数理解析研究所講究録
1535 巻 2007 年 45-56

密行列固有値解法の最近の発展
– マルチシフト QR 法とその収束理論に向けて –

名古屋大学大学院　計算理工学専攻　山本有作
Yusaku Yamamoto, Department of Computational Science & Engineering, Nagoya University

1 はじめに

行列の固有値問題は，科学技術計算における中心的な問題の一つである．本論文では，特に $n \times n$ の実対称密行列に対する標準固有値問題

$$A\mathbf{x} = \lambda\mathbf{x} \tag{1}$$

を考える．このタイプの問題は，分子軌道法，統計計算など様々な分野で現れる．近年では，シミュレーションの大規模化に伴い，解くべき行列も大型化しており，たとえば分子軌道法によるたんぱく質の解析では，$n = 10^5$ 程度の行列の固有値計算が必要とされつつある．このような大規模問題を実用的な時間で解くには，効率的なアルゴリズムに加え，並列計算機をはじめとする高性能な計算機の利用が必須である．

実対称密行列の固有値を求める標準的な方法では，まず行列 A を相似変換により対称三重対角行列 $A^{(0)}$ に変換してから，$A^{(0)}$ の固有値を求める [7][11]．$A^{(0)}$ への相似変換のアルゴリズムとしては，通常，鏡像変換に基づくハウスホルダー法が使われる．一方，$A^{(0)}$ の固有値を計算するアルゴリズムとしては，二分法 [24]，QR 法 [9][10][16]，分割統治法 [12]，dqds 法 [8]，mdLVs アルゴリズム [13] など，多くの方法が提案されている．中でも QR 法は，長い歴史があるとともに収束性の理論的解明も進んでおり [20]，もっとも信頼性の高い解法として広く使われている．

しかし，三重対角行列に対する QR 法は以下の章で述べるように本質的に逐次的なアルゴリズムであり，そのままの形では並列計算機を利用することが困難である．そのため，アルゴリズムの書き換えによる並列化の試みがいくつか行われてきた．たとえば Sameh らは，QR 法の計算に出てくる非線形の漸化式を変数の置き換えで線形化することにより，$O(n)$ 個のプロセッサを使って計算時間を $O(\log n/n)$ にする方法を提案した [21]．しかし，この方法では誤差が n に関して指数的に増加するという問題点があり，実用化は困難である．一方，Bar-on らは三重対角行列をブロックに分割して並列化する方法を提案している [2]．この方法は数値的安定性が証明されており，興味深いが，演算量の詳しい評価は行われておらず，また，並列計算機上での実装と性能評価も行われていない．

本論文では，QR 法の並列化を実現するもう一つの方法であるマルチシフト QR 法について紹介する．マルチシフト QR 法では，QR 法において複数のステップを同時に行えるようにアルゴリズムを変更する．この結果，各ステップをそれぞれ 1 個のプロセッサに割り当てることにより，並列計算機の利用が可能になる．マルチシフト QR 法はもともと Bai ら [1] により非対称行列向けに提案されたが，Kaufman[15]，van de Geijn[22]，宮田 [17][18][19] らの研究により，対称三重対角行列に対しても有効であることがわかってきた．

以下では，第 2 章でまず対称三重対角行列向けの QR 法について述べ，その拡張としてのマルチシフト QR 法を紹介する．また，deferred shifting scheme[22]，fully pipelined shifting scheme[17][18][19] など，並列性能を向上させるために提案された変種についても紹介する．第 3 章では，マルチシフト QR 法に関する既存の収束理論について述べた後，様々な変種の収束性を統一的に解析するために有効と考えられる中心多様体理論について説明する．最後に第 4 章でまとめと今後の課題を述べる．なお，本論文では非対称行列に対するマルチシフト QR 法については扱わないが，これについても small-bulge マルチシフト QR 法 [3]，aggressive early deflation[4] など，最近アルゴリズム上の大きな進展があった．これらについては，たとえばサーベイ論文 [25] などを参照されたい．

2 マルチシフト QR 法のアルゴリズム

2.1 従来の QR 法

いま，$A^{(0)}$ を $n\times n$ の対称三重対角行列とする．QR 法では，行列 $A^{(0)}$ から出発して次のように QR 分解と行列 Q による相似変換を繰り返してゆく．

$$
\begin{aligned}
A^{(0)} &= Q^{(0)}R^{(0)} \\
A^{(1)} &= R^{(0)}Q^{(0)} \quad \left(=\left(Q^{(0)}\right)^{-1}A^{(0)}Q^{(0)}\right) \\
A^{(1)} &= Q^{(1)}R^{(1)} \\
A^{(2)} &= R^{(1)}Q^{(1)} \quad \left(=\left(Q^{(1)}\right)^{-1}A^{(1)}Q^{(1)}=\left(Q^{(0)}Q^{(1)}\right)^{-1}A^{(0)}Q^{(0)}Q^{(1)}\right)
\end{aligned} \tag{2}
$$

このとき，適当な条件の下で行列 $A^{(k)}$ は対角行列に収束する [20]．より詳しくは，$A^{(0)}$ の固有値を絶対値の大きいほうから順に $\lambda_1,\lambda_2,\ldots,\lambda_n$ とするとき，$A^{(k)}$ の副対角要素 $a_{i,i-1}^{(k)}$（$i=2,3,\ldots,n$）は収束率 $|\lambda_i|/|\lambda_{i-1}|$ で 0 に 1 次収束する．

QR 法では，収束を加速するため，ある固有値 λ_i の近似値を $s^{(k)}$ として行列 $A^{(k)}-s^{(k)}I$（I は単位行列）に対して次のように算法を適用するのが普通である．

$$
\begin{aligned}
A^{(k)}-s^{(k)}I &= Q^{(k)}R^{(k)}, \\
A^{(k+1)} &= R^{(k)}Q^{(k)}+s^{(k)}I \quad \left(=\left(Q^{(k)}\right)^{-1}A^{(k)}Q^{(k)}\right).
\end{aligned} \tag{3}
$$

このとき，非対角要素の収束率は $|\lambda_i-s^{(k)}|/|\lambda_{i-1}-s^{(k)}|$ となり，$s^{(k)}$ がよい近似値であれば，この値は 0 に近くなって収束を加速できる．$s^{(k)}$ としては，$A^{(k)}$ の右下の対角要素 $a_{nn}^{(k)}$ を用いる方法（レイリー商シフト），$A^{(k)}$ の右下隅の 2×2 の小行列の固有値のうち，$a_{nn}^{(k)}$ に近い方を用いる方法（Wilkinson シフト）などが使われている [20]．

以下では，次章で使うため，QR 法の各ステップの演算をより詳しく見ていく．シフトなしの場合で説明するが，シフトを用いる場合は最初の $A^{(k)}$ を $A^{(k)}-s^{(k)}I$ に置き換えればよい．まず，行列 $A^{(k)}$ の QR 分解は，$A^{(k)}$ に左から適当な Givens 回転行列

$$
G=\begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \tag{4}
$$

（を $n\times n$ に拡大したもの）をかけ，非対角要素を左上から順に消去していくことにより行う．たとえば最初は $a_{21}^{(k)}$ を消去する．このためには，$A^{(k)}$ の第 1 行，第 2 行に G をかければよい．この結果，$A^{(k)}$ の第 1 行，第 2 行は，

$$
\begin{bmatrix}
a_{11}^{(k)}\cos\theta+a_{21}^{(k)}\sin\theta & a_{12}^{(k)}\cos\theta+a_{22}^{(k)}\sin\theta & a_{23}^{(k)}\sin\theta & 0 & \ldots & 0 \\
-a_{11}^{(k)}\sin\theta+a_{21}^{(k)}\cos\theta & -a_{12}^{(k)}\sin\theta+a_{22}^{(k)}\cos\theta & a_{23}^{(k)}\cos\theta & 0 & \ldots & 0
\end{bmatrix} \tag{5}
$$

となる．そこで，第 $(2,1)$ 要素を 0 にするには，

$$
-a_{11}^{(k)}\sin\theta+a_{21}^{(k)}\cos\theta=0 \tag{6}
$$

となるように θ を決めればよい．すなわち，

$$
\begin{aligned}
\cos\theta &= \frac{a_{11}^{(k)}}{\sqrt{\left(a_{11}^{(k)}\right)^2+\left(a_{21}^{(k)}\right)^2}}, \\
\sin\theta &= \frac{a_{21}^{(k)}}{\sqrt{\left(a_{11}^{(k)}\right)^2+\left(a_{21}^{(k)}\right)^2}}
\end{aligned} \tag{7}
$$

とすればよい．こうして第 $(3,2)$ 要素，第 $(4,3)$ 要素，...，第 $(n,n-1)$ 要素を順に消去してゆき，$n-1$ 個の Givens 回転をかけることで $A^{(k)}$ を上三角行列 $R^{(k)}$ に変形できる．

いま，$A^{(k)}$ を $A^{(k,0)}$ と書き直し，第 k ステップにおける i 番目の Givens 回転をかけた後の行列を $A^{(k,i)}$ と書く．また，i 番目の Givens 回転を

$$G_i = \begin{bmatrix} c_i & s_i \\ -s_i & c_i \end{bmatrix} \quad (c_i^2 + s_i^2 = 1) \tag{8}$$

と書く．$A^{(k)}$ に作用させるのは G_i を $n \times n$ に拡大した行列であるが，以下では混同の恐れのない限り，後者も同じ記号 G_i で表す．このとき，$A^{(k)}$ から $R^{(k)}$ を計算する詳しい手順は次のようになる．

[アルゴリズム 1: $A^{(k)}$ から $R^{(k)}$ の計算]

do $i = 1,\ n-1$

$c_i = a_{ii}^{(k,i-1)} \Big/ \sqrt{\left(a_{ii}^{(k,i-1)}\right)^2 + \left(a_{i+1,i}^{(k,i-1)}\right)^2}$

$s_i = a_{i+1,i}^{(k,i-1)} \Big/ \sqrt{\left(a_{ii}^{(k,i-1)}\right)^2 + \left(a_{i+1,i}^{(k,i-1)}\right)^2}$

$a_{ii}^{(k,i)} = c_i a_{ii}^{(k,i-1)} + s_i a_{i+1,i}^{(k,i-1)}$

$a_{i+1,i}^{(k,i)} = 0$

$a_{i,i+1}^{(k,i)} = c_i a_{i,i+1}^{(k,i-1)} + s_i a_{i+1,i+1}^{(k,i-1)}$

$a_{i+1,i+1}^{(k,i)} = -s_i a_{i,i+1}^{(k,i-1)} + c_i a_{i+1,i+1}^{(k,i-1)}$

if $i \neq n-1$ **then**

$a_{i,i+2}^{(k,i)} = s_i a_{i+1,i+2}^{(k,i-1)}$

$a_{i+1,i+2}^{(k,i)} = c_i a_{i+1,i+2}^{(k,i-1)}$

end if

end do

ただし，アルゴリズム中で陽に更新されていない要素については，$a_{lj}^{(k,i)} = a_{lj}^{(k,i-1)}$ とする．アルゴリズム 1 の終了後に得られる $a_{ij}^{(n-1)}$ が上三角行列 $R^{(k)}$ の要素となる．また，直交行列 $Q^{(k)}$ は

$$\left(Q^{(k)}\right)^T = G_{n-1} \cdots G_2 G_1 \tag{9}$$

（上付きの T は転置を表す）により与えられる．

一方，$A^{(k+1)} = R^{(k)}Q^{(k)}$ の計算は次のようになる．

[アルゴリズム 2: $R^{(k)}$ から $A^{(k+1)}$ の計算]

do $i = 1,\ n-1$

$a_{ii}^{(k,n+i-1)} = c_i a_{ii}^{(k,n+i-2)} + s_i a_{i,i+1}^{(k,n+i-2)}$

$a_{i,i+1}^{(k,n+i-1)} = -s_i a_{ii}^{(k,n+i-2)} + c_i a_{i,i+1}^{(k,n+i-2)}$

$a_{i+1,i}^{(k,n+i-1)} = s_i a_{i+1,i+1}^{(k,n+i-2)}$

$a_{i+1,i+1}^{(k,n+i-1)} = c_i a_{i+1,i+1}^{(k,n+i-2)}$

if $i \neq 1$ **then**

$a_{i-1,i}^{(k,n+i-1)} = c_i a_{i-1,i}^{(k,n+i-2)} + s_i a_{i-1,i+1}^{(k,n+i-2)}$

$a_{i-1,i+1}^{(k,n+i-1)} = 0$

end if

end do

ここで，c_i，s_i はアルゴリズム 1 で与えられる三角関数の値である．また，右から第 i 番目の Givens 回転 G_i^T をかけた後の行列を $A^{(k,n+i-1)}$ と書いた．また，$a_{i-1,i}^{(k,n+i-1)} = 0$ となることは，$A^{(k+1)}$ の下帯幅が 1 になることと，$A^{(k+1)}$ が対称行列であることからわかる．以上により，QR 法の 1 ステップの計算が完了する．

2.2 マルチシフト QR 法

いま，シフトなしの QR 法について考える．上記のアルゴリズム 1 では，do ループの第 i 番目の繰り返しにおいて，行番号が i から $i+1$，列番号が i から $i+2$ の間にある要素のみが更新される．一方，アルゴリズム 2 では，第 i 番目の繰り返しにおいて，行番号が $i-1$ から $i+1$，列番号が i から $i+1$ の間にある要素のみが更新される．このように更新の範囲が局限されているため，アルゴリズム 1 で $i=3$ の繰り返しが終了したら，アルゴリズム 2 を開始することができる．さらに，アルゴリズム 2 で $i=3$ の繰り返しが終了したら，次の k に対するアルゴリズム 1 を開始することができる．このようにして，シフトなしの QR 法では，パイプライン式の並列化を行うことが可能である．

しかし，シフトを導入した場合は状況が異なる．レイリー商シフトも Wilkinson シフトも，シフト $s^{(k)}$ の計算には行列 $A^{(k)}$ の右下隅の要素を用いる．この要素は $k-1$ に対するアルゴリズム 2 の最後の繰り返しで計算されるから，$k-1$ に対するアルゴリズム 2 の最後の繰り返しが終わるまでは，k に対するアルゴリズム 1 は開始できないことになる．このように，従来のシフト付き QR 法では，パイプライン式の並列化を行えない．

この問題点を解決するため，Bai らはマルチシフト QR 法を提案した [1]．Bai らの元々のアルゴリズムは非対称行列向けであるが，考え方は対称三重対角行列の場合でも同じである．マルチシフト QR 法では，$A^{(k)}$ の右下隅の $m \times m$ 行列の m 個の固有値 $s_1^{(k)}, s_2^{(k)}, \ldots, s_m^{(k)}$ を計算し，これらを順にシフトとして用いて $A^{(k)}$ から $A^{(k+m)}$ を計算する．計算式は次のようになる．

$$\begin{aligned}
A^{(k)} - s_1^{(k)} I &= Q^{(k)} R^{(k)} \\
A^{(k+1)} &= \left(Q^{(k)}\right)^{-1} A^{(k)} Q^{(k)} \\
&\vdots \\
A^{(k+m-1)} - s_m^{(k)} I &= Q^{(k+m-1)} R^{(k+m-1)}
\end{aligned}$$

$$A^{(k+m)} = \left(Q^{(k+m-1)}\right)^{-1} A^{(k+m-1)} Q^{(k+m-1)}. \tag{10}$$

マルチシフト QR 法では，最初に m 個のシフトを計算しておくため，従来のシフト付き QR 法における演算の依存関係が解消され，$A^{(k+1)}, A^{(k+2)}, \ldots, A^{(k+m)}$ の計算をパイプライン式に並列化できる．シングルシフト QR 法では，通常，反復を行うことにより右下の 1×1 の対角ブロックが分離され，1 個の固有値が求まる．これに対し，マルチシフト QR 法では，$m \times m$ 程度の大きさの対角ブロックが分離され，このブロックの固有値を求めることにより，元の行列 $A^{(0)}$ の m 個（程度）の固有値が同時に求まる．

2.3 マルチシフト QR 法の変種

マルチシフト QR 法では，$A^{(k+1)}, A^{(k+2)}, \ldots, A^{(k+m)}$ の計算を m 個のプロセッサを用いてパイプライン式に行う．このとき，正しく計算が行われるためには，プロセッサの間で同期を行い，他のプロセッサが更新している領域を同時にアクセスしないようにする必要がある．SIMD（Single Instruction Multiple Data）型の並列計算機では，もともと 1 サイクル毎にプロセッサ間の同期が行われるため，新たなオーバーヘッドは発生しない．この結果，マルチシフト QR 法が有効に働くことが確認されている [15]．しかし，最近主流となっている SPMD（Single Program Multiple Data）型の並列計算機では，プロセッサ間同期に多くの時間がかかり，これがマルチシフト QR 法の性能を制約する要因になる．

この問題点を解決するため，いくつかの改良法が提案されている．たとえば，1 個のプロセッサが演算を開始してから次のプロセッサが演算を開始するまでの間隔を長くすることで，プロセッサの待ち時間を長くする代わりに，同期の回数を削減できる．宮田らは，この方法において，待ち時間と同期オーバーヘッドのトレードオフから，最適な間隔を理論的に求めている [18][19]．一方，van de Geijn は，より古いシフト，たとえば $A^{(k-m)}$ から計算したシフトを $A^{(k+1)}, A^{(k+2)}, \ldots, A^{(k+m)}$ の計算で使うことを提案した [22]．この方式は deferred shifting scheme と呼ばれ，プロセッサの待ち時間を増加させることなく，同期オーバーヘッドを削減できる．しかし，古いシフトを使うことにより，マルチシフト QR 法の収束性は低下する．この問題を解決するため，宮田らは fully pipelined shifting scheme と呼ばれる方式を提案した [18][19]．この方式では，1 回の QR 分解と相似変換が終わる毎に，右下の $m \times m$ 小行列の固有値を計算し，そのうちの 1 個をシフトとして用いる．この方式では，シフトの計算量は従来のマルチシフト QR 法の m 倍となるが，収束性をあまり落とさずに，プロセッサの待ち時間と同期オーバーヘッドの両方を低減できる．これらの改良法の詳細については，たとえば [18][19][25] を参照されたい．また，[18][19] では，前節で述べた従来のマルチシフト QR 法，deferred shifting scheme，fully pipelined shifting scheme の 3 種の解法の性能比較が行われている．

3 収束性の理論的解析

3.1 収束性に関する従来の結果

2.2 節で定義したマルチシフト QR 法については，次の収束定理が知られている [23]．

定理 1 シフト $s_1^{(k)}, s_2^{(k)}, \ldots, s_m^{(k)}$ を $A^{(k)}$ の右下隅の $m \times m$ 小行列の m 個の固有値に取ったとき，マルチシフト QR 法は局所的に 3 次収束する．特に，$Q^{(0)}, Q^{(1)}, Q^{(2)}, \ldots, Q^{(k+m-1)}$ の積を Q_{k+m-1} とするとき，Q_{k+m-1} の最後の m 列の張る部分空間は，$A^{(0)}$ のある不変部分空間に 2 次収束する．

また，deferred shifting scheme では収束性が低下するが，これに関しては van de Geijn が次の定理を示している [22]．

定理 2 h 回だけ古いシフトを使う deerred shifting scheme の収束次数 τ_h は，非線形方程式

$$t^{h+1} - t^t - 2 = 0 \tag{11}$$

の唯一の正の解として与えられる．特に，$\tau_1 = 2$ である．

ここで，シフトの遅れ h は，$A^{(k)}$ の右下隅の小行列の固有値の代わりに $A^{(k-m)}$ の小行列の固有値を使った場合に $h=1$ と数える．定理 2 より，deferred shifting scheme では，プロセッサの待ち時間を増加させずに同期オーバーヘッドを削減する代償として，収束次数が 3 次から 2 次に落ちていることがわかる．

Fully pipelined shifting scheme の収束性については，まだ理論的な結果が知られていない．

3.2 中心多様体理論

QR 法の反復計算は，行列 $A^{(k)}$ から行列 $A^{(k+1)}$ への写像と見なせる．また，従来のマルチシフト QR 法の計算は，行列 $A^{(k)}$ から行列 $A^{(k+m)}$ への写像と見なせる．どちらの場合も，写像の定義域および値域は対称三重対角行列であり，その自由度は $n+(n-1)$ であるから，アルゴリズムの 1 反復（以下，マルチシフトの場合は $A^{(k)}$ から行列 $A^{(k+m)}$ の計算を 1 反復と定義する）は $\mathbf{R}^{n+(n-1)}$ から $\mathbf{R}^{n+(n-1)}$ への写像と見なすことができる．

このような写像の性質を解析する理論として，Carr[5] による中心多様体理論がある．

いま，T を $\mathbf{R}^{l_1+l_2}$ から $\mathbf{R}^{l_1+l_2}$ への写像とし，$\mathbf{x}^{(k)} \in \mathbf{R}^{l_1}$，$\mathbf{y}^{(k)} \in \mathbf{R}^{l_2}$ に対して，

$$(\mathbf{x}^{(k+1)}, \mathbf{y}^{(k+1)}) = T(\mathbf{x}^{(k)}, \mathbf{y}^{(k)}) \tag{12}$$

が成り立つとする．さらに，$l_1 \times l_1$ 行列 A，$l_2 \times l_2$ 行列 B，$\mathbf{R}^{l_1+l_2}$ から $\mathbf{R}^{l_1}$ への写像 f，$\mathbf{R}^{l_1+l_2}$ から $\mathbf{R}^{l_2}$ への写像 g を用いて T を具体的に

$$\begin{aligned} \mathbf{x}^{(k+1)} &= A\mathbf{x}^{(k)} + f(\mathbf{x}^{(k)}, \mathbf{y}^{(k)}) \\ \mathbf{y}^{(k+1)} &= B\mathbf{y}^{(k)} + g(\mathbf{x}^{(k)}, \mathbf{y}^{(k)}) \end{aligned} \tag{13}$$

と書くことができるとする．さらに，A のすべての固有値が絶対値 1 であり，B のすべての固有値の絶対値が 1 より小さく，f，g は $\mathbf{C}^1$ 級で

$$f(0,0) = 0 \tag{14}$$
$$g(0,0) = 0 \tag{15}$$
$$Df(0,0) = 0 \tag{16}$$
$$Dg(0,0) = 0 \tag{17}$$

が成り立つとする．ただし，式 (16)，(17) は，f，g のすべての変数に対する偏導関数が原点において 0 であることを意味する．また，T の不変多様体であって，性質

$$\mathbf{y}^{(k)} = h(\mathbf{x}^{(k)}), \quad h(0) = 0, \quad Dh(0) = 0 \tag{18}$$

を持つものを中心多様体と呼ぶ．

以上の条件の下で，次の 4 つの定理が成り立つ．

定理 3 (Carr) T に対して中心多様体 $h : \mathbf{R}^{l_1} \to \mathbf{R}^{l_2}$ が存在する．

定理 4 (Carr) 中心多様体上でのフローは次の式に従う．

$$\mathbf{u}^{(k+1)} = A\mathbf{u}^{(k)} + f(\mathbf{u}^{(k)}, h(\mathbf{u}^k)) \tag{19}$$

定理 5 (Carr) 写像 (19) の原点における安定性は，写像 (13) の原点における安定性と等価である．特に，(19) において原点が安定であり，$(\mathbf{x}^{(k)}, \mathbf{y}^{(k)})$ が，原点に十分近い点 $(\mathbf{x}^{(0)}, \mathbf{y}^{(0)})$ から出発して写像 (13) を繰り返し適用して得られた点であるとする．このとき，(19) に対して，$n \to \infty$ のとき

$$\begin{aligned} \mathbf{x}^{(k)} &= \mathbf{u}^{(k)} + O(e^{-\mu k}) \\ \mathbf{y}^{(k)} &= h(\mathbf{u}^{(k)}) + O(e^{-\mu k}) \end{aligned} \tag{20}$$

を満たす解 $\mathbf{u}^{(k)}$ が存在する．ここで，$\mu > 0$ は定数である．

定理 6 (Carr) $\phi : \mathbf{R}^{l_1} \to \mathbf{R}^{l_2}$ が $\phi(0) = 0$，$D\phi(0) = 0$ を満たす $\mathbf{C}^1$ 級写像で，写像 M がある $q > 1$ に対して $\mathbf{x}^{(k)} \to 0$ のとき $(M\phi)(\mathbf{x}^{(k)}) = O(|\mathbf{x}^{(k)}|^q)$ を満たすならば，$\mathbf{x}^{(k)} \to 0$ のとき $h(\mathbf{x}^{(k)}) = \phi(\mathbf{x}^{(k)}) + O(|\mathbf{x}^{(k)}|^q)$ が成り立つ．

中心多様体理論を使うことにより，特異値計算のための mdLVs 法の局所的収束性に関する結果が得られている [14]．また，常微分方程式系に対する中心多様体理論を使うことにより，QR 法の連続版と見なせる常微分方程式系である Toda flow の局所的収束性に関する結果が得られている [6]．

3.3 中心多様体理論の QR 法への適用

中心多様体理論を適用することにより，マルチシフト QR 法の種々の変種についても，その局所的収束性を統一的に解析できる可能性がある．そこで本節では，その準備として，シフトなしの QR 法およびレイリー商シフトを使った QR 法について，中心多様体理論を適用できるかどうか検討する．

3.3.1 シフトなしの QR 法の場合

中心多様体理論を適用するには，QR 法における $A^{(k)}$ から $A^{(k+1)}$ への写像を式 (13) の形に表現し，係数行列 A，B および写像 f，g が条件を満たすことを確認する必要がある．

そのための準備として，まず行列 $A^{(0)}$ の固有値を $\lambda_1, \lambda_2, \ldots, \lambda_n$ とし，

$$|\lambda_1| > |\lambda_2| > \cdots > |\lambda_n| > 0 \tag{21}$$

と仮定する．また，$a_{ij}^{(k)}$ から次式で定義される $b_{ij}^{(k)}$ に変数を変更する．

$$b_{ii}^{(k)} = a_{ii}^{(k)} - \lambda_i \quad (i = 1, 2, \ldots, n), \tag{22}$$

$$b_{ij}^{(k)} = a_{ij}^{(k)} \quad (i \neq j). \tag{23}$$

したがって，$n + (n-1)$ 個の変数 $b_{ii}^{(k)}$ $(i = 1, 2, \ldots, n)$ および $b_{i+1,i}^{(k)}$ $(i = 1, 2, \ldots, n-1)$ から $n + (n-1)$ 個の変数 $b_{ii}^{(k+1)}$ $(i = 1, 2, \ldots, n)$ および $b_{i+1,i}^{(k+1)}$ $(i = 1, 2, \ldots, n-1)$ を計算する写像を調べればよい．すると，アルゴリズム 1，2 における各演算が解析的であることより（$|\lambda_i| > 0$ より，平方根を求める関数も解析的となることに注意），この写像も解析的，すなわちテイラー展開可能となる．したがって，条件 (14) 〜 (17) は，写像 f，g を原点の周りでテイラー展開したとき，0 次および 1 次の項がすべて 0 になることと等価である．以上より，中心多様体理論が適用できることを言うには，次のことを示せばよい．

- A のすべての固有値が絶対値 1.
- B のすべての固有値の絶対値が 1 より小さい．
- f および g を原点の周りでテイラー展開したとき，0 次および 1 次の項がすべて 0.

いま，$|b_{ij}^{(k)}| \ll 1$ と仮定すると，帰納法によりすべての l に対して $|b_{ij}^{(k,l)}| \ll 1$ であることが示せる．そこで，まず，アルゴリズム 1 における c_i，s_i を $b_{ij}^{(k,l)}$ についてテイラー展開し，1 次の項までを取ると次のようになる．ただし，$b_{ij}^{(k,l)}$ について 2 次以上の項を $O(b^2)$ と表す．

$$
\begin{aligned}
c_i &= \frac{\lambda_i + b_{ii}^{(k,i-1)}}{\sqrt{\lambda_i^2 + 2\lambda_i b_{ii}^{(k,i-1)} + \left(b_{ii}^{(k,i-1)}\right)^2 + \left(b_{i+1,i}^{(k,i-1)}\right)^2}} \\
&= \frac{1 + \frac{b_{ii}^{(k,i-1)}}{\lambda_i}}{\sqrt{1 + 2\frac{b_{ii}^{(k,i-1)}}{\lambda_i} + \frac{1}{\lambda_i^2}\left\{\left(b_{ii}^{(k,i-1)}\right)^2 + \left(b_{i+1,i}^{(k,i-1)}\right)^2\right\}}} \\
&= \left(1 + \frac{b_{ii}^{(k,i-1)}}{\lambda_i}\right)\left(1 - \frac{b_{ii}^{(k,i-1)}}{\lambda_i} + O(b^2)\right) \\
&= 1 + O(b^2),
\end{aligned} \tag{24}
$$

$$
\begin{aligned}
s_i &= \frac{b_{i+1,i}^{(k,i-1)}}{\sqrt{\lambda_i^2 + 2\lambda_i b_{ii}^{(k,i-1)} + \left(b_{ii}^{(k,i-1)}\right)^2 + \left(b_{i+1,i}^{(k,i-1)}\right)^2}} \\
&= \frac{\frac{b_{i+1,i}^{(k,i-1)}}{\lambda_i}}{\sqrt{1 + 2\frac{b_{ii}^{(k,i-1)}}{\lambda_i} + \frac{1}{\lambda_i^2}\left\{\left(b_{ii}^{(k,i-1)}\right)^2 + \left(b_{i+1,i}^{(k,i-1)}\right)^2\right\}}} \\
&= \frac{b_{i+1,i}^{(k,i-1)}}{\lambda_i}\left(1 - \frac{b_{ii}^{(k,i-1)}}{\lambda_i} + O(b^2)\right) \\
&= \frac{b_{i+1,i}^{(k,i-1)}}{\lambda_i} + O(b^2).
\end{aligned} \tag{25}
$$

アルゴリズム 1，2 におけるこれ以外の計算は，もともと $b_{ij}^{(k,l)}$ について線形なので，変更の必要はない．
式 (24)，(25) を代入して整理すると，アルゴリズム 1 は次のように書き直せる．

[アルゴリズム 1': $A^{(k)}$ から $R^{(k)}$ の計算]

do $i = 1, n-1$

$\quad b_{ii}^{(k,i)} = b_{ii}^{(k,i-1)} + O(b^2)$

$\quad b_{i+1,i}^{(k,i)} = 0$

$\quad b_{i,i+1}^{(k,i)} = b_{i,i+1}^{(k,i-1)} + \frac{\lambda_{i+1}}{\lambda_i} b_{i+1,i}^{(k,i-1)} + O(b^2)$

$\quad b_{i+1,i+1}^{(k,i)} = b_{i+1,i+1}^{(k,i-1)} + O(b^2)$

$\quad$**if** $i \neq n-1$ **then**

$\qquad b_{i,i+2}^{(k,i)} = O(b^2)$

$\qquad b_{i+1,i+2}^{(k,i)} = b_{i+1,i+2}^{(k,i-1)} + O(b^2)$

$\quad$**end if**

end do

すなわち，$b_{ij}^{(k,l)}$ について 1 次の項までを考える範囲では，アルゴリズム 1 によって対角要素は変化しない．
さらに，行列の各要素はアルゴリズム中に高々2 回しか登場しないことに注意すると，最初の行列要素 $b_{ij}^{(k,0)}$

を使って最終的な行列要素 $b_{ij}^{(k,n-1)}$ を次のように陽に書き表すことができる.

$$\begin{aligned}
b_{ii}^{(k,n-1)} &= b_{ii}^{(k,0)} + O(b^2) \quad (i = 1, 2, \ldots, n), && (26)\\
b_{i+1,i}^{(k,n-1)} &= 0 \quad (i = 1, 2, \ldots, n-1), && (27)\\
b_{i,i+1}^{(k,n-1)} &= b_{i,i+1}^{(k,0)} + \frac{\lambda_{i+1}}{\lambda_i} b_{i+1,i}^{(k,0)} + O(b^2) \quad (i = 1, 2, \ldots, n-1), && (28)\\
b_{i,i+2}^{(k,n-1)} &= O(b^2) \quad (i = 1, 2, \ldots, n-2). && (29)
\end{aligned}$$

一方，アルゴリズム 2 は次のようになる.

[アルゴリズム 2': $R^{(k)}$ から $A^{(k+1)}$ の計算]

do $i = 1,\ n-1$

  $b_{ii}^{(k,n+i-1)} = b_{ii}^{(k,n+i-2)} + O(b^2)$

  $b_{i,i+1}^{(k,n+i-1)} = -b_{i+1,i+1}^{(k,n+i-2)} + b_{i,i+1}^{(k,n+i-2)} + O(b^2)$

  $b_{i+1,i}^{(k,n+i-1)} = \frac{\lambda_{i+1}}{\lambda_i} b_{i+1,i}^{(k,i-1)} + O(b^2)$

  $b_{i+1,i+1}^{(k,n+i-1)} = b_{i+1,i+1}^{(k,n+i-2)} + O(b^2)$

  if $i \neq 1$ **then**

    $b_{i-1,i}^{(k,n+i-1)} = b_{i-1,i}^{(k,n+i-2)} + O(b^2)$

    $b_{i-1,i+1}^{(k,n+i-1)} = 0$

  end if

end do

アルゴリズム 2' でも同様に，$b_{ij}^{(k,l)}$ について 1 次の項までを考える範囲では，対角要素は変化しないことがわかる．さらに，行列の各要素がアルゴリズム中に高々2 回しか登場しないことに注意すると，最初の行列要素 $b_{ij}^{k,n-1}$ を使って最終的な行列要素 $b_{ij}^{(k,2n-2)}$ を次のように書き表すことができる.

$$\begin{aligned}
b_{ii}^{(k,2n-2)} &= b_{ii}^{(k,n-1)} + O(b^2) \quad (i = 1, 2, \ldots, n), && (30)\\
b_{i+1,i}^{(k,2n-2)} &= \frac{\lambda_{i+1}}{\lambda_i} b_{i+1,i}^{(k,i-1)} + O(b^2) \quad (i = 1, 2, \ldots, n-1). && (31)
\end{aligned}$$

さらに，これらの式の左辺が $b_{ii}^{(k+1)}$，$b_{i+1,i}^{(k+1)}$ であることに注意し，第 1 式に式 (26) を代入し，第 2 式で $b_{i+1,i}^{(k,i-1)} = b_{i+1,i}^{(k)}$ を使うと次の結果が得られる.

$$\begin{aligned}
b_{ii}^{(k+1)} &= b_{ii}^{(k)} + O(b^2) \quad (i = 1, 2, \ldots, n), && (32)\\
b_{i+1,i}^{(k+1)} &= \frac{\lambda_{i+1}}{\lambda_i} b_{i+1,i}^{(k)} + O(b^2) \quad (i = 1, 2, \ldots, n-1). && (33)
\end{aligned}$$

式 (32)，(33) は，中心多様体理論の定式化において，

$$\begin{aligned}
A &= I, && (34)\\
B &= \mathrm{diag}\left(\frac{\lambda_2}{\lambda_1}, \frac{\lambda_3}{\lambda_2}, \ldots, \frac{\lambda_n}{\lambda_{n-1}}\right). && (35)
\end{aligned}$$

であり，かつ f，g が 2 次以上の項のみからなることを示す．仮定より，$|\lambda_i/\lambda_{i+1}| < 1$ であるから，シフトなしの QR 法に対しては中心多様体理論の前提が成り立ち，定理 3 〜 6 が適用できる．また，式 (35) は，$O(b^2)$ の項が無視できるならば，副対角要素は収束率 $|\lambda_i|/|\lambda_{i+1}|$ で 0 に 1 次収束することを示しており，シフトなし QR 法に対する古典的な収束理論の結果 [20][24] と一致する.

3.3.2 レイリー商シフトを使った QR 法の場合

次に，レイリー商シフトを使った QR 法の場合を考える．ただし，シフト $s^{(k)} = a_{nn}^{(k)}$ は真の固有値と一致しないと仮定する．また，

$$\bar{a}_{ii}^{(k)} = a_{ii}^{(k)} - a_{nn}^{(k)} \quad (i = 1, 2, \ldots, n) \tag{36}$$

$$\bar{a}_{ij}^{(k)} = a_{ij}^{(k)} \quad (i \neq j) \tag{37}$$

$$\bar{\lambda}_i = \lambda_i - a_{nn}^{(k)} \tag{38}$$

とおき，

$$|\bar{\lambda}_1| > |\bar{\lambda}_2| > \cdots > |\bar{\lambda}_n| > 0 \tag{39}$$

と仮定する．式 (39) は，たとえば，元の固有値がすべて正であり，式 (21) が成り立ち，かつ $a_{nn}^{(k)}$ が λ_n の十分よい近似値であれば成り立つ．さらに，式 (22)，(23) にならい，

$$\bar{b}_{ii}^{(k)} = \bar{a}_{ii}^{(k)} - \bar{\lambda}_i \quad (i = 1, 2, \ldots, n), \tag{40}$$

$$\bar{b}_{ij}^{(k)} = \bar{a}_{ij}^{(k)} \quad (i \neq j). \tag{41}$$

と定義する．以上の準備の下で，前項の議論における $b_{ij}^{(k)}$ を $\bar{b}_{ij}^{(k)}$ に，λ_i を $\bar{\lambda}_i$ に置き換えて考えればよい．

前項と異なる点は，レイリー商シフトを行った結果，$\bar{a}_{nn}^{(k)} = \bar{b}_{nn}^{(k)} + \bar{\lambda}_n = 0$ となっていることである．この点に注意して，アルゴリズム 1 を見直すと，まず，c_i，s_i の計算では $\bar{a}_{nn}^{(k)}$ を使わないから影響はない．$\bar{a}_{nn}^{(k)}$ を使うのは，最後のステップ $i = n-1$ における $\bar{a}_{n-1,n}^{(k,n-1)}$ と $\bar{a}_{nn}^{k,n-1}$ の計算だけであり，これらは $\bar{a}_{nn}^{(k)} = 0$ を代入して，

$$\bar{a}_{n-1,n}^{(k,n-1)} = \bar{c}_{n-1}\bar{a}_{n-1,n}^{(k,n-2)} \tag{42}$$

$$\bar{a}_{nn}^{(k,n-1)} = -\bar{s}_{n-1}\bar{a}_{n-1,n}^{(k,n-2)} \tag{43}$$

となる．さらに，両辺を $\bar{b}_{ij}^{(k,l)}$ で書き換え，$\bar{c}_{n-1} = 1 + O(b^2)$，$\bar{s}_{n-1} = \bar{b}_{n,n-1}^{(k)}/\bar{\lambda}_{n-1} + O(b^2)$ を代入し，アルゴリズム 1' より $\bar{b}_{n-1,n}^{(k,n-2)} = \bar{b}_{n-1,n}^{(k)}$ であることに注意すると，次式を得る．

$$\bar{b}_{n-1,n}^{(k,n-1)} = \bar{b}_{n-1,n}^{(k)} + O(b^2), \tag{44}$$

$$\bar{b}_{nn}^{(k,n-1)} = -\frac{\bar{b}_{n,n-1}^{(k)}\bar{b}_{n-1,n}^{(k)}}{\bar{\lambda}_{n-1}} - \bar{\lambda}_n + O(b^2) = -\bar{\lambda}_n + O(b^2). \tag{45}$$

一方，アルゴリズム 2 より，

$$\bar{a}_{n,n-1}^{(2n-2)} = \bar{s}_{n-1}\bar{a}_{nn}^{(k,2n-3)} \tag{46}$$

$$\bar{a}_{nn}^{(2n-2)} = \bar{c}_{n-1}\bar{a}_{nn}^{(k,2n-3)} \tag{47}$$

$$\tag{48}$$

であるが，各式の両辺を $\bar{b}_{ij}^{(k,l)}$ で書き換え，$\bar{s}_{n-1}$，$\bar{c}_{n-1}$ の式を代入して，$\bar{b}_{n,n-1}^{(2n-2)} = \bar{b}_{n,n-1}^{(k+1)}$，$\bar{b}_{nn}^{(2n-2)} = \bar{b}_{nn}^{(k+1)}$，$\bar{b}_{nn}^{(k,2n-3)} = \bar{b}_{nn}^{(k,n-1)}$（アルゴリズム 2 では $\bar{b}_{nn}$ が更新されるのは $i = n-1$ のときが最初）に注意すると，次式を得る．

$$\bar{b}_{n,n-1}^{(k+1)} = \frac{\bar{b}_{n,n-1}^{(k)}}{\bar{\lambda}_{n-1}}\left(\bar{b}_{nn}^{(k,n-1)} + \bar{\lambda}_n\right), \tag{49}$$

$$\bar{b}_{nn}^{(k+1)} = -\bar{\lambda}_n + O(b^2). \tag{50}$$

式 (45)，(49)，(50)，および $\bar{b}_{nn}^{(k)} = \bar{a}_{nn}^{(k)} - \bar{\lambda}_n = -\bar{\lambda}_n$ より，

$$\bar{b}_{n,n-1}^{(k+1)} = \frac{\bar{b}_{n,n-1}^{(k)}}{\bar{\lambda}_{n-1}} \times O(b^2) = O(b^3). \tag{51}$$

$$\bar{b}_{nn}^{(k+1)} = \bar{b}_{nn}^{(k)} + O(b^2). \tag{52}$$

となる．その他の対角要素，副対角要素については前項と同様である．

式 (51)，(52) より，レイリー商シフトを用いた QR 法では

$$
\begin{aligned}
A &= I, & (53)\\
B &= \mathrm{diag}\left(\frac{\bar{\lambda}_2}{\bar{\lambda}_1}, \frac{\bar{\lambda}_3}{\bar{\lambda}_2}, \ldots, \frac{\bar{\lambda}_{n-1}}{\bar{\lambda}_{n-2}}, O(b^2)\right). & (54)
\end{aligned}
$$

であり，かつ f，g が 2 次以上の項のみからなることがわかる．よって，この場合にも中心多様体理論が適用できる．また，式 (51) は，$O(b^2)$ の項が無視できるならば，副対角要素 $a^{(k)}_{n,n-1}$ は 0 に 3 次収束することを示す．ここでも，古典的な収束理論と一致する結果が得られているのは興味深い．

4 おわりに

本論文では，対称三重対角行列の固有値を求めるためのマルチシフト QR 法とその変種について紹介した．また，既存の収束理論について簡単に述べ，様々な変種の収束性を統一的に解析できる可能性がある理論的枠組みとして，中心多様体理論について説明した．シフトなしの QR 法，レイリー商シフトを用いた QR 法に対する予備的検討では，これらの解法に対して中心多様体理論が適用できることがわかった．今後は，様々なマルチシフト QR 法の変種に対して中心多様体理論を適用し，その収束性を調べることが課題である．

謝辞

本研究集会にお招き下さった弘前大学理工学部の中里博教授と，中心多様体理論についてご教示下さった京都大学大学院情報学研究科の岩崎雅史博士に感謝いたします．なお，本研究は名古屋大学 21 世紀 COE プログラム「計算科学フロンティア」および科学研究費補助金基盤研究（C）（課題番号 18560058）の補助を受けている．

参考文献

[1] Z. Bai and J. Demmel: On a block implementation of Hessenberg QR iteration, *Int. J. of High Speed Computing*, Vol. 1, pp. 97–112 (1989).

[2] I. Bar-On and B. Codenotti: A fast and stable parallel QR algorithm for symmetric tridiagonal matrices, *Linear Algebra Appl.*, Vol. 220, pp. 63–95 (1995).

[3] K. Braman, R. Byers and R. Mathias: The multishift QR algorithm. part I: Maintaining well-focused shifts and level 3 performance, *SIAM J. Matrix Anal. Appl.*, Vol. 23, No. 4, pp. 929–947 (2002).

[4] K. Braman, R. Byers and R. Mathias: The multishift QR algorithm. part II: Aggressive early deflation, *SIAM J. Matrix Anal. Appl.*, Vol. 23, No. 4, pp. 948–973 (2002).

[5] J. Carr: *Applications of Centre manifold Theory*, Springer Verlag, New York, 1981.

[6] M. T. Chu: The generalized Toda flow, the QR algorithm and the center manifold theory, *SIAM J. Alg. Disc. Meth.*, Vol. 5, No. 2, pp. 187–201 (1984).

[7] J. W. Demmel: *Applied Numerical Linear Algebra*, SIAM, 1997.

[8] K. Fernando and B. N. Parlett: Accurate singular values and differential qd algorithms, *Numer. Math.*, Vol. 67, No. 2, pp. 191–229 (1994).

[9] J. G. F. Francis: The QR transformation. A unitary analogue to the LR transformation. I, *Comput. J.*, Vol. 4, pp. 265–271 (1961/1962).

[10] J. G. F. Francis: The QR transformation. II, *Comput. J.*, Vol. 4, pp. 332–345 (1961/1962).

[11] G. H. Golub and C. F. van Loan: *Matrix Computations*, Johns Hopkins University Press, Third Edition, 1996.

[12] M. Gu and S. C. Eisenstat: A divide-and-conquer algorithm for the symmetric tridiagonal eigenproblem, *SIAM J. Matrix Anal. Appl.*, Vol. 16, No. 1, pp. 172–191 (1995).

[13] M. Iwasaki and Y. Nakamura: Accurate computation of singular values in terms of shifted integrable schemes, *Japan J. Indust. Appl. Math.*, Vol. 23, No. 3, pp. 239–259 (2006).

[14] M. Iwasaki and Y. Nakamura: Asymptotic analysis of the mdLVs algorithm in terms of center manifolds, 2006.

[15] L. Kaufman: A parallel QR algorithm for the symmetric tridiagonal eigenvalue problem, *J. Parallel and Distributed Comput.*, Vol. 3, pp. 429–434 (1994).

[16] V. N. Kublanovskaya: On some algorithms for the solution of the complete eigenvalue problem, *U.S.S.R. Comput. Math. and Math. Phys.*, Vol. 3, pp. 637–657 (1961).

[17] 宮田考史, 山本有作: 新しいマルチシフト QR 法による実対称三重対角行列の固有値計算の評価, 2005 年応用数学合同研究集会予稿集, pp. 155-158, 龍谷大学, 2005 年 12 月 20 日–22 日.

[18] T. Miyata, Y. Yamamoto and S-L. Zhang: A fully pipelined multishift QR algorithm for the parallel solution of symmetric tridiagonal eigenproblems, in preparation.

[19] 宮田考史: 完全パイプライン化シフト QR 法による実対称三重対角行列の固有値並列計算, 名古屋大学大学院工学研究科計算理工学専攻修士論文, 2007.

[20] B. N. Parlett: *The symmetric Eigenvalue Problem*, SIAM, 1997.

[21] A. H. Sameh and D. J. Kuck: A parallel QR algorithm for symmetric tridiagonal matrices, *IEEE Trans. Comput.*, Vol. C-26, pp. 147–153 (1977).

[22] R. A. van de Geijn: Deferred shifting schemes for parallel QR methods, *SIAM J. Matrix Anal. Appl.*, Vol. 14, No. 1, pp. 180–194 (1993).

[23] D. S. Watkins and L. Elsner: Convergence of algorithms of decomposition type for the eigenvalue problem, *Linear Algebra Appl.*, Vol. 143, pp. 19–47 (1991).

[24] J. H. Wilkinson: *The Algebraic Eigenvalue Problem*, Oxford University Press, 1965.

[25] 山本有作: 密行列固有値解法の最近の発展 (II) – マルチシフト QR 法 –, 日本応用数理学会論文誌, Vol. 16, No. 4, pp. 149–176 (2006).