$ in "multi-preconditioned" Krylov subspace

$$\begin{aligned} &\mathcal{K}_{\sum_{m+1}}(r^0, A) = \mathrm{Span}\{B_{0,1}^{-1} r^0, ..., B_{0,M_0}^{-1} r^0, \\ &AB_{1,1}^{-1} r^1, ..., AB_{1,M_1}^{-1} r^1, ..., A^m B_{m,1}^{-1} r^m, ..., A^m B_{m,M_m}^{-1} r^m\}, \end{aligned} \tag{22}$$

which dimension equals $\sum_{m+1} = M_0 + ... + M_m$. In semi-conjugate gradient methods, i.e. for $\gamma = 0$, the functional $\Phi_m^{(0)}(r^{m+1}) = (A^{-1} r^{m+1}, r^{m+1})$ does not reach its minimum, if A is non-symmetric matrix.

In the considered block semi-conjugate direction methods, we have so called "long" recursion, i.e., it is necessary to save all direction vectors which were computed at the previous iterations, or direction matrices $P_1, ..., P_m$. In order to avoid strong requirements to memory, we will use the known restart procedure: after some number of iterations the residual vector is defined not from usual recursion (13), but from original equation, and Krylov process is formed anew, from the current iterative value of unknown vector.

In order to accelerate the obtained iterative process, at each restart with number n, we will realize the orthogonalization of the residual vector to subspace, which is similar to $\mathcal{S}$ in (2). However, this procedure is fulfilled not by conventional for IDR methods approach but by help of deflation, or aggregation hint, see [24] and papers, citied therein. In this case, in fact, some additional preconditioning matrix is formed on the base of some rectangular matrix $\tilde{R}_n = (\tilde{r}_1 ... \tilde{r}_{s_n})^T \in \mathcal{R}^{N, s_n},\ s_n \ll N$:

$$B_{n,0}^{-1} = \tilde{R}_n (\hat{A}_n)^{-1} \tilde{R}_n^T,\ \ \hat{A}_n = \tilde{R}_n^T A \tilde{R}_n \in \mathcal{R}^{s_n, s_n}. \tag{23}$$

Here, $B_{n,0}^{-1}$ is low rank approximation of the matrix A^{-1}, inversed to original one. It is non-singular, if A is the same, and if deflation matrix has full rank s_n. We suppose here, that the subspaces $\mathcal{S}_n$ and their dimensions s_n depend on the number of restart n. Denoting by m_n the number of iteration ($m_0 = 0$) of MPSCD method (13), implementation of restart and orthogonalization including, we can write the obtained algorithm in the following way:

$$\begin{aligned} &u^0 = u^{-1} + B_{0,0}^{-1} r^{-1},\ \ r^{-1} = f - Au^{-1}, \\ &r^0 = f - Au^0,\ \ p_l^0 = (I - B_{0,0}^{-1} A) B_{0,l}^{-1} r^0,\ \ l = 1, ..., M_0. \end{aligned} \tag{24}$$

In this case, u^{-1} denotes arbitrary initial guess, and u^0 is the "corrected" original value of the solution to be sought. It is easy to check, that formulas (24) provide the conditions

$$\tilde{R}_0^T r^0 = 0,\ \ \tilde{R}_0^T A P_0 = 0. \tag{25}$$

where matrix P_0 consists of M_0 columns p_ℓ^0, which are defined in (24).

The realization of formulas (24) is simple enough, because multiplication by $B_{0,0}^{-1} \in \mathcal{R}^{N,N}$ includes the solution of auxiliary SLAEs with low order dimension $\hat{A}_0 = \tilde{R}_0^T A \tilde{R}_0 \in \mathcal{R}^{s_0, s_0}$. We save especialy the notations for $\tilde{R}_n$ in formulas (23), in order emphasize the analogy with the methods in Sonneveld supspaces from Section 2.

For definition of linear independent columns $\tilde{r}_k$, $k = 1, ..., s_n$, of matrix $\tilde{R}_n$, it is possible to use the simplest piece-wise constant basis vectors: divide the set of indeces $\Omega : i = 1, ..., N$ into equal approximately non-overlapping subsets $\Omega_k^{(n)}$ (these decompositions can be different at the different iterations) and put $\tilde{r}_k(i) = 1$ for $i \in \Omega_k^{(n)}$, and $\tilde{r}_k(i) = 0$ for $i \notin \Omega_k^{(n)},\ k = i, ..., s_n$. Let us remark, that basis functions can be defined and more smooth, as it was done in [26]. In

some papers, the columns of matrix $\tilde{R}_n$ are choosen in random manner, or by means solving auxiliary eigenproblem. In total, the problem of choosing the most informative basis is the topic for special research.

If the iterative value u^{m_n-1} is known at each restart, then formulas (24), (25) are changed evidently:

$$\begin{array}{l} u^{m_n} = u^{m_n-1} + B^{-1}_{m_n,0} r^{m_n-1}, \;\; r^{m_n-1} = f - Au^{m_n-1}, \\ r^{m_n} = f - Au^{m_n}, \;\; p_l^{m_n} = (I - B^{-1}_{m_n,0}A)B^{-1}_{m_n,l} r^{m_n}. \end{array} \tag{26}$$

In this case, the preconditioners $B_{m_n,l}$ are defined by (23), under changing n to m_n, and orthogonalization conditions are changed to the following:

$$\tilde{R}^T_{m_n} r^{m_n} = 0, \;\; \tilde{R}^T_{m_n} A P_{m_n} = 0. \tag{27}$$

Remark 6. It is easy to check the equalities

$$(C_1^{(n)})^2 \equiv (B^{-1}_{m_n,0}A)^2 = C_1^{(n)}, \;\; (C_2^{(n)})^2 \equiv (I - B^{-1}_{m_n,0}A)^2 = C_2^{(n)},$$

which mean that the matrices $C_1^{(n)}$ and $C_2^{(n)}$ are projectors.

To provide orthogonalization conditions (27) not only in restart points, but at the every iteration, the formulas (17) for direction vectors in MPSCD for $m \neq m_n$ should be changed to the following:

$$\begin{array}{l} p_l^{m+1} = \tilde{B}^{-1}_{m+1,l} r^{m+1} - \sum\limits_{k=0}^{m} \sum\limits_{l=1}^{M_k} \beta^{(\gamma)}_{m,k,l} p_l^k, \;\; \tilde{B}^{-1}_{m+1,l} = (I - B^{-1}_{m+1,0}A)B^{-1}_{m+1,l}, \\ l = 1, ..., M_m, \;\; m = m_{n-1}, \;\; m_{n-1}+1, ..., m_n - 1. \end{array} \tag{28}$$

It is easy to check by induction, that the following relations are valid in this case:

$$\tilde{R}^T_m r^m = 0, \;\; \tilde{R}^T_m A P_m = 0, \;\; m = 0, 1, ... \tag{29}$$

But here it is necessary to have in mind, that relations (28), which differ from (17) just in the form of preconditioning matrices ($\tilde{B}^{-1}_{m+1,l}$ instead of $B^{-1}_{m+1,l}$), demand to change orthogonality conditions for recurrent coefficients $\beta^{(\gamma)}_{m,k,l}$: the matrices $B^{-1}_{m+1,l}$ in (18)–(19) should be changed to $\tilde{B}^{-1}_{m+1,l}$.

It is known from the theoretical estimations and from numerical expirements, that the restarts brought the reducing Krylov subspaces, and as consequence, the decreasing convergence rate of iterative processes. But it is inevitable cost for the memory saving for all algorithms with short recursions.

In the considered MPSCD methods with deflation restarts, it is natural to take the corresponding iteration numbers m_n through uniform intervals, i.e. $m_n = m_{n-1} + æ$, with some apriori integer $æ$. At the other hand, it is possible to stay the question about optimal control of the values m_n, using some aposteriori information.

Remark 7. As it is mentioned in Remark 4, for each sequence of Krylov subspaces $\mathcal{K}_{n+1}$, which dimensions increase with growing n, the set of shrinking subspaces is corresponding, and they present the orthogonal complement $\mathcal{K}^{\perp}_{n+1}$ in full space $\mathcal{R}^N$. It is easy to see, that modifications of the Krylov subspaces, which conclude in additional orthogonalization of the computed vectors to the "test" subspace $\mathcal{S}$, do not change the duality relations. The orthogonality conditions (25), (27), and (29) allow to define for MPSCD the subspaces with reducing dimensions, which are similar to $\mathcal{G}_n$ in IDR methods.

In total, the active discussions on the various approaches of the optimizing the iterative processes confirm the perspectivity of the developing methods in Krylov subspaces.

The alternative way to save memory and to shorten recursions is the approach with bounded orthogonality, in the sense, that in constructing iterations for the solving non-symmetric SLAEs, only several last direction vectors in Krylov subspaces are saved and used in recursions. If the number of such vectors is constant and equals $q+1$, then formulas (28) should be changed to following:

$$p_l^{m+1} = \tilde{B}_{m+1,l}^{-1} r^{m+1} - \sum_{k=m-q}^{m} \sum_{l=1}^{M_k} \beta_{m,k,l}^{(\gamma)} p_l^k, \quad m = 0, 1, ... \tag{30}$$

If the vectors p_l^0 are computed by formulas (24) in this case, then orthogonality properties (29) are valid also. The realization of additional orthogonality in this approach can be done by deflation or aggregation method.

4. CONCLUSION

The goal of this paper consists, in some sense, in dethroning the myth about exclusivity of induced dimension reduction methods, which are opposed by misunderstanding to classical iterative processes in Krylov subspaces. On the example of proposed multi-preconditioned semi-conjugate residual methods, it is shown that modified Krylov type algorithms with using additional orthogonalization (deflation, aggregation, etc.) can be associated with the subspaces under decreasing dimensions, which are called by some authors as Sonneveld subspaces. The subspaces of "shadow residuals" in IDR methods and deflation subspaces in "classical" Krylov methods have the close analogy, which did not mention by any authors of various researches. Moreover, it is suprisingly, that publications on these two considered approaches do not include the cross references, although we have the closed areas of computational algebra.

The paper is supported by the Russian Scientific Foundation grant N 15-11-10024, and by RFBR grant N 16-29-15122.

REFERENCES

1. Wesseling P., Sonneveld P. Numerical experiments with a multiple grid and a preconditioned Lanczos type method. Lecture Notes in Math., vol. 771, Springer-Verlag, Berlin, 1980, 543-562.

2. Sonneveld P., Van Gijzen M.B. IDR(s): a family of simple and fast algorithms for solving large nonsymmetric systems of linear equations. SIAM J. Sc. Comput, 31 (2), 2008/09, 1035-1062.

3. Gutkenecht M.H. IDR explained. Electron. Trans. Numer. Anal., 36, 2009/10, 126-148.

4. Onoue Y., Fujino S, Nakashima N. An overview of a family of new iterative methods based on IDR theorem and its estimation, in: Proceedings of the International Multi-Conference of Engineers and Computer Scientists 2009, vol. 11, IMECS 2009, Hong Kong, 2009, 2129-2134.

5. Simoncini V., Szyld D.B. Interpreting IDR as a Petrov–Galerkin method.–SIAM J. Sci. Comput., 32(4), 2010, 1898-1912.

6. Tanio M., Sugihara M. GBi-CGSTAB(s,l): IDR(s) with higher-order stabilization polynomials.– J.Comput. Appl. Math., 235(3), 2010, 765-784.

7. Sleijpen G.L., Sonneveld P., Van Gijzen M.B. Bi-CGStab as an induced dimension reduction method.–Appl. Numer. Math., 60(11), 2010, 1100-1114.

8. Sleijpen G.L., Van Gijzen M.B. Exploiting BiCGStab(l) strategies to induce dimension reduction.–SIAM J. Sci. Comput., 32(5), 2010, 2687-2709.

9. Sonneveld P. On the convergence behavior of IDR(s) and related methods.–SIAM J. Sci. Comput., 34(5), 2012, 2576-2598.

10. Saad Y. Iterative Methods for Sparse Linear Systems (Second edition). Society for Industrial and Applied Mathematics: Philadelphia, PA, 2003.

11. Sleijpen G.L., Fokkema D.R. BiCGStab(l) for linear equations involving unsymmetric matrices with complex spectrum.–Electron. Trans. Numer. Anal., 1, 1993, 11-32.

12. Il'in V.P., Biconjugate direction methods in Krylov subspaces. // J. Appl. Indust. Math. Vol. 4, N. 1, 2010, 65-78.

13. Il'in V.P. Methods and Technologies of Finite Elements. Izd. IVMiMG SO RAN,Novosibirsk, in Russian), 2007.

14. Van Gijzen M.B., Sleijpen G.I., Zemke J.-P. Flexible and multi-shift induced dimension reduction algorithms for solving large sparse linear systems. – Numer. Linear Algebra Appl., v. 22, 2014, 1-25.

15. Van Gijzen M.B., Sonneveld P. Algorithm 913: an elegant IDR(s) variant that efficiently exploits biorthogonality properties.– ACM Trans.Math. Software 38(1), 2011, 5.1-5.19.

16. Rendel O., Rizvanolli A., Zemke J.-P. IDR: A new generation of Krylov subspace methods.– Linear Algebra and its Applications, 439(4), 2013, 1040-1061.

17. Astudillo R., Van Gijzen M.B. A restarted Induced Dimension Reduction method to approximate eigenpairs of large unsymmetric matrices. –J. of Comput. Appl. Math., v. 296, 2016, 24-35.

18. Il'in V.P. Methods of semiconjugate directions.– Russ. J. Numer. Anal.Math. Modelling, v.23, N 4, 2008, 369-387.

19. Il'in V.P., Itskovich E.A. On the semi-conjugate direction methods with dynamic preconditioning.//J.Appl. Ind. Math., v. 3, N 2, 2009, 222-233.

20. Eisenstat S.C., Elman H.C., Schultz M.H. Variational iterative methods for nonsymmetric systems of linear equations.//SIAM J. Numer. Anal., v. 20, N 3, 1983, 345-357.

21. Yuan J.Y., Golub G.H., Plemmons R.J., Cecilio W.A. Semi-conjugate direction methods for real positive definite systems//BIT, v. 44, N 1, 2004, 189-207.

22. Nicolaides R. Deflation of conjugate gradients with applications to boundary value problems.//SIAM J. Numer. Anal., 24, 1987, 355-365.

23. Kolotilina L.Y. Twofold deflation preconditioning of linear algebraic systems. I. Theory. –Journal of Mathematical Sciences, 89: 1998, 1652-1689.

24. Coulaud O., Giraud L., Ramet P., Vasseur X. Deflation and augmentation techniques in Krylov subspace methods for the solution of linear systems. – Inria Bordeaux-Sud Ouest, France, Research Report N 8265, 2013.

25. Gutknecht M.H. Deflated and Augmented Krylov Subspace Methods: A Framework for Deflated BICG and Related Solvers.//SIAM J. Matrix Anal.& Appl., v. 35, ISS. 4, 2014, 1444-1466.

26. Gurieva Y.L., Il'in V.P. Parallel approaches and technologies of domain decomposition methods.–Journal of Mathematical Sciences, v. 207, N 5, 2015, 724-735.

27. Il'in V.P. Problems of parallel solution of large systems of linear algebraic equations .–J. of Mathem. Sci., v. 216, N 6, 2016, 795-804.

28. Bridson R., Greif C. A multipreconditioned conjugate gradient algorithm.–SIAM J. Matrix Anal. Appl., v.27, N 4, 2006, 1056-1068.

29. Lawson C.L., Hanson R.J. Solving Least Squares Problems.–Prentice–Hall, Inc., N.J., 1974.