

Recent Developments of the mdLVs Algorithm for Computing Matrix Singular Values

行列の特異値計算の mdLVs アルゴリズムにおける最近の進展

京都大学情報学研究科 中村 佳正¹(Yoshimasa Nakamura)

Graduate School of Informatics, Kyoto University

京都府立大学人間環境学部 岩崎 雅史 (Masashi Iwasaki)

Faculty of Human Environment, Kyoto Prefectural University

新潟大学自然科学研究科 木村 欣司 (Kinji Kimura)

Graduate School of Science & Technology, Niigata University

奈良女子大学人間文化研究科 高田 雅美 (Masami Takata)

Graduate School of Humanities & Sciences, Nara Women's University

Abstract. Any attempt to develop a numerical linear algebra library superior to the standard library such as LAPACK has dropped off for long time in Japan. Recently a new singular value decomposition (SVD) algorithm named I-SVD and its code have been developed in Japan. Here SVD is one of the most basic mathematical tools in information processing such as data search, signal separation and image compression. The I-SVD algorithm consists of two parts. One is a new stable bidiagonal singular value computing algorithm named the mdLVs. The other is a singular vector computation by a new twisted factorization of dLV-type. The I-SVD is a new $O(N^2)$ algorithm for SVD and has a good orthogonality of singular vectors. In this report the I-SVD algorithm (the mdLVs algorithm and the twisted factorization of dLV-type) and its potential are surveyed.

1 はじめに

米国の標準ライブラリ LAPACK[16] より優れたコードを日本で独自に開発しようという動きは途絶えていた。最近、我々は新しい理論的基盤をもつ特異値分解法 I-SVD (Integrable SVD) を開発し[17], さらに, I-SVD 法を DBDSLVS コードに実装した[23]。DBDSLVS の計算量は $O(N^2)$ であり, LAPACK の既存の特異値分解コードである DBDSQR の $O(N^3)$, DBDSLVS の $O(N^2) \sim O(N^3)$ と比較して多くの行列で圧倒的に高速である。

I-SVD は mdLVs(modified dLV with shift) 法[13] と呼ばれる特異値計算部, dLV 型ツイスト分解法[15] による特異ベクトル計算部からなり, 上 2 重対角行列への前処理部と逆変換部と接続することで, 原理的には任意の長方形行列の特異値分解が計算可能である。アルゴリズムの開発にあたっては我が国で高度に発展した離散可積分系の理論と直交多項式の変形理論が本質的な役割を果たした。I-SVD 法では, mdLVs 法によって高い相対精度で高速に計算された特異値をもとに, dLV 型ツイスト分解で特異ベクトルを高速計算する。この結果, $N \times N$ 上 2 重対角行列の特異値分解の計算量を $O(N^2)$ に止めることに成功している。よほど近接していない限り, 計算された特異ベクトルは十分な直交性をもつ。しかも, 特異ベクトル計算はベクトルごとに並列実行可能であるため, I-SVD 法は高い並列化性能を有し, BLAS(Basic Linear Algebra Subprograms) による高速行列乗算を駆使した前処理・逆変換部の高速化と合わせて, 超大規模特異値分解問題を実用時間で解くことも視野に入っている。

¹ynaka@i.kyoto-u.ac.jp

2 行列の特異値分解

まず、行列の特異値分解 (SVD) についてまとめておく。一般の長方形行列 $A \in \mathbf{R}^{n \times m}$ 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、通常の特異値は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。長さ 1 で互いに直交する適当な次元のベクトル $\mathbf{u}_j, \mathbf{v}_j$ を並べた行列 $U := (\mathbf{u}_1, \dots, \mathbf{u}_r), V := (\mathbf{v}_1, \dots, \mathbf{v}_r)$, 対角行列 $\Sigma := \text{diag}(\sigma_1, \dots, \sigma_r)$ を用いて A は

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

と表される [8]。 (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。 $\text{rank} A = m = n$ 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 [8]。 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 法に基づいている。

- (i) Golub-Kahan 法では、まず、QR 法の計算量を減らすため、前処理として Householder 変換のくり返しで A を上 2 重対角行列 B に変形する。

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

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 に乗することで 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 $2m^3/3 + O(m^2)$ 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回数である。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

$$\sigma_1 > \sigma_2 > \dots > \sigma_m$$

となる[26]。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、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3重対角対称行列 $B^T B = V_H^T A^T A V_H$ に対して固有値計算のQR法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。

- (iii) 右特異ベクトルは $V_0 = V_H Q Q' Q'' \dots$ によって与えられる。左特異ベクトル U_0 は $A V_0$ に対する(列ベクトルの pivoting を伴う)QR分解 $A V_0 P = U_0 R$ 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 $4m^3 + O(m^2)$ flops である。

ステップ(ii)において素朴なQR法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、近接特異値を極力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 法では(ii)の特異値計算部では、 m 次3重対角対称行列 $B^T B$ の3重対角部の2次小行列に対して、順次、シフト付きQR法を適用して、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 $6mflops$ である。通常の停止条件のもとで $6m^2 flops$ 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とされている。

固有値・特異値計算では、大きな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 SIAG/LA 賞を受賞している。Demmel-Kahan 法は LAPACK[16] において DBDSQR コードとして実装されており、MATLAB, Mathematica 等の汎用ソフトウェアで広く使われている。シフト付きQR法による固有値計算の収束次数は最大で3次であり[25]、十分な高速性をもつとされてきた。

(ii) のシフト付きQR法で用いた直交行列 $QQ' \dots$ はそのまま(iii)の右特異ベクトルの元になる。逆にみれば、(ii)は(iii)と不可分であり、(ii)の計算量が少ないからといって、直交行列 $QQ' \dots$ を計算することなく特異値計算が終了するわけではない。密行列の1回の積自体が $O(m^3)$ の計算量を必要とする。このため、実際には、(iii)の計算量が大きく、Demmel-Kahan 法による2重対角行列の特異値分解は $O(m^3)$ アルゴリズムである。このため、通常の計算機環境では、数千次

の行列の特異値分解に数千秒が必要であり [15], QR 法が中規模行列向きとされているのと同様, Demmel-Kahan 法も中規模行列用の特異値分解法といえよう.

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法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 法を定式化し [24, 10, 11, 12], 原点シフトの導入による高速化である mdLVs 法 [13] を開発した. 特異ベクトルについては, dLV 型ツイスト分解による高速な特異ベクトル計算法 [15] に到達し, 全体で $O(m^2)$ の計算量の 2 重対角特異値分解法 I-SVD を実現した [17]. コード開発としては, mdLVs 法を実装した DLVS コードの性能評価 [21, 22] を経て, I-SVD を実装した DBDSLVS コードの性能評価 [23] へと発展している.

3 mdLVs 法と dLV 法による特異値計算

mdLVs 法は特異値に 1 次収束する dLV 法に原点シフトを導入したものである. dLV 法の漸化式は対称な直交多項式の変形方程式であり, モーメントの正値性に起因して, 変数の正値性が理論的に保証され, 相対誤差の意味での高精度性, 指数関数的安定性など, 従来の固有値・特異値計算アルゴリズムにない優れた数値的性質をもっている [17]. dLV 法は低速ではあるが, 変数の正値性を壊さないシフトの導入法は明白であり, その結果, 高速高精度の mdLVs 法の開発 [13] が可能となったものである.

有限桁計算においては, マシンイプシロン未満の小数は零と認識される. このため, シフト量が余りに目標値に近すぎると変数の値が零となり零割の恐れがでてくる. このため, mdLVs 法の実装コードにおいては, シフトなしの dLV 反復を行って, 再度 mdLVs 反復に戻るよう設計されている. 収束の終盤ではこのようなことが頻繁に起きる. ゆえに, シフト付きの高速アルゴリズムにおいても, シフトなしの場合の精度や速度等の基本性能は極めて重要である.

本節では, まず, 高精度の特異値計算法である dLV 法について, その特異値への収束性を中心に解説する. アイデアの源泉となった離散可積分系の側面についても触れたい. 続いて, 高速高精度の特異値計算法である mdLVs 法について述べる. Demmel-Kahan 法で用いられた QR 法は絶対誤差の意味で高精度に過ぎない. mdLVs 法は相対誤差の意味で高精度であるだけでなく, Demmel-Kahan 法よりもかなり高速である. なお, mdLVs 法は理想的なシフト戦略のもとでは 3 次収束するが, どのようにシフト量を選べばよいかは現在もなお研究が継続されている.

dLV 法の漸化式は次の通りである. 直交多項式の変形理論に基づくこの漸化式の導出と変数 $u_k^{(n)}$ の正値性については [17] を参照されたい. $k = 1, 2, \dots, 2m - 1$ について初期値 $u_k^{(0)}$ が既知, $n = 0, 1, \dots$ についてパラメータ $\delta^{(n)}$ が与えられているとすれば, 漸化式

$$\begin{aligned} u_k^{(n+1)}(1 + \delta^{(n+1)}u_{k-1}^{(n+1)}) &= u_k^{(n)}(1 + \delta^{(n)}u_{k+1}^{(n)}), \\ u_0^{(n)} &\equiv 0, \quad u_{2m}^{(n)} \equiv 0, \quad 0 < \delta^{(n)} < M \end{aligned} \quad (3)$$

によって変数 $u_k^{(n)}$ は逐次的に定まる. 簡単のため $\delta^{(n)} \equiv 1$ としてよい. $u_k^{(n)}$ を時刻 $t = \sum_{j=0}^{n-1} \delta^{(j)}$ における変数 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 (3) は $u_k = u_k(t)$ についての微分方程式

$$\frac{du_k}{dt} = u_k(u_{k+1} - u_{k-1}), \quad u_0 = 0, \quad u_{2m} = 0 \quad (4)$$

に移行する。微分方程式 (4) は生物数理モデルに現れる Lotka-Volterra(LV) 方程式に他ならない。この方程式は非線形連立方程式であるが、指数関数を用いて解を書き下せる可積分系である。この意味で、漸化式 (3) は $\delta^{(n)}$ を差分間隔とする離散 Lotka-Volterra(discrete LV, dLV) 方程式である。可積分系研究においてこの方程式は $\delta^{(n)} \neq 0$ なる条件のもとに導出されたが [9, 20], 変数 $u_k^{(n)}$ の正值性のためには条件 $\delta^{(n)} > 0$ が必要である。LV 方程式と行列の特異値計算との関わりについては M. Chu[1] が以下を示している。初期値を $u_k(0) = b_k > 0$ と与えるとき、解は $t \rightarrow \infty$ で

$$\lim_{t \rightarrow \infty} u_{2k-1}(t) = \sigma_k^2 > 0, \quad \lim_{t \rightarrow \infty} u_{2k}(t) = 0 \quad (5)$$

のなる。ここに、 σ_k は B の特異値である。微分方程式の高精度差分法を用いて LV 方程式を差分化し、その漸化式を用いて極限 σ_k^2 を近似的に計算すれば良いようであるが、高精度に計算しようとすれば差分間隔を小さく選ばねばならず、反復回数が極端に増加して丸め誤差が無視できないものとなる。このため M. Chu の発見は新しいアルゴリズムの開発に結びつくことはなかった。

漸化式 (3) による特異値計算において、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 B の特異値が計算できるかである。単純に $u_k^{(0)} = b_k$ と与えたのでは正しい特異値への収束は望めない。

dLV 方程式 (3) の行列表示

$$L^{(n+1)}R^{(n+1)} = R^{(n)}L^{(n)} - \left(\frac{1}{\delta^{(n)}} - \frac{1}{\delta^{(n+1)}} \right) I, \\ L^{(n)} \equiv \begin{pmatrix} J_1^{(n)} & & & 0 \\ & 1 & & J_2^{(n)} \\ & & \ddots & \ddots \\ & & & 1 & J_m^{(n)} \end{pmatrix}, \quad R^{(n)} \equiv \begin{pmatrix} 1 & V_1^{(n)} & & \\ & 1 & \ddots & \\ & & \ddots & V_{m-1}^{(n)} \\ 0 & & & 1 \end{pmatrix} \quad (6)$$

からスタートする。ここに

$$J_k^{(n)} := \frac{1}{\delta^{(n)}} \left(1 + \delta^{(n)} u_{2k-2}^{(n)} \right) \left(1 + \delta^{(n)} u_{2k-1}^{(n)} \right), \\ V_k^{(n)} := \delta^{(n)} u_{2k-1}^{(n)} u_{2k}^{(n)} \quad (7)$$

である。新しい変数 $\bar{w}_k^{(n)}$ を

$$\bar{w}_k^{(n)} := u_k^{(n)} (1 + \delta^{(n)} u_{k-1}^{(n)}) \quad (8)$$

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、dLV 方程式 (6) は $Y^{(n)}$ の相似変形を記述し、表示 $Y^{(n+1)} = R^{(n)}Y^{(n)}(R^{(n)})^{-1}$ 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, $Y^{(n)}$ を対称化して

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

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

定理 1. $\tilde{R}^{(n)} := (D^{(n+1)})^{-1} R^{(n)} D^{(n)}$ とおく. dLV 方程式 (6) は正定値な 3 重対角対称行列 $A^{(n)}$ の相似変形

$$A^{(n+1)} = \tilde{R}^{(n)} A^{(n)} (\tilde{R}^{(n)})^{-1} \quad (9)$$

として表され, $A^{(n)}$ の固有値 dLV 方程式の時間発展 $n \Rightarrow n+1$ のもとで不変である. \square

定理 1 より, $A^{(n)}$ の固有値は差分間隔 $\delta^{(n)}$ 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. $A^{(n)}$ は正定値対称だから Cholesky 分解可能で

$$A^{(n)} = (B^{(n)})^T B^{(n)}, \quad B^{(n)} := \begin{pmatrix} \sqrt{\bar{w}_1^{(n)}} & \sqrt{\bar{w}_2^{(n)}} & & \\ & \sqrt{\bar{w}_3^{(n)}} & \cdots & \\ & & \cdots & \sqrt{\bar{w}_{2m-2}^{(n)}} \\ 0 & & & \sqrt{\bar{w}_{2m-1}^{(n)}} \end{pmatrix} \quad (10)$$

と書ける. $B^{(n)}$ の特異値は $A^{(n)}$ の固有値の正の平方根だから, $B^{(n)}$ の特異値もまた dLV 方程式の時間発展のもとで不変である. このことから, dLV 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,

$$u_0^{(0)} = 0, \quad u_{2m}^{(0)} = 0, \quad u_k^{(0)} = \frac{b_k^2}{1 + \delta^{(0)} u_{k-1}^{(0)}} \quad (11)$$

となるように選ぶ必要がある. $b_k > 0$ であれば $A^{(0)}$ は重複固有値をもつことはなく, $B^{(0)}$ の特異値も互いに相異なる.

次に, dLV 方程式の解の B の特異値への収束を論じる. $\text{trace}(A^{(n)}) = \text{trace}(A^{(0)})$ より, 変数 $\bar{w}_k^{(n)}$ の和は一定で

$$\sum_{k=1}^{2m-1} \bar{w}_k^{(n)} = \sum_{k=1}^m \sigma_k^2 \quad (12)$$

が成り立つ. $u_k^{(0)} > 0$ より $u_k^{(n)} > 0$ 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, (12) より $\bar{w}_k^{(n)}$ の有界性が成り立つ. ゆえに変数 $u_k^{(n)}$ について, ある正数 M_1 が存在して $0 < u_k^{(n)} < M_1, (k=1, 2, \dots, 2m-1)$ となる. 数列 $u_k^{(n)}, (n=0, 1, \dots)$ の単調性と有界性に基づいて以下の定理が証明される [11].

定理 2. dLV 方程式 (6) の初期値を (11) に従って与えらるとする. 上 2 重対角行列 B の第 k 特異値を σ_k とすると, $k = 1, \dots, m$ について

$$\lim_{n \rightarrow \infty} u_{2k-1}^{(n)} = \sigma_k^2, \quad \lim_{n \rightarrow \infty} u_{2k}^{(n)} = 0 \quad (13)$$

が成り立つ. すなわち, dLV 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. \square

以上により, 適切な初期値のもとで, dLV 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. この特異値計算法を dLV 法と名付ける. dLV 法の演算は正の数に加えと乗除算のみで減算や平方根はなく, 変数の正値性が保証されており, 桁落ちのない相対誤差の意味で高精度な特異値計算が可能である [12]. 局所的には指数の数値安定性をもつことも示されている [14].

dLV 法の収束次数は 1 次収束で, その速さは

$$R_1 := \max_k \frac{\sigma_{k+1}^2 + 1/\delta^{(n+1)}}{\sigma_k^2 + 1/\delta^{(n)}} \quad (14)$$

に依存する [10]. 近接特異値 $\sigma_k \approx \sigma_{k+1}$ があると収束はとりわけ遅い. そこで, dLV 法における変数の正値性 $0 < \bar{w}_k^{(n)}$ を壊さない範囲で原点シフトを導入することで, mdLVs 法が開発された [13]. 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.

シフト導入のアイデアは以下の通りである. dLV 法を (8) の変数 $\bar{w}_k^{(n)}$ が変数 $w_k^{(n)} := u_k^{(n-1)}(1 + \delta^{(n-1)}u_{k+1}^{(n-1)})$ によって $\bar{w}_k^{(n)} = w_k^{(n)}$ として定まるとみる. 原点シフトをパラメータ $\theta^{(n)2}$ をもつ全単射

$$\phi_{1;\theta^{(n)}}^{(n)} : (w_{2k-2}^{(n)} + w_{2k-1}^{(n)} - \theta^{(n)2}, w_{2k-1}^{(n)} w_{2k}^{(n)}) \mapsto (\bar{w}_{2k-2}^{(n)} + \bar{w}_{2k-1}^{(n)}, \bar{w}_{2k-1}^{(n)} \bar{w}_{2k}^{(n)}) \quad (15)$$

によって導入する. $\theta^{(n)2} = 0$ のときは恒等写像である. この結果, 変数 $w_k^{(n+1)}$ のなす上 2 重対角行列の特異値 $\sigma_k^{(n+1)}$ は $w_k^{(n)}$ のなす上 2 重対角行列の特異値を用いて $\sigma_k^{(n+1)} = \sqrt{\sigma_k^{(n)} - \theta^{(n)2}}$ と表され, 特異値の近接度が弱まる. 反復毎にシフトを適切に取り直すことで, 収束次数の向上が期待できる.

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は以下が成り立つ [13].

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

$$\theta^{(n)2} < \sigma_m^2(B^{(n)}) \quad (16)$$

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. \square

Table 1: 特異値計算時間 (sec.)

| | 10,000 × 10,000 | |
|--------------|-----------------|-------|
| | DBDSQR | mdLVs |
| B_{random} | 19.48 | 2.801 |

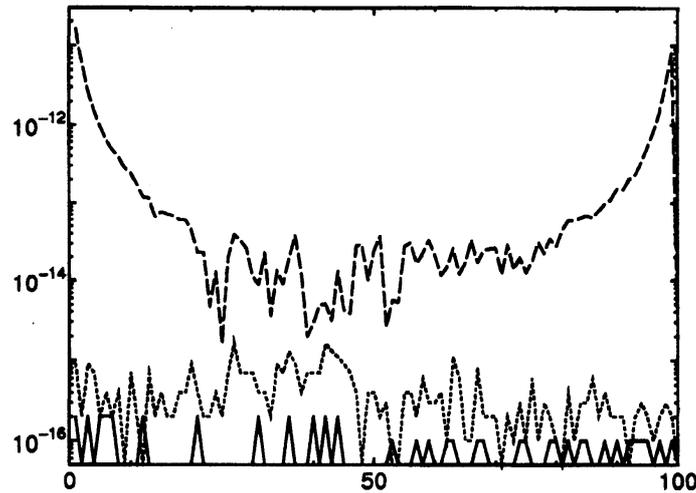


Fig. 1. DBDSQR, dLV, mdLVs 法によって

計算された特異値の相対誤差の分布 ($m = 100$)

定理からは $\theta^{(n)^2} < \sigma_m^2(B^{(n)})$ であれば $w_{2k-1}^{(n)} > 0$ となるが、有限桁計算においては、丸め誤差によって $w_{2k-1}^{(n)} \approx 0$ となることがありえる。そこで、最小特異値のある下界 $\vartheta^{(n)}$ に対して、小さな正定数 ε を用いて、

$$\theta^{(n)^2} = \max\{0, (\vartheta^{(n)})^2 - \varepsilon\} \quad (17)$$

にとって適切なシフト量を与えられることになる。丸め誤差が生じて、零割はおろか精度を悪化させる 0 に近い数での除算はなく、数値安定に高精度な特異値計算が実行可能になる。紙面の都合で mdLVs 法の収束性については省略する。詳細は [13, 17] を参照されたい。

mdLVs 法の数値計算例を与える。

Table 1 での比較対象は、Demmel-Kahan 法を実装した LAPACK の DBDSQR コードの特異値計算部である。収束判定は同一条件を採用する。mdLVs 法においてパラメータは $\delta^{(n)} = 1$ に固定する。数値はランダムに生成した 10,000 次上 2 重対角行列 100 個の平均計算時間 (秒) である。ここでは、シフト量は一般化 Newton 下界を用いて計算している。

Fig. 1 では、100 × 100 次の特異値分布がクラスターをもつ行列について、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 dLV 法の相対誤差 (鎖線) は大きく減少している。DBDSQR による相対誤差 (点線) と比較しても mdLVs 法 (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 $\varepsilon = 2.22 \times 10^{-16}$) 以下、多くの特異値については誤差なしで計算されている。

4 dLV 型ツイスト分解法

与えられた m 次上 2 重対角行列 B に対して高速高精度な特異値計算法である mdLVs 法によって B の全ての特異値 σ_j が求められているとする。本節では σ_j に対する右特異ベクトル v_j , 左特異ベクトル u_j を計算するための dLV 型ツイスト分解法を解説する [17]。 v_j は連立 1 次方程式

$$(B^T B - \sigma_j^2 I)v_j = 0 \quad (18)$$

の解ベクトル $v_j = (v_j(1), v_j(2), \dots, v_j(m))^T$ である。左特異ベクトルは $U = BV\Sigma^{-1}$ によって特異値と右特異ベクトルから定まる。 $(BB^T - \sigma_j^2 I)u_j = 0$ を解いてもよい。

mdLVs 法とはいえ、計算された特異値は一般にいくらかの誤差を含む。逆に、 v_j が真の特異ベクトルならば、近似値 $\hat{\sigma}_j$ に対しては、 $(B^T B - \hat{\sigma}_j^2 I)v_j \neq 0$ となる。そこで、(18) の右辺に適切な残差項 $c_j \neq 0$ を加えた連立 1 次方程式

$$(B^T B - \hat{\sigma}_j^2 I)v_j = c_j \quad (19)$$

を解くことでより真値に近い特異ベクトルを得ることができよう。 $\hat{\sigma}_j$ が高精度である程、係数行列 $B^T B - \hat{\sigma}_j^2 I$ は悪条件となるため、連立 1 次方程式 (19) を解くのに反復解法を利用するのは有効ではない。以下では、 $B^T B - \hat{\sigma}_j^2 I$ の分解に基づく直接法を与える。議論の過程で残差項 c_j の表現は明らかとなる。

正定値実対称行列は下 3 角行列 (または、上 3 角行列) とその転置行列の積に分解可能である。これは Cholesky 分解と呼ばれる。3 角行列の逆行列計算は容易であるから正定値実対称行列を係数行列とする連立方程式は Cholesky 分解を経由して解くことができ、Gauss の消去法より計算量の点で有利である [26]。係数行列 $B^T B - \hat{\sigma}_j^2 I$ は以下の 2 種類の Cholesky 分解が可能である。具体的には、上 2 重対角行列 ((k, k) 成分と $(k, k+1)$ 成分以外がすべて零の上 3 角行列) または下 2 重対角行列 ((k, k) 成分と $(k+1, k)$ 成分以外がすべて零の下 3 角行列) とその転置行列との積の形で

$$B^T B - \hat{\sigma}_j^2 I = \begin{cases} (B^+)^T B^+, \\ (B^-)^T B^-, \end{cases} \quad (20)$$

$$B^+ := \begin{pmatrix} b_1^+ & b_2^+ & & & \\ & b_3^+ & \cdots & & \\ & & \cdots & b_{2m-2}^+ & \\ 0 & & & & b_{2m-1}^+ \end{pmatrix}, \quad B^- := \begin{pmatrix} b_1^- & & & & 0 \\ b_2^- & b_3^- & & & \\ & \cdots & \cdots & & \\ & & & b_{2m-2}^- & b_{2m-1}^- \end{pmatrix} \quad (21)$$

と表現できる。分解 (20) を 2 重 Cholesky 分解と呼ぶことにする。 $\hat{\sigma}_j^2$ が $B^T B$ の最小固有値より大きい場合は $B^T B - \hat{\sigma}_j^2 I$ は正定値ではない。この場合も B^\pm を複素行列とすることで、複素型の 2 重 Cholesky 分解 (20) を導入することができる。

問題は、 $B^T B - \hat{\sigma}_j^2 I$ が悪条件の場合に、Cholesky 分解は数値不安定になりやすく高精度に 3 角行列を計算することが難しいことである。この問題に対する基本アイデアは以下の通りである。近似特異値 (の平方) を

$$\hat{\sigma}_k^2 = \frac{1}{\delta(0)} - \frac{1}{\delta(1)} \quad (22)$$

のように表す。この結果、目標とする Cholesky 分解

$$B^T B - \left(\frac{1}{\delta(0)} - \frac{1}{\delta(1)} \right) I = (B^+)^T B^+ \quad (23)$$

は次の3つのステップに分割することができる。

$$B^T B - \frac{1}{\delta^{(0)}} I = (\mathcal{W}^{(0)})^T \mathcal{W}^{(0)}, \quad (24)$$

$$(\mathcal{W}^{(0)})^T \mathcal{W}^{(0)} = (\mathcal{W}^{(1)})^T \mathcal{W}^{(1)}, \quad (25)$$

$$(\mathcal{W}^{(1)})^T \mathcal{W}^{(1)} + \frac{1}{\delta^{(1)}} I = (B^+)^T B^+, \quad (26)$$

$$\mathcal{W}^{(\ell)} := \begin{pmatrix} \mathcal{W}_1^{(\ell)} & \mathcal{W}_2^{(\ell)} & & & \\ & \mathcal{W}_3^{(\ell)} & \cdots & & \\ & & \cdots & \mathcal{W}_{2m-2}^{(\ell)} & \\ 0 & & & & \mathcal{W}_{2m-1}^{(\ell)} \end{pmatrix}, \quad (27)$$

$$\mathcal{W}_k^{(\ell)} := \sqrt{u_k^{(\ell)} (1 + \delta^{(\ell)} u_{k-1}^{(\ell)})}, \quad \ell = 0, 1. \quad (28)$$

パラメータ $\delta^{(0)}$ の値によっては $B^T B - \frac{1}{\delta^{(0)}} I$ は正定値ではないので, Cholesky 分解 (24) は複素型, すなわち, $\mathcal{W}_k^{(\ell)}$ が純虚数となる場合も含んでいる. 第1式 (24) を書き下すと

$$\begin{aligned} b_{2k-1}^2 &= \frac{1}{\delta^{(0)}} \left(1 + \delta^{(0)} u_{2k-2}^{(0)} \right) \left(1 + \delta^{(0)} u_{2k-1}^{(0)} \right), \\ b_{2k}^2 &= \delta^{(0)} u_{2k-1}^{(0)} u_{2k}^{(0)} \end{aligned} \quad (29)$$

となる. パラメータ $\delta^{(0)}$ を B の特異値の近似値の平方, すなわち, $B^T B$ の固有値の逆数を避けて選ぶ必要がある. そのためには, $B^T B$ の最小固有値の下からの見積りを σ_m^2 とするとき, $\delta^{(0)} > 1/\sigma_m^2$ なる適当な正の値にとり, $u_0^{(0)} \equiv 0$ とおいて, 与えられた b_k から順に $u_k^{(0)}$ を計算する. $B^T B - 1/\delta^{(0)} I$ は一般に悪条件ではないから, このようにして高精度な (複素) Cholesky 分解 (24) を実現できる.

第2式 (25) は

$$u_k^{(0)} (1 + \delta^{(0)} u_{k-1}^{(0)}) = u_k^{(1)} (1 + \delta^{(1)} u_{k-1}^{(1)}) \quad (30)$$

と表される. ここにパラメータ $\delta^{(1)}$ は (22) によって $\delta^{(0)}$ から一意に $\delta^{(1)} = \delta^{(0)} / (1 - \delta^{(0)} \sigma_k^2) < 0$ と定まる. 漸化式 (30) は dLV 方程式 (3) に類似するが添字が異なる. 形式的に $\delta^{(0)} = \delta^{(1)}$ とおけば, $u_k^{(0)} = u_k^{(1)}$ となることから定常離散 Lotka-Volterra (stdLV) 変換と名付けられている [15]. $\delta^{(0)} \neq \delta^{(1)}$ であるから, 一般に $u_k^{(0)} \neq u_k^{(1)}$ である.

第3式 (26) は stLV 変換の変数から 2 重対角行列を再構成するプロセスである. 具体的に書き下すと

$$\begin{aligned} \frac{1}{\delta^{(1)}} \left(1 + \delta^{(1)} u_{2k-2}^{(1)} \right) \left(1 + \delta^{(1)} u_{2k-1}^{(1)} \right) &= b_{2k-1}^{+2}, \\ \delta^{(1)} u_{2k-1}^{(1)} u_{2k}^{(1)} &= b_{2k}^{+2}, \end{aligned} \quad (31)$$

となる

$B^T B - 1/\delta^{(0)} I = B^T B - \sigma_j^2 I - 1/\delta^{(1)} I$ とみれば, (24) は非正則に近い正則行列 $B^T B - \sigma_j^2 I$ に対する原点シフトとみなせる. 適切な $\delta^{(0)}$ の選択により, Cholesky 分解が数値不安定となるのを回避している. 一方, (25), (26) はシフトされた 3 重対角行列の Cholesky 行列 $\mathcal{W}^{(0)}$ をシフトなしの 3 重対角行列 $B^T B - \sigma_j^2 I$ の Cholesky 行列 B^+ に戻す手続きである. 紙面の都合上, もうひとつの Cholesky 分解 $B^T B - \sigma_j^2 I = (B^-)^T B^-$ を計算する rdLV 変換については省略する.

以上によって 2 重 Cholesky 分解 (20) が計算される. stdLV 変換と rdLV 変換を合わせて dLV 型変換と名付けよう. パラメータ $\delta^{(0)}$ を最初にどのように選べば, 再設定が不要になるかは未解明

Table 2: DBDSQR と I-SVD の誤差 ($\times 10^{-9}$)

| $\ B - \hat{U}\hat{\Sigma}\hat{V}^T\ _{sum}$ | | $\ \hat{V} - V\ _{sum}$ | | $\ \hat{V}^T\hat{V} - I\ _{sum}$ | |
|--|-------|-------------------------|-------|----------------------------------|-------|
| DBDSQR | I-SVD | DBDSQR | I-SVD | DBDSQR | I-SVD |
| 0.0690 | 3.98 | 102 | 4.14 | 0.0831 | 0.324 |

Table 3: DBDSQR と I-SVD の計算時間 (sec.)

| $m = 1000$ | | $m = 2000$ | | $m = 6000$ | |
|------------|-------|------------|-------|------------|-------|
| DBDSQR | I-SVD | DBDSQR | I-SVD | DBDSQR | I-SVD |
| 44.92 | 1.13 | 432.12 | 4.91 | 42573.60 | 43.73 |

端の成分に丸め誤差が蓄積しやすいが、途中でねじることで、誤差の蓄積を軽減する効果があると考えられる。2重 Cholesky 分解 (20) が求まると、あとは B の次数 m 回の除算を行うだけでツイスト分解 (35) が完了する。

$D_\rho e_\rho = \gamma_{j,\rho} e_\rho$, $N_\rho e_\rho = e_\rho$, $D_\rho N_\rho e_\rho = N_\rho D_\rho e_\rho$ なので、もしベクトル v_j が

$$N_\rho^T v_j = e_\rho \quad (36)$$

を満たせば、 v_j は $(B^T B - \hat{\sigma}_j^2 I)v_j = \gamma_{j,\rho} e_\rho$ を満たし、 $\hat{\sigma}_j^2$ に対応する $B^T B$ の固有ベクトルである。 D_ρ の対角成分が $D^+(k) \neq 0$ かつ $D^-(k) \neq 0$ ならば、反復解法なしで連立1次方程式 (36) の解 v_j は

$$v_j(k) = \begin{cases} 1, & (k = \rho), \\ -N(k)v_j(k+1), & (k = \rho-1, \rho-2, \dots, 1), \\ -N(k-1)v_j(k-1), & (k = \rho+1, \rho+2, \dots, m) \end{cases}$$

のようにわずかな演算で求められる。一本の固有ベクトルについてツイスト行列の成分を全て求めるのに除算 m 回が必要である。その他の演算と合わせて、 k 本の固有ベクトルについて $O(km)$ 、全固有ベクトルについて $O(m^2)$ の計算量で計算は完了する。

最後に [23] から数値実験の結果を紹介する。Table 2 では $m = 1000$ のランダム行列 100 個の平均によって特異値分解の精度と右特異ベクトルの直交性をみている。Demmel-Kahan の QR 法を実装した LAPACK の DBDSQR コードと最後に逆反復による特異ベクトルの再直交化を 1 反復 ($O(m)$ の計算量) だけ行う I-SVD 法の試作コードとの比較では、DBDSQR が 1~2 桁程度良いものの、特異ベクトルの真値との差 $\|\hat{V} - V\|_{sum}$ の意味では、I-SVD 法の方が高精度である。ここに、 $\|*\|_{sum}$ は行列の要素の絶対値の総和を表す。Table 3 では上 2 重対角行列の特異値分解の計算時間の比較である。 $O(N^3)$ 解法の Demmel-Kahan の QR 法と $O(N^2)$ 解法 I-SVD 法のスケラビリティの違いは明白である。なお、前処理や逆変換も込めた密行列の特異値分解の全計算時間の比較は [23] を参照されたい。

参考文献

- [1] M. T. Chu, A differential equation approach to the singular value decomposition of bidiagonal matrices, *Lin. Alg. Appl.* **80**(1986), 71–79.
- [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] I. S. Dhillon and B. N. Parlett, Orthogonal eigenvectors and relative gaps, *SIAM J. Matrix Anal. Appl.*, **25**(2004), 858–899.
- [5] I. S. Dhillon and B. N. Parlett, Multiple representations to compute orthogonal eigenvectors of symmetric tridiagonal matrices, *Lin. Alg. Appl.*, **387**(2004), 1–28.
- [6] I. S. Dhillon, B. N. Parlett and C. Vömel, Glued matrices and the MRRR algorithm, *SIAM J. Sci. Comput.*, **27**(2005), 496–510.
- [7] K. V. Fernando, B. N. Parlett, Accurate singular values and differential qd algorithms, *Numer. Math.*, **67** (1994), 191–229.
- [8] G. H. Golub and C. F. Van Loan, *Matrix Computation*, Third Edition, The Johns Hopkins Univ. Press, Baltimore, 1996.
- [9] R. Hirota, Conserved quantities of “random-time Toda equatiopn”, *J. Phys. Soc. Japan* **66**(1997), 283–284.
- [10] M. Iwasaki and Y. Nakamura, On a convergence of solution of the discrete Lotka-Volterra system, *Inverse Problems*, **18**(2002), 1569–1578.
- [11] 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.
- [12] 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.
- [13] M. Iwasaki and Y. Nakamura, Accurate computation of singular values in terms of shifted integrable schemes, *Japan J. Indust. Appl. Math.* **23**(2006), 239–259.
- [14] M. Iwasaki and Y. Nakamura, Center manifold approach to discrete integrable systems related to eigenvalues and singular values, *Hokkaido Math. J.*, **36**(2007), 759–775.
- [15] 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.
- [16] LAPACK, <http://www.netlib.org/lapack/>
- [17] 中村佳正, *可積分系の機能数理*, 共立出版, 2006.
- [18] B. N. Parlett and I. S. Dhillon, Fernando’s solution to Wilkinson’s problem: An application of double factorization, *Lin. Alg. Appl.*, **267**(1997), 247–279.
- [19] B. N. Parlett and O. A. Marques, An implementation of the dqds algorithm (positive case), *Lin. Alg. Appl.*, **309**(2000), 217–259.
- [20] V. Spiridonov and A. Zhedanov, Discrete-time Volterra chain and classical orthogonal polynomial, *J Phys. A: Math.*, **30**(1997), 8727–8737.
- [21] 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.
- [22] 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, pp. 410–416.
- [23] 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,, *情報処理学会論文誌*, **47**, No.SIG 7(ACS 14), (2006), 91–104.
- [24] S. Tsujimoto, Y. Nakamura and M. Iwasaki, The discrete Lotka-Volterra system computes singular values, *Inverse Problems*, **17** (2001), 53–58.
- [25] J. H. Wilkinson, *The Algebraic Eigenvalue Problem*, Clarendon Press, Oxford, 1965.
- [26] 山本哲朗, *数値解析入門*, サイエンス社, 1976.
- [27] 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