Available online at www.sciencedirect.com
SCIENCE DIRECT®
Applied Mathematics Letters 17 (2004) 903–907
Applied Mathematics Letters
www.elsevier.com/locate/aml

A Note on the Stability of Toeplitz Matrix Inversion Formulas

YOU-WEI WEN, MICHAEL K. NG AND WAI-KI CHING
Department of Mathematics, The University of Hong Kong
Pokfulam Road, Hong Kong

HONG LIU
Institute of Geology and Geophysics, Chinese Academy of Sciences
Beijing, P.R. China

(Received February 2003, revised and accepted November 2003)

Abstract—In this paper, we consider the stability of the algorithms emerging from Toeplitz matrix inversion formulas. We show that if the Toeplitz matrix is nonsingular and well-conditioned, then they are numerically forward stable.

Keywords—Toeplitz matrix, Inversion formulas, Stability, Forward stable.

1. INTRODUCTION

In this paper, we consider Toeplitz matrix inversion formulas. Toeplitz systems occur in a variety of applications in applied science and engineering, see [1,2]. In these applications, one would like to construct the inverse of a Toeplitz matrix. By exploiting the structure of Toeplitz matrices T_n, Trench [3] and Gohberg and Semencul [4] used the first and last columns of the inverse of a Toeplitz matrix to reconstruct the whole inverse if $[T_n]_{0,0}^{-1} \neq 0$. In [5], an inversion formula was exhibited for every nonsingular Toeplitz matrix. The method requires the solution of two linear systems of equations (the so-called fundamental equations). In [6], Ben-Artzi and Shalom proved that three columns of the inverse of a Toeplitz matrix, when properly chosen, are always enough to reconstruct the inverse. Labahn and Shalom [7], and Ng, Rost and Wen [8] presented modifications of this result. Others' formulas using circulant type matrices were also presented in literature, see for instance [9]. The main aim of this note is to study the stability of the formulas presented in [6–8]. Our results show that they are numerically forward stable.

2. TOEPLITZ INVERSION FORMULAS

In this section, we review the formulas given in [6–8].

THEOREM 1. *(See [6].) Let $T_n = (t_{p-q})_{p,q=0}^{n-1}$ be a Toeplitz matrix. If each of the systems of equations*

$$T_n\mathbf{x} = \mathbf{e}_0, \quad T_n\mathbf{y} = \mathbf{e}_k, \quad \text{and} \quad T_n\mathbf{z} = \mathbf{e}_{k+1} \tag{1}$$

Research supported by RGC Grant Nos. HKU7130/02P and HKU7046/03P.
Research supported by GD-NSF Grant No. 032475 and CAS Grant No. KZC1-SW-18.

0893-9659/04/$ - see front matter doi:10.1016/j.aml.2003.11.002

are solvable and $x_{n-1-k} \neq 0$ *for an integer* k $(0 \leq k \leq n-1)$, *where*

$$\mathbf{x} = (x_0, x_1, \ldots, x_{n-1})^t, \qquad \mathbf{y} = (y_0, y_1, \ldots, y_{n-1})^t, \qquad \mathbf{z} = (z_0, z_1, \ldots, z_{n-1})^t,$$

then T_n *is nonsingular and*

$$T_n^{-1} = \frac{1}{x_{n-k-1}}(L_1U_1 + L_2U_2), \tag{2}$$

where

$$L_1 = \begin{pmatrix} x_0 & 0 & \cdots & 0 \\ x_1 & x_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ x_{n-1} & \cdots & x_1 & x_0 \end{pmatrix}, \quad U_1 = \begin{pmatrix} y_{n-1} & y_{n-2} - z_{n-1} & \cdots & y_0 - z_1 \\ 0 & y_{n-1} & \ddots & \vdots \\ \vdots & \ddots & \ddots & y_{n-2} - z_{n-1} \\ 0 & \cdots & 0 & y_{n-1} \end{pmatrix},$$

$$L_2 = \begin{pmatrix} z_0 & 0 & \cdots & 0 \\ z_1 - y_0 & z_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ z_{n-1} - y_{n-2} & \cdots & z_1 - y_0 & z_0 \end{pmatrix} \quad \text{and} \quad U_2 = \begin{pmatrix} 0 & x_{n-1} & \cdots & x_1 \\ 0 & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & x_{n-1} \\ 0 & \cdots & 0 & 0 \end{pmatrix}.$$

This result was improved by Labahn and Shalom, [7]. In [8], the displacement equation is used to obtain similar results.

THEOREM 2. *(See [7].) The Toeplitz matrix* $T_n = (t_{p-q})_{p,q=0}^{n-1}$ *is invertible if the systems of equations* $T_n\mathbf{x} = \mathbf{e}_0$ *and* $T_n\mathbf{z} = \mathbf{e}_{n-l+1}$ *are solvable, where* l *is such that* $x_{l-1} \neq 0$ *and* $x_q = 0$, *for all* $q \geq l$. *The inverse of* T_n *is then given by (2), where* $k = n - l$ *and*

$$\mathbf{y} = (Z_n - \mathbf{x}[t_{-1}, t_{-2}, \ldots, t_{1-n}, 0])^{n-l}\mathbf{x}, \tag{3}$$

where $Z_n = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_{n-1}, 0)$.

THEOREM 3. *(See [8].) Let* $T_n = (t_{p-q})_{p,q=0}^{n-1}$ *be a nonsingular Toeplitz matrix and* $J_n = (\mathbf{e}_{n-1}, \mathbf{e}_{n-2}, \ldots, \mathbf{e}_0)$ *be the* $n \times n$ *anti-identity matrix. Let* $\mathbf{x}$, $\mathbf{y}$ *be the solution of the following equations* $T_n\mathbf{x} = \mathbf{e}_0$ *and* $T_n\mathbf{y} = \mathbf{e}_{n-l-1}$ *where* l $(0 < l \leq n-2)$ *is the smallest positive integer such that* $x_l \neq 0$ *and* $x_q = 0$, *for all* $q < l$. *Then the inverse of* T_n *is given by (2), where*

$$\mathbf{z} = \frac{1}{\mu}J_n\left[\left(Z_n^t - \frac{1}{\mu}[t_1, t_2, \ldots, t_{n-1}, 0]^t\mathbf{x}^t Z_n J_n\right)^{l-1}\right]^t J_n Z_n^t \mathbf{x}, \tag{4}$$

$k = n - l - 1$ *and* $\mu = [0, t_{n-1}, t_{n-2}, \ldots, t_1]\mathbf{x}$.

3. STABILITY ANALYSIS

In this section, we show that the Toeplitz inversion formulas presented in Section 2 are evaluation forward stable. An algorithm is called *forward stable* if for all well-conditioned problems, the computed solution $\tilde{\mathbf{x}}$ is close to the true solution $\mathbf{x}$ in the sense that the relative error $||\mathbf{x} - \tilde{\mathbf{x}}||_2/||\mathbf{x}||_2$ is small. In the matrix computation, roundoff errors occur. Let $A, B \in \mathbb{C}^{n,n}$, and $\alpha \in \mathbb{C}$, if we neglect the $O(\varepsilon^2)$ terms, then for any floating-point arithmetic with machine-precision ε, there are

$$\begin{aligned} \mathrm{fl}(\alpha A) &= \alpha A + E, & ||E||_F &\leq \varepsilon|\alpha|\,||A||_2 \leq \varepsilon\sqrt{n}|\alpha|\,||A||_2, \\ \mathrm{fl}(A + B) &= A + B + E, & ||E||_F &\leq \varepsilon||A + B||_2 \leq \varepsilon\sqrt{n}||A + B||_2, \\ \mathrm{fl}(AB) &= AB + E, & ||E||_F &\leq \varepsilon n||A||_F||B||_F, \end{aligned}$$

see [10]. According to the floating-point arithmetic, we have the following bound.

LEMMA 1. *Let $A \in \mathbb{C}^{n,n}$ and k is a positive integer. If we neglect the $O(\varepsilon^2)$ terms, then for any floating-point arithmetic with machine-precision ε, there are*

$$\mathrm{fl}\left(A^k\right) = A^k + E, \qquad ||E||_F \leq kn\varepsilon||A||_F^k. \tag{5}$$

THEOREM 4. *Let T_n be a nonsingular Toeplitz matrix and be well conditioned, then formula (2) presented in Theorems 1–3 is forward stable.*

PROOF. Assume that we have computed the solutions $\tilde{\mathbf{x}}$, $\tilde{\mathbf{y}}$, and $\tilde{\mathbf{z}}$ in (1) which are perturbed by the norm-wise relative errors bounded by $\tilde{\varepsilon}$

$$||\tilde{\mathbf{x}}||_2 \leq ||\mathbf{x}||_2(1+\tilde{\varepsilon}), \qquad ||\tilde{\mathbf{y}}||_2 \leq ||\mathbf{y}||_2(1+\tilde{\varepsilon}), \qquad ||\tilde{\mathbf{z}}||_2 \leq ||\mathbf{z}||_2(1+\tilde{\varepsilon}).$$

Therefore, we have $||L_1||_F \leq \sqrt{n}||\mathbf{x}||_2$, $||L_2||_F \leq \sqrt{n}(||\mathbf{y}||_2 + ||\mathbf{z}||_2)$, $||U_1||_F \leq \sqrt{n}(||\mathbf{y}||_2 + ||\mathbf{z}||_2)$, and $||U_2||_F \leq \sqrt{n}||\mathbf{x}||_2$. Using the perturbed solutions $\tilde{\mathbf{x}}$, $\tilde{\mathbf{y}}$, and $\tilde{\mathbf{z}}$, the inversion formula (2) can be expressed as

$$\begin{aligned}
\tilde{T}_n^{-1} &= \mathrm{fl}\left(\frac{1}{x_{n-k-1}}\left(\tilde{L}_1\tilde{U}_1 + \tilde{L}_2\tilde{U}_2\right)\right) \\
&= \mathrm{fl}\left(\frac{1}{x_{n-k-1}}\left(\left(L_1 + \Delta L_1\right)\left(U_1 + \Delta U_1\right) + \left(L_2 + \Delta L_2\right)\left(U_2 + \Delta U_2\right)\right)\right) \\
&= T_n^{-1} + \frac{1}{x_{n-k-1}}\left(\Delta L_1 U_1 + L_1 \Delta U_1 + \Delta L_2 U_2 + L_2 \Delta U_2 + E + F\right) + G,
\end{aligned} \tag{6}$$

where E is the matrix containing the error which results from computing the matrix products, and F contains the error from subtracting the matrices, and G represents the error of the multiplication by $1/x_{n-k-1}$. For the error matrices ΔL_1, ΔU_1, ΔL_2, and ΔU_2, we have $||\Delta U_2||_F \leq ||\Delta L_1||_F \leq \tilde{\varepsilon}||L_1||_F \leq \tilde{\varepsilon}\sqrt{n}||\mathbf{x}||_2$ and $||\Delta U_1||_F \leq \tilde{\varepsilon}\sqrt{n}(||\mathbf{y}||_2 + ||\mathbf{z}||_2)$ and $||\Delta L_2||_F \leq \tilde{\varepsilon}\sqrt{n}(||\mathbf{y}||_2 + ||\mathbf{z}||_2)$. It follows that

$$\begin{aligned}
&||E||_2 \leq ||E||_F \leq \varepsilon n(||L_1||_F||U_1||_F + ||L_2||_F||U_2||_F) \leq 2\varepsilon n^2||\mathbf{x}||_2(||\mathbf{y}||_2 + ||\mathbf{z}||_2), \\
&||F||_2 \leq \varepsilon\sqrt{n}||T_n^{-1}||_2, \\
&||G||_2 \leq ||G||_F \leq \varepsilon\frac{1}{x_{n-k-1}}(||L_1||_F||U_1||_F + ||L_2||_F||U_2||_F) \leq \frac{2\varepsilon n}{x_{n-k-1}}||\mathbf{x}||_2(||\mathbf{y}||_2 + ||\mathbf{z}||_2).
\end{aligned}$$

Adding all these error bounds, we have

$$\left\|\tilde{T}_n^{-1} - T_n^{-1}\right\|_2 \leq \frac{4\tilde{\varepsilon}n + 2\varepsilon n + 2\varepsilon n^2}{x_{n-k-1}}||\mathbf{x}||_2(||\mathbf{y}||_2 + ||\mathbf{z}||_2) + \frac{\varepsilon\sqrt{n}}{x_{n-k-1}}\left\|T_n^{-1}\right\|_2. \tag{7}$$

Note that $T_n\mathbf{x} = \mathbf{e}_0$, then $||\mathbf{x}||_2 \leq ||T_n^{-1}||_2$, thus, the relative error is

$$\frac{\left\|\tilde{T}_n^{-1} - T_n^{-1}\right\|_2}{\left\|T_n^{-1}\right\|_2} \leq \frac{4\tilde{\varepsilon}n + 2\varepsilon n + 2\varepsilon n^2}{x_{n-k-1}}(||\mathbf{y}||_2 + ||\mathbf{z}||_2) + \frac{\varepsilon\sqrt{n}}{x_{n-k-1}}. \tag{8}$$

As $||\mathbf{y}||_2 \leq ||T_n^{-1}||_2$, $||\mathbf{z}||_2 \leq ||T_n^{-1}||_2$, and T_n are well conditioned, thus, $||\mathbf{y}||_2$ and $||\mathbf{z}||_2$ are finite. The formula presented in Theorem 1 is forward stable.

For Theorem 2, we first let $\mathbf{v} = (t_{-1}, t_{-2}, \ldots, t_{1-n}, 0)^t$. Then, we have

$$\begin{aligned}
\mathrm{fl}(\tilde{\mathbf{y}}) = \mathrm{fl}\left[\left(Z_n - \tilde{\mathbf{x}}\,\mathbf{v}^t\right)^{n-l}\tilde{\mathbf{x}}\right] &= \mathrm{fl}\left\{\mathrm{fl}\left[\left(Z_n - \tilde{\mathbf{x}}\,\mathbf{v}^t\right)^{n-l}\right]\tilde{\mathbf{x}}\right\} \\
&= \mathrm{fl}\left\{\left[\left(Z_n - \mathbf{x}\,\mathbf{v}^t\right)^{n-l} + E\right](\mathbf{x} + \Delta\mathbf{x})\right\} \\
&= \mathbf{y} + E\,\mathbf{x} + \left(Z_n - \mathbf{x}\,\mathbf{v}^t\right)^{n-l}\Delta\mathbf{x} + F,
\end{aligned}$$

where E represents the error results from computing the matrix products $(Z_n - \mathbf{x}\,\mathbf{v}^t)^{n-l}$, F contains the error from the products $(Z_n - \mathbf{x}\,\mathbf{v}^t)^{n-l}$ and $\mathbf{x}$. According to (5), we get

$$||E||_2 \le ||E||_F \le (n-l)n\varepsilon(||Z_n||_F + ||\mathbf{x}||_2||\mathbf{v}||_2)^{n-l} \le (n-l)n\varepsilon\left(\sqrt{n} + ||\mathbf{x}||_2||\mathbf{v}||_2\right)^{n-l}$$

and

$$||F||_2 \le ||F||_F \le n\varepsilon \left\|\left(Z_n - \mathbf{x}\mathbf{v}^t\right)^{n-l}\right\|_2 ||\mathbf{x}||_2 \le n\varepsilon\left(\sqrt{n} + ||\mathbf{x}||_2||\mathbf{v}||_2\right)^{n-l} ||\mathbf{x}||_2.$$

It follows that

$$\begin{aligned}
||\tilde{\mathbf{y}} - \mathbf{y}||_2 &\le \left\|E\mathbf{x} + \left(Z_n - \mathbf{x}\mathbf{v}^t\right)^{n-l}\Delta\mathbf{x} + F\right\|_2 \\
&\le ||E||_2||\mathbf{x}||_2 + \left\|\left(Z_n - \mathbf{x}\mathbf{v}^t\right)^{n-l}||_2\right\|\Delta\mathbf{x}||_2 + ||F||_2 \\
&\le (n-l)n\varepsilon\left(\sqrt{n} + ||\mathbf{x}||_2||\mathbf{v}||_2\right)^{n-l}||\mathbf{x}||_2 + \tilde{\varepsilon}\left(\sqrt{n} + ||\mathbf{x}||_2||\mathbf{v}||_2\right)^{n-l}||\mathbf{x}||_2 \\
&\quad + n\varepsilon\left(\sqrt{n} + ||\mathbf{x}||_2||\mathbf{v}||_2\right)^{n-l}||\mathbf{x}||_2 \\
&= ((n-l+1)n\varepsilon + \tilde{\varepsilon})\left(\sqrt{n} + ||\mathbf{x}||_2||\mathbf{v}||_2\right)^{n-l}||\mathbf{x}||_2.
\end{aligned}$$

It is easy to show that $||\mathbf{v}||_2 \le ||T_n||_2$. Since $T_n\mathbf{x} = \mathbf{e}_0$ and $||\mathbf{x}||_2 \le ||T_n^{-1}||_2$, both $||\mathbf{x}||_2$ and $||\mathbf{v}||_2$ are bounded by the condition number $\kappa(T_n)$ of T_n. We obtain

$$||\tilde{\mathbf{y}} - \mathbf{y}||_2 \le ((n-l+1)n\varepsilon + \tilde{\varepsilon})\left(\sqrt{n} + \kappa(T_n)\right)^{n-l}||\mathbf{x}||_2.$$

For the error matrices ΔU_1 and ΔL_2 in (6), we have

$$||\Delta U_1||_F \le \sqrt{n}((n-l+1)n\varepsilon + \tilde{\varepsilon})\left(\sqrt{n} + \kappa(T_n)\right)^{n-l}||\mathbf{x}||_2 + \tilde{\varepsilon}\sqrt{n}||\mathbf{z}||_2$$

and

$$||\Delta L_2||_F \le \sqrt{n}((n-l+1)n\varepsilon + \tilde{\varepsilon})\left(\sqrt{n} + \kappa(T_n)\right)^{n-l}||\mathbf{x}||_2 + \tilde{\varepsilon}\sqrt{n}||\mathbf{z}||_2.$$

Adding all error bounds in (6), we show that the formula presented in Theorem 2 is also forward stable.

Finally, for Theorem 3, we let

$$\mathbf{w} = \frac{1}{\mu}J_n Z_n^t\mathbf{x} \qquad \text{and} \qquad \mathbf{u} = [t_1, t_2, \ldots, t_{n-1}, 0]^t.$$

Then we have $\mathbf{z} = J_n[(Z_n^t - \mathbf{u}\mathbf{w}^t)^{l-1}]^t\mathbf{w}$ with $\mu = [0, t_{n-1}, t_{n-2}, \ldots, t_1]\mathbf{x}$ and $|\tilde{\mu} - \mu| = |\mathrm{fl}(\mu) - \mu| \le \tilde{\varepsilon}n||\mathbf{u}||_2||\mathbf{x}||_2$. Next, we compute

$$\begin{aligned}
||\mathrm{fl}(\mathbf{w}) - \mathbf{w}||_2 = \left\|\frac{1}{\tilde{\mu}}J_n Z_n^t\tilde{\mathbf{x}} - \frac{1}{\mu}J_n Z_n^t\mathbf{x}\right\|_2 &= \left\|\frac{1}{\mu + \Delta\mu}J_n Z_n^t(\mathbf{x} + \Delta\mathbf{x}) - \frac{1}{\mu}J_n Z_n^t\mathbf{x}\right\|_2 \\
&= \left\|\frac{\Delta\mu}{(\mu + \Delta\mu)\mu}J_n Z_N^t\mathbf{x} + \frac{1}{\mu + \Delta\mu}J_n Z_n^t\Delta\mathbf{x}\right\|_2 \\
&\le \left\|\frac{\Delta\mu}{(\mu + \Delta\mu)\mu}J_n Z_n^t\mathbf{x}\right\|_2 + \left\|\frac{1}{\mu + \Delta\mu}J_n Z_n^t\Delta\mathbf{x}\right\|_2 \\
&\le \frac{\Delta\mu}{(|\mu| - |\Delta\mu|)|\mu|}||\mathbf{x}||_2 + \frac{\tilde{\varepsilon}}{|\mu| - |\Delta\mu|}||\mathbf{x}||_2 \\
&\le \frac{\tilde{\varepsilon}|\mu| + \tilde{\varepsilon}n||\mathbf{u}||_2||\mathbf{x}||_2}{(|\mu| - \tilde{\varepsilon}n||\mathbf{u}||_2||\mathbf{x}||_2)|\mu|}||\mathbf{x}||_2.
\end{aligned}$$

Using the fact that $||\mathbf{w}||_2 = (1/|\mu|)\ ||\mathbf{x}||_2$, $||\mathbf{u}||_2 \le ||T_n||_2$, and $||\mathbf{x}||_2 \le ||T_n^{-1}||_2$, we have

$$\begin{aligned}
||\tilde{\mathbf{z}} - \mathbf{z}||_2 &\le \left(\ln\varepsilon||\mathbf{w}||_2 + \frac{\tilde{\varepsilon}|\mu| + \tilde{\varepsilon}n||\mathbf{u}||\,||\mathbf{x}||_2}{(|\mu| - \tilde{\varepsilon}n||\mathbf{u}||_2||\mathbf{x}||_2)|\mu|}||\mathbf{x}||_2\right)\left(\sqrt{n} + ||\mathbf{u}||_2||\mathbf{w}||_2\right)^{l-1} \\
&\le \left(\frac{\ln\varepsilon}{|\mu|} + \frac{\tilde{\varepsilon}|\mu| + \tilde{\varepsilon}n\kappa(T_n)}{(|\mu| - \tilde{\varepsilon}n\kappa(T_n))|\mu|}\right)\left(\sqrt{n} + \frac{1}{|\mu|}\kappa(T_n)\right)^{l-1}||\mathbf{x}||_2.
\end{aligned}$$

We change the bounds for ΔU_1 and ΔL_2 in (6), the formula presented in Theorem 3 can be shown to be forward stable. ∎

REFERENCES

1. R. Chan and M. Ng, Conjugate gradient methods for Toeplitz systems, *SIAM Review* **38**, 427–482, (1996).
2. W. Ching, *Iterative Methods for Queuing and Manufacturing Systems, Springer Mathematics Monograph*, Springer, London, (2001).
3. W. Trench, An algorithm for the inversion of finite Toeplitz matrices, *J. Soc. Indust. Appl. Math.* **12**, 515–522, (1964).
4. I. Gohberg and A. Semencul, On the inversion of finite Toeplitz matrices and their continuous analogs (in Russian), *Math. Issled.* **7** (2), 201–223, (1972).
5. G. Heinig and K. Rost, Algebraic methods for Toeplitz-like matrices and operators, In *Operator Theory: Advances and Applications, Volume 13*, Birkhäuser, Boston, MA, (1984).
6. A. Ben-Artzi and T. Shalom, On inversion of Toeplitz and close to Toeplitz matrices, *Linear Algebra Appl.* **75**, 173–192, (1986).
7. G. Labahn and T. Shalom, Inversion of Toeplitz matrices with only two standard equations, *Linear Algebra Appl.* **175**, 143–158, (1992).
8. M. Ng, K. Rost and Y. Wen, On inversion of Toeplitz matrices, *Linear Algebra Appl.* **348**, 145–151, (2002).
9. G. Ammar and P. Gader, A variant of the Gohberg-Semencul formulas involving circulant matrices, *SIAM J. Matrix Anal. Appl.* **12**, 534–540, (1991).
10. G. Golub and C. van Loan, *Matrix Computations*, Second Edition, Johns Hopkins U.P., Baltimore, MD, (1989).