

複数の格納形式を利用した疎行列ベクトル積の高速化に関する検討

石田 幸輝¹, 三浦 瑛絵¹, 深谷 猛², 岩下 武史², 中島 浩³

¹北海道大学 工学部, ²北海道大学 情報基盤センター, ³京都大学 学術情報メディアセンター
e-mail: fukaya@iic.hokudai.ac.jp

1 はじめに

疎行列ベクトル積は, 多くのシミュレーションにおいて用いられる重要な計算核の一つであり, その高速化に関して大きな需要がある. 本稿では, 対象とする行列から特定の疎構造を自動的に抽出し, その構造を活用した格納形式を部分的に適用することにより疎行列ベクトル積の高速化を行う.

2 提案手法

代表的な疎行列の格納形式として, CRS 形式がある. CRS 形式は, 特定の疎構造を仮定せず, 疎行列一般に適用可能であるが, 非ゼロ要素の位置情報を配列に格納するため, 疎行列ベクトル積の実装において, 乗算ベクトルの各要素への参照が位置情報の配列を経由する間接参照となる. 間接参照は直接参照と比べて一般に低性能であり, その他にも効率的な SIMD 化実装の妨げになる等, 性能上の問題を引き起こす要因となる.

一方, 偏微分方程式を7点差分公式で離散化した際に得られる行列のように, 限られた斜めのラインのみに非ゼロ要素が存在する疎行列が実応用上に多く見られる. このような場合, 各ライン上の非ゼロ要素を連続的に格納する Diagonal 形式[1]と呼ばれる形式を利用することができる. 本形式による疎行列ベクトル積の実装では, 間接参照を回避し, 乗算ベクトルの各要素への参照が連続となる利点がある.

そこで, 入力された疎行列が (部分的にでも) Diagonal 形式を適用できる構造を有している場合, その構造を活用することで疎行列ベクトル積を高速化することを行う[2]. 提案する手法では, 与えられた行列の各斜めのラインにおける非ゼロ要素の割合を算出し, それが一定の閾値 α 以上である場合にはその斜めのライン上の非ゼロ要素は Diagonal 形式で格納し, 残りの非ゼロ要素は CRS 形式で格納することを行う. これにより, Diagonal 形式で格納された部分の演算において計算時間が削減されることが期待される. ただし, Diagonal 形式では, 斜めのライン上のゼロ要素も非ゼロ要素として扱うため, ゼロ要素部分の演算がオーバーヘッドとなる. 従って, Diagonal 形式の適用を判別

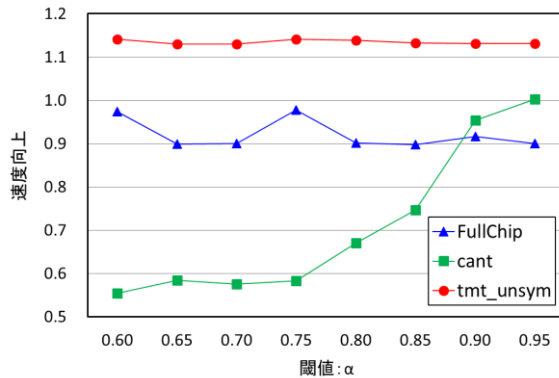


図 1. 提案手法による疎行列ベクトル積の速度向上の閾値による変化

に用いる閾値 α の設定は重要となる.

3 数値実験の概要

提案手法の効果を検証するために, 京都大学学術情報メディアセンターの CRAY CS400 2820XT (Intel Xeon Broadwell) 1 ノードを用いて数値実験を行った. プログラムは C 言語で作成し, OpenMP によるスレッド並列化を行っている.

フロリダ大学提供の疎行列データベースから取得した 3 種類の疎行列に関して, 閾値 α を 6.0 から 9.5 まで変化させて, 提案手法による速度向上 (CRS 形式のみの場合の実行時間を提案手法の実行時間で割った値) を評価した. 数値実験結果を図 1 に示す. 図 1 から分かるように, α に関わらず, tmt_unsym では有効, FullChip では逆効果となっている. 一方, cant に関しては, α によって速度向上が大きく変化した. 以上の結果から, 提案手法により疎行列ベクトル積を高速化できる事例があることが確認できた. 今後, タイリング手法の適用やチューニングより, 更なるプログラムの高性能化を図っていく予定である.

参考文献

- [1] Y. Saad: Iterative Methods for Sparse Linear Systems, SIAM, 2003.
- [2] 三浦瑛絵 他: ステンシル構造を利用した疎行列ベクトル積の高速化に関する検討, 日本応用数学会 若手の会 第 2 回学生研究発表, 東京, 2017 年 3 月 5 日.