 ICCES, vol.9, no.2, pp.99-105

Preconditioners for finite element consolidation

G. Gambolati[1], M. Ferronato[1], C. Janna[1]

Summary

The Finite Element (FE) solution to consolidation equations in large geological settings raises a few numerical issues depending on the actual process addressed by the analysis. There are two basic problems where the solver efficiency plays a crucial role: 1- fully coupled consolidation, and 2- non-linear faulted (uncoupled) consolidation. Using a proper nodal numbering the FE matrices exhibit a block (or multilevel) structure. Krylov subspace solvers are attracting a growing attention, provided that a relatively inexpensive and effective preconditioner is available. For both problems possible preconditioners include the Diagonal Scaling (DS), the Incomplete Triangular Factorization (ILU), the Mixed Constraint Preconditioning (MCP) and the Multilevel Incomplete Factorization (MIF). The present communication provides a review and a critical discussion of DS, ILU, MCP and MIF when used as preconditioners for the numerical solution of a consolidation model.

Introduction

Krylov subspace, or Conjugate Gradient (CG)-like, methods are attracting a growing interest for the iterative solution to large sets of linear equations that generally arise from the numerical integration of partial differential equations [1]. A key-factor, however, for a successful implementation of any Krylov method is the availability of a relatively cheap and efficient preconditioner capable to accelerate properly the convergence to the wanted solution. An important class of problems where CG-like solvers are becoming more and more popular occurs in computational geomechanics, e.g. with the Finite Element (FE) solution to coupled or uncoupled consolidation of faulted porous media. In the former example the FE coefficient matrix is typically symmetric indefinite [2] while in the latter is symmetric positive definite (SPD) [3] with in both cases a block or multilevel structure provided that the FE nodes are numbered in a suitable order. Being the FE matrix either symmetric indefinite or SPD makes a big difference. In the first case the Symmetric Quasi-Minimal Residual (SQMR, [4]) should be used, while for a SPD problem the classical Preconditioned Conjugate Gradient (PCG) is much more attractive. Both SQMR and PCG have to be properly preconditioned to become competitive with direct solvers. Most traditional preconditioners advanced in computational geomechanics include the Diagonal Scaling (DS, [5]) and the Incomplete Triangular Factorization (ILU) with either zero [6] or partial controlled fill-in [7] possibly improved by an ad hoc preliminary scaling [8]. Recently novel preconditioners have been developed that take advantage of the particular block

[1]Dept. of Mathematical Methods and Models for Scientific Applications, University of Padova, Italy

or multilevel structure of the coefficient matrix. These are the Mixed Constraint Preconditioners (MCP, [3] [9]) and the Multilevel Incomplete Factorization (MIF, [10]). The present communication provides a brief review of DS, ILU, MCP and MIF as they are implemented into a FE consolidation model along with a discussion of their respective computational advantages/disadvantages.

The block (multilevel) consolidation problem

Linear coupled consolidation

Coupled consolidation describes the transient process that involves simultaneusly groundwater flow and solid skeleton deformation in saturated porous media and is typically solved in space by the FE method and in time by a finite difference scheme. Although the final coefficient matrix may take on a different form, i.e. symmetric indefinite, unsymmetric indefinite or unsymmetric positive definite [2], the former structure is to be generally preferred and is actually more often used. The corresponding block or two-level matrix $\mathscr{A}$ reads:

$$\mathscr{A} = \begin{bmatrix} K & B \\ B^T & -C \end{bmatrix} \tag{1}$$

where K is a classical elastic stiffnes SPD matrix, C is the classical subsurface flow SPD matrix and B is a rectangular matrix accounting for coupling between fluid flow and medium stress. Matrix C incorporates the time integration step Δt, hence it generally changes as Δt is small at the beginning of the simulation and progressively grows as the steady state is approached.

Non-linear uncoupled faulted consolidation

The FE modeling of geological faults requires the use of special Interface Elements (IE, [11]) with the non-overlapping condition addressed by a penalty formulation yielding an ill-conditioned stiffness matrix. In the simplest two-level case the FE-IE SPD matrix $\mathscr{A}$ takes on the following block form:

$$\mathscr{A} = \begin{bmatrix} K & B \\ B^T & C \end{bmatrix} \tag{2}$$

where K is similar to that in (1), C is a SPD penalty matrix and B is a rectangular matrix accounting for the connection between FE and IE. Matrix C contains all the entries from the nodes lying on the fault surfaces and its coefficients depend on the system solution as well as the structural block K in (2) if a non-linear constitutive law is used also for the FE. Typically we may think of a region where the FE deforms non-linearly with the remaining volume discretized into linear FE and the faults described by the highly non-linear IE. Altogether three levels can be identified, two related to linear and non-linear FE and one to the non-linear IE. If the

unknown displacements associated to the linear FE nodes are numbered first, those associated to the non-linear FE nodes second and those to the IE nodes last, the final FE-IE matrix has the following block form:

$$\mathscr{A} = \begin{bmatrix} K_1 & B_{11} & B_{12} \\ B_{11}^T & K_2 & B_2 \\ B_{12}^T & B_2^T & C \end{bmatrix} \tag{3}$$

Based on the above premise K_1, B_{11} and B_{12} in (3) are invariant during the simulation while K_2, B_2 and C depending on the system unknowns may change during the process. The size s_1 of K_1 is related to the number of linear FE nodes, the size s_2 of K_2 to the number of non-linear FE nodes and the size s_3 of C to the number of IE nodes. Although highly variable from application to application, just to give an idea, s_1 can be $2 \div 5$ times larger than s_2 that in turn might be $2 \div 5$ times larger than s_3.

Preconditioners

With reference to both matrix form (1) and (2), $\mathscr{A}$ can be factorized as follows:

$$\mathscr{A} = \mathscr{L}\,\mathscr{J}\,\mathscr{U} = \begin{bmatrix} I & 0 \\ B^T K^{-1} & I \end{bmatrix} \begin{bmatrix} K & 0 \\ 0 & S = \mp C - B^T K^{-1} B \end{bmatrix} \begin{bmatrix} I & K^{-1}B \\ 0 & I \end{bmatrix} \tag{4}$$

with sign - or + according to (1) and (2), respectively.

Diagonal Scaling (DS)

A diagonal substitute for $\mathscr{A}^{-1}$ can be generated by using the diagonal entries of $\mathscr{J}$ in (4). However, since the Schur complement S is too expensive to calculate, a further approximation is introduced using $\tilde{S} = \mp C - B^T [\mathrm{diag}(K)]^{-1} B$ instead of S. The diagonal preconditioner D^{-1}, labelled as Generalized Jacobi [5], for $\mathscr{A}$ thus becomes:

$$D^{-1} = \begin{bmatrix} \mathrm{diag}(K) & 0 \\ 0 & \varphi\,\mathrm{diag}(\tilde{S}) \end{bmatrix} \tag{5}$$

with φ a user specified parameter whose optimal value is to be found empirically.

Incomplete Triangular Factorization (ILU)

ILU is one of the oldest preconditioning strategy advanced by Kershaw [6] and later improved by a number of authors. In essence, $\mathscr{A}$ being symmetric, only one unit triangular factor is needed with the diagonal terms stored in a separate matrix to avoid the generation of imaginary coefficients. Following [7], a controlled partial fill-in can be allowed for to improve the preconditioner quality. The fill-in degree is determined by two user-specified parameters related to the largest number of new entries allowed for in the incomplete triangular factor (in excess of the original

non-zeroes of $\mathscr{A}$) and the absolute value of any triangular factor entry (that must be above a prescribed tolerance). To avoid instability problems Gambolati *et al.* [8] suggest to properly left and right diagonally scale matrix $\mathscr{A}$ before performing the ILU.

Mixed Constraint Preconditioner (MCP)

With reference to the exact form (4) K and S are replaced by $\hat{K}$ and $\hat{S}$, respectively, where $\hat{K} = L_K L_K^T$ and $\hat{S} = \mp L_S L_S^T$ with L_K and L_S the incomplete triangular factors with partial controlled fill-in of K and $\hat{S} = C \pm B^T Z Z^T B$, respectively, ZZ^T being the factorized approximate inverse (AINV) of K [12]. Thus the MCP reads:

$$\mathscr{M}^{-1} = \begin{bmatrix} L_K^{-T} & -L_K^{-T} R \\ 0 & L_S^{-T} \end{bmatrix} \begin{bmatrix} L_K^{-1} & 0 \\ \pm R^T L_K^{-1} & \mp L_S^{-1} \end{bmatrix} \tag{6}$$

where $R = L_K^{-1} B L_S^{-T}$. Equation (6) is obtained by inverting matrix $\mathscr{M}$ below:

$$\mathscr{A} \simeq \mathscr{M} = \begin{bmatrix} I & 0 \\ B^T L_K^{-T} L_K^{-1} & I \end{bmatrix} \begin{bmatrix} L_K L_K^T & 0 \\ 0 & \mp L_S L_S^T \end{bmatrix} \begin{bmatrix} I & L_K^{-T} L_K^{-1} B^T \\ 0 & I \end{bmatrix} \tag{7}$$

In the case of the faulted uncoupled consolidation matrix (2), the approximate Schur complement $\hat{S}$ is not guaranteed to be positive definite for any Z choice, with the PCG convergence consequently not theoretically ensured. To avoid this possible inconvenience we can define a new approximation for S which is always positive definite. This is readily achieved by dropping Z, taking $\hat{S} = C$ and using L_S as the exact Cholesky factor of C. Factorizing C in an exact way does not require a big effort due to the substantially 2D topology of C in a 3D medium. This new preconditioner can be formally written as $\mathscr{M}^{-1}$ of equation (6), with the understanding that L_S is the exact triangular factor of C instead of the incomplete triangular factor of $\hat{S}$.

Multilevel Incomplete Factorization (MIF)

MIF is particularly suited to problems having a multilevel structure, e.g. equation (3) with three levels or a similar matrix with more than three levels, for instance a 4-level coupled FE-IE model or a 5-level coupled thermo-flow-mechanical FE-IE model. Let us explain the basic MIF concept starting from equation (3). Write (3) in the following form:

$$\mathscr{A} = \begin{bmatrix} K_1 & B_1 \\ B_1^T & C_1 \end{bmatrix} \tag{8}$$

where the definitions of B_1 and C_1 follow immediately comparing (8) with (3). The partial incomplete factorization of (8) can be accomplished as:

$$\mathscr{A} \simeq \mathscr{P}_0 = \begin{bmatrix} L_1 & 0 \\ H_1 & I \end{bmatrix} \begin{bmatrix} D_1 & 0 \\ 0 & S_1 \end{bmatrix} \begin{bmatrix} L_1^T & H_1^T \\ 0 & I \end{bmatrix} = \mathscr{L}_1 \mathscr{D}_1 \mathscr{L}_1^T \tag{9}$$

where L_1 is the incomplete factor of K_1 with unit diagonal, D_1 is the corresponding diagonal matrix, H_1^T is a rectangular matrix approximating $D_1^{-1}L_1^{-1}B_1$, e.g. arising naturally from the factorization process of the first s_1 rows of $\mathscr{A}$, and $S_1 = C_1 - H_1 D_1 H_1^T$. As is well known PCG requires a product between the preconditioner $\mathscr{P}_0^{-1}$ and the residual r at the current iteration which is materially performed as the solution to the system $\mathscr{L}_1\mathscr{D}_1\mathscr{L}_1^T v = r$. Setting $\mathscr{L}_1^T v = y$, we get $\mathscr{L}_1\mathscr{D}_1 y = r$. Setting $\mathscr{D}_1 y = w$, we can write $\mathscr{L}_1 w = r$ from which w is easily derived. The intermediate vector y is computed inverting a block diagonal matrix. The first s_1 components are readily computed as $y_1 = D_1^{-1}w_1$. To solve $S_1 y_2 = w_2$ we recall that:

$$S_1 = C_1 - H_1 D_1 H_1^T = \begin{bmatrix} K_2 & B_2 \\ B_2^T & C \end{bmatrix} - H_1 D_1 H_1^T \tag{10}$$

S_1 is explicitly computed and can be written as:

$$S_1 = \begin{bmatrix} F_{11} & F_{12} \\ F_{12}^T & C_2 \end{bmatrix} \tag{11}$$

Now S_1 can be treated in the same way as equation (8), i.e. we perform the incomplete factorization of the F_{11} block and write the corresponding approximation $\mathscr{P}_1$ of S_1:

$$S_1 \simeq \mathscr{P}_1 = \begin{bmatrix} L_2 & 0 \\ H_2 & I \end{bmatrix} \begin{bmatrix} D_2 & 0 \\ 0 & S_2 \end{bmatrix} \begin{bmatrix} L_2^T & H_2^T \\ 0 & I \end{bmatrix} = \mathscr{L}_2\mathscr{D}_2\mathscr{L}_2^T \tag{12}$$

where $\mathscr{L}_2$, $\mathscr{D}_2$ and $\mathscr{L}_2^T$ have a similar meaning and structure as $\mathscr{L}_1$, $\mathscr{D}_1$ and $\mathscr{L}_1^T$ above and S_2 reads $C_2 - H_2^T D_2 H_2$. Again S_2 is explicitly calculated and the system $S_1 y_2 = w_2$ is solved approximately with y_2 obtained as $\mathscr{L}_2^{-T}\mathscr{D}_2^{-1}\mathscr{L}_2^{-1}w_2$. Again $\mathscr{D}_2^{-1}$ times the vector $\mathscr{L}_2^{-1}w_2$ is regarded as a system with the coeficient matrix $\mathscr{D}_2$. Again the first part of the second level unknowns are determined by solving a diagonal (D_2) system. The S_2 system can now be solved exactly if the S_2 size is small, or, more conveniently, in an approximate way by using the incomplete factorization of S_2. Once the new y is calculated we get v by solving $\mathscr{L}_1^T v = y$. At this point the application of the preconditioner to the residual vector in a PCG iteration is completed. The algorithm may be extended more or less easily to 4, 5, etc., levels.

Careful inspection of MCP and MIF reveals that actually the former may be regarded as a special case of MIF with two levels, where the block K of equation (2) is incompletely factorized, H_1 of equation (9) is equal to $(L_K L_K^T)^{-1}B$ and the (only one) Schur complement is computed with the aid of the approximate inverse (AINV) of K. On summary MCP may be thought of as a preconditioner belonging to the class of MIF preconditioners with a different (more complex) selection of the various approximations.

Conclusions

MCP is natively designed for a two level consolidation problem. Although it can be extended in principle to more than two levels the need for the calculation at each level of both the incomplete factorization and the approximate inverse of the upper left matrix block makes MCP quite cumbersome. In this respect MIF lends itself to an easier and more natural generalization to higher levels, having said that from a computational viewpoint MCP and MIF on two levels are more or less equivalent for a linear simulation. Of course the comparison of the performance of the different preconditioners is to be restricted between MCP and MIF on one side and the classical ILU on the other (DS is a poor preconditioner and except for very special cases it cannot successfully compete with MCP, MIF and ILU). In this respect both MCP and MIF turn out to be superior in the consolidation problems addressed in the present study to ILU implemented in the symmetric form by a factor that based on our recent experience is between two or three [2] [3] [9] [10].

References

1. Saad, Y., Van der Vorst, H. A. (2000): "Iterative solution to linear systems in the 20^{th} century", *J. Comput. Appl. Math.*, Vol. 23, pp. 1-33.
2. Ferronato, M., Pini, G., Gambolati, G. (2008): "The role of preconditioning in the solution to FE coupled consolidation equations by Krylov subspace methods", *Int. J. Numer. Anal. Meth. Geomech.*, DOI:10.1002/nag.729, in press.
3. Ferronato, M., Janna, C., Gambolati, G. (2008): "Mixed constraint preconditioning in computational contact mechanics", *Comp. Meth. Appl. Mech. Eng.*, Vol. 197, pp. 3922-3931.
4. Freund, R. W., Nachtigal, N. M. (1994): "A new Krylov subspace method for symmetric indefinite linear systems", *Proc. 14^{th} IMACS World conf. on Comp. and App. Math.*, pp. 1253-1256.
5. Phoon, K. K., Toh, K. C., Chan, S. H., Lee, F. H. (2002): "An efficient diagonal preconditioner for finite element solution of Biot's consolidation equations", *Int. J. Numer. Meth. Eng.*, Vol. 55, pp. 377-400.
6. Kershaw, D. S. (1978): "The incomplete Cholesky-Conjugate Gradient method for the iterative solution of systems of linear equations", *J. Comput. Phys.*, Vol. 20, pp. 43-65.
7. Saad, Y. (1994): "ILUT: a dual threshold incomlete LU factorization", *Numer. Lin. Alg. Appl.*, Vol. 1(4), pp. 387-402.
8. Gambolati, G. , Pini, G., Ferronato, M. (2003): "Scaling improves stability of preconditioned CG-like solvers for FE consolidation equations", *Int. J. Numer. Anal. Meth. Geomech.*, Vol. 27, pp. 1043-1056.

9. Bergamaschi, L., Ferronato, M., Gambolati, G. (2008): "Mixed Constraint Preconditioners for the iterative solution to FE coupled consolidation equations", *J. Comput. Phys.*, Vol. 227, pp. 9985-9897.

10. Janna, C., Ferronato, M., Gambolati, G. (2008): "Multilevel incomplete factorizations for the iterative solution of non-linear FE problems", *Int. J. Numer. Meth. Eng.*, submitted.

11. Ferronato, M., Gambolati, G., Janna, C., Teatini P. (2008): "Numerical Modelling of regional faults in land subsidence prediction above gas/oil reservoirs", *Int. J. Numer. Anal. Meth. Geomech.*, Vol. 32, pp. 633-657.

12. Benzi, M., Cullum, J., Tůma, M. (2000): "Robust approximate inverse preconditioning for the conjugate gradient method", *SIAM J. Sci. Comp.*, Vol. 22, pp. 1318-1332.