Recent Developments of the mdLVs Algorithm for Singular Values and the I-SVD Algorithm for Singular Value Decomposition

特異値計算の mdLVs アルゴリズムと
特異値分解の I-SVD 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
京都大学情報学研究科 木村 欣司 (Kinji Kimura)
Graduate School of Informatics, Kyoto University
奈良女子大学人間文化研究科 高田 雅美 (Masami Takata)
Graduate School of Humanitics & Sciences, Nara Women's University

Abstract. Recently a new algorithm for computing singular values named the mdLVs (modified discrete Lotka-Volterra with shift) is designed. The first part of this report is a brief survey of the recent developments on the positivity and shifts of the mdLVs algorithm. The second part is an exposition of the I-SVD (Integrable Singular Value Decomposition) algorithm which is a combination of the mdLVs algorithm and the dLV-type transformation for computing singular vectors of bidiagonal matrices. Because of the separation of computation of singular values from that of singular vectors the I-SVD algorithm runs in $O(m^2)$ flops and is rather faster than DBDSQR code of LAPACK.

1. Introduction

An algorithm named the dLV (discrete Lotka-Volterra) algorithm for computing bidiagonal matrix singular values has been discussed in the series of papers [11, 12, 13, 14, 15], where such bidiagonal matrices can be derived from arbitrary nonsingular matrices through the Householder preconditioning process [9]. A general background is furnished in [21, 22]. See also a recent review paper [3]. The recurrence relation of dLV itself is a discrete-time integrable dynamical system. Convergence of the dLV algorithm to singular values is shown in [11]. See §4 of this report. The basic fact is that dLV is a deformation equation of orthogonal polynomials (OPs). The parameter of dLV should be positive. Therefore the recurrence relation of dLV is

---

[1]ynaka@i.kyoto-u.ac.jp

subtraction free. A positivity of Hankel determinants is ensured whose elements are moments associated with OPs and then all the variables of dLV are also positive [14]. This fact will be reviewed in §3. In a recent work [15] an exponential stability of dLV in a local sense is proved by using the existence of a center manifold around the fixed points. This implies that dLV is robust and highly credible algorithm. The convergence rate of the dLV algorithm is linear [11]. Therefore some shifted versions of dLV have been formulated in [12, 13].

The mdLVs (modified dLV with shift) algorithm presented in [13] is shown to satisfy the same positivity of variables and has a higher relative accuracy. Speed of the mdLVs algorithm depends on the choice of shift of origin. The cost for computing shifts wastes more than 30% of total execution time of the mdLVs code [29], where the shift is determined as a lower bound of the minimal singular value of given upper bidiagonal matrices. Therefore a lighter shift based on more accurate bound must be important to accelerate the mdLVs algorithm. The Johnson bound [17], a Gersgorin-type lower bound of symmetric tridiagonal, has been adopted in the mdLVs code [29]. Recently a new lower bound is found which is called the p-th generalized Newton bound. In the first part of this report (§2~§5) we discuss the recent developments on the positivity and shifts of the mdLVs algorithm.

The second part (§6~§7) is an exposition of the I-SVD (Integrable Singular Value Decomposition) algorithm[16] which is a combination of the mdLVs algorithm for singular values and the dLV-type transformation for singular vectors of $m \times m$ bidiagonal matrices. Namely, computations of singular values and vectors are completely separated in I-SVD. Here the dLV-type transformation performs an accurate double Cholesky decomposition of a shifted symmetric tridiagonal matrix which gives rise to a twisted factorization of the same matrix. Each singular vector is computed from the twisted matrix within $O(m)$ flops. Then the I-SVD algorithm solves the bidiagonal SVD problem within $O(m^2)$ flops. It is shown in [30] for some class of test matrices that the I-SVD code is rather faster than the standard SVD code of LAPACK.

2. Orthogonal Polynomials: Preliminary

Let us begin with Favard's theorem [2]. Let $\{s_k\}$, $(k = 1, 2, \dots)$ be a sequences of real numbers. When the bilinear form $\sum_{k=0}^{m} \sum_{\ell=0}^{m} s_{k+\ell} \alpha_k \alpha_\ell$ is positive for any m, then $\{s_k\}$ is called positive. It is known that $\{s_k\}$ is positive if and only if the Hankel determinants

$$
D_{n+1} := \begin{vmatrix} s_0 & s_1 & \cdots & s_n \\ s_1 & s_2 & \cdots & s_{n+1} \\ \vdots & \vdots & & \vdots \\ s_n & s_{n+1} & \cdots & s_{2n} \end{vmatrix}, \quad (n = 0, 1, 2, \dots)
$$

are positive for any $n = 0, 1, \dots$.

Theorem (Favard) Let $\{a_k\}$, $\{b_k\}$, $(k = 1, 1, \dots)$ be sequences of real numbers. Let $\{p_k(\lambda)\}$

be polynomials of λ defined by the three terms recurrence relation

$$p_{k+1}(\lambda) = (\lambda - b_{k+1})p_k(\lambda) - a_k{}^2 p_{k-1}(\lambda),$$
$$p_0(\lambda) = 1, \quad p_1(\lambda) = \lambda - b_1.$$

Then there exists a unique linear functional J such that $s_0 = J[1], J[p_k(\lambda)p_\ell(\lambda)] = 0, (k \neq \ell, k, \ell = 0, 1, \dots)$ for any positive constant s_0. Moreover, $a_k{}^2 > 0$ if and only if the moments $s_k := J[\lambda^k], (k = 0, 1, \dots)$ are positive.

The polynomial $p_k(\lambda)$, $(k = 1, 2, \dots)$ takes the determinant form [27]

$$p_k(\lambda) = \frac{1}{D_k} \begin{vmatrix} s_0 & s_1 & \cdots & s_k \\ s_1 & s_2 & \cdots & s_{k+1} \\ \vdots & \vdots & & \vdots \\ s_{k-1} & s_k & \cdots & s_{2k-1} \\ 1 & \lambda & \cdots & \lambda^k \end{vmatrix}.$$

Then the coefficients $a_k{}^2$ of the recurrence relation are

$$a_k{}^2 = \frac{D_{k-1}D_{k+1}}{D_k{}^2}.$$

It is shown from $a_1{}^2 \cdots a_k{}^2 = D_{k+1}/D_k$ that $D_k > 0$ for any k and the corresponding moments are positive.

Favard's theorem says that the polynomials $\{p_k(\lambda)\}$ defined by the three terms recurrence relation with positive coefficients $a_k{}^2$ are orthogonal with respect to the linear functional J. Namely, $J[p_k(\lambda)p_\ell(\lambda)] = s_0 a_1{}^2 \cdots a_k{}^2 \delta_{k,\ell}$. In this case the corresponding set of moments $\{s_k\}$ is positive and vice verse.

OPs have some special features. One of them is the position of zeros. It is known that zeros of OPs are mutually distinct real numbers and has an interlacing property[1]. For example, let $\lambda_j^{(n-1)}, (j = 1, 2, \dots, n-1)$ and $\lambda_j^{(n)}, (j = 1, 2, \dots, n)$ be zeros of $p_{n-1}(\lambda)$ and $p_n(\lambda)$, respectively. Then

$$\lambda_1^{(n)} < \lambda_1^{(n-1)} < \lambda_2^{(n)} < \lambda_2^{(n-1)} < \cdots < \lambda_{n-1}^{(n-1)} < \lambda_n^{(n)}.$$

This leads to the following statement. The rational function $p_{n-1}(\lambda)/p_n(\lambda)$ of degree n admits a partial fraction expansion

$$\frac{p_{n-1}(\lambda)}{p_n(\lambda)} = \sum_{j=1}^{n} \frac{\rho_j^{(n)}}{\lambda - \lambda_j^{(n)}}, \quad \rho_j^{(n)} = \frac{p_{n-1}(\lambda_j^{(n)})}{p_n'(\lambda_j^{(n)})}.$$

From the interlacing property it follows that the residues $\rho_j^{(n)}$, the Christoffel coefficients, satisfy the positivity condition $\rho_j^{(n)} > 0$.

For the Hermite, Legendre and Chebyshev polynomials every moments with odd order are zero, namely, $s_{2k-1} = 0$. In the linear functionals of those cases the measure $d\mu(\lambda)$ and the

contour C_λ are invariant under the exchange $\lambda \to -\lambda$. The linear functional J satisfying $s_{2k-1} = J[\lambda^{2k-1}] = 0$, $(k = 1, 2, \ldots)$ is called symmetric and the corresponding OP is called symmetric. When $d\mu(\lambda) = w(\lambda)d\lambda$, the weight function $w(\lambda)$ is an even function over the interval $(-\xi, \xi)$. The coefficients b_k of the recurrence relation are zero for symmetric OPs. Namely, $b_k = 0$, $(k = 1, 2, \ldots)$. In the following discussions we restrict ourselves to symmetric OPs.

Let us consider the three terms recurrence relation of symmetric OPs

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

In [14] we obtain the Christoffel-Darboux formula for symmetric OPs as follows.

$$\begin{cases} a_1{}^2 \cdots a_{2m-1}{}^2 \left(\displaystyle\sum_{j=1}^{m} \frac{p_{2j-1}(\lambda)p_{2j-1}(\kappa)}{a_1{}^2 \cdots a_{2j-1}{}^2} \right) \\ = \dfrac{p_{2m-1}(\lambda)p_{2m+1}(\kappa) - p_{2m+1}(\lambda)p_{2m-1}(\kappa)}{\kappa^2 - \lambda^2} \qquad \text{for } k = 2m - 1 \\[2em] a_1{}^2 \cdots a_{2m}{}^2 \left(\displaystyle\sum_{j=1}^{m} \frac{p_{2j}(\lambda)p_{2j}(\kappa)}{a_1{}^2 \cdots a_{2j}{}^2} + p_0(\lambda)p_0(\kappa) \right) \\ = \dfrac{p_{2m}(\lambda)p_{2m+2}(\kappa) - p_{2m+2}(\lambda)p_{2m}(\kappa)}{\kappa^2 - \lambda^2} \qquad \text{for } k = 2m \end{cases}$$

In contrast to the case of usual OPs [1, 2, 27], a parity emerges. The Christoffel-Darboux formula is useful, for example, to discuss the convergence of OP series.

3 Discrete Lotka-Volterra System and Its Positivity

In this section we fist define a kernel polynomial $p_k^*(\lambda)$ of the original symmetric OP $p_k(\lambda)$. To this end we assume $p_k(\kappa) \neq 0$.

$$p_k^*(\lambda) := \begin{cases} -\dfrac{a_1{}^2 \cdots a_{2m-1}{}^2}{p_{2m-1}(\kappa)} \displaystyle\sum_{j=1}^{m} \frac{p_{2j-1}(\lambda)p_{2j-1}(\kappa)}{a_1{}^2 \cdots a_{2j-1}{}^2} & \text{for } k = 2m - 1 \\[2em] -\dfrac{a_1{}^2 \cdots a_{2m}{}^2}{p_{2m}(\kappa)} \left(\displaystyle\sum_{j=1}^{m} \frac{p_{2j}(\lambda)p_{2j}(\kappa)}{a_1{}^2 \cdots a_{2j}{}^2} + p_0(\lambda)p_0(\kappa) \right) & \text{for } k = 2m \end{cases}$$

Then the Christoffel-Darboux formula leads to

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

When $k = 2m - 1$, $p_k(\lambda)$ is an odd function. When $k = 2m$, $p_k(\lambda)$ is even. The poles $\lambda = \pm\kappa$ are apparent poles. Hence $p_k^*(\lambda)$ is a polynomial of degree k. The transformation

$$\{p_k(\lambda)\} \longrightarrow \{p_k^*(\lambda)\}$$

is just the Christoffel transformation for the symmetric OP $\{p_k(\lambda)\}$. Let us introduce a new linear functional J^* by

$$J^*[A(\lambda)] := J[(\kappa^2 - \lambda^2)A(\lambda)],$$

for any polynomial $A(\lambda)$ and a suitable constant $\kappa < 0$. The corresponding weight function is $w^*(\lambda) := (\kappa^2 - \lambda^2)w(\lambda)$. We can generalize a theorem in [2] on the positivity of linear functional.

Theorem [14] Let the linear functional J be positive definite over the interval $[-\xi, \xi,]$ with $\xi > 0$. The J^* is positive, i.e. $\{s_k^* := J^*[\lambda^k]$ is positive over $[-\xi, \xi,]$, if and only if $\kappa \leq -\xi$.

We now consider a successive use of the Christoffel transformations

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

to generate a sequence of kernel polynomials

$$\{p_k^{(0)} := p_k(\lambda)\} \to \{p_k^{(1)} := p_k^*(\lambda)\} \to \{p_k^{(2)}\} \to \cdots,$$

where $p_k^{(n)}(\kappa^{(n)}) \neq 0$ follows from $\kappa^{(n)} > -\lambda_1^{(n)}$.

As the compatibility conditions of the Christoffel transformation and the recurrence relation $p_{k+1}^{(n+1)} = \lambda p_k^{(n+1)} - (a_k^{(n+1)})^2 p_{k-1}^{(n+1)}$ we obtain

$$\begin{aligned}
(a_k^{(n+1)})^2 &= (a_k^{(n)})^2 \frac{A_k^{(n)}}{A_{k-1}^{(n)}} \\
&= (a_k^{(n)})^2 \frac{p_{k+2}^{(n)}(\kappa^{(n)})}{p_k^{(n)}(\kappa^{(n)})} \frac{p_{k-1}^{(n)}(\kappa^{(n)})}{p_{k+1}^{(n)}(\kappa^{(n)})}.
\end{aligned}$$

Let us set

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

It follows from $p_{-1}^{(n)} = 0$ that $\hat{u}_0^{(n)} = 0$. Let $\lambda_j^{(n)}$, $(j = 1, \dots, k)$ be zeros of the OP $p_k^{(n)}(\lambda)$. Note that in the partial fraction expansion

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

the residues $\rho_j^{(n)}$ of are positive. While it follows from the positivity of the linear functional J^* that $\kappa^{(n)} - \lambda_j^{(n)} < 0$. Thus $p_{k-1}^{(n)}(\kappa^{(n)})/p_k^{(n)}(\kappa^{(n)}) < 0$ and then $\hat{u}_k^{(n)} < 0$.

Inserting $\hat{u}_k^{(n)}$ into the three terms recurrence relation we derive

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

Similarly we have

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

We eliminate $(a_k^{(n+1)})^2$ to have

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

This equation was first derived by Hirota (1997, [10]) and Spiridonov-Zhedanov(1997, [26]), independently, where $\kappa^{(n)}$ are arbitrary constant. In our case $\kappa^{(n)}$ should be less than or equal to $-\xi$ to guarantee the positivity of the linear functional, say $J^{(n)}$ and then Hankel determinants $D_k^{(n)}$. Since $\hat{u}_k^{(n)}$ is expressed as a ratio of the Hankel determinants, this property is very important to design stable numerical algorithm. Define

$$\delta^{(n)} := \frac{1}{(\kappa^{(n)})^2} > 0, \quad u_k^{(n)} = \kappa^{(n)}\hat{u}_k^{(n)} > 0.$$

We can introduce a scale change $u_k^{(n)} \to 1/(\xi^2 M)u_k^{(n)}$ to relax the condition $0 < \delta^{(n)} \le 1/\xi^2$ to $0 < \delta^{(n)} \le M$ for some positive constant M. Then we obtain

$$u_k^{(n+1)}(1 + \delta^{(n+1)}u_{k-1}^{(n+1)}) = u_k^{(n)}(1 + \delta^{(n)}u_{k+1}^{(n)}),$$
$$u_k^{(n)} > 0, \quad 0 < \delta^{(n)} \le M, \quad (n = 0, 1, \ldots, \quad k = 1, 2, \ldots).$$

Let us regard $u_k^{(n)}$ as the value of $u_k = u_k(t)$ at the time $t = \sum_{j=0}^{n-1} \delta^{(j)}$. Keeping t to a constant we take a limit $\delta^{(n)} \to +0$ such that $\delta^{(n+1)}/\delta^{(n)} \to 1$. We then derive the system of differential equations

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

for the variable $u_k = u_k(t)$ from the recurrence relation. This process corresponds to the limit $\kappa^{(n)} \to -\infty$ and does not violate the positivity of linear functionals. This system is sometimes called the Lotka-Volterra (LV) system in mathematical biology. In this section it is shown that the successive Christoffel transformations of symmetric OPs induce a deformation of the coefficients $\{a_k^{(n)}\}$ of the three terms recurrence relation. The resulting deformation equation is the dLV system having a positive explicit solution.

4   Convergence of dLV Algorithm

In the series of papers [11, 12, 15] it is shown a solution of dLV converges to the same limit as of the LV for any choice of positive $\delta^{(n)}$. Namely,

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

where σ_k are singular vales of B such that

$$\sigma_1 > \sigma_2 > \cdots > \sigma_m > 0.$$

It is to be remarked that the initial value setting is different from that in the LV case. We should choose

$$u_{2k-1}^{(0)} = \frac{b_{2k-1}{}^2}{1 + \delta^{(0)}u_{2k-2}^{(0)}}, \quad (k = 1, 2, \ldots, m),$$

$$u_{2k}^{(0)} = \frac{b_{2k}{}^2}{1 + \delta^{(0)}u_{2k-1}^{(0)}}, \quad (k = 1, 2, \ldots, m - 1),$$

as well as $u_0^{(0)} = 0$ and $u_{2m}^{(0)} = 0$. We named this procedure the dLV algorithm for computing singular values of bidiagonal matrices.

More important notion is the numerical stability. It is known [25], for example, the qd algorithm is numerically instable because of division by a small amount. One the other hand the Demmel-Kahan QR has been standard algorithm for a long time as a stable algorithm in spite of slow speed. The numerical stability of dLV is proved in [12]. The starting point is a matrix representation of dLV.

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

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

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

where I is the $m \times m$ unit matrix. Note that $1/\delta^{(n)} - 1/\delta^{(n+1)}$ gives a shift of origin for the matrix $R^{(n)}L^{(n)}$. Let us introduce new nonnegative variables $w_k^{(n)}$ defined as

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

and a tridiagonal matrix $Y^{(n)}$ such that

$$Y^{(n)} := L^{(n)}R^{(n)} - \frac{1}{\delta^{(n)}}I.$$

We derive from the matrix form of dLV

$$Y^{(n+1)} = R^{(n)}Y^{(n)}(R^{(n)})^{-1}.$$

It is not hard to see $w_k^{(n)} > 0$ providing $u_k^{(0)} > 0$ and $\delta^{(n)} > 0$ for $k = 1, 2, \cdots, 2m - 1$. Thus $R^{(n)}$ is nonsingular for any n. This similarity transformation implies that the eigenvalues of $Y^{(n)}$ are invariant under the evolution $n \Rightarrow n + 1$ of the dLV system. A symmetrization of $Y^{(n)}$ is introduced in [11] by using a diagonal matrix $G^{(n)}$ as follows:

$$A^{(n)} := (G^{(n)})^{-1}Y^{(n)}G^{(n)}, \quad G^{(n)} := \text{diag}\left(\prod_{j=1}^{m-1}\sqrt{w_{2j-1}^{(n)}w_{2j}^{(n)}}, \ldots, \sqrt{w_{2m-3}^{(n)}w_{2m-2}^{(n)}}, 1\right).$$

Note that $G^{(n)}$ is nonsingular for any n and $|A^{(n)}| = \prod_{j=1}^{m} w_{2j-1}^{(n)}$. We see that the dLV system takes the form of similarity transformation

$$A^{(n+1)} = P^{(n)}A^{(n)}(P^{(n)})^{-1}, \quad P^{(n)} := (G^{(n+1)})^{-1}R^{(n)}G^{(n)}$$

of the positive definite matrix $A^{(n)}$, which implies that the eigenvalues of $A^{(n)}$, for all n, are invariant under the evolution $n \Rightarrow n + 1$ of dLV.

Since the eigenvalues of $A^{(n)}$ are identically equal to those of $Y^{(0)}$, these eigenvalues are independent from the choice of the variable step size $\delta^{(n)}$. Note also that $A^{(n)}$ can be decomposed into

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

Therefore the singular values of $B^{(n)}$ are equal to the positive square roots of the eigenvalues of $A^{(n)}$. Then the singular values of the upper bidiagonal matrix $B^{(n)}$ are invariant under the time evolution $n \Rightarrow n + 1$ of the dLV system.

Numerical stability of the dLV algorithm is proved as follows. The positivity of the parameter $\delta^{(n)}$ and the variable $u_k^{(n)}$ play a key role. In Ref. [12] the condition $0 < \delta^{(n)} \le M$ are assumed. As is shown in §3 of this paper the condition naturally follows from the positivity of the sequence of kernel OPs. By taking trace of the similarity transformation we see

$$\sum_{k=1}^{2m-1} w_k^{(n)} = \sum_{k=1}^{m} \sigma_k^2.$$

Namely $w_k^{(n)}$ are bounded as well as positive. Consequently $u_k^{(n)}$ are also positive and bounded. Let $k = 1$ in the dLV system, then we have $\lim_{n\to\infty} u_1^{(n+1)} = u_1^{(0)} \prod_{n=0}^{\infty}(1 + \delta^{(n)}u_2^{(n)})$. This implies $u_1^{(0)} \le u_1^{(1)} \le \cdots u_1^{(n)} \le \cdots$. Since $u_1^{(n)}$, $n = 0, 1, \cdots$, is monotonically increasing and bounded, $u_1^{(n)}$ converges to some positive constant c_1 as $n \to \infty$. Simultaneously, $\prod_{n=0}^{\infty}(1 + \delta^{(n)}u_2^{(n)})$ converges to some positive constant p_1.

Let us assume that $\prod_{n=1}^{\infty}(1 + \delta^{(n)}u_{2k-2}^{(n)})$ converges to some positive constant p_{k-1}. Let $v_k^{(0)} = u_k^{(0)}(1+\delta^{(0)}u_{k+1}^{(0)})$ and $v_k^{(0)} > 0$. Then, by using $0 < \delta^{(n)} < M$, we see that $(v_{2k-1}^{(0)}/p_{k-1})\prod_{n=1}^{N}(1 + \delta^{(n)}u_{2k}^{(n)})$ converges to $u_{2k-1}^{(N+1)}$ as $N \to \infty$. Hence it follows that $0 < \prod_{n=0}^{\infty}(1 + \delta^{(n)}u_{2k}^{(n)}) < M_3$ for some constant M_3. It is also obvious that $\prod_{n=1}^{N}(1 + \delta^{(n)}u_{2k}^{(n)})$, $N = 1, 2, \cdots$, is monotonically increasing. Therefore it follows that $\prod_{n=1}^{\infty}(1 + \delta^{(n)}u_{2k}^{(n)}) = p_k$. Simultaneously, we see that $\lim_{n\to\infty} u_{2k-1}^{(n)} = v_{2k-1}^{(0)}p_k/p_{k-1} > 0$, namely,

$$\lim_{n\to\infty} u_{2k-1}^{(n)} = c_k$$

where c_k is some positive constant.

Note here that $\sum_{n=0}^{\infty} \delta^{(n)}u_{2k}^{(n)}$ converges to some constant $s_k > 0$ if and only if $\prod_{n=1}^{\infty}(1 +$

$\delta^{(n)}u_{2k}^{(n)}) = p_k$ for $\delta^{(n)}u_{2k}^{(n)} > 0$, $n = 0, 1, \cdots$. Moreover $\lim_{n\to\infty} \delta^{(n)}u_{2k}^{(n)} = 0$ for any positive bounded sequence $\delta^{(n)}$, if $\sum_{n=0}^{\infty} \delta^{(n)}u_{2k}^{(n)} = s_k$. Therefore it follows that

$$\lim_{n\to\infty} u_{2k}^{(n)} = 0.$$

Note here that $\lim_{n\to\infty} A^{(n)} = \text{diag}(c_1, c_2, \cdots, c_m)$. This implies that c_k is one of the eigenvalues of $A^{(n)}$, namely, the square of a singular value of $B^{(n)}$. Singular values of $B^{(n)}$ are equal to those of $B = B^{(0)}$. It is concluded that the dLV algorithm converges to singular values of the bidiagonal matrix B with nonzero diagonals and sub-diagonals in numerically stable way. The Christoffel transformation of symmetric OPs gives rise to the positivity and boundedness of the parameter $\delta^{(n)}$ and the variable $u_k^{(n)}$ of the dLV algorithm. No subtraction appears in dLV. Then a higher accuracy follows.

5 Recent Developments on Shifts of mdLVs Algorithm

On speed the dLV algorithm has a first order convergence providing that $\delta^{(n)} > 0$. It is slow. The shift $1/\delta^{(n)} - 1/\delta^{(n+1)}$ for $R^{(n)}L^{(n)}$ brings an "internal shift" of the dLV algorithm. To accelerate the convergence an "external" shift of origin is introduced in [13] through a mapping $(B^{(n)})^{\top}B^{(n)} \to (\bar{B}^{(n)})^{\top}\bar{B}^{(n)} := (B^{(n)})^{\top}B^{(n)} - (\theta^{(n)})^2 I$. If a summation of shifts $(\theta^{(n)})^2$ is less than the square $\sigma_m{}^2$ of the minimal singular value of B, then

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

is shown. The positivity and boundedness of the variable $u_k^{(n)}$ are not violated by the shift. Such a stable shift is given by using the Johnson bound [17], for example. Then a new stable algorithm with shift for bidiagonal singular value problem results which is named the mdLVs algorithm. The mdLVs has two types of parameters. One is the internal parameter $\delta^{(n)}$. The other is the external shift parameter $\theta^{(n)}$. The mdLVs algorithm is more accurate than the Demmel-Kahan QR algorithm, the Divide & Conquer algorithm and the dqds algorithm which are practically used bidiagonal singular value computing algorithms (cf. [4]) through the present LAPACK codes [20]. The mdLVs algorithm is faster than the Demmel-Kahan QR, Divide & Conquer the as well as the bisection algorithm. On these established algorithms see the book [4] and references therein.

The Johnson bound has been adopted in the original implementation [28, 29] of the mdLVs algorithm [13]. The Johnson shift is

$$\Theta^{(m)} := \min_{k=1,\ldots,m} \left\{ \sqrt{w_{2k-1}} - \frac{1}{2} \left(\sqrt{w_{2k-2}} + \sqrt{w_{2k}} \right) \right\}$$
$$< \sigma_m,$$

where $w_0 = 0$, $w_{2k-1} + w_{2k-2}$ are the (k, k)-elements and $\sqrt{w_{2k}w_{2k-1}}$ are the $(k, k+1)$ and $(k+1, k)$-elements of the tridiagonal matrix $A = B^{\top}B$.

Recently the following lower bound is found by K. Kimura

$$
\begin{aligned}
\Theta_p^{(m)} &:= \left(\operatorname{trace}(B^\top B)^{-p}\right)^{-\frac{1}{2p}} \\
&= \frac{1}{\left(\frac{1}{\sigma_1^{2p}} + \cdots + \frac{1}{\sigma_m^{2p}}\right)^{1/2p}} \\
&< \sigma_m, \quad (p = 1, 2, \dots)
\end{aligned}
$$

which is called the p-th **generalized Newton bound**. The cost for the generalized Newton shift is shown in [18] to be only $O(m)$. Y. Yamamoto [19] proves that the generalized Newton shift performs a weakly $(p+1)$-th order convergence. The mdLVs code with the generalized Newton shift where $p = 2, 3, 4$ is faster and more accurate than the mdLVs code with the Johnson shift. Here

$$
I_p := \operatorname{trace}(B^\top B)^{-p}
$$

are conserved quantities of the dLV system. This strongly suggests that $\Theta_p^{(m)}$ are also expressed only by using positive variables. Recently this conjecture is proved affirmatively for by T. Yamashita and a subtraction free $O(m)$ formula for computing $\Theta_p^{(m)}$ is presented in [32]. The generalized Newton shift will be useful for the dqds algorithm [23, 24] for singular values.

6 Double Cholesky Decomposition and dLV-type Transformation

Let us assume that all of the singular values of an $m \times m$ upper bidiagonal matrix B are positive, simple and are already computed. Let $\hat{\sigma}_j$ be the computed singular value. In this section we introduce the dLV-type transformation for computing the right singular vector v_j and the left singular vector u_j corresponding to each $\hat{\sigma}_j$[22]. The right singular vector v_j of B is a solution vector $v_j = (v_j(1), v_j(2), \cdots, v_j(m))^\top$ of the system of linear equation

$$
(B^\top B - \hat{\sigma}_j^2 I) v_j = 0.
$$

Computed singular value $\hat{\sigma}_j$ usually contains some errors though the mdLVs algorithm has a higher relative accuracy. Conversely, if v_j is a correct singular vector corresponding to the correct singular value σ_j, then $(B^\top B - \hat{\sigma}_j^2 I) v_j \neq 0$ for an approximant $\hat{\sigma}_j$ of σ_j. Therefore let us find more accurate singular vector by solving the linear equation

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

for a suitable constant vector $c_j \neq 0$. A derivation of the residual vector c_j will be described later. As $\hat{\sigma}_j$ is close to σ_j, the coefficient matrix $B^\top B - \hat{\sigma}_j^2 I$ becomes singular. Thus we use a direct method for solving the ill-conditioned linear equation $(B^\top B - \hat{\sigma}_j^2 I) v_j = c_j$.

Any positive definite real symmetric matrix can be decomposed into the product of a lower (or upper) triangular matrix and its transposed. This is called the Cholesky decomposition. Once the coefficient matrix of linear equation is decomposed, the linear equation is solved through an

inversion of the triangular factor at a lower computational cost than the Gaussian elimination. In this case we first consider the so-called "double Cholesky decomposition" of a positive definite matrix $B^\top B - \hat{\sigma}_j^2 I$ as

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

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

where B^+ and B^- are upper and lower bidiagonal matrices, respectively. If $B^\top B - \hat{\sigma}_j^2 I$ is nonsingular but indefinite, namely, if $\hat{\sigma}_j$ is greater than the minimal singular value σ_m of B, the double Cholesky decomposition of complex type can be introduced similarly.

A problem arises. Cholesky decomposition of such an ill-conditioned matrix as $B^\top B - \hat{\sigma}_j^2 I$ is liable to be numerically unstable. It is difficult to compute accurate triangular factors B^+ and B^-. Note that the Cholesky decomposition takes the form of the shift mapping of the mdLVs algorithm $B^\top B - \hat{\sigma}_j^2 I =: \bar{B}^\top \bar{B}$. Whereas an internal shift of the matrix representation of dLV algorithm is caused by the difference $1/\delta^{(0)} - 1/\delta^{(1)}$, where $\delta^{(0)}$ and $\delta^{(1)}$ are parameters of dLV. Therefore, we divide $\hat{\sigma}_j^2$ into two

$$\hat{\sigma}_j^2 = \frac{1}{\delta_+^{(0)}} - \frac{1}{\delta_+^{(1)}}.$$

Consequently the target Cholesky decomposition can be divided into three

$$B^\top B - \frac{1}{\delta_+^{(0)}} I = (\mathcal{W}^{(0)})^\top \mathcal{W}^{(0)},$$

$$(\mathcal{W}^{(0)})^\top \mathcal{W}^{(0)} = (\mathcal{W}^{(1)})^\top \mathcal{W}^{(1)},$$

$$(\mathcal{W}^{(1)})^\top \mathcal{W}^{(1)} + \frac{1}{\delta_+^{(1)}} I = (B^+)^\top B^+,$$

where

$$\mathcal{W}^{(\ell)} := \begin{pmatrix} \mathcal{W}_1^{(\ell)} & \mathcal{W}_2^{(\ell)} & & & \\ & \mathcal{W}_3^{(\ell)} & \ddots & & \\ & & \ddots & \mathcal{W}_{2m-2}^{(\ell)} & \\ \mathbf{0} & & & \mathcal{W}_{2m-1}^{(\ell)} \end{pmatrix}, \quad \mathcal{W}_k^{(\ell)} := \sqrt{u_k^{(\ell)}(1 + \delta_+^{(\ell)} u_{k-1}^{(\ell)})}, \quad (\ell = 0, 1).$$

The first equation reads

$$b_{2k-1}^2 = \frac{1}{\delta_+^{(0)}} \left(1 + \delta_+^{(0)} u_{2k-2}^{(0)}\right) \left(1 + \delta_+^{(0)} u_{2k-1}^{(0)}\right),$$

$$b_{2k}^2 = \delta_+^{(0)} u_{2k-1}^{(0)} u_{2k}^{(0)}, \quad u_0^{(0)} \equiv 0.$$

From the assumption on the given bidiagonal matrix B we see $b_{2k-1} \neq 0$ and $b_{2k} \neq 0$. Thus $u_{2k-1}^{(0)} \neq 0$ and $u_{2k}^{(0)} \neq 0$. The parameter $\delta_+^{(0)}$ should not be equal to the inverse of any eigenvalue of $B^\top B$. Then, since eigenvalues of $B^\top B - 1/\delta_+^{(0)} I$ are simple and nonzero, $1 + \delta_+^{(0)} u_{2k-2}^{(0)} \neq 0$ and $1 + \delta_+^{(0)} u_{2k-1}^{(0)} \neq 0$. Let $\bar{\sigma}_m$ be a lower bound of the minimal singular value of B. We choose a positive number $\delta_+^{(0)}$ so that $\delta_+^{(0)} > 1/\bar{\sigma}_m^2$. Then $u_k^{(0)}$ are computed sequentially. Since $B^\top B - 1/\delta_+^{(0)} I$ is not ill-conditioned, in general, the Cholesky factor $\mathcal{W}^{(0)}$ can be computed stably. It follows from $1/\delta_+^{(0)} < \bar{\sigma}_m^2$ that $B^\top B - 1/\delta_+^{(0)} I$ is positive definite. When B has a tiny singular values, we can choose a positive number $\delta_+^{(0)}$ so that $\delta_+^{(0)} \neq 1/\hat{\sigma}_k^2$, where $\hat{\sigma}_k$ are singular values of B computed by mdLVs. In this case the elements $\mathcal{W}_k^{(\ell)}$ of the Cholesky factor $\mathcal{W}^{(0)}$ becomes pure imaginary. But there is no additional difficulty.

The second equation is written as

$$u_k^{(0)}(1 + \delta_+^{(0)} u_{k-1}^{(0)}) = u_k^{(1)}(1 + \delta_+^{(1)} u_{k-1}^{(1)}), \quad \delta_+^{(1)} := \frac{\delta_+^{(0)}}{1 - \delta_+^{(0)} \hat{\sigma}_j^2}, \quad u_0^{(1)} \equiv 0.$$

This is a transformation from $u_k^{(0)}$ to $u_k^{(1)}$ with the parameter $\delta_+^{(0)}$. Since $u_k^{(0)} \neq 0$ and $(1 + \delta_+^{(0)} u_{k-1}^{(0)}) \neq 0$, $u_k^{(1)}$ in the matrix $\mathcal{W}^{(1)}$ are computed sequentially from given $u_k^{(0)}$ and $\delta_+^{(0)}$. If we set $\delta_+^{(1)} = \delta_+^{(0)}$ temporarily, then $u_k^{(1)} = u_k^{(0)}$, an identity transformation. Since the transformation is very similar to the dLV system, the transformation is named the stationary dLV (stdLV) in [16].

The third equation describes a process to generate bidiagonal matrix B^+ from the variables $u_k^{(1)}$ with parameter $\delta_+^{(1)}$ through

$$\frac{1}{\delta^{(1)}} \left(1 + \delta_+^{(1)} u_{2k-2}^{(1)}\right)\left(1 + \delta_+^{(1)} u_{2k-1}^{(1)}\right) = b_{2k-1}^{+}{}^2,$$

$$\delta_+^{(1)} u_{2k-1}^{(1)} u_{2k}^{(1)} = b_{2k}^{+}{}^2. \tag{1}$$

The left hand side of the first equation can be regarded as a shift of origin of the ill-posed matrix $B^\top B - \hat{\sigma}_j^2 I$. This is because $B^\top B - 1/\delta_+^{(0)} I = B^\top B - \hat{\sigma}_j^2 I - 1/\delta_+^{(1)} I$. By a suitable choice of $\delta_+^{(0)}$ a possible numerical instability in the Cholesky decomposition $B^\top B - \hat{\sigma}_j^2 I = (B^+)^\top B^+$ can be avoided. On the other hand the third equation recovers the factor B^+ of the Cholesky decomposition from the transform $u_k^{(1)}$ of $u_k^{(0)}$ by the stdLV transformation. The relationship is expressed in the following stdLV diagram.
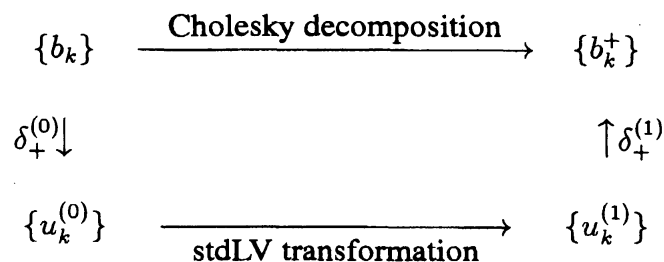

Fig. 1: stdLV diagram

The other Cholesky decomposition $B^\top B - \hat{\sigma}_j{}^2 I = (B^-)^\top B^-$ is also divided into three

$$B^\top B - \frac{1}{\delta_-^{(0)}} I = (\mathcal{W}^{(0)})^\top \mathcal{W}^{(0)},$$

$$(\mathcal{W}^{(0)})^\top \mathcal{W}^{(0)} = (\mathcal{V}^{(-1)})^\top \mathcal{V}^{(-1)},$$

$$(\mathcal{V}^{(-1)})^\top \mathcal{V}^{(-1)} + \frac{1}{\delta_-^{(-1)}} I = (\mathcal{B}^-)^\top \mathcal{B}^-,$$

where

$$\hat{\sigma}_j{}^2 = \frac{1}{\delta_-^{(0)}} - \frac{1}{\delta_-^{(-1)}},$$

$$\mathcal{V}^{(-1)} := \begin{pmatrix} \mathcal{V}_1^{(-1)} & \mathcal{V}_2^{(-1)} & & & \\ & \mathcal{V}_3^{(-1)} & \ddots & & \\ & & \ddots & \mathcal{V}_{2m-2}^{(-1)} & \\ \mathbf{0} & & & \mathcal{V}_{2m-1}^{(-1)} \end{pmatrix}, \qquad \mathcal{V}_k^{(-1)} = \operatorname{sgn}(\mathcal{W}_k^{(0)})\sqrt{u_k^{(-1)}(1 + \delta_-^{(-1)} u_{k+1}^{(-1)})}.$$

The second equation leads to

$$u_k^{(0)}(1 + \delta_-^{(0)} u_{k-1}^{(0)}) = u_k^{(-1)}(1 + \delta_-^{(-1)} u_{k+1}^{(-1)}), \qquad \delta_-^{(-1)} := \frac{\delta_-^{(0)}}{1 - \delta_-^{(0)}\hat{\sigma}_j{}^2}, \qquad u_{2m}^{(-1)} \equiv 0.$$

The mapping from $u_k^{(0)}$ to $u_k^{(-1)}$ is called the reverse-time dLV (rtdLV) transformation. By a suitable choice of $\delta_-^{(0)}$ a possible numerical instability in the Cholesky decomposition $B^\top B - \hat{\sigma}_j{}^2 I = (B^-)^\top B^-$ can be avoided. The relationship between variables is expressed in the rtdLV diagram.
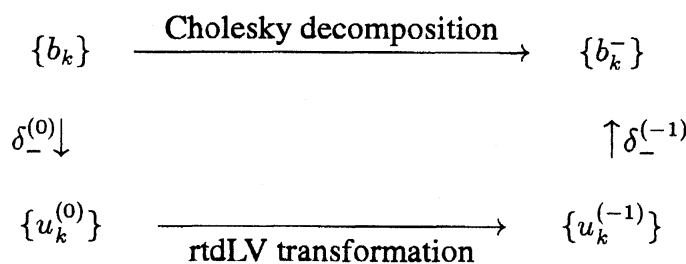

Fig. 2: rtdLV diagram

We name the pair of stdLV and rtdLV as the dLV-type transformation. This performs the double Cholesky decomposition of a wide class of positive definite symmetric tridiagonal matrices $B^\top B$ in a numerical stable way [16] by choosing suitable parameters $\delta_\pm^{(0)}$. On the other hand, the qd-type transformation for computing eigenvectors of symmetric tridiagonals, proposed by Parlett and Dhillon [5, 23] has no such parameter. Indeed there is a 3×3 nonsingular test matrix having a tiny eigenvalue. The qd-type transformation make a serious error though the dLV-type with $\delta_\pm^{(0)} = 1$ gives an accurate Cholesky decomposition. Noted that v_j is an eigenvector of the

symmetric positive definite tridiagonal matrix $B^\top B$. When a relative gap of eigenvalues is very small, both the qd-type and the dLV-type have challenges in orthogonality of resulting eigenvectors and singular vectors. For example, the qd-type fails to compute accurate eigenvectors of the glued Wilkinson matrix [7].

7 Twisted Factorization and Numerical Examples

In this section we explain a procedure for computing accurate right singular vectors v_j of B from the factors $B^\pm$ of double Cholesky decomposition. This part is essentially same as the twisted factorization method by [5, 6, 23]. The residual vector c_j in the linear system $(B^\top B - \hat{\sigma}_j{}^2 I)v_j = c_j$ is set as

$$c_j = \gamma_{j,\rho}e_\rho, \quad \begin{aligned} \gamma_{j,k} &:= b_{2k-1}^{+}{}^2 + b_{2k-1}^{-}{}^2 - (b_{2k-2}{}^2 + b_{2k-1}{}^2 - \hat{\sigma}_j{}^2) \neq 0, \\ e_\rho &:= (0, \cdots, 0, 1, 0, \cdots, 0)^\top, \end{aligned}$$

where $\gamma_{j,k}$ are the residual parameters, the ρ-th element of e_ρ is 1 and ρ is a number such that $|\gamma_{j,k}|$ takes the minimum for $k = \rho$, namely, ρ indicates the "most accurate" point. Then the so-called "twisted matrix" N_ρ is introduced as follows. Set

$$N(k) = \begin{cases} \dfrac{b_{2k}^{+}}{b_{2k-1}^{+}}, & (k = 1, 2, \ldots, \rho - 1), \\[2mm] \dfrac{b_{2k}^{-}}{b_{2k+1}^{-}}, & (k = \rho, \rho + 1, \ldots, m - 1), \end{cases}$$

$$D^+(k) = b_{2k-1}^{+}{}^2, \quad (k = 1, 2, \ldots, \rho - 1),$$

$$D^-(k) = b_{2k-1}^{-}{}^2, \quad (k = \rho + 1, \rho + 2, \ldots, m).$$

If $b_{2k}^{\pm}$ is real, then so is $b_{2k\mp1}^{\pm}$. If $b_{2k}^{\pm}$ is pure imaginary, so is $b_{2k\mp1}^{\pm}$. Therefore, $N(k)$ are always real. Define

$$N_\rho := \begin{pmatrix} 1 & & & & & & \\ N(1) & 1 & & & & & \\ & \ddots & \ddots & & & & \\ & & N(\rho-1) & 1 & N(\rho) & & \\ & & & & 1 & \ddots & \\ & & & & & \ddots & N(m-1) \\ & & & & & & 1 \end{pmatrix},$$

$$D_\rho := \mathrm{diag}(D^+(1), \cdots, D^+(\rho - 1), \gamma_{j,\rho}, D^-(\rho + 1), \cdots, D^-(m)).$$

The coefficient matrix of the linear system $(B^\top B - \hat{\sigma}_j{}^2 I)v_j = \gamma_{j,\rho}e_\rho$ takes the form

$$B^\top B - \hat{\sigma}_j{}^2 I = N_\rho D_\rho (N_\rho)^\top.$$

Table 1: Accuracy and orthogonality of DBDSQR and I-SVD $(\times 10^{-9})$

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

Table 2: Execution time of DBDSQR and I-SVD (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

This is sometimes called the twisted factorization. The substitution for determining $N(k)$ and $D^\pm(k)$ can be done from the twisted point $k = \rho$ as $k = \rho, \rho \pm 1, \rho \pm 2, \ldots$. While various errors may be accumulated if we compute matrix factorizations by the usual one-side substitution. The twisted factorization for each $\hat{\sigma}_j$ can be computed by $O(m)$ times of divisions. Since $D_\rho e_\rho = \gamma_{j,\rho} e_\rho$, $N_\rho e_\rho = e_\rho$, $D_\rho N_\rho e_\rho = N_\rho D_\rho e_\rho$, v_j satisfies the linear system $(B^\top B - \hat{\sigma}_j^2 I)v_j = \gamma_{j,\rho} e_\rho$ and is an eigenvector of $B^\top B$ providing that

$$N_\rho^\top v_j = e_\rho.$$

The solution vector $v_j = (v_j(k))$ of $N_\rho^\top v_j = e_\rho$ is given by

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

which gives rise to the right singular vector v_j. Since each singular vector can be computed within $O(m)$ flops, the computation of k-singular vectors costs $O(km)$ flops for $k = 1, 2, \ldots, m$. The left singular vectors u_j of B are given through $U = BV\Sigma^{-1}$, with $U = (u_1, \ldots, u_m)$, $V = (v_1, \ldots, v_m)$, $\Sigma = \text{diag}(\hat{\sigma}_1, \ldots, \hat{\sigma}_m)$, or by solving the system $(BB^\top - \hat{\sigma}_j^2 I)u_j = 0$ directly.

Finally we quote some numerical examples from [30] on an implementation of the I-SVD algorithm. In Table 1 the accuracy of singular value decomposition, the accuracy of right singular vectors and their orthogonality are considered. Here $\|\hat{V} - V\|_{sum}$ indicates the sum of absolute values of every elements of the matrix $\hat{V} - V$, where $\hat{v}_j$ in $\hat{V} = (\hat{v}_1, \ldots, \hat{v}_m)$ are computed

right singular vectors. The numerical numbers in Table 1 are averages of 100 1000 × 1000 bidiagonal test matrices whose singular values are randomly distributed. Here DBDSQR is the standard LAPACK code for bidiagonal SVD where the Demmel-Kahan QR algorithm is implemented. In the I-SVD code the left singular vectors are computed by $\hat{U} = B\hat{V}\hat{\Sigma}^{-1}$ and a re-orthogonalization of all of the singular vectors is performed by inverse iterations once. It takes extra $O(m^2)$ flops. The parameters $\delta_{\pm}^{(0)}$ are fixed to 1. Table 1 shows that DBDSQR is better than the I-SVD code by $1 \sim 2$ digits on the accuracy of SVD and the orthogonality of singular vectors. However on the accuracy of singular vectors I-SVD is better than DBDSQR.

Table 2 is a comparison of execution time between DBDSQR and the I-SVD code. It is obvious that the I-SVD algorithm needs only $O(m^2)$ flops and is rather faster than DBDSQR of $O(m^3)$ flops. I-SVD also has a better scalability. The I-SVD code with the Householder preconditioning to bidiagonal matrices and an inverse transformation is still sufficiently faster than DBDSQR with Householder [30].

8. Concluding Remarks

This report surveys recent developments of the mdLVs algorithm for singular values and the I-SVD algorithm for bidiagonal SVD. The mdLVs is a shifted version of the dLV algorithm. We first show how the dLV algorithm has a higher accuracy. The Christoffel transformation of symmetric OPs gives rise to the positivity and boundedness of the parameter $\delta^{(n)}$ and the variable $u_k^{(n)}$ of the dLV algorithm. No subtraction appears in dLV. Positivity is also essential in the formulation of the mdLVs algorithm. Namely, if a shift is less than the minimal singular value, then the positivity of mdLVs follows.

The Johnson bound [17] has been adopted in the mdLVs code [29]. Recently a new lower bound is found which is called the p-th generalized Newton bound. The generalized Newton shift costs only $O(m)$ flops where m is the size of given bidiagonal matrix [18]. Y. Yamamoto [19] proves that the generalized Newton shift performs a weakly $(p + 1)$-th order convergence. The mdLVs code with the generalized Newton shift where $p = 2, 3, 4$ is faster and more accurate than the mdLVs code with the Johnson shift. Though lower and upper bounds of matrix eigenvalues have been studied fully [31], exploring for new bound is still an important problem in numerical linear algebra.

The I-SVD algorithm is a combination of the mdLVs algorithm and the dLV-type transformation for singular vectors. Because of the separation of computation of singular values from that of singular vectors the I-SVD algorithm runs in $O(m^2)$ flops. Whereas the Demmel-Kahan QR algorithm requires $O(m^3)$ flops. Thus the I-SVD code is rather faster than DBDSQR code of LAPACK. The I-SVD code has a good orthogonality of singular vectors for the case of random matrices. To improve the orthogonality for clustered matrices the I-SVD algorithm should be investigated more.

参考文献

[1] N. I. Akheiezer, *The Classical Moment Problem and Some Related Questions in Analysis*, Edinburgh: Olver & Boyd, 1965.

[2] T. S. Chihara, *An Introduction to Orthogonal Polynomials*, New York: Gordon & Breach, 1978.

[3] Moody T. Chu, "Linear algebra algorithms as dynamical systems, " *Acta Numerica*, 2008, to appear.

[4] J. Demmel, *Applied Numerical Linear Algebra*, Philadelphia: SIAM, 1997.

[5] I. S. Dhillon and B. N. Parlett, Orthogonal eigenvectors and relative gaps, SIAM J. Matrix Anal. Appl., **25**(2004), 858–899.

[6] I. S. Dhillon and B. N. Parlett, Multiple representations to compute orthogonal eigenvectors of symmetric tridiagonal matrices, Lin. Alg. Appl., **387**(2004), 1–28.

[7] I. S. Dhillon, B. N. Parlett and C. Vömel, Glued matrices and the MRRR algorithm, SIAM J. Sci. Comput., **27**(2005), 496–510.

[8] K. V. Fernando, B. N. Parlett, Accurate singular values and differential qd algorithms, Numer. Math., **67** (1994), 191-229.

[9] G. H. Golub and C. F. Van Loan, *Matrix Computations 3rd edn*, Baltimore: The Johns Hopkins Univ. Press, 1996.

[10] R. Hirota, "Conserved quantities of random-time Toda equation", *J. Phys. Soc. Japan* Vol. 66, pp. 283–284, 1997.

[11] M. Iwasaki and Y. Nakamura, "On a convergence of solution of the discrete Lotka-Volterra system," *Inverse Problems*, vol. 18, pp. 1569–1578, 2002.

[12] M. Iwasaki and Y. Nakamura, "An application of the discrete Lotka-Volterra system with variable step-size to singular value computation," *Inverse Problems*, Vol. 20, pp. 553–563, 2004.

[13] M. Iwasaki and Y. Nakamura, "Accurate computation of singular values in terms of shifted integrable schemes, " *Japan J. Indust. Appl. Math.* Vol. 23, pp. 239–259, 2006.

[14] M. Iwasaki and Y. Nakamura, "Positivity and stability of the dLV algorithm for computing matrix singular values," in:*Proceedings of the 2nd International Conference on Informatics Research for Development of Knowledge Society Infrastructure ICKS 2007, Eds. M. Fukushima, K. Tajima and K. Tanaka*, IEEE Computer Society Press, pp. 103–110, 2007.

[15] M. Iwasaki and Y. Nakamura, "Center manifold approach to discrete integrable systems related to eigenvalues and singular values," *Hokkaido Math. J.*, Vol. 36, pp. 759–775, 2007.

[16] M. Iwasaki, S. Sakano and Y. Nakamura, "Accurate twisted factorization of real symmetric tridiagonal matrices and its application to singular value decomposition," Trans. Japan SIAM, **15**(2005), 461–481. (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).

[17] C. R. Johnson, "A Gersgorin-type lower bound for the smallest singular value," *Lin. Alg. Appl.*, Vol. 112, pp. 1–7, 1989.

[18] K. Kimura, T. Yamashita and Y. Nakamura, "On $O(N)$ formula for the diagonal elements of inverse powers of symmetric positive definite tridiagonal matrices," in preparation.

[19] K. Kimura, Y. Yamamoto, M. Takata and Y. Nakamura, "Generalized Newton shifts: An arbitrarily high order shifting scheme for singular value computation," in preparation.

[20] LAPACK, http://www.netlib.org/lapack/

[21] Y. Nakamura, "A new approach to numerical algorithms in terms of integrable systems," in:*Proceedings of International Conference on Informatics Research for Development of Knowledge Society Infrastructure ICKS 2004, Eds. T. Ibaraki, T. Inui and K. Tanaka,* IEEE Computer Society Press, pp. 194–205, 2004.

[22] Y. Nakamura, *Functionality of Integrable Systems,* Tokyo: Kyoritsu Shuppan, 2006, (中村佳正著「可積分系の機能数理」共立出版).

[23] 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.

[24] B. N. Parlett and O. A. Marques, An implementation of the dqds algorithm (positive case), Lin. Alg. Appl., **309**(2000), 217–259.

[25] H. Rutishauser, "Ein Quotienten-Differenzen-Algorithmus," *Z. angew. Math. Phys.,* vol. 5, pp. 233–251, 1954.

[26] V. Spiridonov and A. Zhedanov, "Discrete-time Volterra chain and classical orthogonal polynomials," *J. Phys. A,* vol. 30, pp. 8727–8737, 1997.

[27] G. Szegö, *Orthogonal Polynomials,* Providence: Amer. Math. Soc., 1939.

[28] M. Takata, M. Iwasaki, K. Kimura and Y. Nakamura, "Implementation and its evaluation of routine for computing singular values with high accuracy," IPSJ Trans. Adv. Comput. Systems, **46**, No.SIG12(2005), 299-311, (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).

[29] M. Takata, M. Iwasaki, K. Kimura and Y. Nakamura, "An evaluation of singular value computation by the discrete Lotka-Volterra system," in:*Proceedings of The 2005 International Conference on Parallel and Distributed Processing Techniques and Applications (PDPTA2005),* Vol. II, pp. 410–416, 2005.

[30] M. Takata, K. Kimura, M. Iwasaki and Y. Nakamura, "Implementation of library for high speed singular value decomposition," IPSJ Trans. Adv. Comput. Systems, **47**, No.SIG 7(ACS 14), (2006), 91-104, (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).

[31] R. Varga, *Geršgorin and His Circles,* Heidelberg: Springer-Verlag, 2004.

[32] T. Yamashita, K. Kimura and Y. Nakamura, "On subtraction free formula for the diagonal elements of inverse powers of positive definite tridiagonal matrices," in preparation.