

なぜ可積分な特異値計算アルゴリズムは高精度か Why so accurate is an integrable SVC algorithm?

京都大学情報学研究科 中村 佳正¹(Yoshimasa Nakamura)

Depart. of Appl. Math. and Phys., Grad. School of Informatics, Kyoto Univ.

Abstract. Recently a new singular value computing (SVC) algorithm with shift is designed by M. Iwasaki and this author which has an ensured cubic convergence and a numerical stability. The algorithm is named the mdLVs (modified dLV with shift) and is more accurate than the dqds algorithm as well as the Demmel-Kahan QR s algorithm. Combining the mdLVs algorithm with a pair of the dLV-type transformations for computing a double Cholesky factorization the I-SVD (Integrable-Singular Value Decomposition) algorithm is presented by the same authors. The I-SVD is a new $O(N^2)$ SVD algorithm having a better orthogonality property of singular vectors than the MRRR algorithm. The I-SVD is now implemented in DBDSL routine which is rather faster than DBDSQR and DBSDC routines, today's standard routines for bidiagonal SVD, of LAPACK. This report is a brief survey of the mdLVs algorithm. Especially it is explained why the mdLVs algorithm is so accurate.

1 はじめに

行列の特異値分解 (singular value decomposition, SVD) は最小 2 乗法を通じて情報処理などに非常に幅広い応用をもつ重要な線形数値演算である。

一般の長方形行列 $A \in \mathbb{R}^{n \times m}$ に対して, 対称行列 $A^T A$ は非負定値, すなわち, 固有値は負になることはない. $A^T A$ の固有値 λ_j の正の平方根 $\sigma_j := \sqrt{\lambda_j}$ を A の特異値という. $A^T A$ の零固有値に対しては零特異値を対応させる. 特異値は対称行列の固有値計算法によって計算可能であるが, $A^T A$ は非負定値であるため, 通常の対称行列の固有値計算法とは異なる高速かつ高精度の特異値計算専用アルゴリズムを考えることができる.

特異値問題では左右の特異ベクトルの計算が重要な課題であり固有値問題にない困難さがある. $r := \text{rank} A \leq \min\{m, n\}$ とする. 簡単のため $A^T A$ の固有値に重根はないとし, 零でない特異値を順に $0 < \sigma_r < \dots < \sigma_1$ と表すことにする. A の Rayleigh 商と $A^T A$ の最大固有値 λ_1 の間には

$$\max_{\|x\|=1} \frac{\|Ax\|^2}{\|x\|^2} = \lambda_1(A^T A)$$

の関係がある [26]. Rayleigh 商を最大とするベクトル x を x_1 とする. $y_1 := Ax_1$ とおくと, y_1 は $A^T y_1 = A^T A x_1 = \lambda_{\max} x_1$ を満たす. ゆえに, x_1 は最大固有値 λ_1 に対応する $A^T A$ の固有ベクトルである. $A_1 := A - y_1 x_1^T$ とおけば, $\text{rank} A_1 = \text{rank} A - 1$ が成り立つ. 以上のプロセスを繰り返せば

$$A - y_1 x_1^T - y_2 x_2^T - \dots - y_r x_r^T = O$$

¹ynaka@i.kyoto-u.ac.jp

となる. x_j は相異なる λ_j に対応する対称行列 $A^T A$ の固有ベクトルだから互いに直交する. さらに, $u_j := y_j / \sigma_j$, $v_j := x_j$ と書けば

$$A = \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \cdots + \sigma_r u_r v_r^T$$

と表される. u_j も長さ 1 でまた互いに直交する. ベクトル u_j , v_j を並べて $U := (u_1, \dots, u_r)$, $V := (v_1, \dots, v_r)$ と書けば, 特異値からなる対角行列 $\Sigma := \text{diag}(\sigma_1, \dots, \sigma_r)$ について

$$A = U \Sigma V^T \quad (1)$$

となる. $0 < \sigma_r \leq \cdots \leq \sigma_1$ のときも含めて, (1) を A の特異値分解, 行列 U, V の各列ベクトルを, それぞれ, 左特異ベクトル, 右特異ベクトルという [5]. $r = m = n$ のときは U, V は直交行列となる.

以下に $\text{rank} A = 2 = m, n = 4$ の具体的な特異値分解の例をあげよう.

$$\begin{aligned} A &= U \Sigma V^T \\ A &= \begin{pmatrix} \sqrt{2} & 2\sqrt{2} \\ -2\sqrt{2} & -\sqrt{2} \\ -2\sqrt{2} & -\sqrt{2} \\ \sqrt{2} & 2\sqrt{2} \end{pmatrix}, \quad \Sigma = \begin{pmatrix} 6 & 0 \\ 0 & 2 \end{pmatrix}, \\ U &= \begin{pmatrix} 1/2 & -1/2 \\ -1/2 & -1/2 \\ -1/2 & -1/2 \\ 1/2 & -1/2 \end{pmatrix}, \quad V = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix} \end{aligned}$$

今日, 行列の特異値計算, 特異値分解で主流となっているのが 1965 年に発表された **Golub-Kahan** 法である [5]. Golub-Kahan 法は, 数値安定で収束証明のある固有値計算法である **QR** アルゴリズム [17, 24, 25] に基づいている. **QR** アルゴリズムは, IEEE Computer Society 誌 (2000) が特集した 20 世紀の “The Top 10 Algorithms” に, シンプレックス法, 高速 Fourier 変換, クイックソートなどとともに選ばれている. Golub-Kahan 法では, まず, 前処理として,

(i) Householder 変換のくり返しで A を上 2 重対角行列 B に変形する.

$$U_H^T A V_H = \begin{pmatrix} B \\ O \end{pmatrix}, \quad B = \begin{pmatrix} b_1 & b_2 & & \\ & b_3 & \ddots & \\ & & \ddots & b_{2m-2} \\ 0 & & & b_{2m-1} \end{pmatrix} \quad (2)$$

ここに, U_H, V_H はある手順で行列 A から定まる直交行列である. 対角行列 $\text{diag}(\pm 1, \dots, \pm 1)$ を右から V_H に乗ずることで B の対角成分 b_{2k-1} は全て正としてよい. 上 2 重対角化に要する計算は数値的に安定で, 有限回で完了し, その計算量は $2m^3/3 + O(m^2)\text{flops}$ であり比較的軽いとされている [2]. ここに, flops とは浮動小数点演算 (floating point operations) 回数の略である. ここで, すべての副対角成分 b_{2k} について $b_{2k} \neq 0$ を仮定する. もし, ある b_{2k} が零であれば, B は小さな 2 つの上 2 重対角行列に分けて考えることができるから, この仮定は一般性を失うものではない. このとき, B の特異値に重複はなく

$$\sigma_1 > \sigma_2 > \cdots > \sigma_m$$

となる [25]. 副対角成分 b_{2k} を $\text{sgn}(b_{2k})b_{2k}$ と置き換えても B の特異値は変わらないから、以下では簡単のため $b_{2k} > 0$ とする. さらに, $\sigma_m > 0$, すなわち, 零特異値はないものとする. 最初から $\text{rank} A = m$ となるように A を選んで計算開始すれば零特異値は現れない.

- (ii) 正定値対称 3 重対角行列 $B^T B = V_H^T A^T A V_H$ に対して固有値計算の QR アルゴリズムを適用して対角化する. 具体的には, $B^T B = QR$ と QR 分解, すなわち, Gram-Schmidt の直交化を行い, $B^T B \rightarrow Q^T B^T B Q$ なる相似変形を繰り返して

$$(QQ'Q''\dots)^T B^T B QQ'Q''\dots =: \Sigma^2$$

のように対角行列に近づけて行く. $Q, Q', \dots$ は m 次直交行列, R は上 3 角行列. (ii) は反復計算であり, 精度を上げようとすれば, 反復回数を増やさねばならないが, 同時に計算量が増加し, 丸め誤差などのもととなる.

- (iii) A の右特異ベクトルは $V = V_H QQ'Q''\dots$ によって与えられる. A の左特異ベクトル U は AV に対する (列ベクトルの pivoting を伴う) QR 分解 $AVP = UR$ によって得られる. P は列ベクトルの置換のための行列. 全特異ベクトルの計算量は $4m^3 + O(m^2)\text{flops}$ である.

Golub-Kahan 法の上 2 重対角化のステップ (i) は (ii) の計算量を減らすために行う前処理である. $(QQ'Q''\dots)^T B^T B QQ'Q''\dots$ が 3 重対角に保たれることからわかるように, 密行列 $A^T A$ に QR アルゴリズムを適用する場合に比べて計算量が少ない. 零特異値がないとすれば, QR 分解の過程で零による除算はない.

ステップ (ii) において素朴な QR アルゴリズムをそのまま適用したのでは, 正規直交化のための大量の平方根計算に起因して収束が遅い. とりわけ近接特異値がある場合は収束が遅く精度も悪化する. そこで, 原点シフトを行い, 近接特異値をなくして特異値の相対的な距離を大きくする. 原点シフトとは行列 $B^T B$ の対角成分から一斉に同じ数 $\theta^2 (> 0)$ を減ずることで, 特異値分布の平行移動を引き起こす. すなわち,

$$B^T B \rightarrow B'^T B' = B^T B - \theta^2 I.$$

ここに θ^2 がシフト量, I は m 次単位行列である.

原点シフト付き Golub-Kahan 法の完成形が **Demmel-Kahan 法** (1990, [2, 3]) である. Demmel-Kahan 法における (ii) の特異値計算部では, m 次対称 3 重対角行列 $B^T B$ の 3 重対角部の 2 次小行列に対して, 順次, シフト付き QR アルゴリズムを適用して, 少ない計算量で特異値を計算する. 計算量は 1 反復について $6m\text{flops}$ である. 通常の停止条件のもとで $6m^2\text{flops}$ の計算量で特異値計算は終了し [2], (i), (iii) に比べて無視できるほど小さい. 固有値・特異値計算では, 大きな原点シフトを選ぶほど加速の効果は大きい, 大きすぎると行列 $B^T B$ の正定値性をこわし, 反復計算が収束しなくなる. Demmel-Kahan 法では Wilkinson シフトと呼ばれる安全な原点シフトの選び方が知られており, 数値安定性と収束性の保証された信頼性の高い標準解法の地位を獲得し, Demmel と Kahan はこの研究で 1991 年 SIAM 賞を受賞している. Demmel-Kahan 法は LAPACK [13] において DBDSQR ルーチンとして実装されており, Matlab, Mathematica 等の汎用ソフトウェアの中などで広く使われている. シフト付き QR アルゴリズムによる固有値計算の収束次数は最大で 3 次であり [24], Newton 法が 2 次収束することと比べていかに高速かがわかる.

(ii) のシフト付き QR アルゴリズムで用いた直交行列の積 $QQ'Q''\dots$ はそのまま (iii) の B の右特異ベクトルを与える. 逆にみれば, (ii) は (iii) と不可分であり, 直交行列の積 $QQ'Q''\dots$ を計算することなく (ii) が終了するわけではない. Demmel-Kahan 法では Q を回転行列にとることで計

算量の低減が図られているが、2重対角行列 B の特異値分解法とみても、Demmel-Kahan 法は $O(m^3)$ アルゴリズムである。また、応用上は一部の特異ベクトルだけが必要とされることもあるが、Demmel-Kahan 法は構造上、いくつかの特異値に対応する特異ベクトルだけを選んで計算することはできない。これも高速計算の妨げとなる。

様々な点で優れた特徴をもつ Demmel-Kahan 法であるが、大規模行列の高速特異値分解が必要とされる昨今では、その性能的限界が明らかとなっており、新しい動作原理に基づくより高速・高精度な特異値分解アルゴリズムの登場が期待されるようになってきた。

筆者のグループでは 2001 年頃から、離散可積分系と直交多項式理論に基づく特異値計算法 dLV アルゴリズム (discrete Lotka-Volterra algorithm) を定式化し [23, 7, 8, 9], 原点シフトの導入による高速特異値計算法である mdLVs アルゴリズム (modified dLV with shift)[10] の開発を経て、2重 Cholesky 分解による高速な特異ベクトル計算法 [11] に到達し、全体で $O(m^2)$ の計算量の 2重対角特異値分解アルゴリズム I-SVD (Integrable-SVD) を実現している。LAPACK の DBDSQR に相当するルーチンの開発としては、mdLVs アルゴリズムを実装した DLVS ルーチン [20, 21] を経て、I-SVD を実装した DBDSLV ルーチン [22] へと発展している。

本レポートでは、高速・高精度の特異値計算法である mdLVs アルゴリズムについて、その不等間隔離散可積分系としての側面、とりわけ高精度性のもととなる変数の正値性を中心に解説する。mdLVs アルゴリズムは最大で 3 次収束し、変数の正値性が保証され、その結果、シフト付きながら収束証明をもつ。高速性は丸め誤差の蓄積を低減し、正値性は桁落ちのない相対誤差の意味での高精度計算を可能とする。さらに、Demmel-Kahan 法による特異値計算との比較数値実験で mdLVs アルゴリズムの優位性を検証する。また、同じように離散可積分系とみなされる Rutishauser の qd (quotient difference) アルゴリズムの改良版である pqds 法や dqds 法との比較を行い、mdLVs アルゴリズムの数値安定性からみた qd 型アルゴリズムの問題点を明らかにする。この考察により、「アルゴリズムとしての可積分性」は離散可積分系の代数的性質に加えて、正値性という解析的性質を必要とする概念であることを主張する。

2 対称な直交多項式の Christoffel 変換

実軸上の直交多項式の Christoffel 変換が引き起こす 3 項漸化式の係数の変換は不等間隔離散半無限戸田方程式となることが知られている [18]。ここでは、Spiridonov-Zhedanov[19] の精密化として、対称な直交多項式の Christoffel 変換が引き起こす 3 項漸化式の係数の変形方程式である不等間隔離散半無限 Lotka-Volterra 方程式 (dLV) を導出する。

対称なモニック直交多項式の 3 項漸化式

$$p_{k+1}(\lambda) = \lambda p_k(\lambda) - a_k^2 p_{k-1}(\lambda), \quad p_0(\lambda) = 1, \quad p_1(\lambda) = \lambda \quad (3)$$

からスタートする。簡単化して

$$y_{k+1} = \lambda y_k - a_k^2 y_{k-1}, \quad y_k = p_k(\lambda), \quad z_{k+1} = \kappa z_k - a_k^2 z_{k-1}, \quad z_k = p_k(\kappa) \quad (4)$$

とかき、2 度用いて漸化式

$$y_{k+2} = (\lambda^2 - a_{k+1}^2) y_k - \lambda a_k^2 y_{k-1}, \quad z_{k+2} = (\kappa^2 - a_{k+1}^2) z_k - \kappa a_k^2 z_{k-1} \quad (5)$$

を準備する。(5) の第 1 式、第 2 式に、それぞれ、 z_k , y_k を乗じて

$$(\lambda^2 - \kappa^2) y_k z_k = y_{k+2} z_k - y_k z_{k+2} + a_k^2 (\lambda y_{k-1} z_k - \kappa y_k z_{k-1}) \quad (6)$$

となるが、右辺の $\lambda y_{k-1} z_k - \kappa y_k z_{k-1}$ は (4) を繰り返し用いると対称な直交多項式の満たす関係式

$$\begin{aligned}\lambda y_{k-1} z_k - \kappa y_k z_{k-1} &= -a_{k-1}^2 (\lambda y_{k-1} z_{k-2} - \kappa y_{k-2} z_{k-1}) \\ &= -a_{k-1}^2 (\lambda^2 - \kappa^2) y_{k-2} z_{k-2} - a_{k-1}^2 a_{k-2}^2 a_{k-3}^2 (\lambda^2 - \kappa^2) y_{k-4} z_{k-4} \\ &\quad + a_{k-1}^2 a_{k-2}^2 a_{k-3}^2 a_{k-4}^2 (\lambda y_{k-5} z_{k-4} - \kappa y_{k-4} z_{k-5})\end{aligned}$$

を得る。一般の実軸上の直交多項式の場合と異なり、 k が奇数か偶数かで扱いを変える必要がある。

(i) $k = 2m - 1$ のとき、 $y_0 = z_0 = 1$, $y_1 = \lambda$, $z_1 = \kappa$ に注意すると、(6) は

$$\begin{aligned}(\lambda^2 - \kappa^2) y_{2m-1} z_{2m-1} &= y_{2m+1} z_{2m-1} - y_{2m-1} z_{2m+1} - a_{2m-1}^2 a_{2m-2}^2 (\lambda^2 - \kappa^2) y_{2m-3} z_{2m-3} \\ &\quad - a_{2m-1}^2 a_{2m-2}^2 a_{2m-3}^2 a_{2m-4}^2 (\lambda^2 - \kappa^2) y_{2m-5} z_{2m-5} \\ &\quad - \dots - a_{2m-1}^2 \dots a_2^2 (\lambda^2 - \kappa^2) y_1 z_1\end{aligned}$$

となるから、まとめて

$$(\lambda^2 - \kappa^2) \sum_{k=1}^m \left(\prod_{j=0}^{2m-2k} a_{2m-j}^2 y_{2k-1} z_{2k-1} \right) = y_{2m+1} z_{2m-1} - y_{2m-1} z_{2m+1}$$

となる。

(ii) $k = 2m$ のとき、 $a_0^2 y_{-1} = 0$, $a_0^2 z_{-1} = 0$ に注意すると、(6) は、(i) と同様にして

$$(\lambda^2 - \kappa^2) \sum_{k=0}^m \left(\prod_{j=0}^{2m-2k-1} a_{2m-j}^2 y_{2k} z_{2k} \right) = y_{2m+2} z_{2m} - y_{2m} z_{2m+2}$$

と書ける。

以上により対称な直交多項式の Christoffel-Darboux の公式は以下ようになる。

$$\begin{cases} \sum_{j=1}^m \frac{p_{2j-1}(\lambda) p_{2j-1}(\kappa)}{a_1^2 \dots a_{2j-1}^2} = \frac{p_{2m+1}(\lambda) p_{2m-1}(\kappa) - p_{2m-1}(\lambda) p_{2m+1}(\kappa)}{(\lambda^2 - \kappa^2) a_1^2 \dots a_{2m-1}^2} & (k = 2m - 1) \\ \sum_{j=1}^m \frac{p_{2j}(\lambda) p_{2j}(\kappa)}{a_1^2 \dots a_{2j}^2} + p_0(\lambda) p_0(\kappa) = \frac{p_{2m+2}(\lambda) p_{2m}(\kappa) - p_{2m}(\lambda) p_{2m+2}(\kappa)}{(\lambda^2 - \kappa^2) a_1^2 \dots a_{2m}^2} & (k = 2m) \end{cases} \quad (7)$$

さらに、 $p_k(\kappa) \neq 0$ と仮定して、対称な直交多項式 $p_k(\lambda)$ に対する核多項式を

$$p_k^*(\lambda) := \begin{cases} \frac{a_1^2 \dots a_{2m-1}^2}{p_{2m-1}(\kappa)} \sum_{j=1}^m \frac{p_{2j-1}(\lambda) p_{2j-1}(\kappa)}{a_1^2 \dots a_{2j-1}^2} & (k = 2m - 1) \\ \frac{a_1^2 \dots a_{2m}^2}{p_{2m}(\kappa)} \left(\sum_{j=1}^m \frac{p_{2j}(\lambda) p_{2j}(\kappa)}{a_1^2 \dots a_{2j}^2} + p_0(\lambda) p_0(\kappa) \right) & (k = 2m) \end{cases} \quad (8)$$

によって定義する。このとき、Christoffel-Darboux の公式 (7) は

$$p_k^*(\lambda) = \frac{1}{\lambda^2 - \kappa^2} (p_{k+2}(\lambda) + A_k p_k(\lambda)), \quad A_k := -\frac{p_{k+2}(\kappa)}{p_k(\kappa)} \quad (9)$$

となる。 $k = 2m - 1$ のとき、 $p_k(\lambda)$ は奇関数、 $k = 2m$ のとき、 $p_k(\lambda)$ は偶関数である。 $\lambda = \pm \kappa$ は $p_k^*(\lambda)$ のみかけの極であり、 $p_k^*(\lambda)$ は λ の k 次多項式である。変換 $\{p_k(\lambda)\} \rightarrow \{p_k^*(\lambda)\}$ は対称な直交多項式に対する Christoffel 変換である。

$p_k^{(0)} = p_k(\lambda)$ とし, Christoffel 変換の反復を

$$p_k^{(n+1)} = \frac{1}{\lambda^2 - (\kappa^{(n)})^2} \left(p_{k+2}^{(n)} + A_k^{(n)} p_k^{(n)} \right), \quad A_k^{(n)} := -\frac{p_{k+2}^{(n)}(\kappa^{(n)})}{p_k^{(n)}(\kappa^{(n)})}, \quad (n = 0, 1, \dots) \quad (10)$$

によって導入する. ただし, $p_k^{(n)}(\kappa^{(n)}) \neq 0$ とする. (10) を 3 項漸化式 $p_{k+1}^{(n+1)} = \lambda p_k^{(n+1)} - (a_k^{(n+1)})^2 p_{k-1}^{(n+1)}$ に代入して $p_{k+1}^{(n)} = \lambda p_k^{(n)} - (a_k^{(n)})^2 p_{k-1}^{(n)}$ を用いると

$$\left(A_{k+1}^{(n)} - (a_{k+2}^{(n)})^2 - A_k^{(n)} + (a_k^{(n+1)})^2 \right) p_{k+1}^{(n)} + \left((a_k^{(n+1)})^2 A_{k-1}^{(n)} - (a_k^{(n)})^2 A_k^{(n)} \right) p_{k-1}^{(n)} = 0$$

を得る. よって, Christoffel 変換と 3 項漸化式の両立条件として,

$$(a_k^{(n+1)})^2 = (a_k^{(n)})^2 \frac{A_k^{(n)}}{A_{k-1}^{(n)}} = (a_k^{(n)})^2 \frac{p_{k+2}^{(n)}(\kappa^{(n)})}{p_k^{(n)}(\kappa^{(n)})} \frac{p_{k-1}^{(n)}(\kappa^{(n)})}{p_{k+1}^{(n)}(\kappa^{(n)})} \quad (11)$$

を得る. ここで変数

$$\hat{u}_k^{(n)} := (a_k^{(n)})^2 \frac{p_{k-1}^{(n)}(\kappa^{(n)})}{p_k^{(n)}(\kappa^{(n)})} \quad (12)$$

を導入する. $p_{-1}^{(n)} = 0$ だから $\hat{u}_0^{(n)} = 0$ が成り立つ. 直交多項式の零点の分布 [1] より $p_{k-1}^{(n)}(\lambda)/p_k^{(n)}(\lambda)$ は部分分数展開できて

$$\frac{p_{k-1}^{(n)}(\kappa^{(n)})}{p_k^{(n)}(\kappa^{(n)})} = \sum_{j=1}^k \frac{\rho_j^{(n)}}{\kappa^{(n)} - \lambda_j^{(n)}}, \quad (\rho_j^{(n)} > 0) \quad (13)$$

となる. ここに $\lambda_j^{(n)}$ は直交多項式 $p_k^{(n)}(\lambda)$ の零点である. 一方, 核多項式の線形汎関数 J^* の正値性が成り立つためには, 言い換えれば, 核多項式が直交多項式であるためには, $\kappa^{(n)}$ は $\kappa^{(n)} < \lambda_1^{(n)} < \dots < \lambda_k^{(n)}$ を満たさねばならない [1]. ここでは, 対称な直交多項式を考えているから $\kappa^{(n)} < 0$ である. 従って, $p_{k-1}^{(n)}(\kappa^{(n)})/p_k^{(n)}(\kappa^{(n)}) < 0$ だから, (12) より $\hat{u}_k^{(n)} < 0$ がわかる.

さて, (12) を 3 項漸化式 $p_{k+1}^{(n)}(\kappa^{(n)}) = \kappa^{(n)} p_k^{(n)}(\kappa^{(n)}) - (a_k^{(n)})^2 p_{k-1}^{(n)}(\kappa^{(n)})$ に代入して

$$(a_{k+1}^{(n)})^2 = \hat{u}_{k+1}^{(n)} \left(\kappa^{(n)} + \hat{u}_k^{(n)} \right) \quad (14)$$

となる. また, (12) を (11) を代入して (14) を用いると

$$(a_k^{(n+1)})^2 = \hat{u}_k^{(n)} \left(\kappa^{(n)} + \hat{u}_{k+1}^{(n)} \right) \quad (15)$$

となる. (14) と (15) において $(a_k^{(n+1)})^2$ を消去すれば変数 $\hat{u}_k^{(n)}$ の満たすべき関係式

$$\hat{u}_k^{(n+1)}(\kappa^{(n+1)} + \hat{u}_{k-1}^{(n+1)}) = \hat{u}_k^{(n)}(\kappa^{(n)} + \hat{u}_{k+1}^{(n)}) \quad (16)$$

を得る. 対称な直交多項式の Christoffel 変換を用いた (16) の導出は Spiridonov-Zhedanov(1997, [19]) による導出の精密化である.

ここで, (16) において

$$\delta^{(n)} := \frac{1}{(\kappa^{(n)})^2}, \quad u_k^{(n)} = \kappa^{(n)} \hat{u}_k^{(n)} \quad (17)$$

とおき, 変形すると

$$u_k^{(n+1)}(1 + \delta^{(n+1)} u_{k-1}^{(n+1)}) = u_k^{(n)}(1 + \delta^{(n)} u_{k+1}^{(n)}) \quad (18)$$

となる. 変数 $\delta^{(n)}$ と $u_k^{(n)}$ について以下が成り立つことに注意する.

命題 1. 漸化式 (18) において, 正值性

$$\delta^{(n)} > 0, \quad u_k^{(n)} > 0 \quad (19)$$

が成り立つ. $\square$

これは本稿で解説する特異値計算アルゴリズムの収束性と数値安定性の証明においてキーとなる重要な性質である.

$u_k^{(n)}$ を時刻 $t = \sum_{j=0}^{n-1} \delta^{(j)}$ における u_k と値とみて, t を一定値に保ったまま $\delta^{(n+1)}/\delta^{(n)} \rightarrow 1$ なる極限 $\delta^{(n)} \rightarrow +0$ をとれば, 漸化式 (18) は $u_k = u_k(t)$ についての微分方程式

$$\frac{du_k}{dt} = u_k(u_{k+1} - u_{k-1}), \quad u_0(t) = 0, \quad (k = 1, 2, \dots) \quad (20)$$

に移行する. この極限操作は Christoffel 変換のパラメータ操作 $\kappa^{(n)} \rightarrow -\infty$ に対応し, 線形汎関数の正值性を破ることはない. 微分方程式 (20) は半無限 Lotka-Volterra 方程式に他ならない. この意味で, 両立条件 (18) は $\delta^{(n)} = 1/(\kappa^{(n)})^2 > 0$ を差分間隔とする, 不等間隔離散の半無限 Lotka-Volterra 方程式である. 直交多項式の変形方程式とみたとき, 通常の不等間隔離散 Lotka-Volterra 方程式 [6, 19] と異なり, (18) には (19) なる付帯条件がついていることに注意しなければならない. 対称な直交多項式の Christoffel 変換の引き起こす 3 項漸化式の係数の変形は

$$\begin{array}{ccc} \{u_k^{(n)}\} & \xrightarrow{\text{dLV 方程式}} & \{u_k^{(n+1)}\} \\ \downarrow \text{3 項漸化式} & & \uparrow \text{3 項漸化式} \\ \{p_k^{(n)}(\lambda)\} & \xrightarrow{\text{Christoffel 変換}} & \{p_k^{(n+1)}(\lambda)\} \end{array}$$

と表されるが, これは ($\delta^{(n)}$ を散乱データとする) 離散の逆散乱法とみなすことができる.

3 dLV アルゴリズムによる特異値計算

一連の研究 [7, 8, 9, 23] において, Demmel-Kahan 法に代わるべき新しい特異値計算法として, 有限 ($k = 1, 2, \dots, 2m-1$) の場合の離散 Lotka-Volterra 方程式

$$\begin{aligned} u_k^{(n+1)}(1 + \delta^{(n+1)}u_{k-1}^{(n+1)}) &= u_k^{(n)}(1 + \delta^{(n)}u_{k+1}^{(n)}), \\ u_0^{(n)} &\equiv 0, \quad u_{2m}^{(n)} \equiv 0, \quad 0 < \delta^{(n)} < M \end{aligned} \quad (21)$$

が考察されている. まず考えるべきは, どのように初期値 $u_k^{(0)}$ を設定すれば上 2 重対角行列 (2) の特異値計算が開始できるかである. 単純に $u_k^{(0)} = b_k$, ($k = 1, 2, \dots, 2m-1$) と与えたのでは正しい特異値への収束は望めない.

離散 Lotka-Volterra 方程式 (21) の離散 Lax 表示

$$L^{(n+1)}R^{(n+1)} = R^{(n)}L^{(n)} - \left(\frac{1}{\delta^{(n)}} - \frac{1}{\delta^{(n+1)}} \right) I, \quad (22)$$

$$L^{(n)} \equiv \begin{pmatrix} J_1^{(n)} & & & 0 \\ 1 & J_2^{(n)} & & \\ & \ddots & \ddots & \\ & & 1 & J_m^{(n)} \end{pmatrix}, \quad R^{(n)} \equiv \begin{pmatrix} 1 & V_1^{(n)} & & \\ & 1 & \ddots & \\ & & \ddots & V_{m-1}^{(n)} \\ 0 & & & 1 \end{pmatrix}$$

からスタートする。ここに

$$J_k^{(n)} := \frac{1}{\delta^{(n)}} \left(1 + \delta^{(n)} u_{2k-2}^{(n)} \right) \left(1 + \delta^{(n)} u_{2k-1}^{(n)} \right), \quad V_k^{(n)} := \delta^{(n)} u_{2k-1}^{(n)} u_{2k}^{(n)} \quad (23)$$

である。もし $1/\delta^{(n)} - 1/\delta^{(n+1)} \geq 0$ であれば, $J_k^{(n)} = q_k^{(n)}$, $V_k^{(n)} = e_k^{(n)}$, さらに,

$$\theta^{(n)2} := \frac{1}{\delta^{(n)}} - \frac{1}{\delta^{(n+1)}} \geq 0 \quad (24)$$

とおくことで (22) は, $q_k^{(n)}$, $e_k^{(n)}$ についてみれば, Rutishauser[15, 16] が発見したシフト付き pqd (progressive qd with shift, pqds) アルゴリズム

$$e_k^{(n+1)} = \frac{e_k^{(n)}}{q_{k+1}^{(n+1)}} q_{k+1}^{(n)}, \quad q_{k+1}^{(n+1)} = q_{k+1}^{(n)} - e_k^{(n+1)} - \theta^{(n)2} + e_{k+1}^{(n)}, \quad e_0^{(n)} \equiv 0 \quad (25)$$

に一致する。 $q_k^{(n+1)}$ による除算に注意する。これは対称な直交多項式に対して $(\kappa^{(n)})^2 \geq (\kappa^{(n+1)})^2$ なる Christoffel 変換を続けることに対応する。特に, $\delta^{(n)} = \delta > 0$, すなわち, $(\kappa^{(n)})^2 = (\kappa^{(n+1)})^2$, ($n = 0, 1, \dots$) とすれば, 離散 Lotka-Volterra 方程式は $\theta^{(n)2} = 0$ のシフトなし pqd 法に帰着する。

新しい変数 $\bar{w}_k^{(n)}$ を

$$\bar{w}_k^{(n)} := u_k^{(n)} (1 + \delta^{(n)} u_{k-1}^{(n)}) \quad (26)$$

により導入する。 $u_k^{(0)} > 0$, ($k = 1, 2, \dots, 2m-1$), かつ $\delta^{(n)} > 0$ であれば $\bar{w}_k^{(n)} > 0$ が成り立つ。さらに, 3 重対角行列

$$Y^{(n)} := L^{(n)}R^{(n)} - \frac{1}{\delta^{(n)}} I = \begin{pmatrix} \bar{w}_1^{(n)} & \bar{w}_1^{(n)}\bar{w}_2^{(n)} & & \\ 1 & \bar{w}_2^{(n)} + \bar{w}_3^{(n)} & \ddots & \\ & \ddots & \ddots & \bar{w}_{2m-3}^{(n)}\bar{w}_{2m-2}^{(n)} \\ & & 1 & \bar{w}_{2m-2}^{(n)} + \bar{w}_{2m-1}^{(n)} \end{pmatrix}$$

を準備する。このとき, 離散 Lotka-Volterra 方程式 (22) は $Y^{(n)}$ の相似変形を記述し, 離散 Lax 表示

$$Y^{(n+1)} = R^{(n)}Y^{(n)}(R^{(n)})^{-1} \quad (27)$$

をもつ。 $\bar{w}_k^{(n)} > 0$ に注意して対角行列

$$D^{(n)} := \text{diag} \left(\prod_{j=1}^{m-1} \sqrt{\bar{w}_{2j-1}^{(n)} \bar{w}_{2j}^{(n)}}, \prod_{j=2}^{m-1} \sqrt{\bar{w}_{2j-1}^{(n)} \bar{w}_{2j}^{(n)}}, \dots, \sqrt{\bar{w}_{2m-3}^{(n)} \bar{w}_{2m-2}^{(n)}}, 1 \right)$$

を準備し, $Y^{(n)}$ を対称行列化して

$$Y_S^{(n)} := (D^{(n)})^{-1} Y^{(n)} D^{(n)} = \begin{pmatrix} \bar{w}_1^{(n)} & \sqrt{\bar{w}_1^{(n)} \bar{w}_2^{(n)}} & & \\ \sqrt{\bar{w}_1^{(n)} \bar{w}_2^{(n)}} & \bar{w}_2^{(n)} + \bar{w}_3^{(n)} & \ddots & \\ & \ddots & \ddots & \sqrt{\bar{w}_{2m-3}^{(n)} \bar{w}_{2m-2}^{(n)}} \\ & & & \bar{w}_{2m-2}^{(n)} + \bar{w}_{2m-1}^{(n)} \end{pmatrix} \quad (28)$$

とかく, $\det(Y_S^{(n)}) = \prod_{j=1}^m \bar{w}_{2j-1}^{(n)}$ が成り立つ. 以上をまとめる.

補題 1. 離散 Lotka-Volterra 方程式 (22) は正定値な対称 3 重対角行列の相似変形

$$Y_S^{(n+1)} = \tilde{R}^{(n)} Y_S^{(n)} (\tilde{R}^{(n)})^{-1}, \quad \tilde{R}^{(n)} := (D^{(n+1)})^{-1} R^{(n)} D^{(n)} \quad (29)$$

として表され, $Y_S^{(n)}$ の固有値は離散 Lotka-Volterra 方程式の時間発展 $n \Rightarrow n+1$ のもとで不変で

ある. $\square$

補題 1 より, $Y_S^{(n)}$ の固有値は差分ステップサイズ $\delta^{(1)}, \delta^{(2)}, \dots, \delta^{(n)}$ の選び方に依存せず, 初期値 $\bar{w}_k^{(0)} = u_k^{(0)}(1 + \delta^{(0)} u_{k-1}^{(0)})$ のみによって定まることがわかる. $Y_S^{(n)}$ は正定値対称だから Cholesky 分解可能で

$$Y_S^{(n)} = (B^{(n)})^\top B^{(n)}, \quad B^{(n)} := \begin{pmatrix} \sqrt{\bar{w}_1^{(n)}} & \sqrt{\bar{w}_2^{(n)}} & & \\ & \sqrt{\bar{w}_3^{(n)}} & \ddots & \\ & & \ddots & \sqrt{\bar{w}_{2m-2}^{(n)}} \\ 0 & & & \sqrt{\bar{w}_{2m-1}^{(n)}} \end{pmatrix} \quad (30)$$

と書ける. $B^{(n)}$ の特異値は $Y_S^{(n)}$ の固有値の正の平方根だから, 補題 1 より, $B^{(n)}$ の特異値もまた離散 Lotka-Volterra 方程式の時間発展のもとで不変である. このことから, 離散 Lotka-Volterra 方程式による上 2 重対角行列 B の特異値計算のためには, 初期値 $u_k^{(0)}$ は $\sqrt{\bar{w}_k^{(0)}} = b_k$, すなわち,

$$u_0^{(0)} = 0, \quad u_{2m}^{(0)} = 0, \quad u_k^{(0)} = \frac{b_k^2}{1 + \delta^{(0)} u_{k-1}^{(0)}}, \quad (k = 1, \dots, 2m-1) \quad (31)$$

となるように選ぶ必要がある. 変数 $\bar{w}_k^{(n)}$ についてみると,

$$\bar{w}_k^{(0)} = b_k^2, \quad \bar{w}_0^{(0)} = 0, \quad \bar{w}_{2m}^{(0)} = 0 \quad (32)$$

である.

次に, 離散 Lotka-Volterra 方程式の解の B の特異値への収束を論じる. 補題 1 より, $Y_S^{(n)}$ の固有値と初期値 $Y_S^{(0)}$ の固有値は完全に一致する. 正定値対称行列 $Y_S^{(0)}$ の固有値 σ_k^2 を

$$\sigma_1 > \sigma_2 > \dots > \sigma_m > 0 \quad (33)$$

であるものとする. $\text{trace}(Y_S^{(n)}) = \text{trace}(Y_S^{(0)})$ より, 変数 $\bar{w}_k^{(n)}$ の和は一定で

$$\sum_{k=1}^{2m-1} \bar{w}_k^{(n)} = \sum_{k=1}^m \sigma_k^2 \quad (34)$$

が成り立つ。初期値選択 (31) より $0 < u_k^{(0)}$ は明らか。離散 Lotka-Volterra 方程式 (21) で時間発展して $0 < u_k^{(n)}$ だから、変数 $\bar{w}_k^{(n)}$ についても正値性 $0 < \bar{w}_k^{(n)}$ が成り立つ。一方、(34) より $\bar{w}_k^{(n)}$ の有界性も成り立つ。ゆえに変数 $u_k^{(n)}$ について、ある正数 M_1 が存在して

$$0 < u_k^{(n)} < M_1, \quad (k = 1, 2, \dots, 2m-1) \quad (35)$$

となる。 $1 + \delta^{(n)} u_k^{(n)} > 1$ を用いて以下の命題と定理が証明されている [8].

命題 2. 初期値について $0 < u_k^{(0)}$ であるならば、離散 Lotka-Volterra 方程式の解は $c_1 > c_2 > \dots > c_m > 0$ なる定数 c_k と 0 に

$$\lim_{n \rightarrow \infty} u_{2k-1}^{(n)} = c_k, \quad \lim_{n \rightarrow \infty} u_{2k}^{(n)} = 0. \quad (36)$$

のように収束する。□

定理 1. 離散 Lotka-Volterra 方程式 (22) の初期値を (31) に従って与えるとする。このとき、 $b_{2k-1} > 0$ を対角成分、 $b_{2k} > 0$ を副対角成分とする上 2 重対角行列 B の第 k 特異値 σ_k は $\sqrt{c_k}$ に一致し

$$\lim_{n \rightarrow \infty} u_{2k-1}^{(n)} = \sigma_k^2, \quad (k = 1, \dots, m) \quad (37)$$

が成り立つ。すなわち、離散 Lotka-Volterra 方程式の解 $u_{2k-1}^{(n)}$ は、 $\delta^{(n)}$ のとり方によらず、 $n \rightarrow \infty$ で、 σ_k^2 に収束する。□

以上により、適切な初期値のもとで、離散 Lotka-Volterra 方程式の解が、与えられた上 2 重対角行列の特異値に平方に収束することが示された。離散 Lotka-Volterra 方程式 (22) による特異値計算法を dLV アルゴリズムと名付ける。dLV アルゴリズムでは命題 1 により変数の正値性が保証されており、正の数の加算と乗算、 $1 + \delta u$ による除算のみで減算や平方根計算はなく、桁落ちのない数値安定な高精度特異値計算が可能である [9].

Fig. 1a はシフトを 0 とした Demmel-Kahan QR 法、pqd 法、dLV アルゴリズムで計算した特異値の相対誤差の比較、Fig. 1b は pqd 法の代わりに dqd 法としたものである。ともに、特異値がやや集積して分布している以下の 100 次上 2 重対角行列 B_3 の特異値の相対誤差を最大特異値から順に並べている。

$$B_3 = \text{bidiag} \left\{ \begin{array}{cccccc} & 0.001 & & 0.002 & \cdots & 0.002 \\ & & & & & \\ 1 & & & 2 & & \\ & & & & \cdots & \\ & & & & & 2 \end{array} \right\}$$

下の段は B_3 の対角成分 b_{2i-1} 、上の段は副対角成分 b_{2i} である。100 次するとき、Maple の多倍長で計算した B_3 の特異値の真値は、順に、 $\hat{\sigma}_1 = 2.001999014 \dots$, $\hat{\sigma}_2 = 2.001996057 \dots$, $\dots$, $\hat{\sigma}_{99} = 1.998000987 \dots$, $\hat{\sigma}_{100} = 0.999999833 \dots$ 等である。

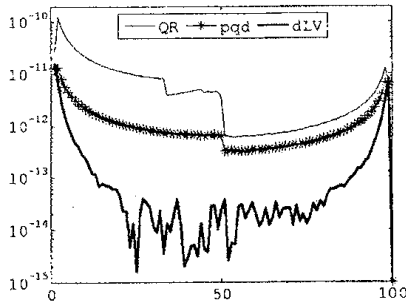


Fig. 1a QR, pqd, dLV の特異値の相対誤差分布

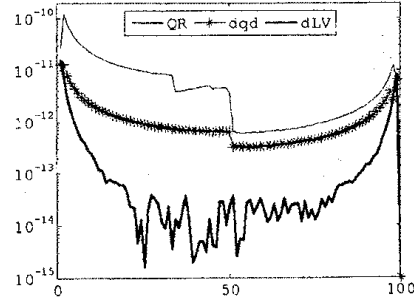


Fig. 1b QR, dqd, dLV の特異値の相対誤差分布

これらの中で dLV アルゴリズムが最も高い相対精度をもつことが確認される。pqd 法より dqd 法の方が僅かながら精度がよい。シフト付きのルーチンにおいても変数の正値性を確保するためにシフト量の 0 として計算することが全反復の 1/3 程度あり、シフトなしの計算における高精度性は重要である。

離散 Lotka-Volterra 方程式と pqds 法の漸化式との間には、直接的な変数変換 (Miura 変換)

$$q_k^{(n)} = \frac{1}{\delta^{(n)}} \left(1 + \delta^{(n)} u_{2k-2}^{(n)} \right) \left(1 + \delta^{(n)} u_{2k-1}^{(n)} \right), \quad e_k^{(n)} = \delta^{(n)} u_{2k-1}^{(n)} u_{2k}^{(n)},$$

$$\theta^{(n)2} = 1/\delta^{(n)} - 1/\delta^{(n+1)} \quad (38)$$

が存在する。一歩進んで、dLV アルゴリズムの収束定理 1 は、変換 (38) を経由して、そのまま pqds 法についても成り立つように見える。しかし、離散時間 Lotka-Volterra 方程式の導出の際に必要な条件 $0 < \delta^{(n)}$ (17), 収束定理の証明で用いた条件 $0 < \delta^{(n)} < M$ (21) を見落としてはならない。 $\delta^{(n)} < 0$ であっても $\delta^{(n)} \leq \delta^{(n+1)}$ であれば条件 (24) は成り立ち、pqds 法 (25) は定式化される。dLV アルゴリズムの収束の速さは最近接特異値の平方と差分ステップサイズから定まる比

$$R_1 := \max_k \frac{\sigma_{k+1}^2 + 1/\delta^{(n)}}{\sigma_k^2 + 1/\delta^{(n)}} < 1 \quad (39)$$

に依存する [7]。この絶対値が小さい程、収束が速い。 R_1 は変換 (38) で dLV と移りあう pqds 法の収束の速さでもある。dLV では $\delta^{(n)} > 0$ でなければならないが、pqds 法では $\delta^{(n)} < 0$ も可能である。pqds 法で原点シフトの効果が顕著となるのは $\delta^{(n)} < 0$ の領域であり、pqds 法はこの領域で高次収束性を実現している。ところが、この場合、 $1 + \delta^{(n)} u_k^{(n)} \leq 0$ となる可能性があり、dLV アルゴリズムの収束性の証明は困難となる。

変数 $\bar{w}_k^{(n)}$ を用いると (38) の第 1 式は、 n が十分大きいとき、

$$q_m^{(n)} = \bar{w}_{2m-1}^{(n)} + u_{2m-2}^{(n)} + \frac{1}{\delta^{(n)}} \approx \sigma_m^2 + \frac{1}{\delta^{(n)}} \quad (40)$$

とかけ、 $\delta^{(n)} < 0$ では $q_m^{(n)} < \sigma_m^2$ となり、シフト量を $\theta^{(n)2} < \sigma_m^2$ と選んでも pqds 法 (25) において $q_m^{(n)}$ の正値性がこわれる可能性がある。このように、pqds 法の高速度性は数値不安定性と隣り合わせである。

別の角度から (38) をみよう。pqds 法で起きていることを、変換 (38) を通じて、dLVs アルゴリズムの変数でみることができる。dLVs の変数 $u_k^{(n)}$, $\delta^{(n)}$ がすべて正であれば、qds の変数 $q_k^{(n)}$, $e_k^{(n)}$ はすべて正となる。しかし、この逆は成り立たない。pqds アルゴリズムにおいて上 2 重対角

行列 B が与えられれば, $q_k^{(0)} > 0, e_k^{(0)} > 0$ なる初期値が定まるが, (38) を通じて対応する dLVs の変数 $u_k^{(0)}$ をみていくと, $\delta^{(0)}$ をどのように選んでも, $u_3^{(0)} < 0$ となることがあり得る. 従って $1 + \delta^{(0)} u_k^{(0)} \leq 0$ となることがある. 一方, 同じ上 2 重対角行列 B について dLVs アルゴリズムの初期値をみると, 必ず, $u_k^{(0)} > 0$, 従って, $1 + \delta^{(0)} u_k^{(0)} > 1$ が成り立つ.

以上が dLV アルゴリズムの変数を用いて解明した pqds 法において数値不安定が発生しうるメカニズムである. 2006 年にリリース予定の LAPACK 4.0 では, 対称 3 重対角行列の高速・高精度な固有値計算に qd 型アルゴリズムである dqds (differential qd with shift) を主力アルゴリズムとして実装すると予告されている. dqds 法の漸化式

$$\begin{aligned} e_0^{(n)} &= 0, \quad d_1^{(n)} = q_1^{(n)} - \theta^{(n)2}, \quad q_k^{(n+1)} = d_k^{(n)} + e_k^{(n)}, \\ e_k^{(n+1)} &= \frac{e_k^{(n)}}{q_k^{(n+1)}} q_{k+1}^{(n)}, \quad d_{k+1}^{(n)} = d_k^{(n)} \frac{q_{k+1}^{(n)}}{q_k^{(n+1)}} - \theta^{(n)2} \end{aligned} \quad (41)$$

は中間変数 $d_k^{(n)}$ を付加した pqds 法の漸化式に他ならない [4]. 従って, dqds 法に期待される高速性もまた数値不安定性と隣り合わせであると言わざるを得ない.

可積分系起源のアルゴリズムであれば, 一般に, 四則演算だけの高速性に恵まれ丸め誤差の蓄積が少ないといえる. しかしながら, 代数的な意味での可積分性だけでは桁落ちや零割の発生は排除できず, 数値安定性は保証されない. dLV アルゴリズムは直交多項式の Christoffel 変換に対する解析的条件である $\delta^{(n)} > 0, u_k^{(n)} > 0$ (19) を制約条件とすることで, 数値安定なアルゴリズムと成り得たのである. この意味で「可積分な特異値計算アルゴリズム dLV」とは, 代数的な関係式 (18) だけでなく, 解析的な条件 (19) も含めての命名でなければならない.

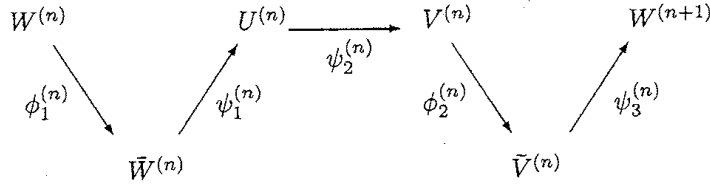
もっとも, dLV アルゴリズムの漸化式の一般項, すなわち, 離散 Lotka-Volterra 方程式のタウ関数解に基づく解の $n \rightarrow \infty$ での漸近的挙動 [7] から明らかなように, dLV アルゴリズムの特異値への収束次数は 1 次には過ぎない. 実用化可能なアルゴリズムとなるためには, まだ高いハードルが残されている.

4 mdLVs アルゴリズム

最近, 筆者のグループで開発した新構想のシフト付き dLV アルゴリズムである mdLVs アルゴリズムを解説する [10]. このアルゴリズムは正值性を壊さない範囲でシフト量を選ぶことが可能で, シフトの効果により, 収束の加速だけでなく, 丸め誤差の蓄積の減少による特異値の精度向上が顕著なものになる. mdLVs アルゴリズムは, 可積分性 $\delta^{(n)} > 0, u_k^{(n)} > 0$ を保つシフトを導入した dLV アルゴリズムと位置づけられる.

mdLVs アルゴリズムを記述するために, 変数の組

$$\begin{aligned} U^{(n)} &= \{u_k^{(n)}\}, \quad V^{(n)} = \{v_k^{(n)}\}, \quad \bar{V}^{(n)} = \{\bar{v}_k^{(n)}\}, \\ W^{(n)} &= \{w_k^{(n)}\}, \quad \bar{W}^{(n)} = \{\bar{w}_k^{(n)}\}, \quad (k = 1, 2, \dots, 2m-1), \end{aligned} \quad (42)$$

Fig. 2. dLV アルゴリズム $W^{(n)} \rightarrow W^{(n+1)}$

および、写像 $\psi_j^{(n)}, \phi_j^{(n)}$

$$\begin{aligned}
 \psi_1^{(n)} : \bar{W}^{(n)} &\rightarrow U^{(n)}, & u_k^{(n)} &= \bar{w}_k^{(n)} / (1 + \delta^{(n)} u_{k-1}^{(n)}), & u_0^{(n)} &\equiv 0, \\
 \psi_2^{(n)} : U^{(n)} &\rightarrow V^{(n)}, & v_k^{(n)} &= u_k^{(n)} (1 + \delta^{(n)} u_{k+1}^{(n)}), & u_{2m}^{(n)} &\equiv 0, \\
 \psi_3^{(n)} : \bar{V}^{(n)} &\rightarrow W^{(n+1)}, & w_k^{(n+1)} &= \bar{v}_k^{(n)}, \\
 \phi_1^{(n)} : W^{(n)} &\rightarrow \bar{W}^{(n)}, & \bar{w}_k^{(n)} &= w_k^{(n)}, \\
 \phi_2^{(n)} : V^{(n)} &\rightarrow \bar{V}^{(n)}, & \bar{v}_k^{(n)} &= v_k^{(n)}, & (k &= 1, 2, \dots, 2m-1)
 \end{aligned} \tag{43}$$

を準備する。具体的には、条件 $u_0^{(n)} \equiv 0$ のもとで、写像 $\psi_1^{(n)}$ によって

$$u_k^{(n)} = \frac{\bar{w}_k^{(n)}}{1} + \frac{\delta^{(n)} \bar{w}_{k-1}^{(n)}}{1} + \dots + \frac{\delta^{(n)} \bar{w}_2^{(n)}}{1 + \delta^{(n)} \bar{w}_1^{(n)}}$$

が定まる。 $\psi_2^{(n)}$ も条件 $u_{2m}^{(n)} \equiv 0$ のもとで、

$$\psi_2^{(n)} : (u_1^{(n)}(1 + \delta^{(n)} u_2^{(n)}), \dots, u_{2m-2}^{(n)}(1 + \delta^{(n)} u_{2m-1}^{(n)}), u_{2m-1}^{(n)}) \rightarrow (v_1^{(n)}, \dots, v_{2m-2}^{(n)}, v_{2m-1}^{(n)})$$

と書ける。 $U^{(n)}$ から $U^{(n+1)}$ への合成写像

$$\Psi_{dLV}^{(n+1)} := \psi_1^{(n+1)} \circ \phi_1^{(n+1)} \circ \psi_3^{(n)} \circ \phi_2^{(n)} \circ \psi_2^{(n)} : U^{(n)} \rightarrow U^{(n+1)} \tag{44}$$

を定義する。 $0 < u_k^{(n)}, 0 < \delta^{(n)} < M$ であれば、 $1 < 1 + \delta^{(n)} \bar{w}_k^{(n)}, 1 < 1 + \delta^{(n)} \bar{w}_k^{(n)}$ だから、 $\psi_j^{(n)}$, $(j = 1, 2, 3)$, $\phi_j^{(n)}$, $(j = 1, 2)$ はすべて $2m-1$ 個の正数の集合の間の全単射であり、合成写像 $\Psi_{dLV}^{(n+1)}$ は不等間隔離散 Lotka-Volterra 方程式 (21) に一致する。

同様に、合成写像

$$\Psi^{(n+1)} := \psi_3^{(n)} \circ \phi_2^{(n)} \circ \psi_2^{(n)} \circ \psi_1^{(n)} \circ \phi_1^{(n)} : W^{(n)} \rightarrow W^{(n+1)}$$

によって $W^{(n)}$ から $W^{(n+1)}$ への写像を定義する。Fig. 2 を参照されたい。 $0 < w_k^{(n)}, w_0^{(n)} \equiv 0, w_{2m}^{(n)} \equiv 0, 0 < \delta^{(n)} < M$ とすれば、 $0 < u_k^{(n)}, u_0^{(n)} = 0, u_{2m}^{(n)} = 0$ は明らかで、写像 $\Psi^{(n+1)}$ は変数 $W^{(n)}$ についてみた不等間隔離散 Lotka-Volterra 方程式である。

前節で示した dLV アルゴリズムの収束定理 1 は写像 $\Psi^{(n+1)}$ を使うと以下のように表現される。
上 2 重対角行列

$$B^{(n)} = \begin{pmatrix} \sqrt{w_1^{(n)}} & \sqrt{w_2^{(n)}} & & \\ & \sqrt{w_3^{(n)}} & \ddots & \\ & & \ddots & \sqrt{w_{2m-2}^{(n)}} \\ 0 & & & \sqrt{w_{2m-1}^{(n)}} \end{pmatrix}, \quad B^{(0)} = B$$

に対して, dLV アルゴリズムの 1 反復は $\Psi^{(n+1)} : W^{(n)} \rightarrow W^{(n+1)}$ と書かれる. $0 < w_k^{(0)}$ であれば, B の特異値 $\sigma_k(B)$ について $\sigma_1(B) > \sigma_2(B) > \dots > \sigma_m(B) > 0$ が成り立つ. このとき, 収束定理 1 より,

$$\begin{aligned} \Psi^{(n)} \circ \Psi^{(n-1)} \circ \dots \circ \Psi^{(1)} : W^{(0)} &= (w_{2k-1}^{(0)}, w_{2k}^{(0)}) \rightarrow W^{(n+1)}, \\ \lim_{n \rightarrow \infty} W^{(n+1)} &= (\sigma_k^2(B), 0) = (\lambda_k(B^\top B), 0) \end{aligned}$$

となる. ここに, $\lambda_k(B^\top B)$ は正定値 3 重対角対称行列 $B^\top B$ の第 k 固有値である. なお, 補題 1 より, $\lambda_k((B^{(n)})^\top B^{(n)})$ は離散 Lotka-Volterra 方程式の時間発展 $\Psi^{(n)}$ のもとで不変である.

さて, dLV アルゴリズムに対する原点シフトを

$$(\bar{B}^{(n)})^\top \bar{B}^{(n)} = (B^{(n)})^\top B^{(n)} - \theta^{(n)2} I, \quad \bar{B}^{(n)} := \begin{pmatrix} \sqrt{\bar{w}_1^{(n)}} & \sqrt{\bar{w}_2^{(n)}} & & \\ & \sqrt{\bar{w}_3^{(n)}} & \ddots & \\ & & \ddots & \sqrt{\bar{w}_{2m-2}^{(n)}} \\ 0 & & & \sqrt{\bar{w}_{2m-1}^{(n)}} \end{pmatrix} \quad (45)$$

によって導入しよう. ここで, $\theta^{(n)2}$ は離散時刻 $\sum_{i=0}^{n-1} \delta^{(i)}$ におけるシフト量である. 後述するよ
うに, 必ず $\bar{w}_k^{(n)} > 0$, $(k = 1, 2, \dots, 2m-1)$ が成り立つように $\theta^{(n)2}$ を選ぶことができる. 原点シフトを写像として表すのに, パラメータ $\theta^{(n)2}$ をもつ全単射

$$\phi_{1;\theta^{(n)}}^{(n)} : (w_{2k-2}^{(n)} + w_{2k-1}^{(n)} - \theta^{(n)2}, w_{2k-1}^{(n)} w_{2k}^{(n)}) \rightarrow (\bar{w}_{2k-2}^{(n)} + \bar{w}_{2k-1}^{(n)}, \bar{w}_{2k-1}^{(n)} \bar{w}_{2k}^{(n)}) \quad (46)$$

を導入する. ここに, 変数 $w_k^{(n)}$ と $\bar{w}_k^{(n)}$ について境界条件 $w_0^{(n)} \equiv 0$ と $\bar{w}_0^{(n)} \equiv 0$ を仮定している. 変数 $\bar{w}_k^{(n)}$, $(k = 1, 2, \dots, 2m-1)$ は, 写像 $\phi_{1;\theta^{(n)}}^{(n)}$ によって, $w_k^{(n)}$ から一意に

$$\begin{aligned} \bar{w}_{2k-1}^{(n)} &= w_{2k-1}^{(n)} + w_{2k-2}^{(n)} - \kappa_{2k-2}^{(n)} - \theta^{(n)2}, \\ \bar{w}_{2k}^{(n)} &= \frac{w_{2k-1}^{(n)} w_{2k}^{(n)}}{\bar{w}_{2k-1}^{(n)}}, \quad \bar{w}_{2k-2}^{(n)} = \kappa_{2k-2}^{(n)}, \\ \kappa_{2k-2}^{(n)} &:= \frac{w_{2k-2}^{(n)} w_{2k-3}^{(n)}}{w_{2k-3}^{(n)} + w_{2k-4}^{(n)} - \theta^{(n)2}} - \frac{w_{2k-4}^{(n)} w_{2k-5}^{(n)}}{w_{2k-5}^{(n)} + w_{2k-6}^{(n)} - \theta^{(n)2}} - \dots - \frac{w_2^{(n)} w_1^{(n)}}{w_1^{(n)} - \theta^{(n)2}} \end{aligned} \quad (47)$$

として計算される. 原点シフトの定義 (45) より, $w_1^{(n)} \geq \bar{w}_1^{(n)}$ だから, $w_{2k-1}^{(n)} \geq \bar{w}_{2k-1}^{(n)}$ および, $w_{2k}^{(n)} \leq \bar{w}_{2k}^{(n)}$ が成り立つ.

dLV アルゴリズム $\Psi^{(n+1)}$ において, 写像 $\phi_1^{(n)}$ をパラメータ付き写像 $\phi_{1;\theta^{(n)}}^{(n)}$ で置き換え, 合成写像

$$\Psi_{\theta^{(n)}}^{(n+1)} := \psi_3^{(n)} \circ \tilde{\phi}_2^{(n)} \circ \psi_2^{(n)} \circ \psi_1^{(n)} \circ \phi_{1;\theta^{(n)}}^{(n)} : W^{(n)} \rightarrow W^{(n+1)} \quad (48)$$

を定義する. もちろんシフト量 $\theta^{(n)2}$ が 0 のときは写像 $\Psi_{\theta^{(n)}}^{(n+1)}$ は dLV アルゴリズム $\Psi^{(n+1)}$ に帰着する. 合成写像

$$\Psi^{(n+1)} \circ \phi_1^{(n)-1} : W^{(n)} \rightarrow W^{(n+1)}$$

のもとで, 固有値は不変, すなわち, $\lambda_k((\bar{B}^{(n)})^\top \bar{B}^{(n)}) = \lambda_k((B^{(n+1)})^\top B^{(n+1)})$ だから,

$$\lambda_k((B^{(n+1)})^\top B^{(n+1)}) = \lambda_k((B^{(n)})^\top B^{(n)}) - \theta^{(n)2}$$

が成り立つ. 固有値はパラメータ $\theta^{(n)2}$ だけシフトを受ける. この意味で写像 $\Psi_{\theta^{(n)}}^{(n+1)}$ に基づく固有値, 特異値計算法を mdLVs アルゴリズムと名付ける.

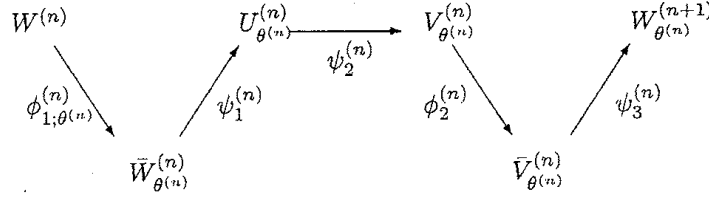


Fig. 3 mdLVs アルゴリズム $W^{(n)} \rightarrow W_{\theta^{(n)}}^{(n+1)}$

定理 2. mdLVs アルゴリズムの反復

$$\Psi_{\theta^{(N)}}^{(N+1)} \circ \Psi_{\theta^{(N-1)}}^{(N)} \circ \dots \circ \Psi_{\theta^{(0)}}^{(1)} : W^{(0)} \rightarrow W^{(N+1)}$$

によって得られる正定値 3 重対角対称行列 $(B^{(N+1)})^\top B^{(N+1)}$ の第 k 固有値は

$$\lambda_k((B^{(N+1)})^\top B^{(N+1)}) = \lambda_k((B^{(0)})^\top B^{(0)}) - \sum_{n=0}^N \theta^{(n)2}. \quad (49)$$

と表される。□

次に、いかにシフト量 $\theta^{(n)2}$ を選べば $\bar{w}_k^{(n)} > 0$, $(k = 1, 2, \dots, 2m-1)$ が成り立つかを論じる。 $\bar{w}_k^{(n)} > 0$ を要請するのは以下の理由による。もしある n と k において $w_{2k-1}^{(n)} = 0$ となれば、写像 $\phi_{1;\theta^{(n)}}^{(n)}$ によって $w_{2k}^{(n)}$ がオーバーフローを起こす。また、 $w_k^{(n)} < 0$ となれば、 $\delta^{(n)}$ の選択によっては $1 + \delta^{(n)} u_k^{(n)} = 0$ となり、写像 $\psi_1^{(n)}$ によって $u_{l+1}^{(n)}$ がオーバーフローとなるからである。従って、あまり大きすぎるシフト量を選ぶと $\bar{w}_k^{(n)} \leq 0$ となり数値計算が破綻する。アルゴリズムとして重要なのは、ある範囲にシフト量 $\theta^{(n)2}$ を選べば、必ず $\bar{w}_k^{(n)} > 0$ となるという保証である。

mdLVs アルゴリズムについて以下が成り立つ [10].

定理 3. $w_k^{(n)} > 0$, $(k = 1, 2, \dots, 2m-1)$ と仮定する。このとき、 $\bar{w}_k^{(n)} > 0$, $(k = 1, 2, \dots, 2m-1)$

であるためには、シフト量 $\theta^{(n)2}$ が不等式

$$\theta^{(n)2} < \sigma_m^2(B^{(n)}) \quad (50)$$

を満たすことが必要十分である。ここに、 $\sigma_m(B^{(n)})$ は $B^{(n)}$ の最小特異値である。□

系 1. $w_{2k-1}^{(n)} \geq \varepsilon_1$, $(k = 1, 2, \dots, m)$, $w_{2k}^{(n)} > 0$, $(k = 1, 2, \dots, m-1)$ と仮定する。このとき、

$$\bar{w}_{2k-1}^{(n)} \geq \varepsilon_2, \quad (k = 1, 2, \dots, m), \quad \bar{w}_{2k}^{(n)} > 0, \quad (k = 1, 2, \dots, m-1) \quad (51)$$

であるためには、シフト量 $\theta^{(n)2}$ が不等式

$$\theta^{(n)2} \leq \sigma_m^2(B^{(n)}) - \varepsilon_3 \quad (52)$$

を満たすことが必要十分である。ここに、 $\varepsilon_1, \varepsilon_2, \varepsilon_3$ は小さな正の定数である。□

最小特異値 $\sigma_m(B^{(n)})$ の推定法として Johnson[12] による Gerschgorin 型境界がある。それによれば $\sigma_m(B^{(n)})$ は

$$\sigma_m(B^{(n)}) \geq \max\{0, \vartheta_1^{(n)}\}, \quad \vartheta_1^{(n)} := \min_k \left\{ \sqrt{w_{2k-1}^{(n)}} - \frac{1}{2} \left(\sqrt{w_{2k-2}^{(n)}} + \sqrt{w_{2k}^{(n)}} \right) \right\} \quad (53)$$

によって下から見積られる。定理 3 と合わせることで mdLVs アルゴリズムが有限桁精度計算においても数値不安定を起こさない安全なシフト量の設定法を得る [10].

定理 4. $w_{2k-1}^{(n)} \geq \varepsilon_1$, ($k = 1, 2, \dots, m$), $w_{2k}^{(n)} > 0$, ($k = 1, 2, \dots, m-1$) と仮定する。このとき、シフト量を

$$\theta^{(n)^2} = \max\{0, (\vartheta_1^{(n)})^2 - \varepsilon\} \quad (54)$$

とすれば、mdLVs アルゴリズムは数値不安定性とならない。ここに、 ε は小さな正の定数である。

□

実際、 $(\vartheta_1^{(n)})^2 \geq \varepsilon$ のとき、 $\vartheta_1^{(n)} > 0$ だから、 $\sigma_m^2(B^{(n)}) - \varepsilon \geq (\vartheta_1^{(n)})^2 - \varepsilon = \theta^{(n)^2}$ となり、系 1 によって $\bar{w}_{2k-1}^{(n)} \geq \varepsilon_2$ が成り立つ。一方、定理 3 でみたように、 $\sigma_m^2(B^{(n)}) > \theta^{(n)^2}$ であれば $\bar{w}_{2k-1}^{(n)} > 0$ となるが、丸め誤差によつては $\bar{w}_{2k-1}^{(n)} \approx 0$ となることがありえる。ゆえに、例えば、定理 4 のようにシフト量を選ぶことで $\bar{w}_{2k-1}^{(n)} \geq \varepsilon_2$ となり、丸め誤差があつても、mdLVs アルゴリズムの漸化式 (47) の除算において零割は発生せず、数値不安定となることはない。なお、 $\bar{w}_{2k-1}^{(n)} \geq \varepsilon_2$ であれば、mdLVs アルゴリズムの定義により、 $u_{2k-1}^{(n)} \geq \varepsilon_4$, $w_{2k-1}^{(n+1)} \geq \varepsilon_5$ が成立することに注意する。定理 4 は一反復において数値不安定が起きないシフト量の設定方法を与えている。定理に従つて毎回シフト量を取り直ししながら、数値安定な反復計算を繰り返すことで、特異値への収束が加速される。

以下では、いよいよ mdLVs アルゴリズムの収束定理を証明する。定理は無限回のシフト付き反復計算を行うという最悪の場合を想定して数学的に証明されるが、その結果、有限桁精度の数値計算では、丸め誤差があつても、必要な精度に有限回の反復で必ず収束することがわかり、収束性の保証された原点シフト付きの新しい特異値計算法であることが明らかとなる。

いくつかの補題を準備する。初期値 $w_k^{(0)}$ に対する mdLVs アルゴリズムの反復計算 $\psi_{\theta^{(N+1)}}^{(N+1)} \circ \psi_{\theta^{(N-1)}}^{(N)} \circ \dots \circ \psi_{\theta^{(0)}}^{(1)} : W^{(0)} \rightarrow W^{(N+1)}$ を、ここでは、 $w_k^{(N+1)} = \psi_{\theta^{(N+1)}}^{(N+1)} \circ \psi_{\theta^{(N)}}^{(N)} \circ \dots \circ \psi_{\theta^{(1)}}^{(1)}(w_k^{(0)})$ とかく。

補題 2. 初期値に関する有界性 $0 < w_k^{(0)} < M_1$ を仮定する。このとき、ある正の定数 M_2 が存在して

$$0 < w_k^{(N+1)} < M_1, \quad 0 < u_k^{(N)} < M_2$$

が成り立つ。□

補題 3. 初期値 $0 < w_k^{(0)} < M_1$ に対する mdLVs アルゴリズムで生成された変数 $w_k^{(n+1)}$ は

$$w_k^{(N+1)} = \prod_{n=0}^N \left(\frac{1}{\gamma_k^{(n)}} \cdot \frac{1 + \delta^{(n)} u_{k+1}^{(n)}}{1 + \delta^{(n)} u_{k-1}^{(n)}} \right) w_k^{(0)} \quad (55)$$

のように表される。ここに、 $\gamma_k^{(n)}$ は $\gamma_{2k-1}^{(n)} \geq 1$, $0 < \gamma_{2k}^{(n)} \leq 1$ なる適当な定数である。□

mdLV アルゴリズムについて顕著なのは、変数 $w_k^{(n)}$ についてみたとき、 $n \rightarrow \infty$ である有限な値に収束する無限積として表されることである。

命題 3. $n \rightarrow \infty$ のとき、変数 $w_k^{(n)}$ は k の偶奇に応じて、それぞれ、変数 $w_{2k-1}^{(n)}$, ($k = 1, 2, \dots, m$) は $c_1 > c_2 > \dots > c_m > 0$ なる正の定数 c_k に、 $w_{2k}^{(n)}$, ($k = 1, 2, \dots, m-1$) は 0 に収束する。すなわち、

$$w_{2k-1}^{(n)} \rightarrow c_k, \quad w_{2k}^{(n)} \rightarrow 0 \quad (56)$$

が成り立つ。□

変数 $w_k^{(n)}$ の極限と上 2 重対角行列の特異値 $\sigma_k^2(B^{(0)})$, および、シフト量の間係をまとめよう。

定理 5. 上 2 重対角行列 $B^{(0)}$ の成分について $b_k > 0$, ($k = 1, \dots, 2m-1$) であるものとする。初期値を $w_k^{(0)} = b_k^2$ とすれば、mdLVs アルゴリズムによって、変数 $w_k^{(n)}$ は、 $n \rightarrow \infty$ で

$$\lim_{n \rightarrow \infty} w_{2k-1}^{(n)} = \sigma_k^2(B^{(0)}) - \sum_{n=0}^{\infty} \theta^{(n)2}, \quad \lim_{n \rightarrow \infty} w_{2k}^{(n)} = 0 \quad (57)$$

のように収束する。もし有限回 (N 回) の反復で収束するときには、

$$w_{2k-1}^{(N)} = \sigma_k^2(B^{(0)}) - \sum_{n=0}^{N-1} \theta^{(n)2}, \quad w_{2k}^{(N)} = 0 \quad (58)$$

となる。ここに、 $\sigma_1 > \sigma_2 > \dots > \sigma_m > 0$ である。□

以上により、シフト付き特異値計算アルゴリズム mdLVs の特異値への収束性が証明された。関係式

$$\bar{w}_k^{(n)} = u_k^{(n)}(1 + \delta^{(n)} u_{k-1}^{(n)}), \quad v_k^{(n)} = u_k^{(n)}(1 + \delta^{(n)} u_{k+1}^{(n)})$$

を用いて変数 $u_k^{(n)}$ を消去すると、mdLVs アルゴリズムの計算は漸化式

$$\bar{w}_{2k}^{(n+1)} = \frac{v_{2k-1}^{(n)} v_{2k}^{(n)}}{\bar{w}_{2k-1}^{(n+1)}}, \quad \bar{w}_{2k-1}^{(n+1)} = v_{2k-2}^{(n)} + v_{2k-1}^{(n)} - \bar{w}_{2k-2}^{(n+1)} - \theta^{(n+1)2} \quad (59)$$

の形で進行する。pqds 法の漸化式 (25) に似ているが、違いは明白である。系 1 でみたように、適切なシフト選択により $\bar{w}_{2k-1}^{(n+1)} \geq \varepsilon_2 > 0$ が保たれ、丸め誤差が生じて、零割はおろか精度を悪化させる 0 に近い数での除算もない。従って、数値安定に高精度な特異値計算が実行可能になる。なお、論文 [9] において、mdLVs アルゴリズムは特異値に通常は 2 次収束し、最高では 3 次収束することが示されている。

Table 1: DBDSQR, dLV, mdLVs による特異値計算時間 (sec.)

	$m = 100$			$m = 1000$	
	DBDSQR	dLV	mdLVs	DBDSQR	mdLVs
B_1	0.02	0.27	0.02	2.20	1.37
B_2	0.03	0.13	0.02	2.27	1.34
B_3	0.02	174	0.02	2.00	1.32

5 mdLVs アルゴリズムによる特異値計算例

mdLVs アルゴリズムの数値計算例を与える。

Table 1 での比較対象は、同様に、収束性と数値安定性が保証され固有値に最大で 3 次収束することがわかっているシフト付き QR 法を特異値計算向けに改良した Demmel-Kahan 法である。具体的には、Demmel-Kahan 法を実装した LAPACK の DBDSQR ルーチン、dLV アルゴリズムと mdLVs アルゴリズムの試作ルーチンを比較する。収束判定は同一条件を採用する。mdLVs アルゴリズムにおいてパラメータ $\delta^{(n)}$ は $\delta^{(n)} = 1$ に固定する。また、シフト量は Johnson 境界を用いて設定し、減次や分割が起きれば小さな行列の特異値計算に帰着させるものとする。数値実験の対象とする行列は以下の 3 種類の 100 次上 2 重対角行列である。

- 1) 各特異値は互いに十分に分離しており、零に近い特異値をもたない行列の例として

$$B_1 = \text{bidiag} \left\{ \begin{array}{ccccc} & 2 & & 2 & \cdots & 2 \\ 2.001 & & 2.001 & & \cdots & 2.001 \end{array} \right\}$$

を取り上げる。100 次のとき B_1 の特異値の真値は、順に、 $\hat{\sigma}_1 = 4.000511306 \dots$, $\hat{\sigma}_2 = 3.999045346 \dots$, $\dots$, $\hat{\sigma}_{99} = 0.094010676 \dots$, $\hat{\sigma}_{100} = 0.031906725 \dots$ 等である。

- 2) 各特異値は互いに分離しており、零に極めて近い特異値一つをもつ行列、すなわち、条件数の大きな行列の例として

$$B_2 = \text{bidiag} \left\{ \begin{array}{ccccc} & 10 & & 10 & \cdots & 10 \\ 1 & & 1 & & \cdots & 1 \end{array} \right\}$$

を考える。100 次のとき B_2 の特異値の真値は、順に、 $\hat{\sigma}_1 = 10.99955222 \dots$, $\hat{\sigma}_2 = 10.99820922 \dots$, $\dots$, $\hat{\sigma}_{99} = 9.000549469 \dots$, $\hat{\sigma}_{100} = 0.000000000 \dots$ 等である。

- 3) 3 節で扱った集積した特異値分布をもつ行列 B_3 。

これらの行列では特異値分布は行列の次数の違いによらず、 $m = 100, 1000$ でほぼ同じ傾向をもつ。Table 1 に計算時間に関する数値実験結果を与える（計算機 CPU: Pentium 3, CPU850MHz, RAM:128MB における倍精度計算）。シフトの導入によりどのタイプでも高速化が実現できていること、行列の次数の増加と計算時間の増加を比較して mdLVs アルゴリズムが優れた scalability をもち、DBDSQR ルーチンの特異値計算部より高速であることがわかる。

次に、Fig. 4 では、集積した特異値分布をもつ 100 次の行列 B_3 について、計算された 100 個の特異値 (x -軸) のもつ相対誤差 (y -軸) を比較する。シフトの導入によって Fig. 1ab でみた dLV アル

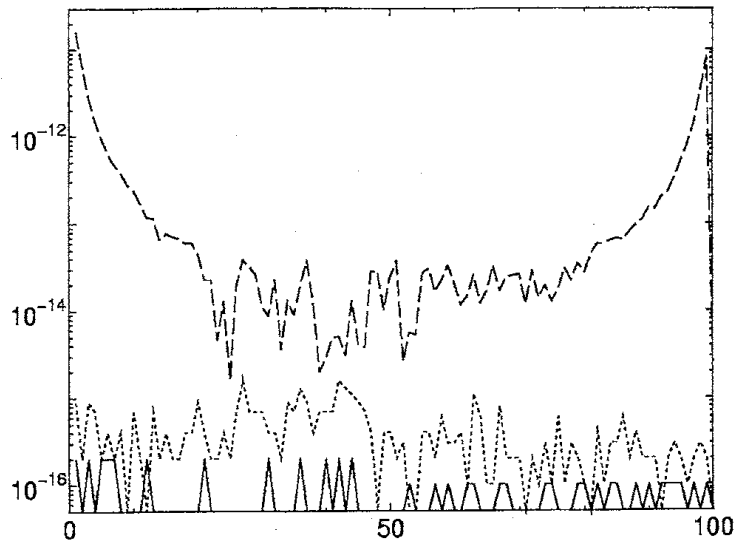


Fig. 4. DBDSQR, dLV, mdLVs による特異値の相対誤差分布 ($m = 100$)

ゴリズムの相対誤差（鎖線）は大きく減少している。シフト付き Demmel-Kahan QR 法を実装した DBDSQR ルーチンによる相対誤差（点線）と比較しても mdLVs アルゴリズム（実線）の高精度性は際だっており、ほぼすべての特異値についてマシンイプシロン（この実験では $\epsilon = 2.22 \times 10^{-16}$ ）以下で計算されている。なお、Fig. 4 では特異値の真値と相対誤差に計算にのみ多倍長計算を行っている。

Fig. 5 は条件数が $10^{10} \sim 10^{58}$ の 1000 次上 2 重対角行列をランダムに 100 個生成し、横軸にその条件数の対数を、縦軸に計算された 1000 個の特異値の相対誤差の総和を、mdLVs アルゴリズムを実装した DLVS ルーチン ($\diamond$)、LAPACK の DLASQ ルーチン ($\square$)、Demmel-Kahan QR 法の LAPACK の DBDSQR ルーチン ($\triangle$) についてプロットしたものである。コンパイラは PGI Fortran である。アルゴリズムの精度を忠実にみるため SSE 機能は使っていない。DLASQ は収束証明はないものの現在の最高精度といわれる dqds 法 [4, 14] を実装した正定値対称 3 重対角行列の固有値計算用ルーチンである。DLASQ では零割が起きそうになるとシフト量を取り直すことで、精度の悪化を防いでいる。条件数の違いによらず、100 個の全てにおいて相対誤差の総和は

$$DLVS(mdLVs) < DLASQ(dqds) \ll DBDSQR(QRs)$$

となり、シフト付き特異値計算における mdLVs アルゴリズムの高精度性が裏付けられる。（さらに多くの例や多様な計算機環境のもとでの精度と計算時間の比較数値実験については [20, 21] を参照されたい）

6 おわりに

pqds 法や dqds 法と異なり、mdLVs アルゴリズムは収束証明をもつ。mdLVs アルゴリズムのシフトの導入プロセスは漸化式で表すと (59) となり、pqds アルゴリズムや dqds アルゴリズムとよく似ている。しかし、定理 3, 系 1 でみたように、mdLVs アルゴリズムでは最小特異値の平方

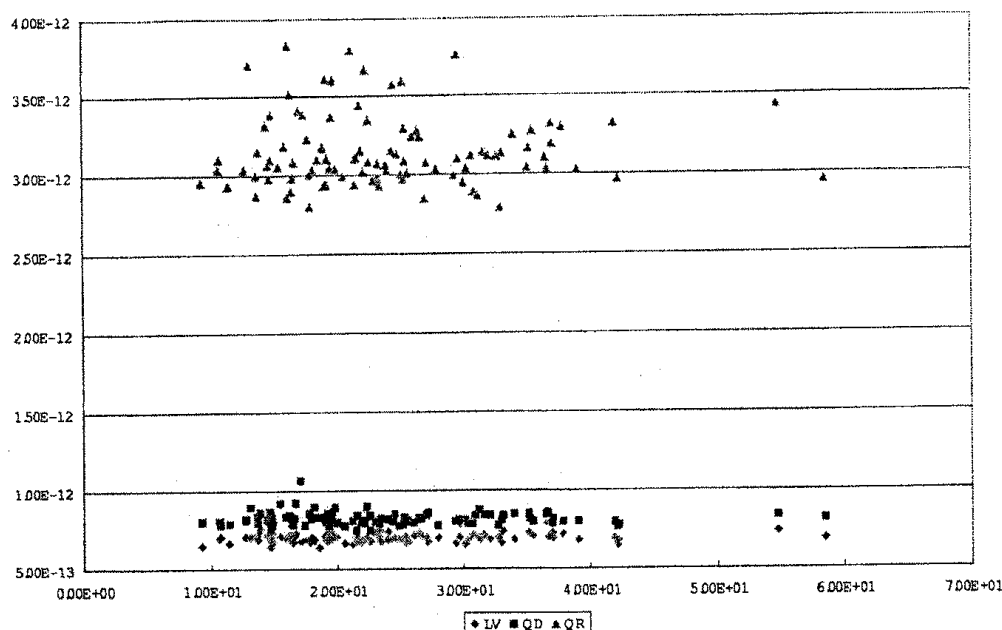


Fig. 5. DBDSQR, DLASQ, DLVS(mdLVs) による特異値の相対誤差の総和 ($m = 1000$)

よりシフト量を小さく選べば、分母に現れる変数 $\bar{w}_k^{(n)}$ の正値性が壊れることはなく、変数は数値安定に特異値に収束する。mdLVs アルゴリズムの収束の速さは

$$R_2 := \max_k \frac{\sigma_{k+1}^2 - \theta^{(n)^2} + 1/\delta^{(n)}}{\sigma_k^2 - \theta^{(n)^2} + 1/\delta^{(n)}} \quad (60)$$

によって支配され、pqds 法や dqds 法のように $\delta^{(n)} < 0$ とすることなく、収束の加速が可能である。シフトの導入の影響は正の値をとる変数 $\bar{w}_k^{(n)}$ と $w_k^{(n)}$ の間のスケール変換 $w_k^{(n)} = \gamma_k^{(n)} \bar{w}_k^{(n)}$ に吸収される。mdLVs アルゴリズムの変数は、dLV の変数である $u_k^{(n)}$ 、不等間隔差分ステップサイズ $\delta^{(n)}$ に加えて、補助変数 $w_k^{(n)}$ と $\bar{w}_k^{(n)}$ である。これらはすべて正の値をとる変数である。この結果、漸化式の分母は常に

$$1 + \delta^{(n)} u_k^{(n)} > 1$$

となり、高い相対精度が保証される。同時に、数列 $\prod_{n=0}^N (1 + \delta^{(n)} u_{k\pm 1}^{(n)})$ の単調性と有界性から変数 $w_k^{(n)}$ の特異値への収束性が従う。mdLVs アルゴリズムは、シフト付き QR アルゴリズムとは異なる動作原理をもつ、収束性の保証された新しい高速・高精度特異値計算アルゴリズムといえることができる。

研究の背景や定理の証明、および、I-SVD アルゴリズムの全体像については中村著「可積分系の機能数理」(共立出版、近刊)を参照されたい。本研究を共に進めてきた岩崎雅史氏、高田雅美氏、木村欣司氏、辻本諭氏、阪野真也氏、松井佑貴夫氏、菅田太朗氏など多くの共同研究者に感謝する。

参考文献

- [1] T. S. Chihara, An Introduction to Orthogonal Polynomials, Gordon & Breach, New York, 1978.
- [2] J. Demmel, Applied Numerical Linear Algebra, SIAM, Philadelphia, 1997.
- [3] J. Demmel and W. Kahan, Accurate singular values of bidiagonal matrices, SIAM J. Sci. Sta. Comput., **11**(1990), 873-912.
- [4] K. V. Fernando, B. N. Parlett, Accurate singular values and differential qd algorithms, Numer. Math., **67**(1994), 191-229.
- [5] G. H. Golub and C. F. Van Loan, Matrix Computation, Third Edition, The Johns Hopkins Univ. Press, Baltimore, 1996.
- [6] R. Hirota, Conserved quantities of "random-time Toda equation", J. Phys. Soc. Japan **66**(1997), 283-284.
- [7] M. Iwasaki and Y. Nakamura, On a convergence of solution of the discrete Lotka-Volterra system, Inverse Problems, **18**(2002), 1569-1578.
- [8] M. Iwasaki and Y. Nakamura, An application of the discrete Lotka-Volterra system with variable step-size to singular value computation, Inverse Problems, **20** (2004), 553-563.
- [9] 岩崎雅史, 中村佳正, 特異値計算アルゴリズム dLV の基本性質について, 日本応用数理学会論文誌, **15**(2005), 287-306.
- [10] M. Iwasaki and Y. Nakamura, Accurate computation of singular values in terms of shifted integrable schemes, preprint, 2005.
- [11] 岩崎雅史, 阪野真也, 中村佳正, 実対称 3 重対角行列の高精度ツイスト分解とその特異値分解への応用, 日本応用数理学会論文誌, **15**(2005), 461-481.
- [12] C. R. Johnson, A Gersgorin-type lower bound for the smallest singular value, Lin. Alg. Appl., **112** (1989), 1-7.
- [13] LAPACK, <http://www.netlib.org/lapack/>
- [14] B. N. Parlett and O. A. Marques, An implementation of the dqds algorithm (positive case), Lin. Alg. Appl., **309**(2000), 217-259.
- [15] H. Rutishauser, Ein Quotienten-Differenzen-Algorithmus, Z. angew. Math. Phys., **5**(1954), 233-251.
- [16] H. Rutishauser, Lectures on Numerical Mathematics, Birkhäuser, Boston, 1990.
- [17] F. シャトラン, (伊理正夫, 伊理由美訳), 行列の固有値, シュプリンガー・フェアラーク東京, 1993.
- [18] V. Spiridonov and A. Zhedanov, Discrete Darboux transformations, the discrete-time Toda lattice, and the Askey-Wilson polynomials, Methods Appl. Anal. **2**(1995), 369-398.
- [19] V. Spiridonov and A. Zhedanov, Discrete-time Volterra chain and classical orthogonal polynomial, J Phys. A: Math., **30** (1997), 8727-8737.
- [20] 高田雅美, 岩崎雅史, 木村欣司, 中村佳正, 高精度特異値計算ルーチンの開発とその性能評価, 情報処理学会論文誌, **46**, No.SIG12(2005), 299-311.
- [21] M. Takata, K. Kimura, M. Iwasaki and Y. Nakamura, An evaluation of singular value computation by the discrete Lotka-Volterra system, Proceedings of The 2005 International Conference on Parallel and Distributed Processing Techniques and Applications, Vol. II, 2005, pp. 410-416.
- [22] 高田雅美, 木村欣司, 岩崎雅史, 中村佳正, 高速特異値分解のためのライブラリ開発, preprint, 2005.
- [23] S. Tsujimoto, Y. Nakamura and M. Iwasaki, The discrete Lotka-Volterra system computes singular values, Inverse Problems, **17**(2001), 53-58.
- [24] J. H. Wilkinson, The Algebraic Eigenvalue Problem, Clarendon Press, Oxford, 1965.
- [25] 山本哲朗, 数値解析入門, サイエンス社, 1976.
- [26] 柳井晴夫, 竹内啓, 射影行列一般逆行列特異値分解, 東京大学出版, 2000.