固有値ソルバの並列化とその性能

片桐 孝洋 [1,2]

[1] 電気通信大学大学院情報システム学研究科, [2] 科学技術振興機構さきがけ

1 はじめに

固有値・固有ベクトルを計算するソルバ (固有値ソルバ) を並列化する時は，近接固有値に対する固有ベクトルの直交精度を保つための再直交化処理が問題となる．

対称行列の固有ベクトル計算には，逆反復法 (Inverse Iteration Method, IIM) を用いることが多い．この場合，たいていは Gram-Schimidt 法による直交化 (以降，G-S 直交化) が用いられる．G-S 直交化を並列 IIM に適用する場合，対象となる行列の数値特性にも依存するが，全体の実行時間に対し IIM の時間が 90%以上を占め，さらにこの時間のうち，G-S 再直交化の時間が 90%以上を占めることが報告されている [5].

そこで本稿ではまず，並列 IIM の再直交化に関する既存の研究を紹介する．つぎに，実際の並列計算機 (PC クラスタ) を用いて各方式の性能を評価した結果を報告する．

2 並列再直交化方式

図 1 に，並列 IIM 法において用いられる既存の再直交化方式をまとめる．

- G-S 直交化を用いる方法：
  - スケジューリングを行う方法：**マルチカラー逆反復法** [3]
  - 古典 G-S 法を混合する方法：HCG-S **法** [6]
  - ソートを付加する方法：CG-SS **法** [7]
  - データ再分散を用いる方法 [4]
- Householder 法を用いる方法：Householder **逆反復法** [2]
- 再直交化を用いない方法：Dhillon **法** [1]

図 1: 並列逆反復法中で利用される再直交化方式

図 1 において，直野らが提案したマルチカラー逆反復法は，G-S 直交化を行う対象の固有ベクトルに関し，(1) 従来の近接固有値と判定する距離の緩和；(2) 近接固有値群の依存関係を解析することによる並列性の抽出；を行う．ただしこの方法でも，理論上重複する固有値に対する固有ベクトルの再直交化は逐次化される．

片桐らが提案した HCG-S 法および CG-SS 法は，並列性の低い修正 G-S 法の利用を制限し，並列性の高い古典 G-S 法を利用する方式である．ただし，古典 G-S 法は修正 G-S 法に比べて直交精度・安定性が悪い．HCG-S 法では，IIM 中は古典 G-S 法を用いるが，収束後 1 回修正 G-S 法を行う．古典 G-S 法のみ利用する方法に比べ，収束性が改善される場合がある．一方 CG-SS 法は，古典 G-S 法での計算中において情報埋没を回避するために，古典 G-S 法の演算順序を内積値をキーとしてソートして変更する方法である．CG-SS 法により，あらゆる入力に対して直交精度が改善されるわけではないが，直交精度が改善される例が報告されている．しかしながら，HCG-S 法，CG-SS 法ともに，実行時間の問題は改善されても直交精度の問題が完全に解決されない．

片桐らが提案したデータ再分散を用いる反復法では，並列性の低い IIM 中の再直交化のためのデータ分散方式を，その都度変更 (データ再分散) することで，実行時間の改善を試みる．この方式は，古典 G-S 法および修正 G-S 法それぞれに適用できる．したがって，各方式の理論限界内での直交精度を保証できる．すなわち HCG-S 法のように，実行時間と直交精度の問題は生じない．しかし性能改善は，並列計算機ハードウエアの性能に依存するので，効果のある計算機と無い計算機がある．

山本らによって提案された Householder 逆反復法は，再直交化に Hoousehokder 法を用いることで高い並列性能と直交精度を達成する．しかし，G-S 直交化に比べ 1.5 倍ほど計算量が多い．

最後に Dhillon 法では，IIM において再直交化をしなくても直交精度を保証する．この方法では，Twisted Decomposition を利用して，IIM で

用いる初期固有ベクトルを見積もる．しかし，二分法による計算結果よりも高い精度の初期固有値が必要となる．また，理論上固有値が重複する場合には適用できない．

以上のように，提案されてきた IIM 中での再直交化方式は，それぞれに長所/短所がある．

3 性能評価

ここでは，電気通信大学大学院情報システム学研究科並列処理学講座が所有する PC クラスタを用いて並列 IIM 中での再直交化方式の性能を評価する．PC クラスタのスペックは以下の通りである．PE(要素計算機) のハードウエアは Intel Pentium4 (2.0GHz)，PE 数は 4，PE 当たりの搭載メモリは 1GB (Direct RDRAM/ECC 256MB*4) である．マザーボードは ASUSTek P4T-E+A (Socket478 対応)，ネットワークカードは Onboard Broadcom Gigabit Ethernet*2 である．OS は Linux 2.4.9-34 である．通信ライブラリは MPICH 1.2.1 である．コンパイラは PGI Fortran90 コンパイラ 4.0-2 で，コンパイラオプションは -fast を用いた．

試験行列として 8000 次元の Frank 行列を用いた．この試験行列の固有ベクトルを，Householder-二分-逆反復法で求める．なお，近接固有値判定法は従来方式による距離 $||T||_1 \times 10^{-3}$ を用いた．また IIM の収束判定方式は，疑収束を防ぐため最低 2 回は行い，その後最大反復回数 40 回に達するか，残差ベクトルが理論限界の $|T|_1 \times \epsilon$ 以下 (ここで ϵ はマシン・イプシロン) になるとき打ち切る方式を採用した．

図 2 は，上記の PC クラスタ 4PE を用いて 8000 次元の行列に対する固有ベクトル計算 (固有ベクトル数：100, 250, 500, 1000, 2000, 4000, 8000) での，各再直交化方式を用いた IIM の実行時間を示している．図 2 中の表記は，MG-S(修正 G-S 法)，CG-S(古典 G-S 法)，RBMG-S(データ再分散を用いた修正 G-S 法)，RBCG-S(データ再分散を用いた古典 G-S 法)，HCG-S(HCG-S 法)，IRCG-S(反復改良 CG-S 法) および SCG-S(PE 内計算も古典 G-S 法で行う方法) である．

図 2 から最も遅いのは RBMG-S である．この理由は，この計算機環境ではデータ再分散のための通信時間がネックとなることによる．また最も高速となるのは，固有ベクトル計算数が 1000 未満の時は CG-S および HCG-S であり，固有ベク

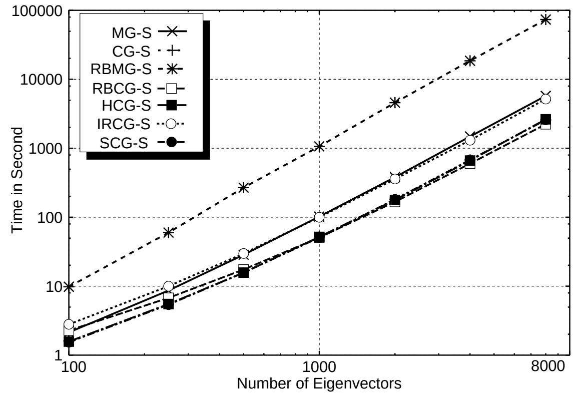

図 2: 各再直交化方式の IIM 実行時間 [秒]

トル計算数が 1000 以上のとき RBCG-S である．固有ベクトル 8000 個計算時，CG-S では 2559 秒に対し，RBCG-S では 2229 秒である．この結果から，(1) 古典 G-S 法による並列性抽出が有効なこと；(2) 固有ベクトル数が増すとデータ再分散する方式が有効となること；がいえる．また，HCG-S 法は古典 G-S 法に加えて収束後 1 回修正 G-S 法を行っているのにもかかわらず，古典 G-S 法と同等もしくはそれ以上高速となる．この理由は，修正 G-S 法を用いることによる収束までの反復回数の短縮があげられる．

なお，解の直交精度を実行時間と共に議論する必要がある．当日の講演にて直交精度検定結果を示し，かつそれに対する考察を述べる．

参考文献

[1] I. Dhillon. A New $O(n^2)$ Algorithm for The Symmetric Tridiagonal Eigenvalue / Eigenvector Problem. *Ph.D Thesis, Computer Science Division, University of California Berkeley*, 1997.

[2] 山本有作, 猪貝光祥, 直野健. 共有メモリ型並列計算機向けの高並列固有ベクトル解法. 2000 年記念並列処理シンポジウム JSPP2000 論文集, pp. 19–26, 2000.

[3] 直野健, 猪貝光祥, 山本有作. 並列固有値ソルバーの開発と性能評価. 並列処理シンポジウム JSPP96 論文集, pp. 9–16, 1996.

[4] 片桐孝洋, 吉瀬謙二, 本多弘樹, 弓場敏嗣. データ再分散を行う並列 Gram-Schmidt 再直交化. 情報処理学会論文誌：コンピューティングシステム, Vol. 45, No. SIG 6 (ACS 6), pp. 75–85, 2004.

[5] 片桐孝洋, 金田康正. 並列固有値ソルバーの実現とその性能. 情報処理学会研究報告, 97-HPC-69, pp. 49–54, 1997.

[6] 片桐孝洋, 金田康正. 並列固有値ソルバーの実現とその並列性の改良. 並列処理シンポジウム JSPP98 論文集, pp. 223–230, 1998.

[7] 片桐孝洋, 金田康正. CGSS:ソートを用いた新しい Gram-Schmidt 直交化法. 情報処理学会研究報告, 1999-HPC-76, pp. 37–42, 1999.