On $O(N)$ formula for the diagonal elements of inverse powers of symmetric positive definite tridiagonal matrices

Kinji Kimura [a], Takumi Yamashita [a,b,§] and Yoshimasa Nakamura [a,b]

Technical Report 2008-010, September 15, 2008

Abstract

On $N \times N$ ($N \geq 2$) non-singular upper bidiagonal matrix $\boldsymbol{B}$ and its transpose $\boldsymbol{B}^T$, matrix products $(\boldsymbol{B}^T\boldsymbol{B})$ and $(\boldsymbol{B}\boldsymbol{B}^T)$ are symmetric positive definite tridiagonal matrices. Let M denote an arbitrary positive integer. We present a formula to compute diagonals of inverse powers of these matrix products such that $((\boldsymbol{B}^T\boldsymbol{B})^M)^{-1}$ and $((\boldsymbol{B}\boldsymbol{B}^T)^M)^{-1}$ in the form of recurrence relations and their initial values. All diagonals of the inverse of $(\boldsymbol{B}^T\boldsymbol{B})^M$ or $(\boldsymbol{B}\boldsymbol{B}^T)^M$ are computed within $O(N)$ flops according to the presented formula. Traces of the inverses of $(\boldsymbol{B}^T\boldsymbol{B})^M$ and $(\boldsymbol{B}\boldsymbol{B}^T)^M$ can be used to compute lower bounds of the minimal singular value of $\boldsymbol{B}$.

1 Introduction

To compute a lower bound of the minimal singular value of a bidiagonal matrix, traces of the inverses of particular matrix products can be utilized. Fernando and Parlett [1] considered a positive upper bidiagonal matrix A where all diagonals and upper subdiagonals of A are positive [¶]. They showed that diagonals of the inverse of the symmetric positive definite tridiagonal matrix AA^T are computed from auxiliary quantities which appear in the oqd (orthogonal quotient difference) algorithm and then a lower bound of the minimal singular value of A can be computed from the trace of the inverse of AA^T. Through this paper, let the suffix T of a matrix denote its transpose. von Matt [2] considered a non-singular upper bidiagonal matrix U where all diagonals and upper subdiagonals of U are nonzero. He presented a method for computing diagonals of the inverses of the symmetric positive definite matrix UU^T and $(UU^T)^2$ and then proposed two different shift strategies for the orthogonal qd-algorithm. Note that this algorithm is different from the algorithm in [1] having the same name. One of the shift strategies requires the trace of the inverse of UU^T and the other one requires the traces both of the inverses of UU^T and $(UU^T)^2$ to compute the shifts. The shifts are determined as lower bounds of the minimal singular value of U. When A or U is an $N \times N$ matrix, diagonals of the inverses of AA^T or UU^T or $(UU^T)^2$ can be computed within $O(N)$ flops.

In this paper, let M denote an arbitrary positive integer. On a non-singular upper bidiagonal matrix $\boldsymbol{B}$ where all diagonals and upper subdiagonals of $\boldsymbol{B}$ are nonzero, we prove a theorem on an $O(N)$ formula for diagonals of inverse of $(\boldsymbol{B}^T\boldsymbol{B})^M$ or $(\boldsymbol{B}\boldsymbol{B}^T)^M$ in a form of recurrence relations and their initial values. If an upper subdiagonal of $\boldsymbol{B}$

[a]Graduate School of Informatics, Kyoto University, Yoshida-Honmachi, Sakyo-ku, Kyoto 606-8501 Japan (kkimur@amp.i.kyoto-u.ac.jp, takumi@amp.i.kyoto-u.ac.jp, ynaka@amp.i.kyoto-u.ac.jp)

[b]SORST, Japan Science and Technology Agency

[§]corresponding author, E-mail : takumi@amp.i.kyoto-u.ac.jp, Tel. : +81-75-753-5511, Fax. : +81-75-753-5497

[¶]In [1], the matrix A is expressed with a symbol B.

is zero, it is readily shown that we can divide the problem to compute these diagonals into the same type of two smaller problems. Then, we can set an additional assumption that all the upper subdiagonals of $\boldsymbol{B}$ are nonzero without losing generality.

Here we fix the notations used in the theorem. Let $\boldsymbol{B} = (B_{i,j})$ denote an $N \times N$ ($N \geq 2$) non-singular upper bidiagonal matrix. Let the diagonal element and the upper subdiagonal element in the $i-$th row of $\boldsymbol{B}$ be denoted by b_i and c_i, respectively. That is,

$$\begin{cases} b_i = B_{i,i} & (1 \leq i \leq N), \\ c_i = B_{i,i+1} & (1 \leq i \leq N-1). \end{cases}$$

Let $N \times N$ matrices $\boldsymbol{V}^{(m)} = (V^{(m)}_{i,j})$, $\boldsymbol{W}^{(m)} = (W^{(m)}_{i,j})$, $\boldsymbol{X}^{(q)} = (X^{(q)}_{i,j})$ and $\boldsymbol{Y}^{(q)} = (Y^{(q)}_{i,j})$ denote

$$\begin{cases} \boldsymbol{V}^{(m)} \equiv ((\boldsymbol{B}^T\boldsymbol{B})^m)^{-1}, \\ \boldsymbol{W}^{(m)} \equiv ((\boldsymbol{B}\boldsymbol{B}^T)^m)^{-1}, \\ \boldsymbol{X}^{(q)} \equiv (\boldsymbol{B}(\boldsymbol{B}^T\boldsymbol{B})^q)^{-1} = ((\boldsymbol{B}\boldsymbol{B}^T)^q\boldsymbol{B})^{-1}, \\ \boldsymbol{Y}^{(q)} \equiv (\boldsymbol{B}^T(\boldsymbol{B}\boldsymbol{B}^T)^q)^{-1} = ((\boldsymbol{B}^T\boldsymbol{B})^q\boldsymbol{B}^T)^{-1} \end{cases} \tag{1}$$

for integers m ($0 \leq m \leq M$) and q ($0 \leq q \leq M-1$), respectively. Next, we simply write diagonals of the matrices $\boldsymbol{V}^{(m)}$, $\boldsymbol{W}^{(m)}$, $\boldsymbol{X}^{(q)}$ and $\boldsymbol{Y}^{(q)}$ as $v^{(m)}_i \equiv V^{(m)}_{i,i}$, $w^{(m)}_i \equiv W^{(m)}_{i,i}$, $x^{(q)}_i \equiv X^{(q)}_{i,i}$ and $y^{(q)}_i \equiv Y^{(q)}_{i,i}$ for $1 \leq i \leq N$, respectively.

Let $z^{(q)}_i$ denote

$$z^{(q)}_i \equiv b_i(x^{(q)}_i + y^{(q)}_i). \tag{2}$$

for $1 \leq i \leq N$ and $0 \leq q \leq M-1$.

Then, under the assumptions that all diagonals and subdiagonals of $\boldsymbol{B}$ are nonzero, the following theorem holds.

Theorem

Let M be an arbitrary positive integer. All the diagonal elements $v^{(M)}_i$ and $w^{(M)}_i$ of inverse matrices $((\boldsymbol{B}^T\boldsymbol{B})^M)^{-1}$ and $((\boldsymbol{B}\boldsymbol{B}^T)^M)^{-1}$, respectively, are computed by the following simple recurrence relations with the initial values. The recurrence relations are

$$v^{(p)}_i = \frac{1}{b_i^2}(c_i^2 v^{(p)}_{i+1} + z^{(p-1)}_i - w^{(p-1)}_i) \qquad (1 \leq i \leq N-1), \tag{3}$$

$$w^{(p)}_i = \frac{1}{b_i^2}(c_{i-1}^2 w^{(p)}_{i-1} + z^{(p-1)}_i - v^{(p-1)}_i) \qquad (2 \leq i \leq N), \tag{4}$$

$$z^{(q)}_i = z^{(q)}_{i-1} + 2(v^{(q)}_i - w^{(q)}_{i-1}) \qquad (2 \leq i \leq N), \tag{5}$$

$$v^{(p)}_N = \frac{1}{b_N^2} w^{(p-1)}_N, \tag{6}$$

$$w^{(p)}_1 = \frac{1}{b_1^2} v^{(p-1)}_1, \tag{7}$$

$$z^{(q)}_1 = 2v^{(q)}_1, \tag{8}$$

for integers p and q such that $1 \leq p \leq M$ and $0 \leq q \leq M-1$. The initial values are

$$v^{(0)}_i = 1, \tag{9}$$

$$w^{(0)}_i = 1, \tag{10}$$

for $1 \leq i \leq N$. $\square$

2 Proof of Theorem

Let $N \geq 2$ unless we specify the range of N. Throughout this section, p is an integer such that $1 \leq p \leq M$ and q is an integer such that $0 \leq q \leq M-1$, respectively. For convenience, let us represent the inverse of $\boldsymbol{B}$ with the notation $\boldsymbol{S} = (S_{i,j})$. Since the matrix product $\boldsymbol{D} \equiv \boldsymbol{BS} = (D_{i,j})$ is the identity matrix $\boldsymbol{I}$,

$$D_{i,j} = \sum_{k=1}^{N} B_{i,k} S_{k,j} = \begin{cases} b_i S_{i,j} + c_i S_{i+1,j} = \delta_{i,j} & (1 \leq i \leq N-1), \\ b_N S_{N,j} = \delta_{N,j} & (i = N), \end{cases} \tag{11}$$

where δ is Kronecker's delta. From Eq. (11), we obtain

$$S_{N,j} = \begin{cases} 0 & (N \geq 2 \text{ and } 1 \leq j \leq N-1), \\ \dfrac{1}{b_N} & (N \geq 2 \text{ and } j = N), \end{cases} \tag{12}$$

$$D_{i,j} = b_i S_{i,j} + c_i S_{i+1,j} = 0 \quad (N \geq 3 \text{ and } 1 \leq j < i \leq N-1). \tag{13}$$

Then, it is derived inductively from Eqs. (12) and (13) that

$$S_{i,j} = 0 \quad (N \geq 3 \text{ and } 1 \leq j < i \leq N-1). \tag{14}$$

From Eqs. (12) and (14), we see that strictly lower triangular elements of $\boldsymbol{S}$ are zero for $N \geq 2$. Since $\boldsymbol{S}$ is an upper triangle matrix, we have

$$D_{i,i} = b_i S_{i,i} = 1 \quad (1 \leq i \leq N)$$

from Eq. (11). Thus the diagonals $S_{i,i}$ of $\boldsymbol{S}$ are the inverses of b_i, respectively,

$$S_{i,i} = \frac{1}{b_i} \quad (1 \leq i \leq N).$$

Since $D_{i,j} = b_i S_{i,j} + c_i S_{i+1,j} = 0$ $(i < j)$ from Eq. (11), we have

$$S_{i+1,j} = -\frac{b_i}{c_i} S_{i,j} \quad (1 \leq i < j \leq N).$$

Summarizing these results, we have

$$\begin{cases} S_{i+1,j} = -\dfrac{b_i}{c_i} S_{i,j} & (1 \leq i < j \leq N), \\ S_{i,j} = \dfrac{1}{b_i} & (1 \leq i = j \leq N), \\ S_{i,j} = 0 & (1 \leq j < i \leq N). \end{cases} \tag{15}$$

On the other hand, since the matrix product $\boldsymbol{SB}$ is $\boldsymbol{I}$ and its (i, j) element is $S_{i,j-1} c_{j-1} + S_{i,j} b_j = 0$ $(i < j)$, we have

$$S_{i,j} = -\frac{c_{j-1}}{b_j} S_{i,j-1} \quad (1 \leq i < j \leq N). \tag{16}$$

Let us set $\boldsymbol{S}' \equiv (\boldsymbol{B}^T)^{-1}$. Since $\boldsymbol{S}'^T \boldsymbol{B} = (\boldsymbol{B}^T \boldsymbol{S}')^T = \boldsymbol{I}^T = \boldsymbol{I}$, we have

$$(\boldsymbol{B}^T)^{-1} = \boldsymbol{S}' = (\boldsymbol{S}'^T)^T = (\boldsymbol{B}^{-1})^T = \boldsymbol{S}^T.$$

Since $(\boldsymbol{B}^T\boldsymbol{B})^{-1} = \boldsymbol{S}\boldsymbol{S}^T$ and $(\boldsymbol{B}\boldsymbol{B}^T)^{-1} = \boldsymbol{S}^T\boldsymbol{S}$, the matrices $\boldsymbol{V}^{(p)}$ and $\boldsymbol{W}^{(p)}$ are expressed by the matrices $\boldsymbol{V}^{(p-1)}$ and $\boldsymbol{W}^{(p-1)}$ as follows

$$\begin{cases} \boldsymbol{V}^{(p)} = (\boldsymbol{S}\boldsymbol{S}^T)^p = \boldsymbol{S}(\boldsymbol{S}^T\boldsymbol{S})^{p-1}\boldsymbol{S}^T = \boldsymbol{S}\boldsymbol{W}^{(p-1)}\boldsymbol{S}^T, \\ \boldsymbol{W}^{(p)} = (\boldsymbol{S}^T\boldsymbol{S})^p = \boldsymbol{S}^T(\boldsymbol{S}\boldsymbol{S}^T)^{p-1}\boldsymbol{S} = \boldsymbol{S}^T\boldsymbol{V}^{(p-1)}\boldsymbol{S}. \end{cases} \tag{17}$$

Next, from the definition (1), the matrices $\boldsymbol{X}^{(q)}$ and $\boldsymbol{Y}^{(q)}$ are expressed by the matrices $\boldsymbol{V}^{(q)}$ and $\boldsymbol{W}^{(q)}$ as follows

$$\begin{cases} \boldsymbol{X}^{(q)} = \boldsymbol{V}^{(q)}\boldsymbol{S} = \boldsymbol{S}\boldsymbol{W}^{(q)}, \\ \boldsymbol{Y}^{(q)} = \boldsymbol{W}^{(q)}\boldsymbol{S}^T = \boldsymbol{S}^T\boldsymbol{V}^{(q)}. \end{cases} \tag{18}$$

From Eqs. (17) and (18), we obtain

$$\begin{cases} \boldsymbol{V}^{(p)} = \boldsymbol{X}^{(p-1)}\boldsymbol{S}^T = \boldsymbol{S}\boldsymbol{Y}^{(p-1)}, \\ \boldsymbol{W}^{(p)} = \boldsymbol{Y}^{(p-1)}\boldsymbol{S} = \boldsymbol{S}^T\boldsymbol{X}^{(p-1)}. \end{cases} \tag{19}$$

Let $\boldsymbol{P} = (P_{i,j})$ and $\boldsymbol{Q} = (Q_{i,j})$ be $N \times N$ matrices having some special relationship to $\boldsymbol{S}$ or $\boldsymbol{S}^T$. In the cases that $\boldsymbol{P}$ and $\boldsymbol{Q}$ hold a relationship such that $\boldsymbol{P} = \boldsymbol{S}\boldsymbol{Q}$, $\boldsymbol{P} = \boldsymbol{Q}\boldsymbol{S}^T$, $\boldsymbol{P} = \boldsymbol{S}^T\boldsymbol{Q}$ or $\boldsymbol{P} = \boldsymbol{Q}\boldsymbol{S}$, the following four lemmas can be proved.

Lemma 1

When $\boldsymbol{P}$, $\boldsymbol{Q}$ and $\boldsymbol{S}$ hold $\boldsymbol{P} = \boldsymbol{S}\boldsymbol{Q}$, then the elements of $\boldsymbol{P}$ and $\boldsymbol{Q}$ satisfy

$$P_{i+1,j} + \frac{b_i}{c_i}P_{i,j} = \frac{1}{c_i}Q_{i,j} \quad (1 \le i \le N-1 \text{ and } 1 \le j \le N). \tag{20}$$

Proof.

In this proof, j is an integer such that $1 \le j \le N$. The element $P_{i,j}$ is expressed with $\alpha_{i,j}$ as

$$\begin{aligned} P_{i,j} &= \sum_{k=1}^{N} S_{i,k}Q_{k,j} = \sum_{k=i}^{N} S_{i,k}Q_{k,j} = S_{i,i}Q_{i,j} + S_{i,i+1}Q_{i+1,j} + \alpha_{i,j} \\ &= \frac{1}{b_i}Q_{i,j} + \left(-\frac{c_i}{b_{i+1}}S_{i,i}\right)Q_{i+1,j} + \alpha_{i,j} \\ &= \frac{1}{b_i}Q_{i,j} - \frac{c_i}{b_{i+1}b_i}Q_{i+1,j} + \alpha_{i,j} \quad (N \ge 2 \text{ and } 1 \le i \le N-1), \end{aligned} \tag{21}$$

where $\alpha_{i,j}$ is defined by

$$\alpha_{i,j} \equiv \begin{cases} \displaystyle\sum_{k=i+2}^{N} S_{i,k}Q_{k,j} & (N \ge 3,\ 1 \le i \le N-2), \\ 0 & (N \ge 2,\ i = N-1). \end{cases}$$

The element $P_{i+1,j}$ is also expressed with $\alpha_{i,j}$. If $N \ge 3$ and $1 \le i \le N-2$, it holds

$$\begin{aligned} P_{i+1,j} &= \sum_{k=i+1}^{N} S_{i+1,k}Q_{k,j} = S_{i+1,i+1}Q_{i+1,j} + \sum_{k=i+2}^{N}\left(-\frac{b_i}{c_i}S_{i,k}\right)Q_{k,j} \\ &= \frac{1}{b_{i+1}}Q_{i+1,j} - \frac{b_i}{c_i}\alpha_{i,j}. \end{aligned} \tag{22}$$

If $N \geq 2$ and $i = N-1$, it holds

$$P_{i+1,j} = \sum_{k=i+1}^{N} S_{i+1,k} Q_{k,j} = S_{i+1,i+1} Q_{i+1,j} = \frac{1}{b_{i+1}} Q_{i+1,j} - \frac{b_i}{c_i}\alpha_{i,j}. \tag{23}$$

From Eqs. (21), (22) and (23), we obtain

$$\begin{aligned} P_{i+1,j} + \frac{b_i}{c_i} P_{i,j} &= \frac{1}{b_{i+1}} Q_{i+1,j} - \frac{b_i}{c_i}\alpha_{i,j} + \frac{1}{c_i} Q_{i,j} - \frac{1}{b_{i+1}} Q_{i+1,j} + \frac{b_i}{c_i}\alpha_{i,j} \\ &= \frac{1}{c_i} Q_{i,j} \quad (1 \leq i \leq N-1 \text{ and } 1 \leq j \leq N). \qquad \square \end{aligned}$$

Lemma 2

When $\boldsymbol{P}$, $\boldsymbol{Q}$ and $\boldsymbol{S}^T$ hold $\boldsymbol{P} = \boldsymbol{Q}\boldsymbol{S}^T$, then the elements of $\boldsymbol{P}$ and $\boldsymbol{Q}$ satisfy

$$P_{i,j+1} + \frac{b_j}{c_j} P_{i,j} = \frac{1}{c_j} Q_{i,j} \quad (1 \leq i \leq N \text{ and } 1 \leq j \leq N-1). \tag{24}$$

Proof.

In this proof, i is an integer such that $1 \leq i \leq N$. The element $P_{i,j}$ is expressed with $\alpha_{i,j}$ as

$$\begin{aligned} P_{i,j} &= \sum_{l=1}^{N} Q_{i,l} S_{j,l} = \sum_{l=j}^{N} Q_{i,l} S_{j,l} = Q_{i,j} S_{j,j} + Q_{i,j+1} S_{j,j+1} + \alpha_{i,j} \\ &= \frac{1}{b_j} Q_{i,j} + Q_{i,j+1}\left(-\frac{c_j}{b_{j+1}} S_{j,j}\right) + \alpha_{i,j} \\ &= \frac{1}{b_j} Q_{i,j} - \frac{c_j}{b_{j+1} b_j} Q_{i,j+1} + \alpha_{i,j} \quad (N \geq 2 \text{ and } 1 \leq j \leq N-1), \end{aligned} \tag{25}$$

where $\alpha_{i,j}$ is defined by

$$\alpha_{i,j} \equiv \begin{cases} \displaystyle\sum_{l=j+2}^{N} Q_{i,l} S_{j,l} & (N \geq 3,\ 1 \leq j \leq N-2), \\ 0 & (N \geq 2,\ j = N-1). \end{cases}$$

The element $P_{i,j+1}$ is also expressed with $\alpha_{i,j}$. If $N \geq 3$ and $1 \leq j \leq N-2$, it holds

$$\begin{aligned} P_{i,j+1} &= \sum_{l=j+1}^{N} Q_{i,l} S_{j+1,l} = Q_{i,j+1} S_{j+1,j+1} + \sum_{l=j+2}^{N} Q_{i,l}\left(-\frac{b_j}{c_j} S_{j,l}\right) \\ &= \frac{1}{b_{j+1}} Q_{i,j+1} - \frac{b_j}{c_j}\alpha_{i,j}. \end{aligned} \tag{26}$$

If $N \geq 2$ and $j = N-1$, it holds

$$P_{i,j+1} = \sum_{l=j+1}^{N} Q_{i,l} S_{j+1,l} = Q_{i,j+1} S_{j+1,j+1} = \frac{1}{b_{j+1}} Q_{i,j+1} - \frac{b_j}{c_j}\alpha_{i,j}. \tag{27}$$

From Eqs. (25), (26) and (27), we obtain

$$\begin{aligned} P_{i,j+1} + \frac{b_j}{c_j} P_{i,j} &= \frac{1}{b_{j+1}} Q_{i,j+1} - \frac{b_j}{c_j}\alpha_{i,j} + \frac{1}{c_j} Q_{i,j} - \frac{1}{b_{j+1}} Q_{i,j+1} + \frac{b_j}{c_j}\alpha_{i,j} \\ &= \frac{1}{c_j} Q_{i,j} \quad (1 \leq i \leq N \text{ and } 1 \leq j \leq N-1). \qquad \square \end{aligned}$$

Lemma 3

When $\boldsymbol{P}$, $\boldsymbol{Q}$ and $\boldsymbol{S}^T$ hold $\boldsymbol{P} = \boldsymbol{S}^T\boldsymbol{Q}$, then the elements of $\boldsymbol{P}$ and $\boldsymbol{Q}$ satisfy

$$P_{i-1,j} + \frac{b_i}{c_{i-1}} P_{i,j} = \frac{1}{c_{i-1}} Q_{i,j} \quad (2 \le i \le N \text{ and } 1 \le j \le N). \tag{28}$$

Proof.

In this proof, j is an integer such that $1 \le j \le N$. The element $P_{i,j}$ is expressed with $\alpha_{i,j}$ as

$$\begin{aligned} P_{i,j} &= \sum_{k=1}^{N} S_{k,i} Q_{k,j} = \sum_{k=1}^{i} S_{k,i} Q_{k,j} = S_{i,i} Q_{i,j} + S_{i-1,i} Q_{i-1,j} + \alpha_{i,j} \\ &= \frac{1}{b_i} Q_{i,j} + \left(-\frac{c_{i-1}}{b_i} S_{i-1,i-1}\right) Q_{i-1,j} + \alpha_{i,j} \\ &= \frac{1}{b_i} Q_{i,j} - \frac{c_{i-1}}{b_i b_{i-1}} Q_{i-1,j} + \alpha_{i,j} \quad (N \ge 2 \text{ and } 2 \le i \le N), \end{aligned} \tag{29}$$

where $\alpha_{i,j}$ is defined by

$$\alpha_{i,j} \equiv \begin{cases} \displaystyle\sum_{k=1}^{i-2} S_{k,i} Q_{k,j} & (N \ge 3,\ 3 \le i \le N), \\ 0 & (N \ge 2,\ i = 2). \end{cases}$$

The element $P_{i-1,j}$ is also expressed with $\alpha_{i,j}$. If $N \ge 3$ and $3 \le i \le N$, it holds

$$\begin{aligned} P_{i-1,j} &= \sum_{k=1}^{i-1} S_{k,i-1} Q_{k,j} = S_{i-1,i-1} Q_{i-1,j} + \sum_{k=1}^{i-2} \left(-\frac{b_i}{c_{i-1}} S_{k,i}\right) Q_{k,j} \\ &= \frac{1}{b_{i-1}} Q_{i-1,j} - \frac{b_i}{c_{i-1}} \alpha_{i,j}. \end{aligned} \tag{30}$$

If $N \ge 2$ and $i = 2$, it holds

$$P_{i-1,j} = \sum_{k=1}^{i-1} S_{k,i-1} Q_{k,j} = S_{i-1,i-1} Q_{i-1,j} = \frac{1}{b_{i-1}} Q_{i-1,j} - \frac{b_i}{c_{i-1}} \alpha_{i,j}. \tag{31}$$

From Eqs. (29), (30) and (31), we obtain

$$\begin{aligned} P_{i-1,j} + \frac{b_i}{c_{i-1}} P_{i,j} &= \frac{1}{b_{i-1}} Q_{i-1,j} - \frac{b_i}{c_{i-1}} \alpha_{i,j} + \frac{1}{c_{i-1}} Q_{i,j} - \frac{1}{b_{i-1}} Q_{i-1,j} + \frac{b_i}{c_{i-1}} \alpha_{i,j} \\ &= \frac{1}{c_{i-1}} Q_{i,j} \quad (2 \le i \le N \text{ and } 1 \le j \le N). \end{aligned}$$

□

Lemma 4

When $\boldsymbol{P}$, $\boldsymbol{Q}$ and $\boldsymbol{S}$ hold $\boldsymbol{P} = \boldsymbol{QS}$, then the elements of $\boldsymbol{P}$ and $\boldsymbol{Q}$ satisfy

$$P_{i,j-1} + \frac{b_j}{c_{j-1}} P_{i,j} = \frac{1}{c_{j-1}} Q_{i,j} \quad (1 \le i \le N \text{ and } 2 \le j \le N). \tag{32}$$

Proof.

In this proof, i is an integer such that $1 \leq i \leq N$. The element $P_{i,j}$ is expressed with $\alpha_{i,j}$ as

$$\begin{aligned}
P_{i,j} &= \sum_{l=1}^{N} Q_{i,l} S_{l,j} = \sum_{l=1}^{j} Q_{i,l} S_{l,j} = Q_{i,j} S_{j,j} + Q_{i,j-1} S_{j-1,j} + \alpha_{i,j} \\
&= \frac{1}{b_j} Q_{i,j} + Q_{i,j-1} \left(-\frac{c_{j-1}}{b_j} S_{j-1,j-1} \right) + \alpha_{i,j} \\
&= \frac{1}{b_j} Q_{i,j} - \frac{c_{j-1}}{b_j b_{j-1}} Q_{i,j-1} + \alpha_{i,j} \quad (N \geq 2 \text{ and } 2 \leq j \leq N),
\end{aligned} \tag{33}$$

where $\alpha_{i,j}$ is defined by

$$\alpha_{i,j} \equiv \begin{cases} \displaystyle\sum_{l=1}^{j-2} Q_{i,l} S_{l,j} & (N \geq 3,\ 3 \leq j \leq N), \\ 0 & (N \geq 2,\ j = 2). \end{cases}$$

The element $P_{i,j-1}$ is also expressed with $\alpha_{i,j}$. If $N \geq 3$ and $3 \leq j \leq N$, it holds

$$\begin{aligned}
P_{i,j-1} &= \sum_{l=1}^{j-1} Q_{i,l} S_{l,j-1} = Q_{i,j-1} S_{j-1,j-1} + \sum_{l=1}^{j-2} Q_{i,l} \left(-\frac{b_j}{c_{j-1}} S_{l,j} \right) \\
&= \frac{1}{b_{j-1}} Q_{i,j-1} - \frac{b_j}{c_{j-1}} \alpha_{i,j}.
\end{aligned} \tag{34}$$

If $N \geq 2$ and $j = 2$, it holds

$$P_{i,j-1} = \sum_{l=1}^{j-1} Q_{i,l} S_{l,j-1} = Q_{i,j-1} S_{j-1,j-1} = \frac{1}{b_{j-1}} Q_{i,j-1} - \frac{b_j}{c_{j-1}} \alpha_{i,j}. \tag{35}$$

From Eqs. (33), (34) and (35), we obtain

$$\begin{aligned}
P_{i,j-1} + \frac{b_j}{c_{j-1}} P_{i,j} &= \frac{1}{b_{j-1}} Q_{i,j-1} - \frac{b_j}{c_{j-1}} \alpha_{i,j} + \frac{1}{c_{j-1}} Q_{i,j} - \frac{1}{b_{j-1}} Q_{i,j-1} + \frac{b_j}{c_{j-1}} \alpha_{i,j} \\
&= \frac{1}{c_{j-1}} Q_{i,j} \quad (1 \leq i \leq N \text{ and } 2 \leq j \leq N). \qquad \square
\end{aligned}$$

Now we derive recurrence relations on $v_i^{(p)}, w_j^{(p)}$ and $z_j^{(q)}$ for $1 \leq i \leq N-1$ or $2 \leq j \leq N$ by applying appropriate lemmas among Lemma 1, 2, 3 and 4 to Eqs. (18) and (19).

Let us derive the recurrence relation (3) on $v_i^{(p)}$. In the following derivation, i is an integer such that $1 \leq i \leq N-1$. The element $V_{i,i+1}^{(p)}$ is expressed in two ways by applying Lemmas 1 and 2 to Eq. (19). Using the lemmas, we obtain

$$V_{i,i+1}^{(p)} = -\frac{c_i}{b_i} v_{i+1}^{(p)} + \frac{1}{b_i} Y_{i,i+1}^{(p-1)}, \tag{36}$$

$$V_{i,i+1}^{(p)} = -\frac{b_i}{c_i} v_i^{(p)} + \frac{1}{c_i} x_i^{(p-1)}. \tag{37}$$

Since the right hand sides of Eqs. (36) and (37) are equal to each other, we have

$$v_i^{(p)} = \frac{c_i^2}{b_i^2} v_{i+1}^{(p)} + \frac{1}{b_i} x_i^{(p-1)} - \frac{c_i}{b_i^2} Y_{i,i+1}^{(p-1)} \tag{38}$$

on the diagonals $v_i^{(p)}$. On $Y_{i,i+1}^{(p-1)}$ in the right hand side of Eq. (38), we obtain

$$Y_{i,i+1}^{(p-1)} = -\frac{b_i}{c_i} y_i^{(p-1)} + \frac{1}{c_i} w_i^{(p-1)} \tag{39}$$

by applying Lemma 2 to Eq. (18). Substituting Eq. (39) into Eq. (38), finally we derive

$$v_i^{(p)} = \frac{1}{b_i^2}(c_i^2 v_{i+1}^{(p)} + z_i^{(p-1)} - w_i^{(p-1)})$$

on the diagonals $v_i^{(p)}$ of the inverse matrix $((\boldsymbol{B}^T\boldsymbol{B})^p)^{-1}$. This is Eq. (3) in Theorem.

Next, let us derive recurrence relations on $w_i^{(p)}$ and $z_i^{(q)}$ in a similar way to $v_i^{(p)}$. In the following derivation, i is an integer such that $2 \le i \le N$. Applying Lemmas 3 and 4 to Eq. (19), we have

$$W_{i,i-1}^{(p)} = -\frac{c_{i-1}}{b_i} w_{i-1}^{(p)} + \frac{1}{b_i} X_{i,i-1}^{(p-1)},$$
$$W_{i,i-1}^{(p)} = -\frac{b_i}{c_{i-1}} w_i^{(p)} + \frac{1}{c_{i-1}} y_i^{(p-1)}.$$

These equations lead to

$$w_i^{(p)} = \frac{c_{i-1}^2}{b_i^2} w_{i-1}^{(p)} + \frac{1}{b_i} y_i^{(p-1)} - \frac{c_{i-1}}{b_i^2} X_{i,i-1}^{(p-1)}.$$

Applying Lemma 4 to Eq. (18), we have

$$X_{i,i-1}^{(p-1)} = -\frac{b_i}{c_{i-1}} x_i^{(p-1)} + \frac{1}{c_{i-1}} v_i^{(p-1)}.$$

Then, we obtain a recurrence relation

$$w_i^{(p)} = \frac{1}{b_i^2}(c_{i-1}^2 w_{i-1}^{(p)} + z_i^{(p-1)} - v_i^{(p-1)})$$

on the diagonals $w_i^{(p)}$ of the inverse matrix $((\boldsymbol{B}\boldsymbol{B}^T)^p)^{-1}$. This is Eq. (4) in Theorem.

Applying Lemmas 1 and 4 to Eq. (18), we have

$$X_{i,i-1}^{(q)} = -\frac{b_{i-1}}{c_{i-1}} x_{i-1}^{(q)} + \frac{1}{c_{i-1}} w_{i-1}^{(q)},$$
$$X_{i,i-1}^{(q)} = -\frac{b_i}{c_{i-1}} x_i^{(q)} + \frac{1}{c_{i-1}} v_i^{(q)}.$$

These equations lead to

$$b_i x_i^{(q)} = b_{i-1} x_{i-1}^{(q)} + v_i^{(q)} - w_{i-1}^{(q)}. \tag{40}$$

Applying Lemmas 2 and 3 to Eq. (18), we have

$$Y_{i-1,i}^{(q)} = -\frac{b_{i-1}}{c_{i-1}} y_{i-1}^{(q)} + \frac{1}{c_{i-1}} w_{i-1}^{(q)},$$
$$Y_{i-1,i}^{(q)} = -\frac{b_i}{c_{i-1}} y_i^{(q)} + \frac{1}{c_{i-1}} v_i^{(q)}.$$

These equations lead to

$$b_i y_i^{(q)} = b_{i-1} y_{i-1}^{(q)} + v_i^{(q)} - w_{i-1}^{(q)}. \tag{41}$$

Summing both hand sides of Eqs. (40) and (41) with the definition (2), we obtain a recurrence relation

$$z_i^{(q)} = z_{i-1}^{(q)} + 2(v_i^{(q)} - w_{i-1}^{(q)})$$

on $z_i^{(q)}$. This is Eq. (5) in Theorem.

Next, let us consider the values of $v_N^{(p)}, w_1^{(p)}$ and $z_1^{(q)}$ which are the end points of the sequence $v_i^{(p)}, w_i^{(p)}$ and $z_i^{(q)}$ for $1 \le i \le N$ on each p or q, respectively. From Eq. (17), we derive

$$\begin{cases} v_N^{(p)} = \displaystyle\sum_{k=1}^{N}\sum_{l=1}^{N} S_{N,k} W_{k,l}^{(p-1)} S_{l,N}^{T} = (S_{N,N})^2 W_{N,N}^{(p-1)} = \frac{1}{b_N^2} w_N^{(p-1)}, \\ w_1^{(p)} = \displaystyle\sum_{k=1}^{N}\sum_{l=1}^{N} S_{1,k}^{T} V_{k,l}^{(p-1)} S_{l,1} = (S_{1,1})^2 V_{1,1}^{(p-1)} = \frac{1}{b_1^2} v_1^{(p-1)}. \end{cases} \tag{42}$$

This is because the matrix $\boldsymbol{S}$ is an upper triangular matrix. These are Eqs. (6) and (7) in Theorem. By a way which is similar to Eq. (42), we obtain

$$\begin{cases} x_1^{(q)} = \displaystyle\sum_{k=1}^{N} V_{1,k}^{(q)} S_{k,1} = V_{1,1}^{(q)} S_{1,1} = \frac{1}{b_1} v_1^{(q)}, \\ y_1^{(q)} = \displaystyle\sum_{k=1}^{N} S_{1,k}^{T} V_{k,1}^{(q)} = S_{1,1} V_{1,1}^{(q)} = \frac{1}{b_1} v_1^{(q)}, \end{cases}$$

from Eq. (18). Summing both hand sides, we have

$$z_1^{(q)} = b_1(x_1^{(q)} + y_1^{(q)}) = 2v_1^{(q)}.$$

This is Eq. (8) in Theorem. Now the recurrence relations in Theorem have been derived.

Finally, let us consider the initial values. The values of $v_i^{(0)}$ and $w_i^{(0)}$ take the following simple values. It follows from the definition of $\boldsymbol{V}^{(0)}$ and $\boldsymbol{W}^{(0)}$ that

$$\begin{cases} v_i^{(0)} = V_{i,i}^{(0)} = I_{i,i}^{-1} = I_{i,i} = 1, \\ w_i^{(0)} = W_{i,i}^{(0)} = I_{i,i}^{-1} = I_{i,i} = 1, \end{cases} \qquad (1 \le i \le N).$$

They are the initial values (9) and (10) in Theorem. □

As a corollary of Theorem, we estimate the computational cost to get all the diagonals.

Corollary

All the diagonals of the inverse matrix $((\boldsymbol{B}^T\boldsymbol{B})^M)^{-1}$ or $((\boldsymbol{B}\boldsymbol{B}^T)^M)^{-1}$ are computed within $O(N)$ flops through the recurrence relations and the initial values presented in Theorem.

Proof.

To compute values of $v_i^{(M)}$ or $w_i^{(M)}$ on all i $(1 \le i \le N)$, values of $v_j^{(r)}, w_j^{(r)}$ and $z_j^{(r)}$ for all j $(1 \le j \le N)$ and r $(0 \le r \le M-1)$ are necessary in addition to themselves. Therefore, $(3M+1)N$ quantities are necessary to determine all the diagonals of the inverse matrix $((\boldsymbol{B}^T\boldsymbol{B})^M)^{-1}$ or $((\boldsymbol{B}\boldsymbol{B}^T)^M)^{-1}$. The initial values $v_i^{(0)}$ and $w_i^{(0)}$ $(1 \le i \le N)$ are given as 1 directly. From the initial values and the recurrence relations (5) and (8) in Theorem, we have

$$z_i^{(0)} = z_{i-1}^{(0)} + 2(v_i^{(0)} - w_{i-1}^{(0)}) = z_{i-1}^{(0)} + 2(1-1) = z_{i-1}^{(0)} = \cdots = z_1^{(0)} = 2v_1^{(0)} = 2 \cdot 1 = 2$$

for $2 \le i \le N$. Then, $z_i^{(0)}$ $(1 \le i \le N)$ are given as 2. Next, let us discuss on the remaining $(3M-2)N$ quantities. Regarding $1/b_i^2$ for $1 \le i \le N$ and c_j^2 for $1 \le j \le N-1$ in the recurrence relations (3), (4), (6) and (7) as new quantities, N times multiplication and division are necessary to compute all $1/b_i^2$ and $N-1$ times multiplication are necessary to compute all c_j^2. Then, each quantity of the remained $(3M-2)N$ quantities can be computed within at most four times of the four basic operations of arithmetic by using one of the recurrence relations as shown in Theorem with the above-mentioned new quantities and known quantities. Then, the total cost of operations to compute all the diagonals of the inverse matrix $((\boldsymbol{B}^T\boldsymbol{B})^M)^{-1}$ or $((\boldsymbol{B}\boldsymbol{B}^T)^M)^{-1}$ is $O(N)$ flops. □

3 Concluding Remarks

On the $N \times N$ $(N \ge 2)$ non-singular upper bidiagonal matrix $\boldsymbol{B}$ where all the diagonals and the upper subdiagonals of $\boldsymbol{B}$ are nonzero, the matrix products $\boldsymbol{B}^T\boldsymbol{B}$ and $\boldsymbol{B}\boldsymbol{B}^T$ are symmetric positive definite tridiagonal matrices. We present a theorem on a simple formula in the form of the recurrence relations with the initial values for computing all the diagonals of the inverse powers of these symmetric positive definite tridiagonal matrices such that $((\boldsymbol{B}^T\boldsymbol{B})^M)^{-1}$ and $((\boldsymbol{B}\boldsymbol{B}^T)^M)^{-1}$. Total cost of operations is within $O(N)$ flops. In Section 1, we referred to the preceding works by Fernando and Parlett [1] and von Matt [2] to compute such diagonals within $O(N)$ flops. Here we discuss a relationship to the methods in [1, 2].

The method by Fernando and Parlett is for computing the diagonals of $(AA^T)^{-1}$ where A is an $N \times N$ upper bidiagonal matrix whose diagonals and upper subdiagonals are positive [**]. The method by von Matt is for computing the diagonals of $(UU^T)^{-1}$ and $((UU^T)^2)^{-1}$ where U is an $N \times N$ upper bidiagonal matrix whose diagonals and upper subdiagonals are nonzero. In their methods, the quantities which are necessary to compute the diagonals are obtained by applying sequential Givens rotations to A^T or U or U^T.

On the other hand, the approach in this paper is somewhat different from those in [1, 2]. Namely, we derived the recurrence relations and the initial values for the desired diagonals $v_i^{(M)}$ and $w_i^{(M)}$ $(1 \le i \le N)$ based on an idea of utilizing the following inherent properties of the inverse $\boldsymbol{S}$ of $\boldsymbol{B}$. Indeed, $\boldsymbol{S}$ is an upper triangular matrix whose diagonals are directly determined from the diagonals of $\boldsymbol{B}$. Moreover, the neighboring elements among the diagonals and the off-diagonal elements in the upper triangular part of $\boldsymbol{S}$ are mutually related in a simple formula. These properties are shown in Eqs. (15) and (16).

As an important application of Theorem, lower bounds of the minimal singular value of $\boldsymbol{B}$ can be computed. When all the diagonals and the upper subdiagonals of $\boldsymbol{B}$ are nonzero, then the singular values of $\boldsymbol{B}$ are simple. Let singular values of $\boldsymbol{B}$ be $\sigma_1, \cdots, \sigma_N$ such that $\sigma_1 > \cdots > \sigma_N > 0$. Let us consider applications of the well-known Newton method and the Laguerre method [2] to the characteristic equation $f(\lambda) \equiv \det((\boldsymbol{B}^T\boldsymbol{B})^M - \lambda \boldsymbol{I}) = 0$. Note that $(\boldsymbol{B}^T\boldsymbol{B})^M$ is positive definite. Let $\tilde{\lambda}_N$ and

[**] In [1], the matrix A is expressed with a symbol B.

λ_N denote the minimal eigenvalues of $(\boldsymbol{B}^T\boldsymbol{B})^M$ and $\boldsymbol{B}^T\boldsymbol{B}$, respectively. It holds that $\tilde{\lambda}_N = \lambda_N^M = \sigma_N^{2M}$. The following quantity θ_M which gives a lower bound of the minimal singular value σ_N can be readily derived by applying the Newton method to $f(\lambda) = 0$ starting from $\lambda = 0$. Namely,

$$\theta_M \equiv (\mathrm{Tr}(((\boldsymbol{B}^T\boldsymbol{B})^M)^{-1}))^{-\frac{1}{2M}} = (\mathrm{Tr}(((\boldsymbol{B}\boldsymbol{B}^T)^M)^{-1}))^{-\frac{1}{2M}} = \left(\frac{1}{\sigma_1^{2M}} + \cdots + \frac{1}{\sigma_N^{2M}}\right)^{-\frac{1}{2M}} < \sigma_N.$$

It is to be noted that θ_M increases monotonically and converges to σ_N when M goes to infinity, that is, $\theta_1 < \theta_2 < \cdots < \sigma_N$. When $M = 1$, θ_1^2 corresponds to the well-known Newton shift discussed in [1, 2]. Then, θ_M^2 for $M \geq 2$ can be used as a better shift of origin in algorithms proposed in [1, 3] which compute singular values of $\boldsymbol{B}$. The following quantity φ_M which gives a lower bound of the minimal singular value σ_N can be derived through the Laguerre method for the characteristic equation. Namely,

$$\varphi_M \equiv \left(\frac{1}{\mathrm{Tr}((F^M)^{-1})} \cdot \frac{N}{1 + \sqrt{(N-1)\left(N\dfrac{\mathrm{Tr}((F^{2M})^{-1})}{(\mathrm{Tr}((F^M)^{-1}))^2} - 1\right)}}\right)^{\dfrac{1}{2M}} < \sigma_N,$$

where F is the matrix product $\boldsymbol{B}^T\boldsymbol{B}$ or $\boldsymbol{B}\boldsymbol{B}^T$. When $M = 1$, φ_1^2 corresponds to the Laguerre shift presented in [2]. Theorem presented in this paper helps us to compute the quantities θ_M and φ_M within $O(N)$ flops.

References

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

[2] U. von Matt, The orthogonal qd-algorithm, SIAM J. Sci. Comput. 18 (1997) 1163-1186.

[3] 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.