

平成 30 年 5 月 23 日現在

機関番号：10101

研究種目：若手研究(B)

研究期間：2015～2017

課題番号：15K16000

研究課題名(和文) 通信回避型行列分解の実用性向上に資する基盤技術の研究

研究課題名(英文) Fundamental research for improving the practicality of communication-avoiding matrix factorization algorithms

研究代表者

深谷 猛 (Fukaya, Takeshi)

北海道大学・情報基盤センター・助教

研究者番号：30633846

交付決定額(研究期間全体)：(直接経費) 3,100,000円

研究成果の概要(和文)：大規模並列計算における通信コストを削減するために、通信回避型の行列分解アルゴリズムが注目されている。本研究では、これらのアルゴリズムの実装方法やチューニング手法に主眼を置き、実際の計算機上でより高い性能を得るための基盤技術进行研究した。具体的には、通信回避型アルゴリズムで必要となる計算カーネルの実装方法や、異なる通信回避型アルゴリズムの性能比較などを実施した。また、これらの研究を支える、並列計算機上でのアルゴリズムの性能モデルの構築方法についても検討を行った。

研究成果の概要(英文)：In order to reduce the communication cost in large-scale parallel computations, so-called communication-avoiding (CA) algorithms for matrix factorization have been actively studied. In this research, we aimed for improving the practicality of CA algorithms in realistic situations. We investigated fundamental techniques for CA algorithms: for example, those for implementing and tuning CA algorithms. We also compared different CA algorithms. In addition, we studied the performance modeling of CA algorithms.

研究分野：高性能計算

キーワード：行列計算 高性能計算 並列計算 通信回避

1. 研究開始当初の背景

京コンピュータをはじめとする、大規模並列計算機の普及が進み、今後のエクサスケール時代には、大規模並列計算機をより手軽に利用できるようになることが予想される。大規模並列計算機を用いた計算では、ネットワークで結合されたノード間におけるデータの通信時間を削減することが、高い性能を実現する上で重要な課題となる。

LU 分解や QR 分解に代表される行列分解は、科学技術計算で用いられる様々なアルゴリズムを構成する重要なパーツの一種である。しかし、既存の行列分解の計算手法は、主に演算回数削減の観点で研究・開発されたものであり、近年の大規模並列計算機に適しているとは言いがたい。事実、標準的なライブラリとして利用されてきた ScaLAPACK の性能が、京コンピュータ等において著しく低いことが指摘されている。

このような状況から、近年、通信回避 (CA: Communication-Avoiding) と呼ばれる、通信回数の削減に主眼を置いた、新しい計算手法の研究が活発に行われるようになった。そして、従来の計算手法よりも、通信回避型の計算手法の方が、大規模並列計算機において高い性能を示す事例が多数報告され、今後、通信回避型の計算手法が重要な役割を果たすであろう、という認識が広まっている。

2. 研究の目的

1 で述べたように、通信回避型の行列分解アルゴリズムへの期待が高まっているが、実際にその性能を現実の計算機上で評価してみると、いろいろな課題が存在することが分かった。例えば、プログラムの実装方法やチューニング方法に関する議論が不十分であり、それらが適切ではない場合、プログラムの性能が大きく損なわれ、従来の計算手法の方が高性能となることがあり得る。また、異なる特徴を持つ複数の通信回避型アルゴリズムが提案されており、状況により適切な使い分けを行うことが求められる。これらを踏まえて、本研究課題では、通信回避型の行列分解アルゴリズムを実際の計算機上で使用する場合において生じる課題を明確にするとともに、それらの解決方法を研究することで、通信回避型行列分解アルゴリズムの実用性の向上に貢献することを目的とした。

3. 研究の方法

本研究課題では、まず、個別の通信回避型行列分解アルゴリズムを対象に、実用面の課題を明らかにすることを行った。具体的には、行列の QR 分解の計算を行うアルゴリズムについて、実際にプログラムを実装し、その性能を分析することで、課題の洗い出しを行った。また、今後の計算機環境における性能等を議論するための道具として、実際の状況を考慮した、プログラムの性能モデルの構築方法についても研究を行った。

上記の内容で得られた知見に基づいて、QR 分解以外の行列計算に関する通信回避技術についても、同様のアプローチで手法の分析等を実施した。

4. 研究成果

本研究の主な成果は以下の通りである。

(1) 再直交化付きコレスキーQR 分解の研究

コレスキーQR 分解は通信回避型の QR 分解アルゴリズムの一つであり、高い実行性能を得ることができる反面、計算精度面の不安定性により、非実用的であると認知されていた。これに対して、再直交化を付与したアルゴリズム (CholeskyQR2) が、一定の条件下では精度面の問題が解決され、安定に計算を行うことが可能となることを理論的・実験的に示した。さらに、京コンピュータにおいて、他の種類の通信回避型の QR 分解アルゴリズムよりも、依然として高性能であることを実証した (図 1)。コレスキーQR 分解は、行列積や MPI の集約通信のみを用いて簡単にプログラムを実装することが可能であり、新しい計算機環境においても、ベンダー等による最適化済みのライブラリを活用することが可能である。そのため、本成果により、(再直交化付き)コレスキーQR 分解の有効性が示されたことは、実際のシミュレーションプログラムにおける本手法の活用を促し、プログラムの高性能化に貢献することが期待される。

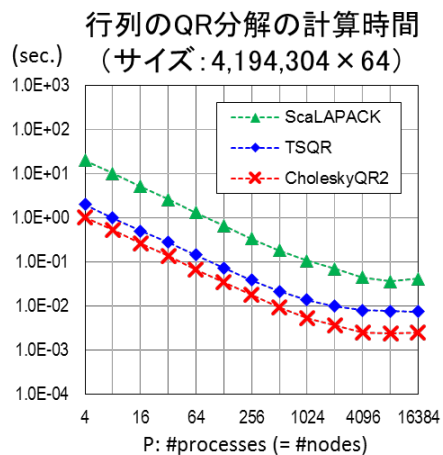


図 1: 京コンピュータ上での性能評価結果

(2) TSQR アルゴリズムの高性能化

TSQR アルゴリズムは非常に高い安定性を持つ、通信回避型 QR 分解アルゴリズムである。本研究では、大規模分散並列計算において、TSQR アルゴリズムの実行時間の大部分が、特殊な構造を有する小規模な行列に対する処理 (Structured QR 分解) の繰り返しになる点に着目し、その処理の実装をより最適化することで、TSQR アルゴリズム全体の性能向上が可能となることを確認した。本成果により、通信回避型の行列分解アルゴリズムにおいて、従来、あまり考慮がなされていなかっ

た小規模な行列に対する処理が実は重要であることが示唆された(図2)。

また,TSQR アルゴリズムにおいて,各プロセスで計算された結果を集約する処理が必要となるが,従来の実装では,二分木に従う形でこれを行っていた。これに対して,異なる構造の木に従うことで,計算時間の短縮が可能となるかどうか,検討を行った。様々な観点から議論を行った結果,使用する計算機環境がある条件を満たす場合においては二つではなく,それ以上の数の小行列をまとめて一度に扱う形の方が,効率が良いことが明らかになった。

本研究は,QR 分解を行う TSQR アルゴリズムに関する内容ではあるが,小規模の行列計算を高速化することが,通信回避型アルゴリズム全体の高性能化において重要である,ということは,QR 分解に限らず,一般的な課題である。そのため,本成果は,今後,様々な通信回避型アルゴリズムの実装方法やチューニング方法の研究につながる事が期待できる。

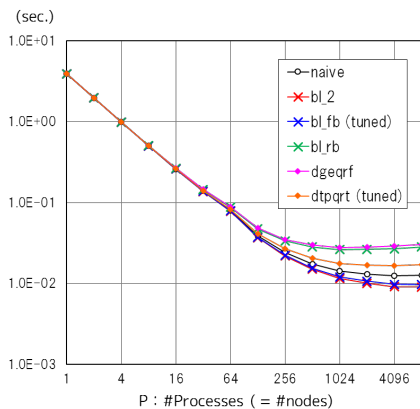


図2: Structured QR 分解の実装方法による TSQR アルゴリズムの性能の違い

(3) 通信回避型の CG 法の研究

代表的な反復法である CG 法に対して 様々な通信回避技術を適用し,その効果を検証した。特に,安定性に優れた,Chebyshev 基底を用いた通信回避 CG 法(CBCG 法)に関して,性能モデルを構築し,通信回避の効果が有意となる条件(並列数)を推測することが可能であることを確認した。また,京コンピュータを用いた性能評価により,実際に CBCG 法が従来の CG 法よりも高性能となり得ることを実証した。

(4) Hybrid 格納形式を用いた疎行列ベクトル積に関する研究

Hybrid 格納形式を用いることで,疎行列ベクトル積における疎行列データのメモリアクセスコストを削減することが可能となる。しかし,従来の実装方法では,計算結果を書き込むベクトルに対する参照が増加してしまうため,本研究では,ベクトルデータに関

するキャッシュブロッキング手法を研究した。さらに,キャッシュブロッキングを前提として,Hybrid 格納形式を改良することを行った。最新のマルチコア/メニーコア CPU を用いた性能評価の結果,データベースから取得した,いくつかの疎行列に関して,提案手法により,疎行列ベクトル積の性能が向上することが確認された。また,他の疎行列に対しても,特に性能低下の弊害が生じないことも併せて確認された。この結果を踏まえて,今後,反復法ライブラリ等への本手法の組み込みが期待できる。

(5) 大規模並列計算機における密行列計算プログラムの性能モデリングの研究

現状,大規模並列計算機を利用できる機会が限られており,エクサスケール時代に向けたアルゴリズムの研究を行う上で,プログラムの性能を予測するモデルを構築することは有益である。モデルを構築する際,プログラムに関する情報や対象とする計算機に関する情報は状況に応じて異なるため,利用可能な情報から構築したモデルの精度を検証することは重要な課題である。本研究では,京コンピュータ上で開発が進められている最新の固有値計算ライブラリの主要ルーチンを対象として,異なる計算機の情報から構築されたモデルによる性能の予測精度の比較を調査した。実験の結果,想定する計算機のごく一部のリソースのみを用いた場合でも,対象とするプログラムを考慮したベンチマークを行うことで,十分実用的な精度でプログラムの性能を予測できる可能性があることを示すことができた(図3)。



図3: 計算機の一部のリソースを用いたベンチマーク結果に基づく性能予測の精度(実測データと予測結果の誤差を評価)

5. 主な発表論文等

(研究代表者,研究分担者及び連携研究者には下線)

[雑誌論文](計4件)

Takeshi Fukaya, Toshiyuki Imamura, and Yusaku Yamamoto, A case study on modeling the performance of dense matrix computation: tridiagonalization in the EigenExa

eigensolver on the K computer, Proceedings of the 13th International Workshop on Automatic Performance Tuning (2018 IEEE International Parallel and Distributed Processing Symposium Workshop), pp. 1113-1122, 2018 (査読有).

(DOI: 10.1109/IPDPSW.2018.00171)

熊谷 洋佑, 藤井 昭宏, 田中 輝雄, 深谷 猛, 須田 礼仁, 共役勾配法へ種々の通信削減手法の適用と評価, 情報処理学会論文誌コンピューティングシステム (ACS), Vol. 9, pp. 1-13, 2016 (査読有).

(https://ipsj.ixsq.nii.ac.jp/ej/index.php?active_action=repository_view_main_item_detail&page_id=13&block_id=8&item_id=174113&item_no=1)

Yusaku Yamamoto, Yuji Nakatsukasa, Yuka Yanagisawa, and Takeshi Fukaya, Roundoff error analysis of the CholeskyQR2 algorithm in an oblique inner product, JSIAM Letters, Vol. 8, pp. 5-8, 2016 (査読有). (DOI: 10.14495/jsiaml.8.5)

Yusuke Kumagai, Akihiro Fuji, Teruo Tanaka, Yusuke Hirota, Takeshi Fukaya, Toshiyuki Imamura, and Reiji Suda, Performance Analysis of the Chebyshev Basis Conjugate Gradient Method on the K Computer, Lecture Notes in Computer Science, Vol. 9573, pp. 74-85, 2016 (査読有).

(DOI: 10.1007/978-3-319-32149-3_8)

[学会発表](計 16 件)

Takeshi Fukaya, Koki Ishida, Akie Miura, Takeshi Iwashita, and Hiroshi Nakashima, An Approach to Accelerating the SpMV Kernel by Exploiting Specific Sparse Structures, 2018 SIAM Conference on Parallel Processing for Scientific Computing (SIAM PP18), 2018 年 3 月 10 日, 早稲田大学 (東京都).

Yuka Yanagisawa, Takeshi Fukaya, Yuji Nakatsukasa, Yusaku Yamamoto, and Ramseshan Kannan, Shifted Cholesky QR for Computing the QR Factorization for Ill-conditioned Matrices, 2018 SIAM Conference on Parallel Processing for Scientific Computing (SIAM PP18), 2018 年 3 月 9 日, 早稲田大学 (東京都).

深谷 猛, TSQR アルゴリズムにおける三角行列のリダクション処理に関する考察, 日本応用数理学会 2017 年度年会, 2017 年 9 月 6 日, 武蔵野大学 (東京都).

Takeshi Fukaya and Toshiyuki Imamura, An Impact of Tuning the Kernel of the Structured QR Factorization in the

TSQR, 2016 SIAM Conference on Parallel Processing for Scientific Computing (SIAM PP16), 2016 年 4 月 14 日, Paris (France).

Takeshi Fukaya, Yuji Nakatsukasa, Yuka Yanagisawa, and Yusaku Yamamoto, Performance Evaluation of the CholeskyQR2 Algorithm, 2015 SIAM Conference on Applied Linear Algebra (SIAM LA15), 2015 年 10 月 27 日, Atlanta (USA).

深谷 猛, 中務 祐治, 山本 有作, これスケーQR 分解を用いたブロック直交変換の生成, 第 44 回数値解析シンポジウム, 2015 年 6 月 9 日, ぶどうの丘 (山梨県甲州市).

6. 研究組織

(1) 研究代表者

深谷 猛 (FUKAYA, Takeshi)

北海道大学・情報基盤センター・助教

研究者番号: 30633846