

特異値計算アルゴリズム dqds 法の収束定理

Convergence Theorems of the dqds Algorithm for Singular Values

¹⁾ 相島 健助 (Kensuke Aishima), 松尾 宇泰 (Takayasu Matsuo),

室田 一雄 (Kazuo Murota), 杉原 正顕 (Masaaki Sugihara)

東京大学大学院 情報理工学系研究科 数理情報学専攻

Graduate School of Information Science and Technology, University of Tokyo

¹⁾E-mail: Kensuke_Aishima@mist.i.u-tokyo.ac.jp

概要 本論文では、行列の特異値計算のための dqds (differential quotient difference with shifts) 法に関して、最近得られた収束定理について概観する。まず、著者らによって証明された dqds 法の大域収束定理について述べる。ここでは、Rutishauser による有名な正定値対称行列に対する固有値計算アルゴリズム コレスキー LR 法に対する大域収束定理との関係も明らかにする。つぎに、dqds 法の収束速度に関する様々な定理について述べる。

Abstract In this paper we survey convergence theorems of the dqds algorithm for computing singular values of matrices. First, we discuss a global convergence theorem obtained recently by the authors, mentioning its relation with famous Rutishauser's convergence theorem of the Cholesky LR method with shifts for the eigenvalue computation of positive definite matrices. Second, we survey a variety of theorems on convergence rate of the dqds algorithm.

1 はじめに

行列の特異値は、最小二乗法などさまざまなデータ処理で重要な役割を果たし、その数値計算法の開発は実用的意義も大きい。行列 A の特異値の 2 乗は $A^T A$ の非零固有値に等しく、したがって特異値の数値計算には反復計算が必要になる。このとき、計算量の観点から、まず与えられた行列 A を直交変換によって上 2 重対角化しておくのが標準的であり、上 2 重対角行列から反復計算によって特異値を計算する。

上 2 重対角行列の特異値計算アルゴリズムとしては、近年、dqds (differential quotient difference with shifts) 法が特に注目されている [6]。従来よく用いられてきた QR 法に基づく方法 [5] に比べて数値的安定性に優れており、適切な原点シフトにより十分な収束速度を達成できるため、現在 LAPACK にも DLASQ として実装され広く利用されている [10]。

この dqds 法の理論的側面について、任意の初期行列に対する収束性（大域的収束性と呼ぶ）の証明が、最近、著者らにより与えられた [2, 3]。しかしながら、dqds 法は、数学的には、正定値対称既約 3 重対角行列に対する固有値計算アルゴリズムであるシフト付きコレスキー LR 法と同値であり、一方、正定値対称行列に対するシフト付きコレスキー LR 法に関しては、その大域的収束性の証明が、1960 年に Rutishauser によって与えられている [14]。このため、研究者の一部には、dqds 法の大域的収束性はすでに証明済みとの認識もあったようである。本論文では、著者らが与えた大域収束定理について

述べると同時に、この Rutishauser の大域収束定理との関係について述べる。より詳しくは、シフト付きコレスキー LR 法に対する Rutishauser の大域収束定理からは、直接的には dqds 法の大域収束性は導かれないこと、しかし、対称既約 3 重対角行列の特性を加味することによって、dqds 法の大域的収束性が導かれることを指摘する。

先に述べたように、著者らは [2, 3] において dqds 法の大域収束定理を与えたが、その証明は直接的かつ初等的なものであり、さらに、すべての要素の振舞いに関する情報を与えるものであった。そのため、様々な具体的なシフト戦略を用いた場合の dqds 法の収束速度を評価することが可能となった。実際、著者らは [2, 3] において、シフト量を Johnson 下界 [7] とよばれる Gershgorin 型定理に基づいて決定した場合に、収束次数が 1.5 次となることを示した。その後、様々なシフト戦略を用いた dqds 法の収束次数が明らかにされてきている [4, 1, 19, 20]。本論文では、これらについても概観する。

2 dqds 法のアルゴリズム

$m \times m$ の上 2 重対角行列 B の各成分を

$$B = \begin{pmatrix} a_1 & b_1 & & 0 \\ 0 & a_2 & \ddots & \\ \vdots & \ddots & \ddots & b_{m-1} \\ 0 & \dots & 0 & a_m \end{pmatrix}$$

とおく。ここで [6] に従って一般性を失わず、次のことを仮定する。

仮定 (A) $a_k > 0$ ($k = 1, \dots, m$), $b_k \neq 0$ ($k = 1, \dots, m-1$).

このとき特異値はすべて異なり $\sigma_1 > \dots > \sigma_m > 0$ とおくことができる [11].

dqds 法のアルゴリズムは以下のとおりである。

アルゴリズム 1 dqds 法

初期設定: $q_k^{(0)} = (a_k)^2$ ($k = 1, 2, \dots, m$); $e_k^{(0)} = (b_k)^2$ ($k = 1, 2, \dots, m-1$); $t^{(0)} := 0$

1: **for** $n := 0, 1, \dots$ **do**

2: シフト量 $s^{(n)} (\geq 0)$ の設定

3: $d_1^{(n+1)} := q_1^{(n)} - s^{(n)}$

4: **for** $k := 1, \dots, m-1$ **do**

5: $q_k^{(n+1)} := d_k^{(n+1)} + e_k^{(n)}$; $e_k^{(n+1)} := \frac{e_k^{(n)} q_{k+1}^{(n)}}{q_k^{(n+1)}}$; $d_{k+1}^{(n+1)} := \frac{d_k^{(n+1)} q_{k+1}^{(n)}}{q_k^{(n+1)}} - s^{(n)}$

6: **end for**

7: $q_m^{(n+1)} := d_m^{(n+1)}$

8: $t^{(n+1)} := t^{(n)} + s^{(n)}$

9: **end for**

このアルゴリズムにおいて、収束判定のために十分小さい $\epsilon > 0$ を設定し、ある n で $|e_{m-1}^{(n)}| \leq \epsilon$ となったときに、

$$\sigma_m^2 \approx q_m^{(n)} + t^{(n)} \left(= q_m^{(n)} + \sum_{l=0}^{n-1} s^{(l)} \right)$$

が成り立つので、 σ_m の近似値を $\sqrt{q_m^{(n)} + t^{(n)}}$ と定める。そして減次を行うことで、すべての特異値を $\sigma_m, \sigma_{m-1}, \dots, \sigma_1$ の順に計算できる。

ここで以下の議論のために、アルゴリズム中の変数を用いて $B^{(n)}$ を

$$B^{(n)} = \begin{pmatrix} \sqrt{q_1^{(n)}} & \sqrt{e_1^{(n)}} & & 0 \\ 0 & \sqrt{q_2^{(n)}} & \cdots & \\ \vdots & \ddots & \ddots & \sqrt{e_{m-1}^{(n)}} \\ 0 & \dots & 0 & \sqrt{q_m^{(n)}} \end{pmatrix} \quad (1)$$

と定義し、この $B^{(n)}$ の最小特異値を $\sigma_{\min}^{(n)}$ と定義する。このときアルゴリズム 1 は

$$(B^{(n+1)})^T B^{(n+1)} = B^{(n)} (B^{(n)})^T - s^{(n)} I \quad (2)$$

という（シフト付き）コレスキー分解を繰り返すことに等しい。ただし $B^{(0)} = B$ である。

3 dqds 法の大域収束定理

この節では、任意の初期行列に対する dqds 法の収束性（大域的収束性）を保証する定理について議論する。特に、dqds 法の大域収束定理 [2, 3] について述べると同時に、正定値対称行列に対するシフト付きコレスキー LR 法に対する大域収束定理 [14] との関係について述べる。

3.1 相島 他による dqds 法の大域収束定理

[2, 3] に与えられた dqds 法の大域収束定理を示す。証明は初等的なものである。

定理 1 (dqds 法の大域的収束性 (相島 他 [2, 3])) 仮定 (A) を満たす任意の初期行列 B に対して、シフト量 $s^{(n)}$ を

$$0 \leq s^{(n)} < (\sigma_{\min}^{(n)})^2 \quad (n = 0, 1, 2, \dots) \quad (3)$$

を満たすように選べば、dqds 法は収束する。より詳しくは

$$\begin{aligned} \lim_{n \rightarrow \infty} e_k^{(n)} &= 0 \quad (k = 1, 2, \dots, m-1), \\ \lim_{n \rightarrow \infty} q_k^{(n)} + t^{(n)} \left(= \lim_{n \rightarrow \infty} q_k^{(n)} + \sum_{l=0}^{n-1} s^{(l)} \right) &= \sigma_k^2 \quad (k = 1, 2, \dots, m) \end{aligned}$$

が成り立つ。 ■

この定理の条件 (3) は, (2) から分かるように, コレスキー分解の可能性を保証する条件, つまり, dqds 法が破綻なく実行できる条件であり, 甚だ自然な条件である. 定理の主張は, この自然な条件さえ満たされれば, dqds 法は収束するというものである.

3.2 dqds 法の大域収束定理と Rutishauser によるシフト付きコレスキー LR 法の大域収束定理との関係

ここでは, Rutishauser による正定値対称行列に対するシフト付きコレスキー LR 法の大域収束定理 [14] について述べ, dqds 法の大域収束定理 (定理 1) との関係进行を述べる. より詳しくは, Rutishauser による正定値対称行列に対するシフト付きコレスキー LR 法の大域収束定理からは, 直接的には dqds 法の大域収束性は導かれないこと, しかし, 対称既約 3 重対角行列の特性を加味することによって, dqds 法の大域的収束性が導かれることを指摘する.

3.2.1 シフト付きコレスキー LR 法

まず, 正定値対称行列に対するシフト付きコレスキー LR 法の説明といくつかの記号の定義をする.

一般の $m \times m$ 正定値対称行列を A とするとき, コレスキー LR 法のアルゴリズムは次の通りである.

アルゴリズム 2 シフト付きコレスキー LR 法

初期設定: $A^{(0)} := A, \quad t^{(0)} := 0$

- 1: for $n := 0, 1, \dots$ do
 - 2: シフト量 $s^{(n)} (\geq 0)$ の設定
 - 3: $A^{(n)} - s^{(n)}I$ をコレスキー分解: $(R^{(n)})^T R^{(n)} = A^{(n)} - s^{(n)}I$ ($R^{(n)}$ は上三角行列)
 - 4: $A^{(n+1)} := R^{(n)}(R^{(n)})^T$
 - 5: $t^{(n+1)} := t^{(n)} + s^{(n)}$
 - 6: end for
-

また, 後の議論のために

$$A^{(n)} = \left(\begin{array}{c|c} U^{(n)} & \boldsymbol{v}^{(n)} \\ \hline (\boldsymbol{v}^{(n)})^T & w^{(n)} \end{array} \right) \quad (4)$$

によって $U^{(n)}, \boldsymbol{v}^{(n)}, w^{(n)}$ を定義しておく. $U^{(n)}$ は $A^{(n)}$ の $(m-1) \times (m-1)$ 首座小行列, $\boldsymbol{v}^{(n)}$ は $m-1$ 次元ベクトル, $w^{(n)}$ は $A^{(n)}$ の (m, m) 成分である. さらに A の固有値を $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m$ と定義し, $A^{(n)}$ の最小固有値を $\lambda_{\min}^{(n)}$ とする.

上記のシフト付きコレスキー LR 法を，シフトを適切に設定して実行すれば，実用上ほとんどの場合に (4) の $v^{(n)}$ が零ベクトルに， $A^{(n)} + t^{(n)}I$ の対角の右下の成分 $w^{(n)} + t^{(n)}$ が λ_m に収束する．したがって，ある程度 $v^{(n)}$ が零ベクトルに近づいた段階で，近似的に $w^{(n)} + t^{(n)} \approx \lambda_m$ が成り立ち，固有値が 1 つ求まる．そして減次を繰り返すことで， $\lambda_m, \lambda_{m-1}, \dots, \lambda_1$ の順ですべての固有値を計算できる．

3.2.2 dqds 法の大域収束定理 (定理 1) のシフト付きコレスキー LR 法への書換え

dqds 法の行列表現 (2) とシフト付きコレスキー LR 法のアルゴリズムを見れば容易に分かるように，仮定 (A) を満たす上 2 重対角行列に対する dqds 法は，正定値対称既約 3 重対角行列に対するシフト付きコレスキー LR 法と同値である．そこで，dqds 法の大域収束定理 (定理 1) と正定値対称行列に対するシフト付きコレスキー LR 法の大域収束定理との関係について論じるために，dqds 法の大域収束定理 (定理 1) をつぎのような正定値対称既約 3 重対角行列に対するシフト付きコレスキー LR 法の大域収束定理に読み換えておく．

定理 1' (正定値対称既約 3 重対角行列に対するシフト付きコレスキー LR 法の大域的収束性) 行列 A を正定値対称既約 3 重対角行列とする．各 n でシフト $s^{(n)}$ を

$$0 \leq s^{(n)} < \lambda_{\min}^{(n)} \quad (n = 0, 1, 2, \dots)$$

を満たすように選んでシフト付きコレスキー LR 法を実行するとき，

$$\lim_{n \rightarrow \infty} (A^{(n)} + t^{(n)}I) = \text{diag}(\lambda_1, \dots, \lambda_m)$$

が成り立つ． ■

3.2.3 Rutishauser によるシフト付きコレスキー LR 法の大域収束定理

まず，Rutishauser によるシフト無しコレスキー LR 法の大域収束定理 [13] を示す．証明は，行列の正定値性を巧妙に用いたもので，初等的かつ簡単なものである ([13, 15, 16] を参照) ．

定理 2 (シフト無しコレスキー LR 法の大域的収束性 (Rutishauser [13])) 正定値対称行列 A に対して，シフト無しコレスキー LR 法を実行したとき， $A^{(n)}$ は固有値を並べた対角行列に必ず収束する． ■

なお，固有値が対角要素に並ぶ順番は，通常は $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m$ となるが，そうならない場合もあり，Rutishauser [13] は，具体例をあげて，このような正定値対称行列 A を “disorder of latent roots” がある行列と呼んでいる．

例 1 (“disorder of latent roots” がある行列 (Rutishauser[13])) 正定値対称行列

$$A = \begin{pmatrix} 5 & 4 & 1 & 1 \\ 4 & 5 & 1 & 1 \\ 1 & 1 & 4 & 2 \\ 1 & 1 & 2 & 4 \end{pmatrix}$$

(固有値は 10, 5, 2, 1) を考える. この行列 A に対してシフト無しのコレスキー LR 法を実行すると, その収束先は

$$\begin{pmatrix} 10 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}$$

となる.

次が, Rutishauser によるシフト付きコレスキー LR 法の大域収束定理 [14] である. 証明は技術的に込み入ったものである.

定理 3 (シフト付きコレスキー LR 法の大域的収束性 (Rutishauser [14])) 行列 A は “disorder of latent roots” がない正定値対称行列とし, さらに, λ_m は単根で, $\lambda_m < \lambda_k$ ($k = 1, \dots, m-1$) とする. 各 n でシフト $s^{(n)}$ を

$$0 \leq s^{(n)} < \lambda_{\min}^{(n)}$$

を満たすように選んで, シフト付きコレスキー LR 法を実行するとき,

$$\lim_{n \rightarrow \infty} (w^{(n)} + t^{(n)}) = \lambda_m$$

が成立する. また, 式 (4) の $v^{(n)}$ は零ベクトルに収束する. ■

dqds 法の大域収束定理と同値な正定値対称既約 3 重対角行列に対するシフト付きコレスキー LR 法の大域収束定理 (定理 1') が上記の Rutishauser の定理 3 からただちに得られるものでないことは明らかである. 定理 3 は, “disorder of latent roots” がない行列の場合の収束保証でしかないからである.

しかしながら, 正定値対称既約 3 重対角行列 A に対して, シフト無しのコレスキー LR 法を実行した場合, $A^{(n)}$ は対角行列に収束し, その対角成分には固有値が降順に並ぶことが知られている. (この事実は, 古くから知られていたようであるが, 著者らはそれを明言した文献を知らない. 証明も, 間接的なもので, dqds 法の文脈での証明 [6], QD 法の文脈での証明 [15, 16] を知るのみである—もっとも, 専門家にとってはこれで十分であったと思われる.) したがって, 次の補題が成立する.

補題 1 正定値対称既約 3 重対角行列は “disorder of latent roots” がない行列である. ■

結局、この補題と、よく知られた事実「対称既約 3 重対角行列の固有値はすべて単根である」に注意すると、上記の Rutishauser による定理 3 より、次の正定値対称既約 3 重対角行列に対するコレスキー LR 法の大域収束定理が得られる。

定理 4 (正定値対称既約 3 重対角行列に対するシフト付きコレスキー LR 法の大域的収束性 (Rutishauser)) 行列 A を正定値対称既約 3 重対角行列とする。各 n でシフト $s^{(n)}$ を

$$0 \leq s^{(n)} < \lambda_{\min}^{(n)}$$

を満たすように選んで、シフト付きコレスキー LR 法を実行するとき、

$$\lim_{n \rightarrow \infty} (w^{(n)} + t^{(n)}) = \lambda_m$$

が成立する。また、式 (4) の $v^{(n)}$ は零ベクトルに収束する。 ■

ここで、念のために、上記の定理 4 から得られる dqds 法の大域収束定理を記しておく。

定理 5 (dqds 法の大域的収束性 (Rutishauser)) 仮定 (A) を満たす任意の初期行列 B に対して、シフト量 $s^{(n)}$ を

$$0 \leq s^{(n)} < (\sigma_{\min}^{(n)})^2 \quad (n = 0, 1, 2, \dots)$$

を満たすように選んで、dqds 法を実行すれば、

$$\lim_{n \rightarrow \infty} e_{m-1}^{(n)} = 0, \quad \lim_{n \rightarrow \infty} q_m^{(n)} + t^{(n)} \left(= \lim_{n \rightarrow \infty} q_m^{(n)} + \sum_{l=0}^{n-1} s^{(l)} \right) = \sigma_m^2$$

が成り立つ。 ■

かくして、Rutishauser によるシフト付コレスキー LR 法の収束性解析を中心とした複雑な議論により、dqds 法の大域的収束性の証明が与えられたことになる。残念ながら、ここで見たような詳細な議論がはっきり述べられた文献は見当たらない。

最後に、dqds 法の大域収束定理 (定理 1) と Rutishauser による dqds 法の大域収束定理 (定理 5) を比較しておく。まず、証明に関しては、定理 5 の証明は結構複雑なものであったが、これに対して、著者らの定理 1 の証明は単純で直接的なものである。また、定理 1 ではすべての対角、副対角成分が収束することを示しているが、定理 5 では、対角と副対角の一番右下の成分 $e_{m-1}^{(n)}$, $q_m^{(n)}$ の収束のみを示すにとどまっている。もちろん、減次を行うことを考慮すれば、実用的にはこれで十分な結果ではあるが、次節に述べるような dqds 法の理論的な収束速度を明らかにするためには、右下以外も含めてすべての対角、副対角成分の収束先を特定することが必要である。

4 dqds 法の様々なシフト戦略とその収束次数

前節の dqds 法の大域的収束性を示す定理 1 における収束の条件 (3) を満たすようなシフトの設定の仕方はいくつか考えられる。最近の研究により、この収束の条件を満たす様々なシフト戦略を用いる場合に対し、その収束次数が明らかにされてきている。この節では、これらの研究結果をまとめることにする。

なお、ここで、dqds 法の収束次数とは、 $e_{m-1}^{(n)}$ が 0 に収束するときの収束次数をいう。dqds 法では、収束判定を $|e_{m-1}^{(n)}| \approx 0$ で行うので、この定義は自然なものである。

4.1 Johnson 下界を用いる dqds 法とその 1.5 次収束性

定理 1 における収束の条件 (3): $0 \leq s^{(n)} < (\sigma_{\min}^{(n)})^2$ を満たすようなシフト量を選ぶために、 $\sigma_{\min}^{(n)}$ の下からの評価を考える。ここでは、行列の最小特異値の下からの評価として、よく知られた Johnson 下界 [7] を用いる。具体的には、 $\sigma_{\min}^{(n)}$ に対する Johnson 下界 $\tau_J^{(n)}$ は、

$$\tau_J^{(n)} = \min_{k=1, \dots, m} \left\{ \sqrt{q_k^{(n)}} - \frac{1}{2} \left(\sqrt{e_{k-1}^{(n)}} + \sqrt{e_k^{(n)}} \right) \right\} \quad (5)$$

で与えられる。もちろん、 $\tau_J^{(n)} < \sigma_{\min}^{(n)}$ が成り立つが、 $\tau_J^{(n)}$ は負にもなりうるので、 $0 \leq s^{(n)} < (\sigma_{\min}^{(n)})^2$ が成り立つように、シフト量を

$$s^{(n)} = \left(\max\{\tau_J^{(n)}, 0\} \right)^2 \quad (6)$$

と定める。このとき、定理 1 より dqds 法は必ず収束する。さらに、次の定理のように、dqds 法は 1.5 次収束することが示される。

定理 6 (Johnson 下界を用いる dqds 法の 1.5 次収束性 (相島 他 [2, 3])) Johnson 下界を用いて、シフト量を (6) とする dqds 法において、

$$\lim_{n \rightarrow \infty} \frac{e_{m-1}^{(n+1)}}{(e_{m-1}^{(n)})^{3/2}} = \frac{1}{\sqrt{\sigma_{m-1}^2 - \sigma_m^2}}$$

が成り立つ。すなわち、dqds 法は 1.5 次収束する。 ■

4.2 Ostrowski 型下界を用いる dqds 法とその 1.5 次収束性

山本 他 [20] は、 $\sigma_{\min}^{(n)}$ の下からの評価に、Johnson 下界より強い下界である Ostrowski 型下界 [8] を用いることを提案した。この Ostrowski 型下界 $\tau_O^{(n)}$ の具体形は

$$\tau_O^{(n)} = \min_{k=1, \dots, m} \left\{ \sqrt{q_k^{(n)} + \frac{1}{4} \left(\sqrt{e_{k-1}^{(n)}} - \sqrt{e_k^{(n)}} \right)^2} - \frac{1}{2} \left(\sqrt{e_{k-1}^{(n)}} + \sqrt{e_k^{(n)}} \right) \right\} \quad (7)$$

で与えられる。 $\tau_O^{(n)} < \sigma_{\min}^{(n)}$ が成り立つが¹, $\tau_O^{(n)}$ は負にもなりうるので, $0 \leq s^{(n)} < (\sigma_{\min}^{(n)})^2$ が成り立つように, シフト量を

$$s^{(n)} = \left(\max\{\tau_O^{(n)}, 0\} \right)^2 \quad (8)$$

と定める。このとき, 定理 1 より dqds 法は必ず収束する。さらに, 次の定理のように, dqds 法は 1.5 次収束することが示される。

定理 7 (Ostrowski 型下界を用いる dqds 法の 1.5 次収束性 (山本 他 [20])) Ostrowski 型下界を用いて, シフト量を (8) とする dqds 法において,

$$\lim_{n \rightarrow \infty} \frac{e_{m-1}^{(n+1)}}{(e_{m-1}^{(n)})^{3/2}} = \frac{1}{\sqrt{\sigma_{m-1}^2 - \sigma_m^2}}$$

が成り立つ。すなわち, dqds 法は 1.5 次収束する。 ■

4.3 Brauer 型下界を用いる dqds 法とその超 1.5 次収束性

山本 他 [20] は, Ostrowski 型下界の他に, Brauer 型下界 [8] を用いることも提案した。Brauer 型下界 $\tau_B^{(n)}$ の具体形は

$$\tau_B^{(n)} = \min_{1 \leq j \leq k \leq m} \frac{1}{2} \left\{ \sqrt{q_k^{(n)}} + \sqrt{q_j^{(n)}} - \sqrt{\left(q_k^{(n)} - q_j^{(n)} \right) + \left(\sqrt{e_{k-1}^{(n)}} + \sqrt{e_k^{(n)}} \right) \left(\sqrt{e_{j-1}^{(n)}} + \sqrt{e_j^{(n)}} \right)} \right\} \quad (9)$$

で与えられる。Brauer 型下界 $\tau_B^{(n)}$ も負の値をとり得るので, シフト量を

$$s^{(n)} = \left(\max\{\tau_B^{(n)}, 0\} \right)^2 \quad (10)$$

と定める。このとき, $0 \leq s^{(n)} < (\sigma_{\min}^{(n)})^2$ が成り立ち, 定理 1 より, dqds 法は必ず収束する。さらに, 次の定理のように, dqds 法は超 1.5 次収束することが示される。

定理 8 (Brauer 型下界を用いた dqds 法の超 1.5 次収束性 (山本 他 [20])) Brauer 型下界を用いて, シフト量を (10) とする dqds 法において,

$$\lim_{n \rightarrow \infty} \frac{e_{m-1}^{(n+1)}}{(e_{m-1}^{(n)})^{3/2}} = 0$$

が成り立つ。すなわち, dqds 法は超 1.5 次収束する。 ■

¹正確には $\sigma_{\min}^{(n)} \leq \tau_O^{(n)}$ であるが, 等号が成り立つことは極めて稀であり (正確な条件は [20] を参照されたい), また, 等号成立は固有値が求められた状況に対応するので, ここでは狭義の不等号の場合を考察する。

4.4 超2次収束シフト戦略

ここでは dqds 法において超2次収束を実現するシフト戦略を示す [4]. 以下のようなシフト戦略を考える動機, 収束証明等々, 詳細に関しては [4] を参照されたい.

まず

$$X^{(n)} = q_{m-1}^{(n)} + q_m^{(n)} - e_{m-2}^{(n)} + e_{m-1}^{(n)}, \quad Y^{(n)} = 4q_m^{(n)}(q_{m-1}^{(n)} - e_{m-2}^{(n)})$$

として, $\tau_Q^{(n)}$ を

$$\tau_Q^{(n)} = \frac{1}{2} \left(X^{(n)} - \sqrt{(X^{(n)})^2 - Y^{(n)}} \right)$$

と定める. そして次のようなシフト戦略を考える.

シフト戦略 (Q)

$$s^{(n)} = \begin{cases} \tau_Q^{(n)} & (\tau_Q = \text{実数, かつ, } 0 < \tau_Q^{(n)} < (\sigma_{\min}^{(n)})^2 \text{ の場合}) \\ 0 & (\text{その他の場合}). \end{cases}$$

条件分岐は, 定理 1 の条件 (3) を満たすためにある. ここで $\sigma_{\min}^{(n)}$ は未知の値なので, $\tau_Q^{(n)} < (\sigma_{\min}^{(n)})^2$ の判定は困難に思えるが, この問題は以下のように解決できる. ある n において, 試しにシフト量を $\tau_Q^{(n)}$ (実数) として dqds 法の 1 反復を計算し, $q_k^{(n+1)} > 0$ ($k = 1, \dots, m$) となれば, $\tau_Q^{(n)} < (\sigma_{\min}^{(n)})^2$ が成立する [6]. したがって実際には, dqds 法の 1 反復を次のように進めることで, シフト戦略 (Q) を実行することができる.

1. 仮にシフトを $s^{(n)} = \tau_Q^{(n)}$ (実数) として dqds 法の 1 反復を試行.
2. $q_k^{(n+1)} > 0$ ($k = 1, \dots, m$) なら, その結果を採用. さもなくばその結果は棄却して, シフトを $s^{(n)} = 0$ として再計算.

以上で, 実装可能な形で具体的なシフト戦略が与えられた.

シフト戦略 (Q) を用いた dqds 法では, 十分大きなすべての n で, シフト $s^{(n)}$ として, $\tau_Q^{(n)}$ が選択されることが示される. さらに, $s^{(n)} = \tau_Q^{(n)}$ として, dqds 法のアルゴリズムの漸近的振舞いを解析することで, つぎの超2次収束定理が得られる.

定理 9 (シフト戦略 (Q) を用いる dqds 法の超2次収束性 (相島 他 [4])) シフト戦略 (Q) を用いる dqds 法において,

$$\lim_{n \rightarrow \infty} \frac{e_{m-1}^{(n+1)}}{(e_{m-1}^{(n)})^2} = 0$$

が成り立つ. すなわち, dqds 法は超2次収束する. ■

4.5 Rutishauser によるシフト戦略に基づく 3 次収束シフト戦略

Rutishauser は、十分収束した段階で、dqds 法のシフト設定（アルゴリズム 1 の 2）をつぎのようにすることを提案した [14].

シフト戦略 (R)

- 1: $e_0^{(n)} := 0, \hat{d}_0^{(n)} := 1$
- 2: **for** $k := 1, \dots, m-1$ **do**
- 3: $\hat{d}_k^{(n+1)} := \hat{d}_{k-1}^{(n+1)} q_k^{(n)} / (\hat{d}_{k-1}^{(n+1)} + e_{k-1}^{(n)}) - q_m^{(n)}$
- 4: **end for**
- 5: シフト設定 $s^{(n)} := \hat{d}_{m-1}^{(n+1)} q_m^{(n)} / (\hat{d}_{m-1}^{(n+1)} + e_{m-1}^{(n)})$
- 6: **return**

さらに、条件 $0 < \varepsilon < \sigma_{m-1}^2 - \sigma_m^2$ を満たす ε に関する“ある条件”（十分収束した時点では成立すると期待される条件）が n 反復目において満たされるとすると、上記のシフトを用いて $n+1$ 反復目を計算したとき、

$$|e_{m-1}^{(n+1)} q_m^{(n+1)}| \leq \frac{1}{(\sigma_{m-1}^2 - \sigma_m^2 - \varepsilon)^4} |e_{m-1}^{(n)} q_m^{(n)}|^3 \quad (11)$$

が成り立つこと（局所 3 次収束性）を示した。

しかし、この Rutishauser の収束解析は、通常の意味での 3 次収束性の証明というには不完全である。なぜなら、 n 反復目で“ある条件”が満たされたとしても、つぎの $n+1$ 反復目でもその条件が満たされるとは限らないからである。また、このシフトを使うのは、“十分収束した段階で”ということであるが、この点も不明確である。

これに対して、著者らは、[1]において、この Rutishauser のシフト戦略 (R) を基礎に、通常の意味で 3 次収束する dqds 法のシフト戦略を提案した。具体的なシフト戦略は以下の通りである：

シフト戦略 (C)

- 1: $e_0^{(n)} := 0, \hat{d}_0^{(n)} := 1$
- 2: **for** $k := 1, \dots, m-1$ **do**
- 3: $\hat{d}_k^{(n+1)} := \hat{d}_{k-1}^{(n+1)} q_k^{(n)} / (\hat{d}_{k-1}^{(n+1)} + e_{k-1}^{(n)}) - q_m^{(n)}$
- 4: **if** $\hat{d}_k^{(n+1)} \leq 0$ **then**
- 5: シフト設定 $s^{(n)} := 0$
- 6: **return**
- 7: **end if**
- 8: **end for**
- 9: シフト設定 $s^{(n)} := \hat{d}_{m-1}^{(n+1)} q_m^{(n)} / (\hat{d}_{m-1}^{(n+1)} + e_{m-1}^{(n)})$
- 10: **return**

このシフト戦略において、実は定理 1 のシフトの条件 (3) が成立するので、大域的収束性が保証される。さらに、十分大きい n では常に、Rutishauser により提案されたシフト量が選択されることが示され、次のように、3 次収束性が証明される。

定理 10 (シフト戦略 (C) を用いる dqds 法の 3 次収束性 (相島 他 [1])) シフト戦略 (C) を用いる dqds 法において,

$$\lim_{n \rightarrow \infty} \frac{e_{m-1}^{(n+1)}}{(e_{m-1}^{(n)})^3} = \frac{1}{(\sigma_{m-1}^2 - \sigma_m^2)^2}$$

が成り立つ. すなわち, dqds 法は 3 次収束する. ■

なお, (11) から, 期待されるように, $q_m^{(n)}$ は 0 に収束し,

$$\lim_{n \rightarrow \infty} \frac{q_m^{(n+1)}}{(q_m^{(n)})^3} = \frac{1}{(\sigma_{m-1}^2 - \sigma_m^2)^2}$$

が成り立つ.

4.6 一般化 Newton シフトを用いる dqds 法とその収束次数

最近, 木村 他 [9] は, 対称 3 重対角行列の固有値計算で用いられる Newton シフト [11, 17] のアイデアを一般化したシフト ([9] では一般化 Newton シフトと呼んでいる) を dqds 法で用いることを提案した [9].

Newton シフトは, dqds 法のアルゴリズム (2) に即して述べると, $B^{(n)}(B^{(n)})^T$ の特性多項式 $\varphi(\lambda) = \det(B^{(n)}(B^{(n)})^T - \lambda I)$ から, $\tau_N^{(n)} = -\varphi(0)/\varphi'(0)$ と定義される. 容易に分かるように, $-\varphi(0)/\varphi'(0) = [\text{Tr}(B^{(n)}(B^{(n)})^T)^{-1}]^{-1}$ であり, この表式から, Newton シフト $\tau_N^{(n)}$ は dqds 法の収束条件 (3) を満たすことが分かる. また, 計算量に関しても, $O(m)$ であることが知られている [17].

この Newton シフトに対して, 木村らは, 任意の正の整数 p に対し,

$$\tau_{p,N}^{(n)} = [\text{Tr}(B^{(n)}(B^{(n)})^T)^{-p}]^{-1/p}$$

を定義し, p 次一般化 Newton シフトと呼んだ. 明らかに,

$$0 < \tau_{1,N}^{(n)} (= \text{Newton シフト}) < \tau_{2,N}^{(n)} < \cdots < (\sigma_{\min}^{(n)})^2, \quad \lim_{p \rightarrow \infty} \tau_{p,N}^{(n)} = (\sigma_{\min}^{(n)})^2$$

が成り立ち, 高次の一般化 Newton シフトは, dqds 法の収束条件 (3) を満たすと同時に, 性能の良いシフトとなることが期待される. また, 一般化 Newton シフト $\tau_{p,N}^{(n)}$ の計算は一見, 計算量が多そうであるが, $O(pm)$ で済む算法が考案されている [21].

この p 次一般化 Newton シフト $\tau_{p,N}^{(n)}$ を用いる dqds 法の収束性について, 次の収束定理が示されている.

定理 11 (p 次一般化 Newton シフトを用いる dqds 法の超 $p + 1 - \varepsilon$ 次収束性 (山本 他 [19])) p 次一般化 Newton シフトを用いる dqds 法において, 任意の $\varepsilon > 0$ に対して

$$\lim_{n \rightarrow \infty} \frac{e_{m-1}^{(n+1)}}{(e_{m-1}^{(n)})^{p+1-\varepsilon}} = 0$$

が成り立つ. ■

4.7 DLASQ ルーチンの超 2 次収束性

dqds 法は LAPACK に DLASQ ルーチンとして実装されている。DLASQ ルーチンのアルゴリズムは、現実的な収束速度、高精度を達成するために、数種類のシフトを使い分ける複雑なものとなっている ([12] を参照)。しかし、実は収束条件 (3) を満たすようにシフトは設定されており、DLASQ ルーチンの大域的収束性は保証される。さらに、詳細な解析により、超 2 次収束であることも示される。

定理 12 (DLASQ の超 2 次収束性) DLASQ ルーチンにおいて、

$$\lim_{n \rightarrow \infty} \frac{e_{m-1}^{(n+1)}}{(e_{m-1}^{(n)})^2} = 0$$

が成り立つ。すなわち、DLASQ ルーチンは超 2 次収束である。 ■

5 おわりに

本論文では、特異値計算のための dqds 法の大域的収束性と収束次数に関して、これまでの研究によって明らかにされていることについて概観した。しかしながら、ここで概観した収束次数に関する定理は、すべて漸近的な収束の状況 ($n \rightarrow \infty$ における状況) について述べたもので、有限の n に関する情報を与えるものではない。今後、Newton 法における Kantorovich の定理 ([18] を参照されたい) のような有限の n に対する定量的な評価を与える結果が得られることを期待したい。

謝辞

本研究は科学研究費補助金の援助を受けた。

参考文献

- [1] K. Aishima, T. Matsuo, K. Murota: Rigorous Proof of Cubic Convergence for the dqds Algorithm for Singular Values, *Japan Journal of Industrial and Applied Mathematics*, vol. 25 (2008), pp. 63–81.
- [2] 相島 健助, 松尾 宇泰, 室田 一雄, 杉原 正顕: 特異値計算のための dqds 法と mdLVs 法の収束性について, *日本応用数理学会論文誌*, 17 (2007), 97–131.
- [3] K. Aishima, T. Matsuo, K. Murota, M. Sugihara: On Convergence of the dqds Algorithm for Singular Value Computation, *SIAM Journal on Matrix Analysis and Applications*, vol. 30 (2008), pp. 522–537.
- [4] K. Aishima, T. Matsuo, K. Murota, M. Sugihara: A Shift Strategy for Superquadratic Convergence in the dqds Algorithm for Singular Values, *Mathematical Engineering Technical Reports*, METR 2007-12, University of Tokyo, March 2007. (<http://www.keisu.t.u-tokyo.ac.jp/Research/techrep.0.html>)

- [5] J. Demmel: *Applied Numerical Linear Algebra*, SIAM, Philadelphia, 1997.
- [6] K. V. Fernando, B. N. Parlett: Accurate Singular Values and Differential qd Algorithms, *Numerische Mathematik*, vol. 67 (1994), pp. 191–229.
- [7] C. R. Johnson: A Gersgorin-type Lower Bound for the Smallest Singular Value, *Linear Algebra and Its Applications*, vol. 112 (1989), pp. 1–7.
- [8] C. R. Johnson, T. Szulc: Further Lower Bounds for the Smallest Singular Value, *Linear Algebra and Its Applications*, vol. 272 (1998), pp. 169–179.
- [9] 木村欣司, 高田雅美, 坪井洋明, 中村佳正: 特異値計算ライブラリのための最小特異値見積もり公式について, 日本応用数理学会 2007 年度年会予稿集, pp. 230–231, 2007.
- [10] LAPACK, <http://www.netlib.org/lapack/>
- [11] B. N. Parlett: *The Symmetric Eigenvalue Problem*, Prentice-Hall, Englewood Cliffs, New Jersey, 1980; SIAM, Philadelphia, 1998.
- [12] B. N. Parlett, O. Marques: An Implementation of the dqds Algorithm (Positive Case), *Linear Algebra and Its Applications*, vol. 309 (2000), pp. 217–259.
- [13] H. Rutishauser: Solution of Eigenvalue Problems with the LR-Transformation, *National Bureau of Standards, Applied Mathematics Series*, vol. 49 (1958), pp. 47–81.
- [14] H. Rutishauser: Über eine kubisch konvergente Variante der LR-Transformation, *Zeitschrift für Angewandte Mathematik und Mechanik*, vol. 40 (1960), pp. 49–54.
- [15] H. Rutishauser: *Lectures on Numerical Mathematics*, Birkhäuser, Boston, 1990.
- [16] H. R. Schwarz: *Numerical Analysis of Symmetric Matrices*, Prentice-Hall, 1973.
- [17] J. H. Wilkinson, C. Reinsch: *Handbook for Automatic Computation. Vol. II, Linear Algebra*, Springer-Verlag, 1971.
- [18] T. Yamamoto: Historical developments in convergence analysis for Newton's and Newton-like methods, *Journal of Computational and Applied Mathematics*, vol. 124 (2000), pp. 1–23.
- [19] 山本有作, 木村欣司, 中村佳正: 特異値計算のためのさまざまなシフト戦略に対する漸近的収束性解析, 日本応用数理学会 2007 年度年会予稿集, pp. 136–137, 2007.
- [20] 山本有作, 宮田考史: Ostrowski 型下界と Brauer 型下界をシフトとして用いた dqds 法の収束性について, 日本応用数理学会論文誌, vol. 18 (2008), pp. 107–134.
- [21] 山下巧, 木村欣司, 中村佳正: 上二重対角行列の最小固有値に関する Newton 下界のための $O(N)$ アルゴリズム, 第 37 回数値解析シンポジウム講演予稿集, pp. 57–59, 2008.