CSMA 2017
13ème Colloque National en Calcul des Structures
15-19 Mai 2017, Presqu'île de Giens (Var)

Limited memory preconditioners for nonsymmetric systems

Sylvain Mercier[1], Serge Gratton[2], Nicolas Tardieu[1], Xavier Vasseur[3]

[1] *IMSIA, UMR 9219 EDF-CNRS-CEA-ENSTA, Saclay*
[2] *INPT et ENSEEIHT, Toulouse*
[3] *ISAE-SUPAERO, Toulouse*

Résumé —

This paper presents a class of limited memory preconditioners (LMPs) for solving linear systems of equations with multiple nonsymmetric matrices and multiple right-hand sides. These preconditioners based on limited memory quasi-Newton formulas require a small number k of linearly independent vectors. They may be used to improve an existing first-level preconditioner and are especially worth considering when the solution of a sequence of linear systems with slowly varying left-hand sides is addressed.

Mots clés — preconditioning, Krylov methods.

1 Introduction

The numerical solution of sequences of linear algebraic systems is frequently required in many applications in computational science and engineering. For small to medium-scale problems, algorithms related to (sparse or dense) direct methods based on Gaussian elimination are usually employed. When the coefficient matrix is fixed, these methods are especially relevant since the factorization can be performed once for all and reused all along the sequence. In the general case where both the left-hand and right-hand sides are changing, preconditioned Krylov subspace methods are the methods of choice for large-scale problems. Indeed the operators in subsequent linear systems have most often similar spectral properties. Hence, a first possible approach to design efficient numerical methods is to extract information generated during the solution of a given linear system to improve the convergence rate of the Krylov subspace method during the subsequent solutions. Deflated and augmented Krylov subspaces [1, 2, 3, 4] or Krylov subspace methods with recycling [5, 6, 7, 8] have been proposed in this setting. We refer the reader to [9, 10, 11, 12] for a comprehensive theoretical overview on these methods and to references therein for a summary of applications, where the relevance of these methods has been shown. An alternative consists in exploiting information generated during the solution of a given linear system to improve a preconditioner when solving the next linear system in the sequence. This is the main subject that we want to address in this paper.

When the coefficient matrices in the sequence are symmetric positive definite, Morales and Nocedal [13] have proposed a preconditioner which has the form of a limited memory quasi-Newton matrix. More recently, Gratton, Sartenaer and Tshimanga [14] have defined a class of limited memory preconditioners (LMP) based on limited memory quasi-Newton formulas that ensures good spectral properties of the preconditioned matrix. This family can be seen as a block variant of the BFGS updating formula for quadratic problems [15, 16].

In [17], an extension of this class of limited memory preconditioners for the solution of linear systems with symmetric indefinite matrices and multiple right-hand sides has been proposed. The spectrum of the preconditioned operator has been characterized and it has been shown that the eigenvalues of the preconditioned operator are real-valued (with at least k eigenvalues equal to 1). Nevertheless, in some cases, this approach cannot be applied. It indeed implies the use of symmetric matrices for both the linear system to be solved and the first-level preconditionner. This limitation on the first-level preconditionner is strong since it forbids the use of right-only or left-only preconditioning, that are the most-commonly implemented in numerical linear algebra libraries. Furthermore, many modelling features in solid mechanics such as large deformation or follower forces lead to the definition of a tangent stiffness matrix that is no longer symmetric. Thus, our objective in this paper is to propose an extension of the limited memory preconditioners to be used when the coefficient matrices are nonsymmetric, to show their efficiency by both a theoretical study of the spectrum of the preconditionned operator and real-life engineering applications.

2 Limited memory preconditioners for nonsymmetric matrices

2.1 Problem setting

We address the solution of a given linear system $Ax = b$ with $A \in \mathbb{R}^{N\times N}$ being nonsingular, $x \in \mathbb{R}^N$ and $b \in \mathbb{R}^N$, respectively. In this setting, our main interest will be to analyze the class of limited memory preconditioners defined next.

Definition 2.1. *Let $A \in \mathbb{R}^{N\times N}$ be a general nonsingular matrix and $S \in \mathbb{R}^{N\times k}$ of full rank k, with $k \leq N$. The matrix $H \in \mathbb{R}^{N\times N}$ defined by*

$$H = (I_N - AS(S^TA^TAS)^{-1}S^TA^T) + S(S^TA^TAS)^{-1}S^TA^T \qquad (1)$$

is called the limited memory preconditioner in the nonsymmetric case (LMP$_{ns}$).

The motivation behind this definition can be briefly explained. In numerical optimization, the application of certain quasi-Newton methods to the numerical solution of $F(x) = 0$ with $F(x) = Ax - b$ provides an approximation of the inverse of the Jacobian (here A^{-1}). Hence, we aim at using this approximation as a candidate preconditioner for Krylov subspace methods when solving $Ax = b$. This idea has been notably exploited in [18, 19], where the construction of the preconditioner was based on rank- 1 updates issued from either the Broyden's method [20] or the Eirola and Nevanlinna's method [21]. The relation (1) in Definition (2.1) corresponds to a block rank-k extension of one of these updates and has been derived in [[22], Proposition 3.1.1] ; see also [[22], Section 3.1] for further details. Here, our main purpose is to study and investigate the potential of the LMP matrix H when considered as a preconditioner of A.

To conclude this section, we shall underline the particular case of $k = N$ where it can be shown that $H = A^{-1}$ [22]. Obviously, it is of no practical use, since it would necessitate to deal with the inverse of the matrix $S^TA^TAS \in \mathbb{R}^{N\times N}$. Nevertheless, it shows that H corresponds to A^{-1} in such a case and strengthens the idea to use this class as a preconditioner to improve the convergence rate of the GMRES method.

2.2 Spectral characterization of the preconditioned matrix

The main contribution of this section concerns the characterization of the spectrum of AH, where H is defined by relation (1). Based on the analysis of the LMP$_{ns}$ as an orthogonal projection onto $(A\mathcal{S})^\perp$ [22], where $\mathcal{S}$ denotes the range of S, the following theorem can be shown :

Theorem 2.1. *Let $A \in \mathbb{R}^{N\times N}$ be a general nonsingular matrix and H defined by relation* (1). *Let the columns of $W \in \mathbb{R}^{N\times k}$ and $W_\perp \in \mathbb{R}^{N\times(N-k)}$ form an orthonormal basis for $A\mathcal{S}$ and $(A\mathcal{S})^\perp$, respectively. The spectrum of AH is then given by*

$$\Lambda(AH) = \{1\} \cup \Lambda(W_\perp^T A W_\perp).$$

We deduce from this theorem that 1 is an eigenvalue of AH of multiplicity at least k. The remaining ones are characterized by the spectrum of $W_\perp^T A W_\perp$, revealing the theoretical interest of choosing $\mathcal{S}$ as an A-invariant subspace. The following corollary gives a result in this sense.

Corollary 2.2. *Assume that $A \in \mathbb{R}^{N\times N}$ is nonsingular with $\Lambda(A) = \{\lambda_1, ..., \lambda_N\} \in \mathbb{C}^N$. Let us set $S \in \mathbb{R}^{N\times k}$ where $\mathcal{S}$ spans the eigenspace associated to $\{\lambda_1, ..., \lambda_k\}$ which contains some eigenvalues and their conjugate. Then*

$$\Lambda(AH) = \{1, ...1, \lambda_{k+1}, ..., \lambda_N\}.$$

Unlike the situation of Corollary 2.2, only approximations to invariant subspaces can be usually used in practice, as illustrated in [17]. The strategy to select S in the numerical experiments will be based either on Ritz vectors. This is also used in some existing methods based for instance on deflation, which is an acceleration technique for the solution of linear system [3]. LMP$_{ns}$ and deflation enjoy strong connections and can exhibit similar convergence [22] ; we shall illustrate in the sequel the interest of our approach. We briefly recall that Ritz vectors are approximations of eigenvectors that are particularly useful when considering the solution of large-scale linear systems or eigenproblems ; see, e.g., [23, Section 5.7.1], [24] and the references therein. For the sake of clarity, we recall their definition next.

Definition 2.2. *A scalar θ is called a Ritz value of A with respect to a subspace $\mathcal{L}$ if there exists a nonzero vector $z \in \mathcal{L}$, called a Ritz vector, such that $Az - \theta z \perp \mathcal{L}$, where orthogonality is considered with respect to the canonical inner product.*

Ritz pairs (θ, z) can be cheaply obtained from the Lanczos process and we refer the reader to [25] for further details related to their computation.

3 Numerical experiments

The purpose of this section is to illustrate the efficiency of the class of LMP$_{ns}$ on systems of saddle point structure in solid mechanics. The method has been implemented in Code_Aster via the PETSc library (version 3.4.5).

3.1 Sequence of preconditioned saddle point systems

We consider, in the next numerical experiments, mechanical problems leading to sequences of symmetric indefinite linear systems, already defined before in this paper. The sequences are of the form

$$\mathcal{K}_i\, y_i = c_i \Longleftrightarrow \begin{pmatrix} G_i & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} u_i \\ \lambda_i \end{pmatrix} = \begin{pmatrix} f_i \\ g_i \end{pmatrix}, \quad i = 1, \cdots, I. \tag{2}$$

Contrary to the study led in [17], the class of LMP$_{ns}$ is adapted to solve nonsymmetric problems and we can solve (2) using any first level preconditioner. Two such techniques are used next :

— **A block upper triangular preconditioner**
In Section 3.2, we consider the specific block upper triangular preconditioner based on the augmented Lagrangian method :

$$\mathcal{M}_t = \begin{pmatrix} G_1 + \gamma B^T B & 2B^T \\ 0 & -\dfrac{1}{\gamma} I_m \end{pmatrix}, \quad \gamma > 0. \tag{3}$$

$\mathcal{M}_t$ is computed from the first matrix $\mathcal{K}_1$ in (2) and is fixed all along the sequence. Since inverting exactly $G_1 + \gamma B^T B$ is too demanding in terms of both computational operations and memory requirements for large-scale problems, we consider a factorized approximate preconditioner of the form $\mathcal{M}_1 \approx \mathcal{L}\mathcal{L}^T$ based on the incomplete Cholesky factorization of $G_1 + \gamma B^T B$ written as $G_1 + \gamma B^T B \approx LL^T$ [26]. Finally, the associated matrix is used as a left preconditioner [1] and we obtain the preconditioned linear systems $\mathcal{A}_i x_i = b_i$, denoted as

$$\mathcal{A}_i\, x_i = b_i \Longleftrightarrow \begin{pmatrix} LL^T & 2B^T \\ 0 & -\frac{1}{\gamma} I_m \end{pmatrix}^{-1} \begin{pmatrix} G_i & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} u_i \\ \lambda_i \end{pmatrix} = \begin{pmatrix} LL^T & 2B^T \\ 0 & -\frac{1}{\gamma} I_m \end{pmatrix}^{-1} \begin{pmatrix} f_i \\ g_i \end{pmatrix} \quad i = 1, \cdots, I. \tag{4}$$

In practice, the preconditioner is applied using the relation

$$\begin{pmatrix} LL^T & 2B^T \\ 0 & -\frac{1}{\gamma} I_m \end{pmatrix}^{-1} = \begin{pmatrix} L^{-T}L^{-1} & 0 \\ 0 & -\gamma I_m \end{pmatrix} \begin{pmatrix} I_n & 2\gamma B^T \\ 0 & I_m \end{pmatrix}.$$

— **Factorization in single precision arithmetic**
In Section 3.3, we consider another first level preconditioner. In these experiments, we use the default preconditioning technique in Code_Aster, named *LDLT_SP* [27]. This method consists in

[1] To be in agreement with all preconditioning strategies in the host finite element software Code_Aster.

computing a complete factorization of the saddle point matrix $\mathcal{K}_1$ in single precision arithmetic using the MUMPS library. We use the computed matrix, denoted by $\mathcal{M}_{sp}$, as a left preconditioner which leads to the sequence of preconditioned linear systems

$$\mathcal{A}_i\, x_i = b_i \Longleftrightarrow \mathcal{M}_{sp}^{-1}\mathcal{K}_i x_i = \mathcal{M}_{sp}^{-1} b_i \quad i = 1, \cdots, I. \tag{5}$$

$\mathcal{M}_{sp}^{-1}$ is applied by successive back and forward substitutions and, unless otherwise stated, $\mathcal{M}_{sp}$ is fixed all along the sequence (5).

GMRES(30) combined with LMP$_{ns}$ as a right preconditioner will be used to solve the sequence of nonsymmetric linear systems (4) or (5). To select $S \in \mathbb{R}^{N\times k}$, we extract k Ritz vectors corresponding to the smallest in modulus Ritz values, when solving the first linear system of the sequence. The limited memory preconditioner H_{ns} is then defined for all the remaining linear systems as

$$H_{ns} = I_{n+m} - \mathcal{A}_1 S(S^T\mathcal{A}_1^T\mathcal{A}_1 S)^{-1}S^T\mathcal{A}_1^T + S(S^T\mathcal{A}_1^T\mathcal{A}_1 S)^{-1}S^T\mathcal{A}_1^T. \tag{6}$$

To summarize, we use GMRES(30) to solve :

$$\mathcal{A}_1 x_1 = b_1 \quad \text{and} \quad \begin{cases} \mathcal{A}_i H_{ns}\tilde{x}_i = b_i \\ x_i = H_{ns}\tilde{x}_i \end{cases} \quad i = 2, \cdots, I.$$

3.2 Containment building of a nuclear reactor

In this section, we investigate the mechanical properties of the containment building of a nuclear reactor of a Pressurized Water Reactor power plant. This building protects both the reactor from external aggressions and the environment if an internal accident occurs. Robust and accurate numerical simulations are thus required for both design and safety analysis. We consider an advanced mechanical modeling that takes into account numerous prestressing tendons, whose role is to improve the global stiffness of the structure (see Figure 1). The containment building is subject to gravity and to an internal pressure. The whole loading is gradually applied into 4 successive steps. Each pitch of loading then corresponds to a specific linear system in the sequence, where only the right-hand side has changed (i.e. $\mathcal{A}_1 = ... = \mathcal{A}_4$). The introduction of Lagrange multipliers stems from the imposition of kinematic relations modeling perfect adhesion between the prestressing tendons and the concrete [28] and to the dualization of the essential boundary conditions. In this setting, B admits either five or one nonzero entries per row, respectively. This study is known to be complex for different reasons. First, from a mechanical point of view, the modeling is rather advanced with a mixing of three-dimensional elements for the concrete, two-dimensional elements for the metal shell placed on the intern wall of the building (to insure the sealing if an accidental leak occurs), and of one-dimensional elements for the tendons. Moreover, since the prestressing tendons are attached to the concrete thanks to dualized linear relations, the number of Lagrange multipliers is really important ($m = 158928$ for a global size of $N = 442725$). The number of nonzero entries of G_1 and $G_1 + \gamma B^T B$ is 7079238 and 8343685, respectively. Secondly, the occurrence of a large number of prestressing tendons (more than 600 here) induces a nullspace of large dimension for the stiffness matrix. Actually, it is known that this dimension is related to the number of rigid body motions of the subbodies of materials contained within the finite element mesh [29]. This numerical study is thus challenging and serves as a relevant realistic test case in solid mechanics to investigate the efficiency of preconditioners for Krylov subspace methods.

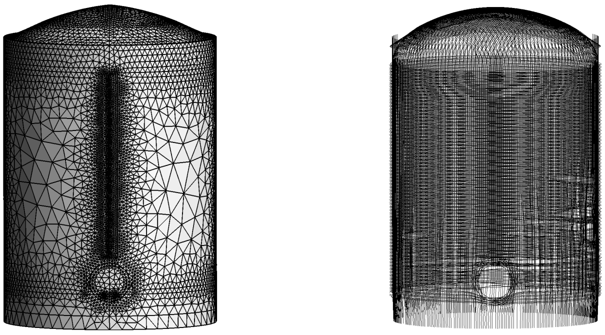

FIGURE 1 – Containment building : three-dimensional mesh (left part) and location of the prestressing tendons on the surface (right part).

In this experiment, we set γ to $2,4684 \times 10^{11}$ and consider a level of fill equal to 8 in the incomplete Cholesky factorization of the $(1,1)$ block of $\mathcal{M}_d$ with the use of the block triangular augmented Lagrangian first level preconditioner. Actually, with a lower level of fill the preconditioned Krylov subspace method can hardly converge. However, even with this value of fill, the required memory is around 7 Go, while state-of-the-art sparse direct solvers require at least 10 Go for the complete factorization of the $(1,1)$ block of $\mathcal{M}_d$.

Figure 2 illustrates the convergence history of GMRES(30) for two linear systems in the sequence $(i = 2,3)$, with or without LMP$_{ns}$. In this experiment, we consider also LMP$_{ns}$ with a varying number of Ritz vectors ($k = 5$, 20, 30). In addition, we show in Table 1 the cumulative iteration count over the last three linear systems, the CPU time and the memory requirements provided by PETSc, respectively. We note that the smallest number of iterations is always obtained when selecting a large value of Ritz vectors ($k = 30$). More precisely, in this case, the iteration count decrease is equal to 53%, associated to a gain in terms of CPU time of 51%. This satisfactory result comes at a price of a low increase in memory requirements (4%).

	No LMP$_{ns}$	LMP$_{ns}$, $k = 5$	LMP$_{ns}$, $k = 20$	LMP$_{ns}$, $k = 30$
Total iteration count	135	74	66	64
Iteration count decrease (%)	×	45	51	**53**
CPU time (sec)	75.6	42.9	40.2	36.9
CPU time decrease (%)	×	43	47	**51**
Memory (Mo)	6895	6950	7071	7172
Memory increase (%)	×	**0.8**	2.6	4

TABLE 1 – Containment building : cumulative iteration count for the last three linear systems in the sequence, CPU time and memory requirements for different limited memory preconditioners. Case of $k = 5$, 20 or 30 Ritz vectors.

The matrix being fixed in this numerical experiments, we next show the action of the LMP$_{ns}$ on sequences arising from Newton's method with slowly varying matrices. From now on, we combine the second level improving techniques with the *LDLT_SP* preconditioner, which is the default method in Code_Aster.

3.3 Shut down nuclear reactor cooling loop

This problem models the thermal fatigue of a shut down nuclear reactor cooling loop. The heat from the nuclear core is extracted by circulation of pressurized water in the primary circuit. When the reactor is shut down, the studied cooling loop allows to evacuate the heat of the primary circuit and the residual power of the nuclear fuel, using cold water. Some oscillations of the temperature can occur where both cold and hot water are in contact and it is necessary to model the resistance of the component to the thermal fatigue. The mesh of the studied material is illustrated in Figure 3.

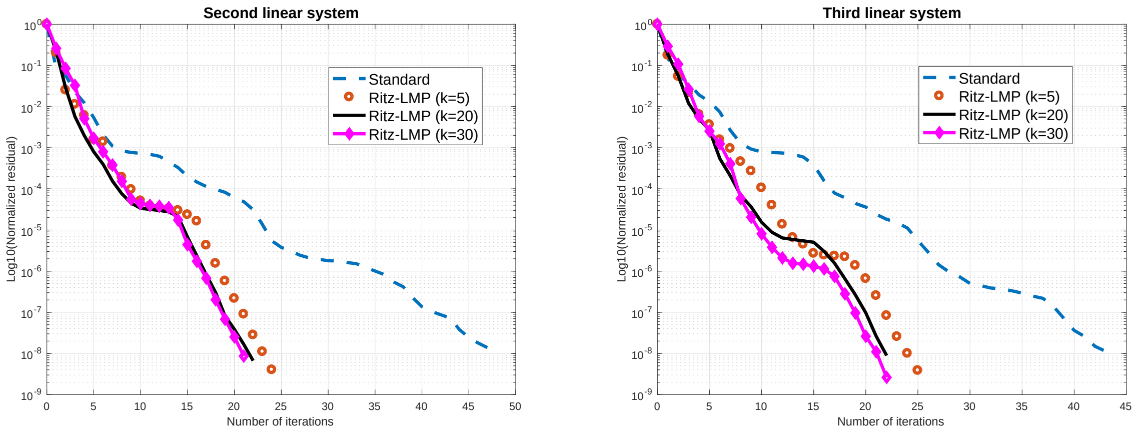

FIGURE 2 – Containment building : convergence history of preconditioned GMRES(30) for two linear systems in the sequence. Case of limited memory preconditioners with $k = 5$, 20 or 30 Ritz vectors.

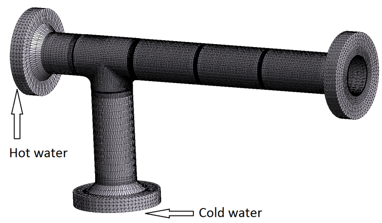

FIGURE 3 – Mesh of the shut down nuclear reactor cooling loop.

The discretized problem involves $N = 571281$ unknowns, with less than 1% of Lagrangian multipliers, corresponding to the essential dualized boundary conditions. Newton's method is employed because of the nonlinearity of the constitutive law of the structure (Chaboche elasto-visco-plastic law [30]) and leads to the sequence (5) of 67 linear systems. This solution is challenging, not only from its dimension, but also from the condition number of the saddle point matrices $\mathcal{K}_i$ which is of order of 10^{10} [2]. Indeed, it is known that this value can be notably related to the properties of the mesh ; in particular, when there are strong spatial variations of the size of the mesh elements, or when some of them are flattened, the condition number increases. The former case arises in this problem, at the intersection of both pipes (see Figure 3). Generally, when the condition number is moderate, the GMRES(30) method, preconditioned by the *LDLT_SP* technique, requires very few iterations to converge. With a large condition number, this iteration count is often larger (about 20 in this case), and we expect that a second level preconditioning technique, such as the LMP$_{ns}$, decreases this count.

We want to compare here the statistics related to the solution of the complete Newton's sequence, also including the *LDLT_SP* factorization step. Table 2 collects the results obtained with three different solution methods : without any second level improving technique, using the LMP$_{ns}$ or the deflation for only $k = 5$. Indeed, in practice, the value $k = 5$ seems to be the better choice, since the number of GMRES(30) iterations with the first level preconditioner remains usually relatively low (i.e. less than 30) and a larger k does not imply a significant additional gain. We first note that the use of either the limited memory preconditioner or the deflation method is very efficient in terms of iteration count with a gain of 53.3% and 58.3%, respectively. Concerning the CPU time, the decrease is largely in favour of the improvement preconditioning technique with a gain of 39.9%, against 9.2% for the deflation. This phenomenon can be explained by the fact that the deflation operator needs to be updated at each new linear system, using the new matrix $\mathcal{A}_i$. In addition, the value 410 in Table 2 corresponds to the cumulative iteration count of the deflated GMRES(30) to solve the 67 linear systems, which means that each system requires in average about 6 iterations. Hence, the computational cost related to the update of the deflation operators is not negligible.

[2] This estimate has been obtained using MUMPS in double precision arithmetics as a direct solver on the first linear system.

	No LMP$_{ns}$	LMP$_{ns}$, $k=5$	Deflation, $k=5$
Total iteration count	983	460	410
Iteration count decrease (%)	×	53.3	**58.3**
CPU time (sec)	961	578	873
CPU time decrease (%)	×	**39.9**	9.2
Memory (Mo)	14074	14117	14095
Memory increase (%)	×	0.03	**0.014**

TABLE 2 – Shut down nuclear reactor cooling loop : cumulative iteration count over the complete Newton's sequence, CPU time and memory requirements for limited memory preconditioner and deflation with $k=5$ Ritz vectors.

4 Conclusions

In this paper, we have proposed a class of limited memory preconditioners adapted to the solution of nonsymmetric linear systems. This preconditioner can be interpreted as a block generalization of update formulas proposed in both EN and Broyden optimization methods.

We have recalled that the spectrum of the preconditioned operator contains the eigenvalue 1 (with multiplicity at least k) and the remaining part of the spectrum has been characterized.

The numerical part of this work deals with solid mechanics problems leading to sequences of symmetric saddle point linear systems. Here, any first level preconditioner can be used here, making the successive preconditioned operators nonsymmetric. Numerical experiments emphasize the efficiency of the class of LMP$_{ns}$. On the one hand, using a block upper triangular first level preconditioner, a saving of up to 51% in terms of computational time is obtained on a large-scale application. On the other hand, if we define the first level preconditioner as a complete factorization in single precision arithmetic with MUMPS, we obtain a gain of 39.9% on a nuclear safety analysis corresponding to the study of a shut down reactor cooling loop. Simultaneously, we have compared the action of the LMP$_{ns}$ with the action of the deflated GMRES method. In terms of iteration count, the gains are close, even if the deflation method gives slightly better results on the large-scale systems illustrated here. However, when the left-hand sides change during the sequence, the deflation operators need to be updated, which makes it less competitive in terms of computational time.

Références

[1] A. Chapman and Y. Saad. Deflated and augmented Krylov subspace techniques. *Numer. Linear Algebra Appl.*, 4(1) :43–66, 1997.

[2] Z. Dostál. Conjugate gradient method with preconditioning by projector. *Int J. Comput. Math.*, 23(3) :315–323, 1988.

[3] R. A. Nicolaides. Deflation of conjugate gradients with applications to boundary value problems. *SIAM J. Numer. Anal.*, 24(2) :355–365, 1987.

[4] Y. Saad, M. Yeung, J. Erhel, and F. Guyomarc'h. A deflated version of the conjugate gradient algorithm. *SIAM J. Sci. Comput.*, 21(5) :1909–1926, 2000.

[5] M. Kilmer and E. de Sturler. Recycling subspace information for diffuse optical tomography. *SIAM J. Sci. Comput.*, 27(6) :2140–2166, 2006.

[6] M. Parks, E. de Sturler, G. Mackey, D. Johnson, and S. Maiti. Recycling Krylov subspaces for sequences of linear systems. *SIAM J. Sci. Comput.*, 28(5) :1651–1674, 2006.

[7] K. M. Soodhalter, D. B. Szyld, and F. Xue. Krylov subspace recycling for sequences of shifted linear systems. *Appl. Num. Math.*, 81 :105–118, 2014.

[8] S. Wang, E. de Sturler, and G. Paulino. Large-scale topology optimization using preconditioned Krylov subspace methods with recycling. *Internat. J. Numer. Methods Engrg.*, 69(12) :2441–2468, 2007.

[9] A. Gaul. *Recycling Krylov Subspace Methods for Sequences of Linear Systems : Analysis and Applications*. PhD thesis, TU Berlin, Germany, 2014.

[10] A. Gaul, M.H. Gutknecht, J. Liesen, and R. Nabben. A framework for deflated and augmented Krylov subspace methods. *SIAM J. Matrix Anal. Appl.*, 34(2) :495–518, 2013.

[11] M. H. Gutknecht. Deflated and augmented Krylov subspace methods : A framework for deflated BiCG and related solvers. *SIAM J. Matrix Anal. Appl.*, 35 :1444–1466, 2014.

[12] V. Simoncini and D.B. Szyld. Recent computational developments in Krylov subspace methods for linear systems. *Numer. Linear Algebra Appl.*, 14(1) :1–59, 2007.

[13] J. L. Morales and J. Nocedal. Automatic preconditioning by limited memory Quasi-Newton updating. *SIAM J. Opt.*, 10(4) :1079–1096, 2000.

[14] S. Gratton, A. Sartenaer, and J. Tshimanga. On a class of limited memory preconditioners for large scale linear systems with multiple right-hand sides. *SIAM J. Opt.*, 21(3) :912–935, 2011.

[15] J. Nocedal and S. J. Wright. *Numerical Optimization*. Springer Ser. Oper. Res., Springer-Verlag, Berlin, 1999.

[16] R. B. Schnabel. Quasi-Newton methods using multiple secant equations. Technical report, Technical report CU-CS-247-83, Department of Computer Sciences, University of Colorado, Boulder, CO, 1983.

[17] S. Gratton, S. Mercier, N. Tardieu, and X. Vasseur. Limited memory preconditioners for symmetric indefinite problems with application to structural mechanics. *Numer. Linear Algebra Appl.*, 23(5) :865–887, 2016.

[18] U. M. Yang. *A family of preconditioned iterative solvers for sparse linear systems*. PhD thesis, University of Illinois at Urbana-Champaign, 1995.

[19] U. M. Yang and K. A. Gallivan. A new family of preconditioned iterative solvers for nonsymmetric linear systems. *Appl. Num. Math.*, 19(3) :287–317, 1995.

[20] C. G. Broyden. A new method of solving nonlinear simultaneous equations. *The Computer Journal*, 12(1) :94–99, 1969.

[21] T. Eirola and O. Nevanlinna. Accelerating with rank-one updates. *Linear Algebra Appl.*, 121 :511–520, 1989.

[22] S. Mercier. *Fast nonlinear solvers in solid mechanics*. PhD thesis, University of Toulouse, http ://thesesups.ups-tlse.fr/3000/1/2015TOU30305.pdf, 2015.

[23] J. Liesen and Z. Strakoš. *Krylov Subspace Methods Principles and Analysis*. Oxford University Press, Oxford, UK, 2013.

[24] Y. Saad. *Numerical Methods for Large Eigenvalue Problems*. SIAM, Philadelphia, USA, 2011. Second edition.

[25] C. C. Paige, B. N. Parlett, and H. A. van der Vorst. Approximate solutions and eigenvalue bounds from Krylov subspaces. *Numer. Linear Algebra Appl.*, 2 :115–134, 1995.

[26] Y. Saad. *Iterative Methods for Sparse Linear Systems*. SIAM, Philadelphia, USA, 2003.

[27] N. Tardieu and E. Cheignon. A Newton–Krylov Method for Solid Mechanics. *European Journal of Computational Mechanics*, 21(3-6) :374–384, 2012.

[28] S. Michel-Ponnelle. Modeling of the pretension cables. Technical report, EDF R&D, AMA, 2013. R7.01.02.

[29] J. Nečas and I. Hlavácek. *Mathematical Theory of Elastic and Elastoplastic Bodies*. Elsevier, 1980.

[30] J. M. Proix. Relations de comportement élasto-visco-plastique de Chaboche. Technical Report, EDF R&D, AMA, 2013. R5.03.04.