古典的アルゴリズムによる大規模対称帯行列の固有値計算

EIGENVALUE COMPUTATION FOR LARGE SYMMETRIC BAND MATRICES WITH LEGACY ALGORITHMS

長谷川 秀彦[1)]
Hidehiko HASEGAWA

[1)]博士（工学）筑波大学 図書館情報メディア研究科（〒 305-8550 つくば市春日 1-2, hasegawa@slis.tsukuba.ac.jp）

The author shows that the eigenvalue analysis for large symmetric band matrices, whose dimension is less than or equal to 10^5, can be performed on a PC in acceptable time. Some legacy but stable algorithms are used for this analysis. The programs waste computation time and storage area for zero elements, however this apporoach can be applicable for any symmetric sparse metrices as a band matrix. The listed timing result will be assumed to be a refarencce data. Beacuase the eigenvalue analysis for matrices strongly depends on their mathematical property.
Key Words : *eigenvalue computation, symmetric matrices, band matrices*

1. はじめに

これまで，メモリ容量と計算時間の制約から，大規模固有値解析は問題ごとに個別なプログラムを使用する必要があった．コンピュータの大容量化と高速化によって，素朴なプログラムを用いた大規模固有値解析が現実的な時間で実行できるようになりつつある．

本報告では，対称帯行列に対する古典的な固有値解析アルゴリズムが現在のコンピュータ，1 CPU でどの程度の時間で実行できるか示す．対象とするアルゴリズムは，対称帯行列に対する 3 重対角化，2 分法・逆反復法を組み合わせた村田法と対称帯行列に対して直接 2 分法と逆反復法を適用するスツルム逆反復法である．

2. 対称行列帯に対するアルゴリズム

対称行列に対する固有値解析アルゴリズムは，直交変換（多くの場合はハウスホルダー変換）を用いて 3 重対角化を行い， 3 重対角行列に対して 2 分法あるいは QR 法によって，固有値と固有ベクトルを求める．元の行列と 3 重対角行列の固有値は等しいので，固有値だけが必要ならここで計算を打ち切ればよい．特定の区間の固有値あるいは最大・最小から p 個の固有値が必要な場合は 2 分法を用いれば効率的に求められる．固有ベクトルは 3 重対角行列の固有ベクトルに逆変換を施せば求められる．大規模な密行列に対する固有値解析，特に並列計算環境を意識した高速化の研究は，片桐 [1]，直野 [2] などが精力的に進めている．しかし，このような密行列を扱う算法は，

(1) N * N の行列を格納するメモリ容量

(2) $O(N^4)$ 以上の演算

が必要であり，行列の次元 N が 10^5 を越えると並列計算環境でも実行が難しい．

一方，対称帯行列に対しても，特殊なアルゴリズムを用いれば帯幅を大きく増大させることなく 3 重対角化できる．この 3 重対角化アルゴリズムは，直交変換の回数が異なるだけで，プログラム上の変更（データ構造，変換順序など）はあっても本質的なアルゴリズムに違いはない．3 重対角行列が得られてしまえば 2 分法または QR 法によって固有値が求められる．変換の記憶コストのため，逆変換によって元の行列の固有ベクトルを求めるのは不可能なので，2 分法で得られた固有値シフト点として帯行列に対して逆反復法を行う．これが「村田法 (mura2)」である．2 分法と逆反復法は，求める固有値・固有ベクトルの数に応じたメモリ容量と演算時間ですむが，固有値を一つだけ求める場合でも 3 重対角化は必要である．いずれのステップでも計算に必要なメモリ容量は 3*m1*N であり，この他に行列の保持用に m1*N と固有ベクトル用の領域が必要となる (m1 は帯半幅， N は次元)．

2 分法では，与えられた値 t に対して t 以下（あるいは以上）の固有値がいくつあるかを知る必要がある．3 重対角行列に対しては，スツルム列の性質を利用した簡単な固有値カウントアルゴリズムが知られているが，同様な固有値カウントを帯行列に対して直接行うのが「スツルム 2 分法 (eigv2)」である．帯行列に対して部分軸選択付きガウスの消去法と同等な方法で固有値カウントをするのが Martin-Wilkinson の特殊ガウスで，多少の精度を犠牲にして軸選択なしの対称ガウスと同等な方法を用いれば演算量を 1/4 に削減できる．帯行列の固有値の存在区間がわかれば，あとは帯行列に対して 2 分法で求められた値をシフト点とした逆反復法を行えばよい．

3. 大規模固有値解析をパソコンで

固有値解析の計算時間は，数学的性質，求める固有値・固有ベクトルの範囲，精度などに強く依存する．テスト問題は場の物理定数に違いがある長方形領域における拡散方程式を差分法（Control Volume 法）で離散化して得られる対称正定値行列である．分割を m1 とすると，帯半幅 m1, 帯幅 2*m1+1, 次元数 N = m1*(2*m1+3) の帯行列が得られる．テストでは最小固有値から 2*m1 組の固有値・固有ベクトルの計算にかかる時間を計測した．今回の目的は，計算時間のおおよその目安を得ることなので，精度評価は行わず文献 [3] の Fortran プログラムをそのまま用いた．

計算環境はパソコンまたは並列計算環境の 1 CPU で，メモリ容量は 2GB とした．具体的には PowerPC 2GHz を二つ搭載した Apple Machintosh G5, Intel Xeon Dual 3.2GHz を搭載したクラスタの 1 CPU 相当の部分，Intel Itanium2 1.3GHz(Madison) から構成される cc-NUMA の共有メモリ方式並列コンピュータ SGI Altix 3700 の 1 CPU, POWER5 1.9GHz から構成される共有メモリ型並列コンピュータ IBM eServer p5 モデル 595 の 1 CPU を用いた．Intel の CPU では Intel コンパイラ， PowerPC と POWER5 では IBM 系のコンパイラを用いた．詳細な評価には，実行方式，キャッシュの構成方式，コンパイラオプションなども必要だが，今回は標準的な最適化オプションとした．

4. 実測結果

実測結果を表-1, 表-2 に示す．m1 = 200, N = 80,600 の帯行列に対して，小さいほうから 400 組の固有値・固有ベクトルを求めるのに，POWER5 1.9GHz では 3 時間， PowerPC 2GHz では 4 時間， Xeon 3.2GHz と Itanium2 1.3GHz では 5 時間かかる．PowerPC 2GHz では m1 = 250, N = 125,750 にすると約 4 倍（それでも 1 日以内!)，m1 = 300, N = 180,900 にすると約 15 倍の計算時間になる．もし 10 組の固有値・固有ベクトルが計算したいのなら，スツルム逆反復法の計算時間は 40 分の 1 になり， m1 = 200 の問題では 570 秒（10 分以内！）である．これから，ざっと 1 日の計算時間で古典的なアルゴリズムを用いた 10^5 次元の固有値解析ができることがわかる．問題はメモリサイズで， m1 = 300, 180,900 の場合にほぼ 2GB になる．

共有メモリ方式の並列コンピュータの使用は，並列実行されなくても，使用できるメモリ容量の増大というメリットがある．並列化が効果的なのは，帯行列に対する 3 重対角化，帯行列に対する 2 分法，帯行列に対する逆反復法である．いろいろな並列化方針が考えられるが，帯行列に対する 2 分法と逆反復法では，内部で使われている部分軸選択付きのガウスの消去法とそれによく似た Martin-Wilkinson の特殊ガウスの並列化が考えられる．しかし Martin-Wilkinson の特殊ガウスと帯行列に対する 3 重対角化はデータ構造とアルゴリズムが複雑なため，並列化による性能向上が難しい．分散メモリ方式の場合，問題サイズが 2 GB を越えるような場合は問題分割が必須となり，並列化が非常に困難となる．

一方，スツルム同時逆反復法を用いれば，問題依存性が強くなるために精度劣化の可能性があるが，高速化は可能である．数値的な安定性は要検討だが，スツルム逆反復法は一般固有値問題 $Ax = \lambda Mx$ （ただし，A, M は帯行列）に容易に拡張できる．

表–1 村田法の計算時間 m1=200, N=80,600, 秒

CPU	3 重対角化	2 分法	逆反復	合計
PoowePC	13,000	11.6	1,310	14,300
Xeon	14,100	10.2	1,240	15,400
Itanium2	14,000	15.8	1,470	15,500
POWER5	9,350	10.4	1,070	10,400

表–2 スツルム逆反復法の時間 m1=200, N=80,600, 秒

CPU	2 分法	逆反復	合計
PowerPC	15,200	7,630	22,800
Xeon	10,700	5,230	16,000
Itanium2	11,000	4,240	15,200
POWER5	8,090	2,810	10,900

5. おわりに

疎行列の非ゼロ構造を意識せず，多少の無駄を許して帯行列とみなすことで，古典的で安定な対称帯行列用固有解析プログラムが利用できる．コンピュータの大容量化＆高速化によって，2GB のメモリと 1 日程度の計算時間で帯半幅 300, 10^5 次元の帯行列の固有値解析が実行できる．したがって，疎構造を活用した最新の（問題依存の強い，安定性が定かではない，高速化可能な）固有値解析アルゴリズムのメインターゲットは 100 万次元以上の問題といえるだろう．

謝辞: 科学技術振興事業団戦略的創造研究推進事業 (CREST) 「大規模シミュレーション向け基盤ソフトウェアの開発」プロジェクトの支援を得た．

参考文献

1) 片桐孝洋, 金田康正 :超並列処理に向く効果的な並列固有値計算法, 情報処理学会論文誌, Vol. 41, No. 5, pp. 1558-1566, 2000
2) 直野健, 今村俊幸 :自動チューニング型の固有値ソルバーについて, 情報処理学会研究報告, Vol. HPC-91-9, No. 80, pp. 49-54, 2002
3) 小国力編著: 行列計算ソフトウェア, 丸善, 1991.