Title	[課題研究報告書]前処理付きGMRES法のGPUに対する適合性調査
Author(s)	伊藤, 健一
Citation	
Issue Date	2022-09
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/18054
Rights	
Description	Supervisor: 井口 寧, 先端科学技術研究科, 修士(情報科学)

概 要

数値シミュレーションでは、物理現象を記述する偏微分方程式を有限差分法や有限要素法などで離散化することで得られる**大規模かつ疎な連立一次方程式** $A\boldsymbol{x} = \boldsymbol{b}$ を求解することになる。クラスタやスーパーコンピュータといった大規模な計算機ではなく、小規模ながら高い演算性能を持つ**GPU**を活用して、オンサイト環境での数値シミュレーションを行う試みが行われている。例えば、医療分野では情報管理やシミュレーションのコントロールの点から外部組織の計算機よりも自組織の計算機での計算が望まれる場合がある。

アクセラレータとしてGPUを用いるが、CPUとGPUでは特性が異なるため、適したアルゴリズムも異なることが予想される。そのため、本研究では大規模かつ疎な連立一次方程式のGPUに適した数値解法を明らかにすべく、GPU (NVIDIA A100 PCIe) と CPU (HPC System "KAGAYAKI") で前処理付き GMRES 法の評価を行った。

連立一次方程式の求解手法である反復解法は、適当に選んだ初期値から出発して、真の解に収束していく近似解の列を逐次作成していく手法である。反復解法として**クリロフ (Krylov) 部分空間**に基づく**一般化最小残差 (GMRES) 法**がある。GMRES 法は右辺ベクトル $\boldsymbol{b}$ から出発してクリロフ部分空間を拡大しながら正規直交基底 $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_{j+1}$ を生成し、各ステップにおいて残差ノルム $\|\boldsymbol{r}_j\|_2 = \|\boldsymbol{b} - A\boldsymbol{x}_j\|_2$ が最小になるように近似解 $\boldsymbol{x}_j$ を更新していく手法である。

クリロフ部分空間法の収束性は一般に係数行列 A の固有値分布に依存し、固有値分布が少なくかつ1 (単位行列) に近いほど収束が早い。収束性や安定性の改善を目的として反復に入る前にあらかじめ係数行列 A に対し**オーダリング**、**前処理**を施す。

GPU上でGMRES法での反復中に使われる直交化について**古典グラムシュミット2 (CGS2) 法**と**修正グラムシュミット (MGS) 法**の比較を行った。CUDAによる評価の結果、GPU上でのCGS2法はMGS法と比べても収束性は問題なく処理時間が大幅に短いということが分かった。

不完全 LU 分解 (ILU) 前処理付き GMRES 法について ILU(0) or ILU(1) or ILU(2) の評価を行った結果、悪条件の問題の場合は手厚い前処理 (ILU(2) など) が必要になってくることがわかった。また、ILUのフィルインを許すレベルについては、非ゼロ要素数の増加量や収束性などに応じて処理時間が最短となる適切なレベルが異なることがわかった。

オーダリングについては**逆 Cuthill-McKee (RCM)** と **Nested Dissection (ND)** についてGMRES法で評価を行った。求解対象の行列によるが概ね、CPUの場合はRCMを適用することで収束性が向上し処理時間が短縮でき、GPUの場合はRCMを適用することで収束性が向上したが処理時間は伸びた。また、GPU

の場合はオーダリングにNDを用いると求解対象の行列によるが概ね、GMRESの処理時間を短縮できることが分かった。