Analytical inversion of general tridiagonal matrices

1997 J. Phys. A: Math. Gen. 30 7919

(http://iopscience.iop.org/0305-4470/30/22/026)

J. Phys. A: Math. Gen. **30** (1997) 7919–7933. Printed in the UK PII: S0305-4470(97)85180-3

Analytical inversion of general tridiagonal matrices

Y Huang† and W F McColl‡

Oxford University Computing Laboratory, Wolfson Building, Parks Road, Oxford OX1 3QD, UK

Received 19 June 1997

Abstract. In this paper we give a complete analysis for general tridiagonal matrix inversion for both non-block and block cases, and provide some very simple analytical formulae which immediately lead to closed forms for some special cases such as symmetric or Toeplitz tridiagonal matrices.

1. Introduction

The need to find the inverse of tridiagonal matrices arises in many scientific and engineering applications. This problem has been investigated for a few decades with an attempt to find a simple and explicit analytic expression for the inverse. However, most of their efforts ended up with formulae for some special cases where the tridiagonal matrix is symmetric Toeplitz, see [1, 7] for example, or some fairly sophisticated formulae that rely on some strict requirements such as that all the entries of the lower/upper diagonals must not be zero, see [5, 3, 6] for example. For a review of the symmetric tridiagonal matrix inverse please refer to [4]. In this paper, we relate general tridiagonal matrix inversion to second-order linear recurrences and provide a set of very simple analytical formulae for both scalar and block cases. These formulae can immediately lead to closed forms for certain tridiagonal matrices such as general (block) Toeplitz tridiagonal matrices. The properties of the inverse are also discussed. To our knowledge, this is the first complete analysis for the general tridiagonal matrix inversion problem.

The paper is organized as follows. In section 2, we give an analytical formula for a general scalar tridiagonal matrix inversion and discuss some properties of the inverse. In section 3, the result is applied to the case of a general Toeplitz tridiagonal matrix and a closed-form expression for the inverse is obtained. In section 4, a formula for a general block tridiagonal matrix is presented. An extension to the block Toeplitz case is given in section 5. Part of the proof of the results have appeared in [2] and are included in the appendix.

† E-mail address: Yuguang.Huang@comlab.ox.ac.uk
‡ E-mail address: Bill.McColl@comlab.ox.ac.uk

2. Non-block case

We consider the inverse of a tridiagonal matrix

$$A = \begin{bmatrix} b_1 & c_1 & & & & & \\ a_2 & b_2 & c_2 & & & & \\ & \ddots & \ddots & \ddots & & & \\ & & a_j & b_j & c_j & & \\ & & & \ddots & \ddots & \ddots & \\ & & & & a_{n-1} & b_{n-1} & c_{n-1} \\ & & & & & a_n & b_n \end{bmatrix}.$$

2.1. Main results

Theorem 2.1. Define the second-order linear recurrences

$$z_i = b_i z_{i-1} - a_i c_{i-1} z_{i-2} \qquad i = 2, 3, \ldots, n \tag{1}$$

where $z_0 = 1$, $z_1 = b_1$, and

$$y_j = b_j y_{j+1} - a_{j+1} c_j y_{j+2} \qquad j = n-1, n-2, \ldots, 1 \tag{2}$$

where $y_{n+1} = 1$, $y_n = b_n$. The inverse matrix $A^{-1} = \{\phi_{i,j}\}$ $(1 \leqslant i, j \leqslant n)$ can be expressed as

$$\phi_{j,j} = \frac{1}{b_j - a_j c_{j-1} \frac{z_{j-2}}{z_{j-1}} - a_{j+1} c_j \frac{y_{j+2}}{y_{j+1}}} \tag{3}$$

where $j = 1, 2, \ldots, n$, $a_1 = 0$, $c_n = 0$ and

$$\phi_{i,j} = \begin{cases} -c_i \dfrac{z_{i-1}}{z_i} \phi_{i+1,j} & i < j \\ -a_i \dfrac{y_{i+1}}{y_i} \phi_{i-1,j} & i > j. \end{cases} \tag{4}$$

Proof. See appendix A. □

Corollary 2.1. The inverse matrix $A^{-1} = \{\phi_{i,j}\}$ can be expressed as

$$\phi_{j,j} = \frac{1}{b_j - a_j c_{j-1} \frac{z_{j-2}}{z_{j-1}} - a_{j+1} c_j \frac{y_{j+2}}{y_{j+1}}}$$

where $j = 1, 2, \ldots, n$, $a_1 = 0$, $c_n = 0$ and

$$\phi_{i,j} = \begin{cases} (-1)^{j-i} \left(\displaystyle\prod_{k=1}^{j-i} c_{j-k} \right) \dfrac{z_{i-1}}{z_{j-1}} \phi_{j,j} & i < j \\ (-1)^{i-j} \left(\displaystyle\prod_{k=1}^{i-j} a_{j+k} \right) \dfrac{y_{i+1}}{y_{j+1}} \phi_{j,j} & i > j. \end{cases} \tag{5}$$

Theorem 2.2. The inverse of the tridiagonal matrix A can be computed in $n^2 + 7n - 7$ arithmetic operations.

Proof. See appendix B. □

2.2. Properties of the inverse matrix

From theorem 2.1, we can easily obtain the following results.

Theorem 2.3. If any element of the lower-diagonal of A is zero, i.e. if $a_k = 0$ ($2 \leqslant k \leqslant n$), then

$$\phi_{i,j} = 0 \qquad \text{where } i = k \sim n,\ \ j = 1 \sim k-1. \tag{6}$$

If any element of the super-diagonal of A is zero, i.e. if $c_k = 0$ ($1 \leqslant k \leqslant n-1$), then

$$\phi_{i,j} = 0 \qquad \text{where } i = 1 \sim k,\ \ j = k+1 \sim n. \tag{7}$$

Lemma 2.1. If A is strictly diagonally dominant, i.e. $|b_i| > |a_i| + |c_i|$ for all $1 \leqslant i \leqslant n$, then

$$|z_i| > |c_i z_{i-1}| \qquad i = 1, \ldots, n \tag{8}$$

and

$$|y_j| > |a_j y_{j+1}| \qquad j = n, \ldots, 1. \tag{9}$$

Proof. We use induction on i.

When $i = 1$, (8) holds since

$$|z_1| = |b_1| > |c_1 z_0| = |c_1|.$$

Suppose (8) holds for all $1 \leqslant i \leqslant k-1 < n-1$, then we have

$$\begin{aligned} |z_k| &= |b_k z_{k-1} - a_k c_{k-1} z_{k-2}| \\ &\geqslant |b_k||z_{k-1}| - |a_k||c_{k-1} z_{k-2}| \\ &\geqslant |b_k||z_{k-1}| - |a_k||z_{k-1}| \\ &> |c_k||z_{k-1}|. \end{aligned}$$

Thus (8) holds for all $1 \leqslant i \leqslant n$. Similarly, we have $|y_j| > |a_j y_{j+1}|$ for all $1 \leqslant j \leqslant n$. $\square$

Theorem 2.4. If A is strictly diagonally dominant, i.e. $|b_i| > |a_i| + |c_i|$ for all $1 \leqslant i \leqslant n$, then theorem 2.1 will not break down.

Proof. All we need to do is to show $z_i \neq 0$, $y_i \neq 0$ and $b_j - a_j c_{j-1} \frac{z_{j-2}}{z_{j-1}} - a_{j+1} c_j \frac{y_{j+2}}{y_{j+1}} \neq 0$ for all $1 \leqslant i \leqslant n$.

From lemma 2.1, since $|z_1| = |b_1| \neq 0$, it is obvious that $|z_i| > 0$ for all $1 \leqslant i \leqslant n$. Similarly, since $|y_n| = |b_n| \neq 0$, we have $|y_j| > 0$ for all $1 \leqslant j \leqslant n$. Also, we have

$$\begin{aligned} \left| b_j - a_j c_{j-1} \frac{z_{j-2}}{z_{j-1}} - a_{j+1} c_j \frac{y_{j+2}}{y_{j+1}} \right| &\geqslant |b_j| - \left| a_j c_{j-1} \frac{z_{j-2}}{z_{j-1}} \right| - \left| a_{j+1} c_j \frac{y_{j+2}}{y_{j+1}} \right| \\ &\geqslant |b_j| - |a_j| - |c_j| > 0. \end{aligned}$$

Thus the theorem holds. $\square$

In general, we have the following theorem.

Theorem 2.5. If the tridiagonal matrix A satisfies any one of the conditions below, then theorem 2.1 will not break down.

(i) $a_i \neq 0$ $(i = 2 \sim n)$, $c_j \neq 0$ $(j = 1 \sim n-1)$ and $|b_i| \geqslant |a_i| + |c_i|$ $(i = 1 \sim n)$, and there exists at least one j $(1 \leqslant j \leqslant n)$ such that $|b_j| > |a_j| + |c_j|$.

(ii) $b_i > 0$ $(i = 1 \sim n)$ or $b_i < 0$ $(i = 1 \sim n)$, and $a_{i+1}c_i \leqslant 0$ $(i = 1 \sim n-1)$.

(iii) $b_i \geqslant 0$ $(i = 1 \sim n)$ or $b_i \leqslant 0$ $(i = 1 \sim n)$ with $b_1, b_n \neq 0$, and $a_{i+1}c_i < 0$ $(i = 1 \sim n-1)$.

Proof. These can be proved by induction, in a manner similar to theorem 2.4. □

Theorem 2.6. If A is strictly diagonally dominant, then the sequence $\phi_{i,j}$ is a strictly increasing function of i for $i < j$, and a strictly decreasing function of i for $i > j$.

Proof. This conclusion can be drawn easily from theorem 2.4 and lemma 2.1. □

3. An example

As a simple example, we consider a general Toeplitz tridiagonal matrix T,

$$T = \begin{bmatrix} 1 & -u & & & & & \\ -l & 1 & -u & & & & \\ & \ddots & \ddots & \ddots & & & \\ & & -l & 1 & -u & & \\ & & & \ddots & \ddots & \ddots & \\ & & & & -l & 1 & -u \\ & & & & & -l & 1 \end{bmatrix}. \tag{10}$$

For this matrix, we have the following theorem.

Theorem 3.1. The inverse of T can be expressed in the following explicit way:

$$(T^{-1})_{i,j} = \frac{(\lambda_+^i - \lambda_-^i)(\lambda_+^{n-j+1} - \lambda_-^{n-j+1})}{(\lambda_+ - \lambda_-)(\lambda_+^{n+1} - \lambda_-^{n+1})} u^{j-i} \qquad i < j \tag{11}$$

$$(T^{-1})_{i,j} = \frac{(\lambda_+^j - \lambda_-^j)(\lambda_+^{n-i+1} - \lambda_-^{n-i+1})}{(\lambda_+ - \lambda_-)(\lambda_+^{n+1} - \lambda_-^{n+1})} l^{i-j} \qquad i \geqslant j \tag{12}$$

where

$$\lambda_+ = \frac{1 + \sqrt{1 - 4lu}}{2} \qquad \lambda_- = \frac{1 - \sqrt{1 - 4lu}}{2}. \tag{13}$$

Proof. According to theorem 2.1, we have:

$$z_0 = 1 \qquad z_1 = 1 \qquad z_i = z_{i-1} - luz_{i-2} \qquad i = 2, \ldots, n \tag{14}$$

and

$$y_{n+1} = 1 \qquad y_n = 1 \qquad y_j = y_{j+1} - luy_{j+2} \qquad j = n-1, \ldots, 1. \tag{15}$$

It is obvious that

$$y_j = z_{n+1-j} \qquad j = 1, \ldots, n.$$

It can be shown that z_i can be expressed as

$$z_i = \beta_0 \lambda_+^{i+1} + \beta_1 \lambda_-^{i+1} \tag{16}$$

where λ_+ and λ_- are the roots of the quadratic equation $\lambda^2 - \lambda + lu = 0$.

$$\lambda_+ = \frac{1+\sqrt{1-4lu}}{2} \qquad \lambda_- = \frac{1-\sqrt{1-4lu}}{2}.$$

From the initial conditions $z_0 = z_1 = 1$, we have $\beta_0 = -\beta_1$. Hence (16) becomes

$$z_i = \beta_0(\lambda_+^{i+1} - \lambda_-^{i+1}). \tag{17}$$

Now, we have

$$\frac{z_{i-2}}{z_{i-1}} = \frac{\lambda_+^{i-1} - \lambda_-^{i-1}}{\lambda_+^{i} - \lambda_-^{i}} \tag{18}$$

$$\frac{y_{i+2}}{y_{i+1}} = \frac{z_{n-i-1}}{z_{n-i}} = \frac{\lambda_+^{n-i} - \lambda_-^{n-i}}{\lambda_+^{n-i+1} - \lambda_-^{n-i+1}}. \tag{19}$$

Combine (18) and (19) with theorem 2.1, we have

$$\begin{aligned}
(T^{-1})_{i,i} &= \frac{1}{1 - lu\left(\frac{\lambda_+^{i-1}-\lambda_-^{i-1}}{\lambda_+^{i}-\lambda_-^{i}} + \frac{\lambda_+^{n-i}-\lambda_-^{n-i}}{\lambda_+^{n-i+1}-\lambda_-^{n-i+1}}\right)} \\
&= \frac{1}{1 - \lambda_+\lambda_-\left(\frac{\lambda_+^{n}-\lambda_+^{n-i+1}\lambda_-^{i-1}-\lambda_+^{i-1}\lambda_-^{n-i+1}+\lambda_-^{n}+\lambda_+^{n}-\lambda_+^{i}\lambda_-^{n-i}-\lambda_+^{n-i}\lambda_-^{i}+\lambda_-^{n}}{\lambda_+^{n+1}-\lambda_+^{i}\lambda_-^{n-i+1}-\lambda_-^{i}\lambda_+^{n-i+1}+\lambda_-^{n+1}}\right)} \\
&= \frac{1}{1 - \lambda_+\lambda_-\left(\frac{2\lambda_+^{n}-\lambda_+^{n-i}\lambda_-^{i-1}(\lambda_++\lambda_-)-\lambda_+^{i-1}\lambda_-^{n-i}(\lambda_-+\lambda_+)+2\lambda_-^{n}}{\lambda_+^{n+1}-\lambda_+^{i}\lambda_-^{n-i+1}-\lambda_-^{i}\lambda_+^{n-i+1}+\lambda_-^{n+1}}\right)} \\
&= \frac{1}{1 - \lambda_+\lambda_-\left(\frac{2\lambda_+^{n}-\lambda_+^{n-i}\lambda_-^{i-1}-\lambda_+^{i-1}\lambda_-^{n-i}+2\lambda_-^{n}}{\lambda_+^{n+1}-\lambda_+^{i}\lambda_-^{n-i+1}-\lambda_-^{i}\lambda_+^{n-i+1}+\lambda_-^{n+1}}\right)} \\
&= \frac{(\lambda_+^{i}-\lambda_-^{i})(\lambda_+^{n-i+1}-\lambda_-^{n-i+1})}{\lambda_+^{n+1}-\lambda_+^{i}\lambda_-^{n-i+1}-\lambda_-^{i}\lambda_+^{n-i+1}+\lambda_-^{n+1}-2\lambda_+^{n+1}\lambda_-+\lambda_+^{n-i+1}\lambda_-^{i}+\lambda_+^{i}\lambda_-^{n-i+1}-2\lambda_+\lambda_-^{n+1}} \\
&= \frac{(\lambda_+^{i}-\lambda_-^{i})(\lambda_+^{n-i+1}-\lambda_-^{n-i+1})}{(\lambda_+-\lambda_-)(\lambda_+^{n+1}-\lambda_-^{n+1})}.
\end{aligned} \tag{20}$$

From (20) and corollary 2.1, it is obvious that equations (11) and (12) hold. □

3.1. Discussion

Based on theorem 3.1, we now discuss three cases.

Case 1. $lu = \frac{1}{4}$.

If $lu = \frac{1}{4}$, from (13) we have $\lambda_+ = \lambda_- = \lambda = \frac{1}{2}$. For any $i > 0$ we have

$$\lambda_+^{i} - \lambda_-^{i} = (\lambda_+ - \lambda_-)\sum_{k=0}^{i-1}(\lambda_+^{k}\lambda_-^{i-1-k}) = (\lambda_+ - \lambda_-)i\lambda^{i-1}. \tag{21}$$

Substituting (21) into (11) and (12), we have

$$(T^{-1})_{i,j} = \frac{i(n-j+1)}{2(n+1)}\left(\frac{u}{l}\right)^{(j-i)/2} \qquad i < j \tag{22}$$

$$(T^{-1})_{i,j} = \frac{j(n-i+1)}{2(n+1)}\left(\frac{l}{u}\right)^{(i-j)/2} \qquad i \geqslant j. \tag{23}$$

Case 2. $lu < \frac{1}{4}$.

If $lu < \frac{1}{4}$, then $1 - 4lu > 0$, λ_+ and λ_- are both real numbers. If we set

$$\lambda_+ = \frac{1+\sqrt{1-4lu}}{2} = \gamma \mathrm{e}^{\theta} \tag{24}$$

$$\lambda_- = \frac{1-\sqrt{1-4lu}}{2} = \gamma \mathrm{e}^{-\theta} \tag{25}$$

then from (24) and (25) we have

$$2\gamma \cosh\theta = 1$$

and

$$\gamma^2 = \frac{1+\sqrt{1-4lu}}{2}\,\frac{1-\sqrt{1-4lu}}{2} = lu.$$

Hence

$$2\sqrt{lu}\cosh\theta = 1. \tag{26}$$

For any $i > 0$ we have

$$\lambda_+^i - \lambda_-^i = \gamma^i \mathrm{e}^{i\theta} - \gamma^i \mathrm{e}^{-i\theta} = 2\gamma^i \sinh i\theta = 2(lu)^{i/2}\sinh i\theta. \tag{27}$$

Substituting (27) into (11) and (12), we have

$$(T^{-1})_{i,j} = \frac{\sinh i\theta \sinh(n-j+1)\theta}{\sinh\theta \sinh(n+1)\theta}\left(\frac{u}{l}\right)^{(j-i)/2} \qquad i < j \tag{28}$$

$$(T^{-1})_{i,j} = \frac{\sinh j\theta \sinh(n-i+1)\theta}{\sinh\theta \sinh(n+1)\theta}\left(\frac{l}{u}\right)^{(i-j)/2} \qquad i \geqslant j \tag{29}$$

where θ is defined in (26).

Case 3. $lu > \frac{1}{4}$.

If $lu > \frac{1}{4}$, then $1 - 4lu < 0$, λ_+ and λ_- are both complex numbers. If we set

$$\lambda_+ = \frac{1}{2} + \frac{\sqrt{4lu-1}}{2}\mathrm{i} = \gamma \mathrm{e}^{\mathrm{i}\theta} \tag{30}$$

$$\lambda_- = \frac{1}{2} + \frac{\sqrt{4lu-1}}{2}\mathrm{i} = \gamma \mathrm{e}^{-\mathrm{i}\theta} \tag{31}$$

where $\mathrm{i} = \sqrt{-1}$. From (30) and (31) we have

$$2\gamma \cos\theta = 1$$

and

$$\gamma^2 = \left(\frac{1}{2} + \frac{\sqrt{4lu-1}}{2}\mathrm{i}\right)\left(\frac{1}{2} + \frac{\sqrt{4lu-1}}{2}\mathrm{i}\right) = lu.$$

Hence

$$2\sqrt{lu}\cos\theta = 1. \tag{32}$$

For any $i > 0$ we have

$$\lambda_+^i - \lambda_-^i = \gamma^i e^{i\theta} - \gamma^i e^{-i\theta} = 2\gamma^i \sin i\theta = 2(lu)^{i/2} \sin i\theta. \tag{33}$$

Substituting (33) into (11) and (12), we have

$$(T^{-1})_{i,j} = \frac{\sin i\theta \sin(n-j+1)\theta}{\sin\theta \sin(n+1)\theta} \left(\frac{u}{l}\right)^{(j-i)/2} \qquad i < j \tag{34}$$

$$(T^{-1})_{i,j} = \frac{\sin j\theta \sin(n-i+1)\theta}{\sin\theta \sin(n+1)\theta} \left(\frac{l}{u}\right)^{(i-j)/2} \qquad i \geqslant j \tag{35}$$

where θ is defined in (32).

This completes our discussion about general Toeplitz tridiagonal matrix inversion. It is straightforward to show that the result in [1] is a special case of our formula.

4. Block case

We now extend our algorithm to the block tridiagonal case where

$$\mathbf{A} = \begin{bmatrix} B_1 & C_1 & & & & & \\ A_2 & B_2 & C_2 & & & & \\ & \ddots & \ddots & \ddots & & & \\ & & A_j & B_j & C_j & & \\ & & & \ddots & \ddots & \ddots & \\ & & & & A_{n-1} & B_{n-1} & C_{n-1} \\ & & & & & A_n & B_n \end{bmatrix}$$

and A_j, B_j and C_j are $m \times m$ matrices.

Theorem 4.1. Define the second-order block recurrences

$$C_i Z_i = B_i Z_{i-1} - A_i Z_{i-2} \qquad i = 2, 3, \ldots, n \tag{36}$$

where $Z_0 = I$, $Z_1 = C_1^{-1} B_1$ and

$$A_j Y_j = B_j Y_{j+1} - C_j Y_{j+2} \qquad j = n-1, n-2, \ldots, 1 \tag{37}$$

where $Y_{n+1} = I$, $Y_n = A_n^{-1} B_n$. The inverse matrix $\mathbf{A}^{-1} = \{\Phi_{i,j}\}$ $(1 \leqslant i, j \leqslant n)$ can be expressed as

$$\Phi_{j,j} = (B_j - A_j Z_{j-2} Z_{j-1}^{-1} - C_j Y_{j+2} Y_{j+1}^{-1})^{-1} \tag{38}$$

where $j = 1, 2, \ldots, n$, $A_1 = 0$, $C_n = 0$ and

$$\Phi_{i,j} = \begin{cases} -Z_{i-1} Z_i^{-1} \Phi_{i+1,j} & i < j \\ -Y_{i+1} Y_i^{-1} \Phi_{i-1,j} & i > j. \end{cases} \tag{39}$$

Proof. Consider the jth block column of the inverse matrix $\mathbf{A}^{-1}$. We have

$$\begin{bmatrix} B_1 & C_1 & & & & & \\ A_2 & B_2 & C_2 & & & & \\ & \ddots & \ddots & \ddots & & & \\ & & A_j & B_j & C_j & & \\ & & & \ddots & \ddots & \ddots & \\ & & & & A_{n-1} & B_{n-1} & C_{n-1} \\ & & & & & A_n & B_n \end{bmatrix} \begin{bmatrix} \Phi_{1,j} \\ \vdots \\ \Phi_{j-1,j} \\ \Phi_{j,j} \\ \Phi_{j+1,j} \\ \vdots \\ \Phi_{n,j} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ I \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \tag{40}$$

For $i < j$, we use induction on i.

Basis step. When $i = 1$, the first equation of (40) is

$$B_1\Phi_{1,j} + C_1\Phi_{2,j} = 0$$

from which we have

$$\Phi_{1,j} = -B_1^{-1}C_1\Phi_{2,j} = -Z_0 Z_1^{-1}\Phi_{2,j}.$$

Equation (39) holds.

Induction step. We assume equation (39) holds for all i in the interval $0 < i \leqslant k-1 < j-1$, then for $i = k-1$, we have

$$\Phi_{k-1,j} = -Z_{k-2}Z_{k-1}^{-1}\Phi_{k,j}. \qquad (41)$$

Also, from the kth equation of (40), we have

$$A_k\Phi_{k-1,j} + B_k\Phi_{k,j} + C_k\Phi_{k+1,j} = 0.$$

Substituting the $\Phi_{k-1,j}$ in the above equation using (41), we have

$$A_k(-Z_{k-2}Z_{k-1}^{-1})\Phi_{k,j} + B_k\Phi_{k,j} + C_k\Phi_{k+1,j} = 0$$

or

$$\Phi_{k,j} = -(B_k - A_k Z_{k-2}Z_{k-1}^{-1})^{-1}C_k\Phi_{k+1,j}. \qquad (42)$$

From equation (36) we have

$$\begin{aligned} C_k &= B_k Z_{k-1}Z_k^{-1} - A_k Z_{k-2}Z_k^{-1} \\ &= B_k Z_{k-1}Z_k^{-1} - A_k Z_{k-2}Z_{k-1}Z_{k-1}^{-1}Z_k^{-1} \\ &= (B_k - A_k Z_{k-2}Z_{k-1}^{-1})Z_{k-1}Z_k^{-1}. \end{aligned}$$

Thus equation (42) becomes

$$\Phi_{k,j} = -Z_{k-1}Z_k^{-1}\Phi_{k+1,j}.$$

Hence equation (39) holds when $i < j$. Similarly, we can show that when $i > j$, equation (39) also holds.

Now, the only thing left is to determine $\Phi_{j,j}$. From the jth equation of (40), we have

$$A_j\Phi_{j-1,j} + B_j\Phi_{j,j} + C_j\Phi_{j+1,j} = 1. \qquad (43)$$

From equation (39) we have

$$\begin{cases} \Phi_{j-1,j} = -Z_{j-2}Z_{j-1}^{-1}\Phi_{j,j} \\ \Phi_{j+1,j} = -Y_{j+2}Y_{j+1}^{-1}\Phi_{j,j}. \end{cases} \qquad (44)$$

Substituting equation (44) into (43), we obtain

$$A_j(-Z_{j-2}Z_{j-1}^{-1}\Phi_{j,j}) + B_j\Phi_{j,j} + C_j(-Y_{j+2}Y_{j+1}^{-1}\Phi_{j,j}) = 1$$

which is

$$(B_j - A_j Z_{j-2}Z_{j-1}^{-1} - C_j Y_{j+2}Y_{j+1}^{-1})\Phi_{j,j} = I.$$

Thus equation (38) holds. □

Corollary 4.1. The inverse matrix $\mathbf{A}^{-1} = \{\Phi_{i,j}\}$ can be expressed as

$$\Phi_{j,j} = (B_j - A_j Z_{j-2}Z_{j-1}^{-1} - C_j Y_{j+2}Y_{j+1}^{-1})^{-1}$$

where $j = 1, 2, \ldots, n$, $A_1 = 0$, $C_n = 0$ and

$$\Phi_{i,j} = \begin{cases} (-1)^{j-i} Z_{i-1}Z_{j-1}^{-1}\Phi_{j,j} & i < j \\ (-1)^{i-j} Y_{i+1}Y_{j+1}^{-1}\Phi_{j,j} & i > j. \end{cases} \qquad (45)$$

5. An example (block Toeplitz case)

As a straightforward extension of theorem 4.1, we now consider the inversion of a block matrix

$$\mathbf{T} = \begin{bmatrix} B & C & & & & & \\ A & B & C & & & & \\ & \ddots & \ddots & \ddots & & & \\ & & A & B & C & & \\ & & & \ddots & \ddots & \ddots & \\ & & & & A & B & C \\ & & & & & A & B \end{bmatrix}$$

where A, B and C are $m \times m$ matrices.

Let Λ_1 and Λ_2 be $m \times m$ matrices such that

$$\begin{cases} \Lambda_1 + \Lambda_2 = C^{-1}B \\ \Lambda_1\Lambda_2 = C^{-1}A. \end{cases} \tag{46}$$

Let Δ_1 and Δ_2 be $m \times m$ matrices such that

$$\begin{cases} \Delta_1 + \Delta_2 = A^{-1}B \\ \Delta_1\Delta_2 = A^{-1}C. \end{cases} \tag{47}$$

Define sequences Z_i and Y_i ($i = 0, \ldots, n$) to be

$$Z_i = \sum_{k=0}^{i} (\Lambda_2^{i-k}\Lambda_1^k) \tag{48}$$

and

$$Y_i = \sum_{k=0}^{n+1-i} (\Delta_2^{n+1-i-k}\Delta_1^k) \tag{49}$$

then we have the following result.

Theorem 5.1. The inverse of matrix $\mathbf{T}$ can be expressed in the following explicit way:

$$(\mathbf{T}^{-1})_{j,j} = (B - AZ_{j-2}Z_{j-1}^{-1} - CY_{j+2}Y_{j+1}^{-1})^{-1} \tag{50}$$

where $j = 1, 2, \ldots, n$ and

$$(\mathbf{T}^{-1})_{i,j} = \begin{cases} (-1)^{j-i}Z_{i-1}Z_{j-1}^{-1}(T^{-1})_{j,j} & i < j \\ (-1)^{i-j}Y_{i+1}Y_{j+1}^{-1}(T^{-1})_{j,j} & i > j\ . \end{cases} \tag{51}$$

Proof. To prove the theorem, we only need to show the sequences Z_i and Y_i defined in (48) and (49) satisfy the second-order block linear recurrences

$$CZ_i = BZ_{i-1} - AZ_{i-2} \qquad i = 2, 3, \ldots, n \tag{52}$$

where $Z_0 = I$, $Z_1 = C^{-1}B$ and

$$AY_j = BY_{j+1} - CY_{j+2} \qquad j = n-1, n-2, \ldots, 1 \tag{53}$$

where $Y_{n+1} = I$, $Y_n = A^{-1}B$.

From (46) and (52), we have

$$Z_i = (\Lambda_1 + \Lambda_2)Z_{i-1} - \Lambda_1\Lambda_2 Z_{i-2} \tag{54}$$

from which we have

$$\begin{aligned} Z_i - \Lambda_2 Z_{i-1} &= \Lambda_1(Z_{i-1} - \Lambda_2 Z_{i-2}) = \Lambda_1^2(Z_{i-2} - \Lambda_2 Z_{i-3}) \\ &= \cdots = \Lambda_1^{i-1}(Z_1 - \Lambda_2 Z_0) = \Lambda_1^i. \end{aligned}$$

Thus

$$Z_i = \Lambda_2 Z_{i-1} + \Lambda_1^i \qquad i = 1, \ldots, n$$

$$Z_i = \Lambda_2^i + \Lambda_2^{i-1}\Lambda_1 + \cdots + \Lambda_2\Lambda_1^{i-1} + \Lambda_1^i = \sum_{k=0}^{i} \Lambda_2^{i-k}\Lambda_1^k.$$

Similarly, we have

$$Y_i = \Delta_2^{n+1-i} + \Delta_2^{n-i}\Lambda_1 + \cdots + \Delta_2\Delta_1^{n-i} + \Lambda_1^{n+1-i} = \sum_{k=0}^{n+1-i} \Delta_2^{n+1-i-k}\Delta_1^k.$$

□

Now we make some discussions about the matrix. When $A = \varepsilon C$, where ε is a constant factor, we have the following results.

Theorem 5.2. If $A = \varepsilon C$, the inverse of the matrix $\mathbf{T}$ can be determined by the following explicit formulae.

$$(\mathbf{T}^{-1})_{i,j} = (-1)^{i-j}(\Psi_1^j - \Psi_2^j)(\Psi_1^{n+1-i} - \Psi_2^{n+1-i})(\Psi_1 - \Psi_2)^{-1}(\Psi_1^{n+1} - \Psi_2^{n+1})^{-1}C^{-1} \quad i < j \tag{55}$$

$$(\mathbf{T}^{-1})_{i,j} = (-\varepsilon)^{i-j}(\Psi_1^i - \Psi_2^i)(\Psi_1^{n+1-j} - \Psi_2^{n+1-j})(\Psi_1 - \Psi_2)^{-1}(\Psi_1^{n+1} - \Psi_2^{n+1})^{-1}C^{-1} \quad i \geqslant j \tag{56}$$

where Ψ_1 and Ψ_2 are $m \times m$ matrices that satisfy

$$\begin{cases} \Psi_1 + \Psi_2 = C^{-1}B \\ \Psi_1\Psi_2 = \varepsilon I. \end{cases} \tag{57}$$

Proof. Since $A = \varepsilon C$, from (46) and (47) we have $\Lambda_1\Lambda_2 = \varepsilon I$. Thus Λ_1 and Λ_2 commute. Similarly Δ_1 and Δ_2 commute. Also we have $\Lambda_1 + \Lambda_2 = \varepsilon(\Delta_1 + \Delta_2) = C^{-1}B$, and $\Lambda_1\Lambda_2 = \varepsilon^2\Delta_1\Delta_2$, hence

$$\Lambda_1 = \varepsilon\Delta_1 = \Psi_1 \qquad \Lambda_2 = \varepsilon\Delta_2 = \Psi_2.$$

Equation (48) now becomes

$$Z_i = \sum_{k=0}^{i} \Lambda_1^{i-k}\Lambda_2^k = (\Psi_1^{i+1} - \Psi_2^{i+1})(\Psi_1 - \Psi_2)^{-1} \tag{58}$$

and (49) becomes

$$Y_i = \sum_{k=0}^{n+1-i} \Delta_1^{n+1-i-k}\Delta_1^k = \varepsilon^{i-n-1}(\Psi_1^{n+2-i} - \Psi_2^{n+2-i})(\Psi_1 - \Psi_2)^{-1}. \tag{59}$$

Substituting (58) and (59) into (50), we have

$$
\begin{aligned}
(\mathbf{T}^{-1})_{j,j} &= (B - AZ_{j-2}Z_{j-1}^{-1} - CY_{j+2}Y_{j+1}^{-1})^{-1} \\
&= [B - A(\Psi_1^{j-1} - \Psi_2^{j-1})(\Psi_1^{j} - \Psi_2^{j})^{-1} \\
&\quad + \varepsilon C(\Psi_1^{n-j} - \Psi_2^{n-j})(\Psi_1^{n+1-j} - \Psi_2^{n+1-j})^{-1}]^{-1} \\
&= (\Psi_1^{j} - \Psi_2^{j})(\Psi_1^{n+1-j} - \Psi_2^{n+1-j})[B(\Psi_1^{j} - \Psi_2^{j})(\Psi_1^{n+1-j} - \Psi_2^{n+1-j}) \\
&\quad + A(\Psi_1^{j-1} - \Psi_2^{j-1})(\Psi_1^{n+1-j} - \Psi_2^{n+1-j}) + A(\Psi_1^{n-j} - \Psi_2^{n-j})(\Psi_1^{j} - \Psi_2^{j})]^{-1} \\
&= (\Psi_1^{j} - \Psi_2^{j})(\Psi_1^{n+1-j} - \Psi_2^{n+1-j})[B(\Psi_1^{j} - \Psi_2^{j})(\Psi_1^{n+1-j} - \Psi_2^{n+1-j}) \\
&\quad + A(2\Psi_1^{n} + 2\Psi_2^{n} - (\Psi_1 + \Psi_2)(\Psi_1^{n-j}\Psi_2^{j-1} + \Psi_2^{n-j}\Psi_1^{j-1}))]^{-1} \\
&= (\Psi_1^{j} - \Psi_2^{j})(\Psi_1^{n+1-j} - \Psi_2^{n+1-j})[B(\Psi_1^{n+1} + \Psi_2^{n+1}) - A(2\Psi_1^{n} + 2\Psi_2^{n})]^{-1} \\
&= (\Psi_1^{j} - \Psi_2^{j})(\Psi_1^{n+1-j} - \Psi_2^{n+1-j}) \\
&\quad \times [C(\Psi_1 + \Psi_2)(\Psi_1^{n+1} + \Psi_2^{n+1}) - \varepsilon C(2\Psi_1^{n} + 2\Psi_2^{n})]^{-1} \\
&= (\Psi_1^{j} - \Psi_2^{j})(\Psi_1^{n+1-j} - \Psi_2^{n+1-j}) \\
&\quad \times [\Psi_1^{n+2} + \Psi_2^{n+2} - (2\varepsilon\Psi_1^{n} - \Psi_2\Psi_1^{n+1}) - (2\varepsilon\Psi_2^{n} - \Psi_1\Psi_2^{n+1})]^{-1}C^{-1} \\
&= (\Psi_1^{j} - \Psi_2^{j})(\Psi_1^{n+1-j} - \Psi_2^{n+1-j})(\Psi_1 - \Psi_2)^{-1}(\Psi_1^{n+1} - \Psi_2^{n+1})^{-1}C^{-1}. \qquad (60)
\end{aligned}
$$

Using (58), (59) and theorem 5.1, we can obtain equation (51). □

5.1. Discussion

Based on theorem 5.2, we now discuss two special cases.

Case 1. $\varepsilon = 0$.

When $\varepsilon = 0$, $A = \varepsilon C = 0$. (57) becomes

$$
\begin{cases} \Psi_1 + \Psi_2 = C^{-1}B \\ \Psi_1\Psi_2 = 0 \end{cases} \qquad (61)
$$

which has solutions $\Psi_1 = C^{-1}B$, $\Psi_2 = 0$. Hence from theorem 5.2 we have

$$
\begin{aligned}
(\mathbf{T}^{-1})_{i,j} &= (-1)^{i-j}\Psi_1^{j}\Psi_1^{n+1-i}\Psi_1^{-1}\Psi_1^{-(n+1)}C^{-1} \\
&= (-1)^{i-j}(C^{-1}B)^{j-i}B^{-1} \qquad i \leqslant j \qquad (62)
\end{aligned}
$$

$$
(\mathbf{T}^{-1})_{i,j} = 0 \qquad i > j. \qquad (63)
$$

Case 2. $A = C = 0$.

When $A = C = 0$. the inverse matrix is extremely simple,

$$
(\mathbf{T}^{-1})_{i,j} = B^{-1} \qquad i = j \qquad (64)
$$

$$
(\mathbf{T}^{-1})_{i,j} = 0 \qquad i \neq j. \qquad (65)
$$

5.2. Solution of equation (57)

We now find the explicit solution of equation (57) in theorem 5.2. For simplicity, suppose $C^{-1}B$ is nondefective. Let $\lambda_1, \lambda_2, \ldots, \lambda_m$ be the eigenvalues of $C^{-1}B$ and Q be the matrix of the eigenvectors such that

$$
C^{-1}B = Q\,\mathrm{diag}_m(\lambda_k)Q^{-1} \qquad (66)
$$

where $\mathrm{diag}_m(\lambda_k)$ is defined to be an $m \times m$ diagonal matrix with λ_k $(k = 1 \ldots m)$ being the diagonal element. Set the solution of (57) to be

$$\Psi_1 = Q\,\mathrm{diag}(\lambda_k^+)Q^{-1} \tag{67}$$

$$\Psi_2 = Q\,\mathrm{diag}(\lambda_k^-)Q^{-1} \tag{68}$$

then we have

$$\begin{cases} \lambda_k^+ + \lambda_k^- = \lambda_k \\ \lambda_k^+ \lambda_k^- = \varepsilon \end{cases} \tag{69}$$

or in other words, λ_k^+ and λ_k^- are the roots of $r^2 - \lambda_k r + \varepsilon = 0$. Using the same analysis as in section 3, we have the following results.

When $\lambda_k^2 > 4\varepsilon$, we have

$$(\lambda_k^+)^i - (\lambda_k^-)^i = 2\varepsilon^{i/2}\sinh i\theta_k \qquad \text{where } 2\varepsilon^{1/2}\cosh\theta_k = \lambda_k. \tag{70}$$

When $\lambda_k^2 \leqslant 4\varepsilon$, we have

$$(\lambda_k^+)^i - (\lambda_k^-)^i = 2\varepsilon^{i/2}\sin i\theta_k \qquad \text{where } 2\varepsilon^{1/2}\cos\theta_k = \lambda_k. \tag{71}$$

Theorem 5.3. If $A = \varepsilon C$, the inverse of the matrix $\mathbf{T}$ can be determined by the following explicit formulae.

$$(\mathbf{T}^{-1})_{i,j} = (-1)^{j-i}\varepsilon^{(j-i-1)/2} Q\,\mathrm{diag}_m \left\{ \frac{\sinh(j\theta_k)\sinh(n+1-i)\theta_k}{\sinh\theta_k \sinh((n+1)\theta_k)} \right\} Q^{-1}C^{-1} \qquad i < j \tag{72}$$

$$(\mathbf{T}^{-1})_{i,j} = (-1)^{i-j}\varepsilon^{(i-j-1)/2} Q\,\mathrm{diag}_m \left\{ \frac{\sinh(i\theta_k)\sinh(n+1-j)\theta_k}{\sinh\theta_k \sinh(n+1)\theta_k} \right\} Q^{-1}C^{-1} \qquad i \geqslant j \tag{73}$$

where Q and θ_k satisfy

$$C^{-1}B = Q\,\mathrm{diag}_m\{\lambda_k\}Q^{-1} \qquad \text{and} \qquad 2\varepsilon^{1/2}\cosh\theta_k = \lambda_k. \tag{74}$$

If $\lambda_k^2 \leqslant 4\varepsilon$ $(1 \leqslant k \leqslant m)$, the hyperbolic sines and cosines in (72)–(74) become sines and cosines, respectively.

Appendix A. Proof of theorem 2.1

Proof. Consider the jth column of the inverse matrix A^{-1}, we have the equations

$$\begin{bmatrix} b_1 & c_1 & & & & & \\ a_2 & b_2 & c_2 & & & & \\ & \ddots & \ddots & \ddots & & & \\ & & a_j & b_j & c_j & & \\ & & & \ddots & \ddots & \ddots & \\ & & & & a_{n-1} & b_{n-1} & c_{n-1} \\ & & & & & a_n & b_n \end{bmatrix} \begin{bmatrix} \phi_{1,j} \\ \vdots \\ \phi_{j-1,j} \\ \phi_{j,j} \\ \phi_{j+1,j} \\ \vdots \\ \phi_{n,j} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \tag{A1}$$

For $i < j$, we use induction on i.

Basis step. When $i = 1$, the first equation of (A1) is

$$b_1\phi_{1,j} + c_1\phi_{2,j} = 0$$

from which we have

$$\phi_{1,j} = -\frac{c_i \phi_{2,j}}{b_1} = -c_1 \frac{z_0}{z_1} \phi_{2,j}.$$

Equation (4) holds.

Induction step. We assume equation (4) holds for all i in the interval $0 < i \leqslant k-1 < j-1$, then for $i = k-1$, we have

$$\phi_{k-1,j} = -c_{k-1} \frac{z_{k-2}}{z_{k-1}} \phi_{k,j}.$$

Also, from the kth equation of equations (A1), we have

$$a_k \phi_{k-1,j} + b_k \phi_{k,j} + c_k \phi_{k+1,j} = 0. \tag{A2}$$

Substituting $\phi_{k-1,j}$ into equation (A2), we have

$$a_k \left(-c_{k-1} \frac{z_{k-2}}{z_{k-1}}\right) \phi_{k,j} + b_k \phi_{k,j} + c_k \phi_{k+1,j} = 0$$

or

$$\phi_{k,j} = -c_k \frac{1}{b_k - a_k c_{k-1} \frac{z_{k-2}}{z_{k-1}}} \phi_{k+1,j}. \tag{A3}$$

From equation (1), we have

$$\frac{z_k}{z_{k-1}} = b_k - a_k c_{k-1} \frac{z_{k-2}}{z_{k-1}}$$

thus equation (A3) becomes

$$\phi_{k,j} = -c_k \frac{z_{k-1}}{z_k} \phi_{k+1,j}.$$

Hence equation (4) holds when $i < j$. Similarly, we can show that when $i > j$, equation (4) also holds.

Now, the only thing left is to determine $\phi_{j,j}$. From the jth row of equation (A1), we have

$$a_j \phi_{j-1,j} + b_j \phi_{j,j} + c_j \phi_{j+1,j} = 1. \tag{A4}$$

From equation (4), we have

$$\begin{cases} \phi_{j-1,j} = -c_{j-1} \dfrac{z_{j-2}}{z_{j-1}} \phi_{j,j} \\ \phi_{j+1,j} = -a_{j+1} \dfrac{y_{j+2}}{y_{j+1}} \phi_{j,j}. \end{cases} \tag{A5}$$

Substituting equation (A5) into equation (A4), we have

$$a_j \left(-c_{j-1} \frac{z_{j-2}}{z_{j-1}} \phi_{j,j}\right) + b_j \phi_{j,j} + c_j \left(-a_{j+1} \frac{y_{j+2}}{y_{j+1}} \phi_{j,j}\right) = 1$$

which gives

$$\left(b_j - a_j c_{j-1} \frac{z_{j-2}}{z_{j-1}} - a_{j+1} c_j \frac{y_{j+2}}{y_{j+1}}\right) \phi_{j,j} = 1.$$

Thus equation (3) holds. □

Appendix B. Proof of theorem 2.2

Proof. The idea is that since only y_{i+1}/y_i and z_{i-1}/z_i are required in theorem 2.1, we do not need to compute every y_j ($j = n-1, \ldots, 1$) and z_i ($i = 2, \ldots, n$) explicitly. We denote

$$\frac{z_i}{z_{i-1}} = \zeta_i. \tag{B1}$$

Then equation (1) becomes

$$\zeta_i = b_i - a_i c_{i-1}/\zeta_{i-1} \qquad \text{with } i = 2, \ldots, n, \ \zeta_1 = \frac{z_1}{z_0} = b_1. \tag{B2}$$

From (B1), (3) and (4) become

$$\phi_{j,j} = \frac{1}{\zeta_j - a_{j+1}c_j \frac{y_{j+2}}{y_{j+1}}} \tag{B3}$$

where $j = 1, 2, \ldots, n$, $c_n = 0$ and

$$\phi_{i,j} = \begin{cases} -\frac{c_i}{\zeta_i}\phi_{i+1,j} & i < j \\ -a_i \frac{y_{i+1}}{y_i}\phi_{i-1,j} & i > j. \end{cases} \tag{B4}$$

From (2), we have

$$a_{j+1}c_j \frac{y_{j+2}}{y_{j+1}} = b_j - \frac{y_j}{y_{j+1}} \tag{B5}$$

where $j = n-1, n-2, \ldots, 1$, $y_{n+1} = 1$ and $y_n = b_n$. If we define

$$\gamma_j = b_j - \frac{y_j}{y_{j+1}} \tag{B6}$$

where $j = n-1, \ldots, 1$ and $\gamma_n = b_n - y_n/y_{n+1} = 0$, then (B5) becomes

$$\gamma_j = \frac{a_{j+1}c_j}{b_{j+1} - \gamma_{j+1}} \qquad j = n-1, \ldots, 1. \tag{B7}$$

Hence, the inverse of the matrix A can be expressed as

$$\phi_{j,j} = \frac{1}{\zeta_j - \gamma_j} \qquad j = 1, 2, \ldots, n \tag{B8}$$

and

$$\phi_{i,j} = \begin{cases} -\frac{c_i}{\zeta_i}\phi_{i+1,j} & i < j \\ -\frac{a_i}{b_i - \gamma_i}\phi_{i-1,j} & i > j. \end{cases} \tag{B9}$$

Now, the computation of $\{\phi_{i,j}\}$ can be carried out in the following steps.
(i) Compute ζ_i and $\frac{c_i}{\zeta_i}$ ($i = 1, \ldots, n$) using equation (B2).
(ii) Compute γ_j and $\frac{a_j}{b_j - \gamma_j}$ ($j = n, \ldots, 1$) using equation (B7).
(iii) Compute $\phi_{j,j}$ ($j = 1, \ldots, n$) using equation (B8).
(iv) Compute $\phi_{i,j}$ ($i \neq j$) using equation (B9).

Computational cost. In step 1, $3(n-1)$ floating point operations are required to calculate the values of ζ_i ($i = 1, \ldots, n$). Note that during the computation of ζ_i, we already have the values of c_i/ζ_i, which will be used in step 4. Similarly, $3(n-1)$ arithmetic operations are required to obtain γ_j and $a_j/(b_j - \gamma_j)$ $(j = n, \ldots, 1)$. In step 3, all $\phi_{j,j}$ $(j = 1, \ldots, n)$ can be obtained in $2n - 1$ operations. Note that one operation is saved by using $\gamma_n = 0$. Since we have all the values of c_i/ζ_i and $a_i/(b_i - \gamma_i)$ $(i = 1 \sim n)$, step 4 can be completed with only $n^2 - n$ floating point operations. Thus, the total cost of the method is $n^2 + 7n - 7$. $\Box$

References

[1] Hu G Y and O'Connell R F 1996 Analytical inversion of symmetric tridiagonal matrices *J. Phys. A: Math. Gen.* **29** 1511–13

[2] Huang Y and McColl W F 1997 Generalised tridiagonal matrix inversion *Proc. 15th IMACS World Congress on Scientific Computation, Modelling and Applied Mathematics (vol 1: Computational Mathematics)* ed A Sydow (Berlin: Wissenschaft & Technik) pp 487–91

[3] Lewis J W 1982 Inversion of tridiagonal matrices *Num. Math.* **38** 333–45

[4] Meurant G 1992 A review on the inverse of symmetric tridiagonal and block tridiagonal matrices *SIAM J. Matrix Anal. Appl.* **13** 707–28

[5] Usmani R A 1994 Inversion of jacobi's tridiagonal matrix *Comput. Math. Appl.* **27** 59–66

[6] Yamamoto T 1979 Inversion of band matrices *Linear Algebra and Its Applications* **24** 105–11

[7] Yamani H A and Abdelmonem M S 1997 The analytic inversion of any finite symmetric tridiagonal matrix *J. Phys. A: Math. Gen.* **30** 2889–93