

# 京における密行列固有値ソルバ Eigen-K の性能評価と性能モデリング

深谷 猛<sup>†,†††</sup> 今村 俊幸<sup>††,†††</sup> 山本 有作 <sup>†,†††</sup>

## 1. はじめに

密行列の固有値計算は、量子化学計算や統計計算など、様々な科学技術計算の分野で必要とされる重要な行列計算の一つである。代表的な密行列向け固有値計算ソルバとしては ScaLAPACK<sup>1)</sup> が有名であるが、近年、京コンピュータをはじめとしたスーパーコンピュータ上での ScaLAPACK の性能限界が指摘されている。そのため、ScaLAPACK にかわる新しいソルバの開発が望まれており<sup>2)</sup>、国内では今村らによる Eigen-K<sup>3)</sup>、海外では DPLASMA<sup>4)</sup> や ELPA<sup>5)</sup>、といったソルバの開発が進んでいる。そこで、本論文では開発中の Eigen-K に関して、

- 京コンピュータ上で行った性能評価結果を、ScaLAPACK と比較しながら報告する。特に、両者の性能差の主な要因となっている三重対角化部分の性能について詳しく述べる。
- 固有値ソルバの開発に活用するために構築中の性能予測モデルを紹介する。そして、実際にモデルにより性能予測をした結果を示し、課題について議論する。

## 2. 対称密行列の固有値計算

対称密行列の固有値計算の一般的な手順は、(1) 行列の三重対角化、(2) 三重対角行列の固有値・固有ベクトル計算、(3) 固有ベクトルの逆変換、である。このうち、性能ネックになっているのは (1) の三重対角化で、その主な原因は、計算が逐次的であることと演算の半分程度が実効性能が低い行列ベクトル積によって行われることにある。そのため、先に挙げた DPLASMA<sup>4)</sup> や ELPA<sup>5)</sup> では、帯行列を経由する 2 ステップ型の三重対角化を採用し、実効性能の高い行列積を有効利用することで性能向上を図っている。しかしながら、2 ステップ型の三重対角化を行うと、(3) の固有ベクトルの逆変換の計算量が 2 倍になってしまうため、求め

る固有ベクトルの本数が多い場合は逆効果になってしまう可能性がある。

## 3. Eigen-K の概要

Eigen-K<sup>3)</sup> は、今村らによって開発中の固有値ソルバで、対称密行列の全固有値・全固有ベクトル計算を主な対象としている。前述のように、2 ステップ型の三重対角化を採用すると、全固有ベクトルを計算する場合には、固有ベクトルの逆変換のコスト増が無視できなくなるので、Eigen-K では、ScaLAPACK と同様の、1 ステップ型の三重対角化の方法を採用している。本論文で用いた Eigen-K のプログラムでは、(1) と (3) はハウスホルダー変換による三重対角化と逆変換を開発者が独自に実装したのになっており、一方、(2) は ScaLAPACK の分割統治法のルーチンを京向けに移植したものとなっている。

## 4. 性能評価

Eigen-K と ScaLAPACK (PDSYEVD) を用いて 100,000 次元の対称密行列の固有値計算を京コンピュータ上で実行した際の、強スケーリングの様子と 4,096 ノード使用時の各計算ステップの占める時間を図 1 に示す。この図より、開発中の Eigen-K は ScaLAPACK よりも高性能であり、性能の主な差は三重対角化部分にあることが確認できる。次に、Eigen-K の三重対角化における各 BLAS レベルの計算時間と演算・通信の占める割合を図 2 に示す。これより、計算の主要部である Lv.2, 3 BLAS 演算がよくスケーリングしている一方で、並列数を増やすと通信時間がボトルネックになっていることが確認できる。

## 5. 性能モデリング

今回は、「並列実行時間が演算時間と通信時間の和で表される」という仮定の下、演算時間と通信時間について以下の方針でモデリングを行う。

### ● 演算時間

三重対角化の計算では行列の成分の値は計算時間に影響を与えないので、プログラムから MPI 通信

† 神戸大学 大学院システム情報学研究所  
†† 理化学研究所 計算科学研究機構  
††† 独立行政法人科学技術振興機構, CREST

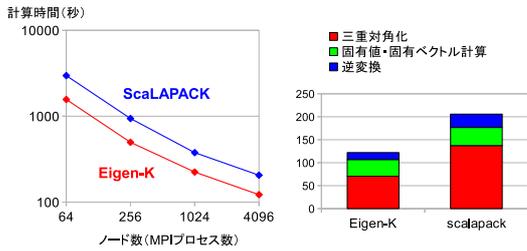


図 1 Eigen-K と ScaLAPACK による京コンピュータ上での 100,000 次元の行列の固有値計算時間の強スケーリング (左) と 4,096 ノードでの並列計算時の計算時間の内訳 (右)

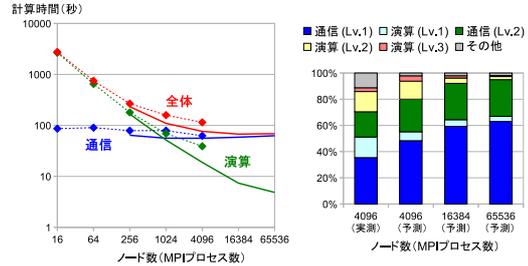


図 3 Eigen-K による京コンピュータ上での 100,000 次元の行列の三重対角化の計算時間について、モデルによりスケールの様子を予測した結果 (左, 実線: 予測, 点線: 実測) と時間の内訳を予測した結果 (右)

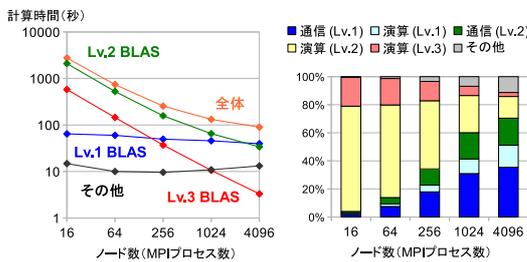


図 2 Eigen-K による京コンピュータ上での 100,000 次元の行列の三重対角化において各 BLAS 演算ごとの時間 (左) と全体に対する演算時間・通信時間が占める割合 (右)

の部分を除き、仮想的に並列計算時と同様の計算 (計算結果は全くの無意味) を 1 プロセスで行うベンチマークプログラムを作成する。そして、このベンチマークプログラムの実行結果を用いて、並列計算時の演算時間を求める。

● 通信時間

Eigen-K では行列データを二次元分散で各プロセスが保持しており、通信は二次元のうちの一次元方向の集団通信を並列に行う形となっている。つまり、一つの通信に参加するプロセス数は高々全プロセス数の平方根程度となるので、MPI のベンチマークが実行可能だと判断できる。そこで、今回はベンチマークの実行結果を利用して通信時間のモデル式のパラメータを決定する。

上記の方法でモデリングを行い、実際に 100,000 次元の行列の三重対角化の計算時間について予測した結果を図 3 に示す。図 3 よりスケールの様子については、ある程度の誤差はあるものの、おおよその傾向は予測できているのではないかと考えている。また、右の割合を示したグラフから、Lv.2 および 3 の BLAS に関しては比較的精度良く予測できているが、Lv.1 BLAS については予測精度があまり良いとは言えない。この誤差の原因については、現在調査中である。

6. おわりに

本論文では、現在開発が進められている密行列固有値ソルバ Eigen-K の京コンピュータ上での性能評価と性能モデリングに関する結果について報告した。ポスター発表当日は、モデリングの詳細および、本論文で紹介できなかった性能評価結果とモデルによる予測結果についても発表する予定である。また、発表までに今回示したモデルの誤差の原因について詳しく調査し、発表当日に議論を行いたい。

謝辞 本研究の一部は JST CREST プロジェクト「ポストペタスケールに対応した階層モデルによる超並列固有値解析エンジンの開発」(領域名: ポストペタスケール高性能計算に資するシステムソフトウェア技術の創出) および東京大学情報基盤センターの「若手・女性利用者推薦」の援助を受けています。

参考文献

- 1) ScaLAPACK, <http://www.netlib.org/scalapack/>
- 2) 今村 俊幸ほか: ポスト・ペタスケール時代の密固有値計算ソルバについて, 日本応用数理学会 2012 年度年会 予稿集, pp. 279-280 (2012).
- 3) Imamura, T.: Eigen-K: High Performance Eigenvalue Solver for Symmetric Matrices Developed for K Computer, PMAA2012, Londo (2012).
- 4) Bosilca, D. et al.: Distributed Dense Numerical Linear Algebra Algorithms on Massively Parallel Architectures: DPLASMA, University of Tennessee Computer Science Technical Report, UT-CS-10-660 (2010).
- 5) Auckenthaler, A.: Parallel Solution of Partial Symmetric Eigenvalue Problems from Electronic Structure Calculations, Parallel Computing, Vol. 27, 12, pp. 783-794 (2011).