

Algorithmes de raffinement itératif modernes pour la résolution de systèmes linéaires creux

B. Vieublé, P. AMESTOY, A. BUTTARI, N. HIGHAM, J. L'EXCELLENT, T. MARY
INPT, MUMPS Technologies, CNRS IRIT, University of Manchester, MUMPS Technologies, CNRS LIP6

Email : bastien.vieuble@irit.fr

Mots Clés : iterative refinement, GMRES, sparse linear system, mixed precision, rounding error analysis, performance

Biographie – Financée par une bourse MESRI, ma thèse s'intéresse à l'utilisation d'algorithmes à précision mixte pour la résolution de systèmes linéaires creux de grande taille. Les méthodes proposées sont mises en œuvre au sein de la plateforme de recherche MUMPS. De cette manière, ces nouvelles méthodes sont validées dans un contexte applicatif réaliste sur divers problèmes issus de la communauté d'utilisateurs de MUMPS.

Resumé :

Il y a un demi-siècle, James Wilkinson a introduit le raffinement itératif, une méthode améliorant la précision de la solution d'un système linéaire en lui appliquant des itérations de raffinement. Ces itérations sont composées du calcul du résidu, de la résolution de l'équation de correction, et de la mise à jour de la solution. Plus tard, des variantes ont été proposées permettant de calculer la factorisation de la matrice (l'opération la plus lourde en terme de complexité) en précision arithmétique basse, ce qui sur les architectures modernes permet d'accélérer les calculs, potentiellement d'un facteur deux si nous passons de la précision double (fp64) à la précision simple (fp32) par exemple, et de réduire la consommation mémoire.

Les méthodes de raffinement itératif voient leurs popularités grandir dernièrement avec les promesses de support matériel pour la précision half (fp16) (ex. avec les tensor cores sur GPU). Dans ce contexte, Carson et Higham ont récemment proposé de nouvelles variantes de raffinement itératif qui utilisent jusqu'à trois précisions arithmétiques différentes [2] et se démarquent des approches précédentes en étant plus robustes au conditionnement du problème [1].

Dans cette présentation nous étudierons l'utilisabilité de ces techniques pour la résolution de systèmes linéaires creux de grande taille. Ainsi, nous devons considérer un certain nombre de contraintes pratiques liées à la consommation mémoire et la performance. En particulier, nous évoquerons le raffinement itératif basé sur un solveur GMRES à trois précisions proposé par Carson et Higham et sa principale limitation pratique qui est de devoir appliquer le préconditionneur en haute précision arithmétique. Nous montrerons que cette condition peut être relaxer donnant ainsi une variante du raffinement itératif utilisant jusqu'à 5 précisions arithmétiques différentes. Nous présenterons ensuite le potentiel de ces différents algorithmes pour l'accélération de la résolution de système linéaire creux en faisant tourner nos implémentations parallèles basées sur le solveur MUMPS sur des problèmes industriels.

Références

- [1] Erin Carson and Nicholas J. Higham. A new analysis of iterative refinement and its application to accurate solution of ill-conditioned sparse linear systems. 39(6):A2834–A2856, 2017.
- [2] Erin Carson and Nicholas J. Higham. Accelerating the solution of linear systems by iterative refinement in three precisions. 40(2):A817–A847, 2018.