氏　　　　名	菱沼利彰
学位の種類	博士（情報学）
学位記番号	博甲第　９６０９　号
学位授与年月日	令和２年３月２５日
学位授与の要件	学位規則第４条第１項該当
審査研究科	図書館情報メディア研究科
学位論文題目	SIMD 演算を用いた高精度疎行列計算ソフトウェアの高速化

主　　査	筑波大学	教授	理学博士	森継　修一
副　　査	筑波大学	教授	博士（工学）	佐藤　哲司
副　　査	筑波大学	教授	工学博士	松本　紳
副　　査	筑波大学	准教授	博士（工学）	中井　央
副　　査	筑波大学	教授	博士（理学）	高橋　大介

論　文　の　要　旨（2,000 字程度）

本学位論文は、線形数値計算の中でも、疎行列に対する演算を高精度に行うためのソフトウェアを開発し、その高速化の手法について研究した成果を述べたものである。物理シミュレーションでは、「疎行列を係数とする連立一次方程式の解法」が主要な課題であり、一般に Krylov 部分空間法とよばれる反復解法アルゴリズムの中から、行列の性質に応じたアルゴリズムが選択される。その際、丸め誤差の影響で解の収束が遅くなったり、近似解が最終的に求まらなかったりする現象が発生することが問題とされている。

第１章では、上記の問題の背景を述べるとともに、収束を改善する試みとして、通常の倍精度演算よりも高精度な演算を利用する手法に着目している。高精度演算は計算時間やメモリデータ量が多く必要になることや、標準的なプログラミング環境から利用できるソフトウェアが少ないことを指摘したうえで、本研究では、4 倍精度演算の実現を目標として、Double-Double 精度(DD)演算に着目している。これは、２つの倍精度数を用いて４倍精度相当の演算を実現する手法で、倍精度演算 10～20 回を組み合わせることで実行できるが、単純には計算時間も 10～20 倍になることを意味し、実用化には計算速度の向上が不可欠となる。

そのため本論文では、現在普及しているマルチコア CPU において、Krylov 部分空間法の実装に必要な DD の基本演算を高速化することを研究目的と定めている。その中でも、疎行列ベクトル積(SpMV)および転置疎行列ベクトル積(TSpMV)という２つの基本演算に対象を絞って検討を行っている。高速化のためのアプローチとしては、現在ほとんどの汎用プロセッサでサポートされている FMA(Fused Multiply-Add)命令、SIMD(Single Instruction Multiple Data Streaming)命令、マルチスレッドを用いて、ハードウェアの性能を最大限引き出す方法を探索している。結

果として、倍精度と DD を組み合わせて SpMV および TSpMV を簡単に扱えるソフトウェアを開発し、その特性を明らかにした点が本研究の成果であるとしている。

第２章では、関連研究としての DD 演算の高速化に関する研究、疎行列とベクトルに対する演算の高速化に関する研究、Krylov 部分空間法に対する高精度演算の適用に関する研究との比較を行い、本研究の新規性を明らかにしている。

第３章では、開発するソフトウェアの設計と戦略について考察している。その結果、ユーザが Krylov 部分空間法を実装する際に、本ソフトウェアが提供すべきインタフェースを議論し、必要な機能を明らかにしている。

第４章において、本研究の成果である DD 演算および疎行列演算の高速化の手法について、計算対象ごとに詳述されている。第１節では、マルチコア CPU において、FMA 命令および SIMD 命令を利用した演算に基づく高速化手法を示している。FMA 演算とは積和演算 x × y+z を １命令で行う演算、SIMD 演算とは１命令で複数のデータに対して同時処理を行う演算であり、CPU のハードウェア性能を最大限引き出す手法について詳細な検討を行っている。

続いて、第２節では DD のスカラー演算アルゴリズム、第３節では DD を要素とするベクトルの演算の実装方法について述べている。

さらに第４節において、疎行列の格納形式について検討を行っている。疎行列の場合、メモリデータ量を節約するため、一般に零要素を除いて非零要素のみを特殊な形式で格納する。最も一般的な方法は、CRS(Compressed Row Storage)とよばれる形式であるが、SIMD 命令で実行すると計算処理性能が十分に出ないことを本論文は明らかにしている。この性能劣化要因を回避するため、４つのデータを連続して処理できる格納形式である BCRS(Block CRS)形式に着目し、比較のため、それぞれの実装を行っている。

第５章では、前章で述べた手法の有効性を確認するため、スカラー演算・ベクトル演算・疎行列ベクトル積(SpMV)・転置疎行列ベクトル積(TSpMV)のそれぞれについて、実験・測定の結果を示している。CRS 形式による行列要素の格納がもたらす性能劣化の原因を明らかにすることで、最終的に、SpMV および TSpMV において DD 演算をそれぞれ高速に実行するために最適な BCRS 格納形式(BCRS4x1)を見出すことに成功している。また、Krylov 部分空間法の代表的なアルゴリズムである BiCG 法のプログラムにおいて、収束特性とは独立に、純粋に 100 回の反復計算にかかる時間を計測した結果、DD 演算は倍精度演算の 1.1～1.3 倍程度の時間で実行できることを示し、十分実用的であると結論付けている。

最後に第６章において、本研究の成果のまとめを述べ、今後の研究の発展の方向性についても議論を行っている。付録として、開発したソフトウェア(DD-AVX v3)の機能の詳細と、これを利用した BiCG 法のプログラムなどが示されている。

審　査　の　要　旨 (2,000 字以上)

【批評】

物理シミュレーションなどにおける線形数値計算の中で、「疎行列を係数とする連立一次方程式の解法」は核となる問題である。CAD データなどの解析対象の領域を格子分割して離散化すると、次元の大きな疎行列（零要素の多い行列）が得られる。この行列を係数とする連立一次方程

式に対しては、一般に Krylov 部分空間法とよばれるカテゴリの反復解法が用いられ、行列の性質に応じて、CG 法、BiCG 法、GMRES 法などの様々なアルゴリズムが存在する。しかしながら、反復解法の性質上、丸め誤差の影響で解の収束が遅くなったり、さらに最終的に収束しなかったりする場合もあり、安定で高速な解法の追求は線形数値計算分野における主要な研究テーマである。

本学位論文は、上記の Krylov 部分空間法アルゴリズムの中で基本演算として用いられている疎行列ベクトル積(SpMV)および転置疎行列ベクトル積(TSpMV)という２つの演算を対象とし、演算プロセッサのハードウェア性能を引き出しながら高精度の演算を効率よく実行する手法を探求したものである。本論文第１章では、これらの問題の背景や高速化の基本方針、開発すべきソフトウェアの要件が述べられており、線形数値計算分野における重要な研究テーマ設定であると認められる。ただし、SpMV および TSpMV に着目した理由がやや天下り的であり、より丁寧な説明を行えば、線形数値計算分野全体において本研究の占める位置が明確になったと思われる。

第２章では、関連研究との比較が述べられている。本研究では、高精度化のため、ソフトウェアによる４倍精度演算の実現手法である Double-Double 精度(DD)演算に着目している。比較対象の関連研究としては、DD 演算の高速化に関する研究、疎行列とベクトルに対する演算の高速化に関する研究、Krylov 部分空間法に対する高精度演算の適用に関する研究などが挙げられている。いずれも本研究とは対象が異なっていることから、本研究の新規性が確認できる。

第３章では、開発するソフトウェアの設計と戦略について述べている。ユーザが Krylov 部分空間法を実装する場合を想定し、提供すべきインタフェースを示している。やや技術文書的な記述ではあるが、ソフトウェア開発の方針を定めるためには必要な項目であるといえる。

第４章第１～３節では、マルチコア CPU における DD 演算の高速化手法を論じている。本研究以前には、同時処理数４のプロセッサにおける SIMD(Single Instruction Multiple Data Streaming)命令を利用した DD 演算の高速化は行われておらず、その効率的な実装方法は未知であったため、これを解明したことは本研究における重要な成果のひとつと認められる。

さらに、第４章第４節では疎行列とベクトルに対する演算について論じている。特に、疎行列の要素を格納する方式として、従来の CRS(Compressed Row Storage)形式 と BCRS(Block CRS)形式の比較を行うため、対象とする演算である疎行列ベクトル積(SpMV)および転置疎行列ベクトル積(TSpMV)のそれぞれに対して、すべての機能を実装して緻密な検討がなされていることが認められる。

以上、第４章全体に、計算対象（スカラー・ベクトル・疎行列ベクトル積）と、高速化のための手法（FMA 命令の利用・SIMD 命令の利用・行列の格納形式の選択）がすべて盛り込まれているため、各項目の記述がやや未整理に映る点に難があり、第１～３節と第４節は独立した章として記述する方法もあったのではないかと思われる。ただし、現状でも各節ごとに検討すれば、それぞれの手法の妥当性は確認が可能である。

第５章では、第４章の記述に対応する形で、計算対象ごとに行った実験の結果を示していて、DD 演算に対する FMA 命令・SIMD 命令の適用効果を確認できる。また、疎行列とベクトルの演算に関して、最終的に、SpMV および TSpMV において DD 演算をそれぞれ高速に実行するために最適な BCRS 格納形式(BCRS4x1)を見出すことに成功したのは、本研究におけるもう一つの新規の成果である。さらに、Krylov 部分空間法の一つである BiCG 法のプログラムにおいて、100 回の反復計算から、DD による 4 倍精度演算が倍精度演算の 1.1～1.3 倍程度の計算時間で実行できたことは、本研究で開発されたソフトウェアの有用性を裏付けているといえる。これを受けて、

今後は、反復計算の収束改善効果も含めた性能評価に発展させることが期待される。

　最後に第６章において、本研究の成果のまとめが詳細に述べられており、「SIMD 演算を利用した高精度疎行列計算の高速化」という研究目的を達成していることが認められる。なお、本研究は SIMD の同時処理数が４の場合を対象としてソフトウェア開発を行ったものであるが、技術の進化に伴い、同時処理数８のプロセッサが普及しはじめている。それに合わせて、本研究の成果を移行・発展させるための検討が今後必要不可欠であるが、現時点までに解明された高精度疎行列演算の高速化手法の特性は、その基盤として重要な学術的意義をもつものと認められる。

　以上から、本論文に示された著者の研究は、線形数値計算における「疎行列計算の高精度かつ高速な実行手法」の開発に大きく貢献するものと高く評価できる。したがって、総合的に判断して、本論文は、情報学の学位論文として十分な内容をもつものと認められる。

【最終試験結果】

　2020 年 1 月 27 日、図書館情報メディア研究科学位論文審査委員会において、審査委員全員出席のもと、本論文について著者に説明を求めた後、関連事項について質疑応答を行った。引き続き、「図書館情報メディア研究科博士後期課程（課程博士）の学位論文審査に関する内規」第 23 項第 3 号に基づく最終試験を行い、審議の結果、審査委員全員一致で合格と判断された。

【結論】

　よって、著者は博士（情報学）の学位を受けるに十分な資格を有するものと認める。