

Enrichissement de préconditionneurs pour la résolution de systèmes type point-selle.

Application aux problèmes non-linéaires en mécanique des solides.

Sylvain Mercier

EDF R&D, CERFACS

Séminaire LAMSID, 7 janvier 2014

Thèse CIFRE

Solveurs non-linéaires rapides en mécanique des solides

Directeurs de thèse

Serge Gratton

Xavier Vasseur

Encadrement industriel

Nicolas Tardieu



Thèse débutée en novembre 2012

Cadre de travail

- Problème quasi-statique non-linéaire
- Dualisation des conditions limites (déplacements imposés, liaisons entre degrés de liberté...)
- Méthode de type Newton
⇒ Séquence de systèmes linéaires symétriques à résoudre:

$$\begin{pmatrix} K_i & B_i^T \\ B_i & 0 \end{pmatrix} \begin{pmatrix} u_i \\ \lambda_i \end{pmatrix} = \begin{pmatrix} f_i \\ g_i \end{pmatrix} \iff A_i x_i = b_i$$

- $u_i, f_i \in \mathbb{R}^n$ et $\lambda_i, g_i \in \mathbb{R}^m$
- $K_i \in \mathbb{R}^{n \times n}$ symétrique positive (matrice de rigidité)
- $B_i \in \mathbb{R}^{m \times n}$

Objectif: Récupérer de l'information de la résolution du premier système afin d'accélérer celle des suivants.

- 1 Résolution de systèmes linéaires
 - Méthodes directes/itératives
 - Méthodes de Krylov
- 2 Limited Memory Preconditioner (LMP)
 - Principe général
 - Cas Symétrique Défini Positif
 - Cas Symétrique Indéfini
- 3 Mise en oeuvre pour une séquence
- 4 Résultats numériques
 - Problème non-linéaire avec comportement élastoplastique
 - Courbes de convergence
- 5 Conclusions & Perspectives

- 1 Résolution de systèmes linéaires
 - Méthodes directes/itératives
 - Méthodes de Krylov
- 2 Limited Memory Preconditioner (LMP)
 - Principe général
 - Cas Symétrique Défini Positif
 - Cas Symétrique Indéfini
- 3 Mise en oeuvre pour une séquence
- 4 Résultats numériques
 - Problème non-linéaire avec comportement élastoplastique
 - Courbes de convergence
- 5 Conclusions & Perspectives

Résolution de $Ax = b$

- **Solveurs directs**

Factorisation de A (LDL^T , MultFront...) puis algorithme de remontée-descente.

- Résolution exacte, robuste.
- Prohibitif en terme de mémoire et temps CPU en 3D.

- **Solveurs itératifs**

Construction d'une suite (x_k) qui converge vers la solution.

- Nécessite seulement l'application de A sur un vecteur.
- Permet de récupérer de l'information sur A lors de la résolution.

- 1 Résolution de systèmes linéaires
 - Méthodes directes/itératives
 - Méthodes de Krylov
- 2 Limited Memory Preconditioner (LMP)
 - Principe général
 - Cas Symétrique Défini Positif
 - Cas Symétrique Indéfini
- 3 Mise en oeuvre pour une séquence
- 4 Résultats numériques
 - Problème non-linéaire avec comportement élastoplastique
 - Courbes de convergence
- 5 Conclusions & Perspectives

Présentation

Résolution de $Ax = b$

Definition

Soit x_0 la valeur initiale et $r_0 = Ax_0 - b$ le vecteur résidu initial.
Espace de Krylov d'ordre p :

$$\mathcal{K}_p = \text{Vect}\{r_0, Ar_0, \dots, A^{p-1}r_0\}$$

Idee générale: A l'itération p , on cherche $x_p \in x_0 + \mathcal{K}_p$ sous une condition portant sur le résidu $r_p = Ax_p - b$. Par exemple:

- $r_p \perp \mathcal{K}_p$ (ex: gradient conjugué)
- $\|r_p\| = \min_{x \in x_0 + \mathcal{K}_p} \|Ax - b\|_2$ (ex: GMRES)

Présentation

2 étapes principales à l'itération p :

- Construction d'une base de \mathcal{K}_p à partir de celle de \mathcal{K}_{p-1}
 \Rightarrow colonnes de $V_p \in \mathbb{R}^{n \times p}$
- On travaille avec $V_p^T A V_p = H_p \in \mathbb{R}^{p \times p}$ (Hessenberg)
NB: A sym $\implies H_p$ tridiagonale

Remarques:

- Pour éviter un stockage trop important, on peut effectuer m itérations de la méthode, puis recommencer en prenant x_m comme vecteur initial (ex: GMRES(restart)).
- Les éléments propres de H_p approchent ceux de A . Ils sont appelés éléments de Ritz.

Préconditionnement

Vitesse de convergence liée:

- au conditionnement de A : $\kappa(A) = \|A^{-1}\|_2 \|A\|_2 = \frac{\max_{\lambda \in Sp(A)} |\lambda|}{\min_{\lambda \in Sp(A)} |\lambda|}$
- à la répartition des valeurs propres.

Préconditionnement: $Ax = b \iff M^{-1}Ax = M^{-1}b$

- M facile à "inverser".
- $M^{-1}A$ a de meilleures propriétés spectrales que A .
- L'utilisation de M requiert peu de mémoire.

Préconditionnement

Variantes:

- $M^{-1}Ax = M^{-1}b$ (gauche)

- $\begin{cases} AM^{-1}\tilde{x} = b \\ x = M^{-1}\tilde{x} \end{cases}$ (droite)

- $\begin{cases} M_1^{-1}AM_2^{-1}\tilde{x} = M_1^{-1}b \\ x = M_2^{-1}\tilde{x} \\ M = M_1M_2 \end{cases}$ (centré)

Exemples dans Code_Aster:

- Jacobi: $M = \text{diag}(A)$
- Factorisation LDL^T incomplète de A (sym.)
- Factorisation LDL^T en simple précision (sym.)

- 1 Résolution de systèmes linéaires
 - Méthodes directes/itératives
 - Méthodes de Krylov
- 2 Limited Memory Preconditioner (LMP)
 - Principe général
 - Cas Symétrique Défini Positif
 - Cas Symétrique Indéfini
- 3 Mise en oeuvre pour une séquence
- 4 Résultats numériques
 - Problème non-linéaire avec comportement élastoplastique
 - Courbes de convergence
- 5 Conclusions & Perspectives

- 1 Résolution de systèmes linéaires
 - Méthodes directes/itératives
 - Méthodes de Krylov
- 2 Limited Memory Preconditioner (LMP)
 - Principe général
 - Cas Symétrique Défini Positif
 - Cas Symétrique Indéfini
- 3 Mise en oeuvre pour une séquence
- 4 Résultats numériques
 - Problème non-linéaire avec comportement élastoplastique
 - Courbes de convergence
- 5 Conclusions & Perspectives

Objectif: Résolution de la séquence $A_i x_i = b_i$, $A_i \approx A_{i+1}$

Cadre:

- Méthode de Krylov
- On dispose déjà d'un préconditionneur M (1^{er} niveau)
- Durant la $1^{ère}$ résolution, on peut récupérer des directions spectrales approchées de $M^{-1}A_1$, notamment celles qui ralentissent la convergence.

Idée: A partir de ces directions, on construit un préconditionneur de 2^{nd} niveau. Il est appliqué aux autres systèmes de la séquence pour accélérer la convergence.

- 1 Résolution de systèmes linéaires
 - Méthodes directes/itératives
 - Méthodes de Krylov
- 2 Limited Memory Preconditioner (LMP)
 - Principe général
 - Cas Symétrique Défini Positif
 - Cas Symétrique Indéfini
- 3 Mise en oeuvre pour une séquence
- 4 Résultats numériques
 - Problème non-linéaire avec comportement élastoplastique
 - Courbes de convergence
- 5 Conclusions & Perspectives

Définition

Soit $Ax = b$, avec $A \in \mathbb{R}^{n \times n}$ SDP et $x, b \in \mathbb{R}^n$.

On suppose que $M \in \mathbb{R}^{n \times n}$ est SDP $\implies Sp(M^{-1}A) \subset \mathbb{R}^{+*}$

Definition

Soit $S \in \mathbb{R}^{n \times k}$ dont les colonnes sont linéairement indépendantes.

$$H = (I_n - S(S^T A S)^{-1} S^T A) M^{-1} (I_n - A S (S^T A S)^{-1} S^T) + S (S^T A S)^{-1} S^T$$

est appelé Limited Memory Preconditioner (LMP)

Interprétation

$$H = (I_n - S(S^T AS)^{-1} S^T A) M^{-1} \overbrace{(I_n - AS(S^T AS)^{-1} S^T)}^P + \overbrace{S(S^T AS)^{-1} S^T}^Q$$

$$H = P^T M^{-1} P + Q$$

- P est un projecteur sur $\text{Im}(S)^\perp$
- Q terme de décalage $\implies H$ inversible et $HAS = S$

Idée: Remplir les colonnes de S avec des directions qui ralentissent la convergence

Propriétés

Proposition (Gratton et al., th. 3.4)

- H est SDP
- $HAS = S$, i.e les k vecteurs colonnes de S sont vecteurs propres de HA associés à 1.
- Le reste du spectre est contracté par rapport à celui de $M^{-1}A$.

- 1 Résolution de systèmes linéaires
 - Méthodes directes/itératives
 - Méthodes de Krylov
- 2 Limited Memory Preconditioner (LMP)
 - Principe général
 - Cas Symétrique Défini Positif
 - Cas Symétrique Indéfini
- 3 Mise en oeuvre pour une séquence
- 4 Résultats numériques
 - Problème non-linéaire avec comportement élastoplastique
 - Courbes de convergence
- 5 Conclusions & Perspectives

Généralisation

Rappel: $H = (I_n - S(S^T A S)^{-1} S^T A) M^{-1} (I_n - A S(S^T A S)^{-1} S^T) + S(S^T A S)^{-1} S^T$

Problème: Il n'y a plus contraction du spectre pour A symétrique indéfinie

Idée: découpler le préconditionneur H afin de travailler séparément dans les espaces invariants positif (I^+) et négatif (I^-) de A

Mise en place: Projection d'une partie de S dans I^+ et du reste dans I^- . Pour ceci, il faut utiliser la matrice signe de A , notée X_A , et les projecteurs:

- $X_A = A(A^2)^{-\frac{1}{2}}$
- $P_+ = \frac{I_n + X_A}{2}$ projecteur sur I_+
- $P_- = \frac{I_n - X_A}{2}$ projecteur sur I_-

Résultats théoriques

Dans la suite, on suppose que $M = LL^T$ est SDP. On travaille sur le système:

$$\underbrace{L^{-1}AL^{-T}}_{\mathcal{A}} \tilde{x} = L^{-1}b$$

Definition

Supposons que $S = [S_+, S_-]$ avec $S_{\pm} = P_{\pm}S \in I_{\pm}(K)$, dont les colonnes sont linéairement indépendantes.

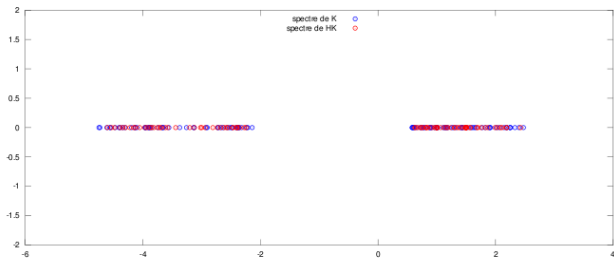
Le LMP associé est défini par:

$$H = (I_n - S(S^TAS)^{-1}S^T\mathcal{A})(I_n - \mathcal{A}S(S^TAS)^{-1}S^T) + S(S^TAS)^{-1}S^T$$

Résultats théoriques

Proposition

- H est symétrique.
- $HAS = S$, i.e les k vecteurs colonne de S sont vecteurs propres de HA associés à 1.
- Le reste du spectre est contracté de chaque côté de l'axe des imaginaire par rapport à celui de A .



Matrice signe et projection

problème: Le calcul de la matrice signe est très coûteux!

Choix des vecteurs colonnes de $S = [S_+, S_-]$

- Idéal: vecteurs propres correspondant aux petites valeurs propres en module

- Calcul des éléments de Ritz (u, θ)

Rappel: Ce sont les éléments propres de $H_p = V_p^T \mathcal{A} V_p$
 \implies facilement accessibles

- On choisit les vecteurs de Ritz associés aux plus petites valeurs de Ritz en module:

Si $\theta > 0$, $u \in S_+$, sinon $u \in S_-$

- On montre que ces vecteurs sont proches de leurs projections
 \implies Pas besoin de la matrice signe

- 1 Résolution de systèmes linéaires
 - Méthodes directes/itératives
 - Méthodes de Krylov
- 2 Limited Memory Preconditioner (LMP)
 - Principe général
 - Cas Symétrique Défini Positif
 - Cas Symétrique Indéfini
- 3 Mise en oeuvre pour une séquence
- 4 Résultats numériques
 - Problème non-linéaire avec comportement élastoplastique
 - Courbes de convergence
- 5 Conclusions & Perspectives

Résolution de $A_i x_i = b_i$ *hypotheses:*

- $A_i \in \mathbb{R}^{n \times n}$ sym., $A_i \approx A_{i+1}$
- $M = LL^T \approx A_1 \in \mathbb{R}^{n \times n}$ SDP (1^{er} niveau)

Résolution du premier système

- Utilisation de GMRES(30) en préconditionnant par M :

$$A_1 x_1 = b_1 \iff \underbrace{L^{-1} A_1 L^{-T}}_{\mathcal{A}_1} \tilde{x}_1 = L^{-1} b_1$$

- Au dernier cycle complet de GMRES, on calcule les éléments de Ritz de \mathcal{A}_1 (éléments propres de $H_{30} = V_{30}^T \mathcal{A}_1 V_{30}$)
- On sélectionne les k éléments de Ritz associés aux valeurs de Ritz de plus petit module $\implies S = [S_+, S_-]$
- Construction de

$$H = (I_n - S(S^T \mathcal{A}_1 S)^{-1} S^T \mathcal{A}_1)(I_n - \mathcal{A}_1 S(S^T \mathcal{A}_1 S)^{-1} S^T) + S(S^T \mathcal{A}_1 S)^{-1} S^T$$

Résolution des autres systèmes

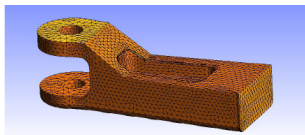
On résout les systèmes $A_i x_i = b_i$, $i \geq 2$ en utilisant les préconditionneurs M et H :

$$\begin{cases} L^{-1} A_i L^{-T} H \tilde{x} = L^{-1} b \\ x = L^{-T} H \tilde{x} \end{cases}$$

Pour ceci, on utilise également GMRES(30)

Remarque: En terme de mémoire, on stocke $2kn$ vecteurs en plus.

- 1 Résolution de systèmes linéaires
 - Méthodes directes/itératives
 - Méthodes de Krylov
- 2 Limited Memory Preconditioner (LMP)
 - Principe général
 - Cas Symétrique Défini Positif
 - Cas Symétrique Indéfini
- 3 Mise en oeuvre pour une séquence
- 4 Résultats numériques
 - Problème non-linéaire avec comportement élastoplastique
 - Courbes de convergence
- 5 Conclusions & Perspectives



Séquence:

$$\underbrace{\begin{pmatrix} K_i & B^T \\ B & 0 \end{pmatrix}}_{A_i} \underbrace{\begin{pmatrix} u_i \\ \lambda_i \end{pmatrix}}_{x_i} = \underbrace{\begin{pmatrix} f_i \\ g_i \end{pmatrix}}_{b_i}, i = 1, \dots, 4$$

- $n = 7305$ inconnues physiques
- $m = 228$ multiplicateurs de Lagrange
- Encastrement à l'extrémité droite
- Pression imposée à gauche

Préconditionneur de 1^{er} niveau:

Proposition

Soit $M = \begin{pmatrix} K_1 + B^T B & 0 \\ 0 & I_m \end{pmatrix}$. Alors $M^{-1}A_1$ admet comme spectre:

- 1 de multiplicité n
- les m restantes appartenant à $[-1, 0)$

De plus, $K_1 + B^T B$ est SDP \implies **M est SDP**

$$\tilde{M} = \begin{pmatrix} RR^T & 0 \\ 0 & I_m \end{pmatrix} \approx M$$

RR^T est une décomposition de Cholesky incomplète de $K_1 + B^T B$.

$$\tilde{M} = \underbrace{\begin{pmatrix} R & 0 \\ 0 & I_m \end{pmatrix}}_L \underbrace{\begin{pmatrix} R^T & 0 \\ 0 & I_m \end{pmatrix}}_{L^T}$$

Résolution de la séquence avec GMRES(30) et $k = 5$

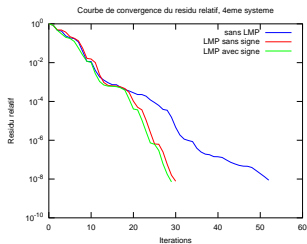
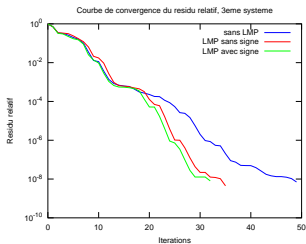
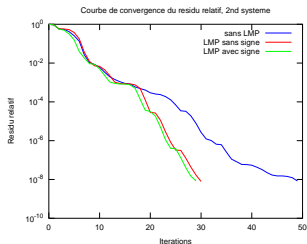
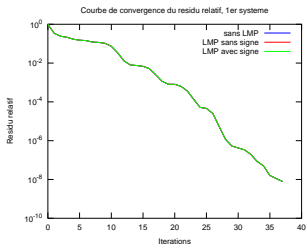


Tableau de résultats

Le tableau suivant représente le nombre d'itérations total de la résolution de la séquence, le temps CPU et les rapports de complexité.

	sans LMP	LMP sans signe	LMP avec signe
Iter total	187	132	127
Tps CPU (s)	3.2	2.59	>3600
Rapport Complexité	1	0.86	>>1

Conclusions

- Proposition d'extension du LMP au cas symétrique indéfini
- Contribution théorique sur la contraction spectrale
- Méthode applicable grâce à l'absence de la matrice signe
- Candidat approprié pour le préconditionnement de *1^{er} niveau*
- Implémentation sous Matlab et premiers résultats encourageants

Perspectives

- Implémentation dans *Code_Aster* via la bibliothèque *PETSc*
- Tests sur des problèmes de grande taille issus des études d'EDF
- Publication des résultats théoriques et numériques
- Généralisation au cas non symétrique

Merci de votre attention !

Bibliography I



M. Benzi, G.H. Golub, J.Liesen.
Numerical Solution of saddle point problems.
Acta Numerica, 1-137, 2005.



P. Krzyzanowski.
On block preconditioners for saddle point problems with singular or indefinite (1,1) block.
Numerical Linear Algebra with Applications 18: 123-140, 2011.



M.F. Murphy, G.H. Golub, A.J. Wathen.
A note on preconditioning for indefinite linear systems.
SIAM J. Sci. Comput., 21-6: 1969-1972, 2000.



A. Klawonn.
An optimal preconditioner for a class of saddle point problems with a penalty term.
SIAM J. Sci. Comput. 19-2: 540-552, 1998.



H.S. Dollar, N.I.M. Gould, M. Stoll, A.J. Wathen.
Preconditioning saddle point systems with applications in optimization.
SIAM J. Sci. Comput. 32-1: 249-270, 2012.



C. Siefert, E. de Sturler.
Preconditioners for generalized saddle point problems.
SIAM J. Numer. Anal. 44-3: 1275-1296, 2006.

Bibliography II



M. Benzi, V. Simoncini.

On the eigenvalues of a class of saddle point matrices.

Numerische Mathematik 103: 173-196, 2006.



L. Bergamaschi, J. Gondzio, M. Venturin, G. Zilli.

Inexact constraint preconditioners for linear systems arising in interior point methods.

Computational Optimization and Applications, 36: 137-147, 2007.



J. Pestana, A.J. Wathen.

Combination preconditioning of saddle point systems for positive definiteness.

Numerical Linear Algebra with Applications, 2012.



M. Stoll, A.J. Wathen.

Combination preconditioning and the Bramble-Pasciak⁺ preconditioner.

SIAM J. Matrix Anal. Appl. 30-2: 582-608, 2008.



S. Bellavia, D. Bertaccini, B. Morini.

Nonsymmetric preconditioner updates in Newton-Krylov methods for nonlinear systems.

SIAM J. Sci. Comput. 33-5: 2595-2619, 2011.



S. Gratton, A. Sartenaer, J. Tshimanga.

On a class of limited memory preconditioners for large scale linear systems with multiple right-hand sides.

SIAM J. Optim. 21-3: 912-935, 2011.



J. Tshimanga.

On a class of limited memory preconditioners for large-scale nonlinear least-squares problems.

PhD Thesis, 2007.



G. Fasano, M. Roma.

AINVK: a class of approximate inverse preconditioners based on Krylov subspace methods, for large indefinite linear systems.

Preprint, 2012.



Y. Saad.

Iterative methods for sparse linear systems - second edition.

SIAM, Philadelphia, PA, 2003.



N.J. Higham.

Functions of matrices: theory and computation.

SIAM, chap. 5, 2008.

Bibliography IV



A. Frommer, K. Kahl, T. Lippert, H. Rittich.
Error bounds for the sign function.
IMACM, 2011.



A. Frommer, K. Kahl, T. Lippert, H. Rittich.
2-norm error bounds and estimates for Lanczos approximations to linear systems and rational matrix functions.
IMACM, 2011.



J.C.R. Bloch, S. Heybrock.
A nested Krylov subspace method to compute the sign function of large complex matrices.
Computer Physics Communications 182: 878-889, 2011.



R.W. Freund, N.M. Nachtigal.
Software for simplified Lanczos and QMR algorithms.
Applied Numerical Mathematics 19: 319-341, 1995.



R.W. Freund, T. Szeto.
A quasi minimal residual squared algorithm for non hermitian linear systems.



H. Elman, D. Silvester, A. Wathen.
Finite elements and fast iterative solvers with applications in incompressible fluid dynamics.
Numerical Mathematics and Scientific Computation, 2005.



H. Elman, A. Ramage, D. Silvester.

IFISS, a matlab toolbox for modelling incompressible flow.
ACM Transactions on Mathematical Software 33-2, 2007.



E. de Sturler, J. Liesen.

Block-diagonal and constraint preconditioners for nonsymmetric indefinite linear systems. Part I: theory.
SIAM J. Sci. Comput. 26-5: 1598-1619, 2005.