ON THE SOLUTION OF LINEAR SYMMETRIC TRIDIAGONAL TOEPLITZ SYSTEM

A.A.A. Moustafa and M.A. Shalaby
Department of Engineering Mathematics and Physics, Faculty of Engineering,
Alexandria University, Alexandria, Egypt.

ABSTRACT

A procedure for solving linear symmetric tridiagonal Toeplitz systems is presented. Without inverting the coefficient matrix, an efficient way to find the solution is obtained by transforming the system into another one which permits the application of Woodbury formula. The proposed method possesses very good stability and is quite competitive with Gaussian elimination and with the modified double sweep method [4].

Keywords: Linear systems, Toeplitz matrices. 1991 *MSC categories*: 15A06, 65F10.

1. INTRODUCTION

Real symmetric Toeplitz matrices arise frequently from many sources and play an important role in many problems in system theory and signal processing [1,2]. The solution of a linear system of equations having a Toeplitz coefficient matrix is a very common problem in practice. Applied to general Toeplitz systems, current fast and superfast solvers are numerically unstable [7]. For some of superfast Toeplitz solvers, stability can be expected for positive definite matrices.

Circulant matrices form a special subclass of Toeplitz matrices. Linear systems of equations involving circulant coefficient matrices appear in many applications [4]. In particular they occur in finite difference solution of one-dimensional elliptic equation subject to periodic boundary conditions [4,5]. For example, the finite difference approximate solution of elliptic equations over a rectangle with periodic boundary conditions yields to the solution of linear system of equations with symmetric circulant tridiagonal coefficient matrix [5]. Symmetric tridiagonal circulant linear systems are always solved using the Gaussian elimination and the modified double sweep method [3,4]. In this paper, we suggest a proposal to solve such a system by transformation into another system applicable to Woodbury formula [6].

Consider the following special class of Toeplitz matrices

$$A = \begin{bmatrix} a_o & a_1 & a_2 & \cdots & a_{n-1} \\ a_{n-1} & a_o & a_1 & \cdots & a_{n-2} \\ a_{n-2} & a_{n-1} & a_o & \cdots & a_{n-3} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_1 & a_2 & a_3 & \cdots & a_o \end{bmatrix} \tag{1.1}$$

which are called circulant matrices. This type of matrices are completely defined by their first row, and thus frequently denoted by

$$A = circ(a_o, a_1, a_2, \ldots, a_{n-1}) \tag{1.2}$$

An $n \times n$ real symmetric circulant tridiagonal matrix S is a matrix of the form

$$S = circ(a_o, a, o, o, \ldots, a)$$

$$= \begin{bmatrix} a_o & a & 0 & \cdots & \cdots & 0 & a \\ a & a_o & a & 0 & \cdots & 0 & 0 \\ 0 & a & a_o & a & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \cdots & \cdots & \cdots & a_o & a \\ a & 0 & 0 & \cdots & \cdots & a & a_o \end{bmatrix}, \quad a \neq 0$$

In this paper, we assume that the matrix S is strictly diagonally dominant, that is

$$|a_o| > 2|a| \qquad (1.3)$$

The following section is devoted to solve the linear system

$$Sx = f \qquad (1.4)$$

where x and f are the unknowns and data vector, respectively.
The algorithmic procedure and stability discussion are given in section 3.

2. THE PROPOSED METHOD

It is easy to see that the system (1.4) is equivalent to the system

$$Ax = b \qquad (2.1)$$

where

$$A = \begin{bmatrix} \alpha & -1 & 0 & \dots & \dots & -1 \\ -1 & \alpha & -1 & 0 & \dots & 0 \\ 0 & -1 & \alpha & -1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \dots & \alpha & -1 \\ -1 & \dots & \dots & \dots & -1 & \alpha \end{bmatrix},$$

$$b = -f/a, \quad \alpha = -a_o/a \qquad (2.2)$$

The condition (1.3) implies that $|\alpha| > 2$. 2.1.

Transformation of the system

An appropriate partition of the system (2.1) to eliminate the last unknown x_n leads to a system which can be solved using the Woodbury formula. The system (2.1) can be partitioned into

$$\begin{bmatrix} \alpha & -1 & 0 & \dots & \dots & 0 \\ -1 & \alpha & -1 & \dots & \dots & 0 \\ 0 & -1 & \alpha & \dots & \dots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \dots & \alpha & -1 \\ 0 & \dots & \dots & \dots & -1 & \alpha \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ \vdots \\ \vdots \\ x_{n-1} \end{bmatrix} - x_n \begin{bmatrix} 1 \\ 0 \\ \vdots \\ \vdots \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ \vdots \\ \vdots \\ b_{n-1} \end{bmatrix}, \qquad (2.3)$$

$$-[1 \ 0 \ \dots \ 0 \ 1] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \end{bmatrix} + \alpha x_n = b_n, \qquad (2.4)$$

Combining (2.3) and (2.4) by premultiplying (2.4) by $\frac{1}{\alpha}[1 \ 0 \ \dots \ 0 \ 1]^T$ and adding the result to (2.3), this yields to

$$\left\{ \begin{bmatrix} \alpha & -1 & & & \\ -1 & \alpha & -1 & \mathbf{0} & \\ & & \ddots & & \\ & \mathbf{0} & & \alpha & -1 \\ & & & -1 & \alpha \end{bmatrix} - \frac{1}{\alpha} \begin{bmatrix} 1 \\ 0 \\ \vdots \\ \vdots \\ 0 \\ 1 \end{bmatrix} [1 \ 0 \ \dots \ 0 \ 1] \right\} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ \vdots \\ x_{n-2} \\ x_{n-1} \end{bmatrix} = \begin{bmatrix} b_1 + b_n/\alpha \\ b_2 \\ \vdots \\ \vdots \\ b_{n-2} \\ b_{n-1} + b_n/\alpha \end{bmatrix}, \qquad (2.5)$$

with $|\alpha| > 2$.

By this partitioning, a symmetric tridiagonal matrix is revealed. To make use of Woodbury formula in the solution, we first change the first element of the tridiagonal symmetric matrix into β and solve the system

$$B\bar{x} = \bar{b} \qquad (2.6)$$

where

$$B = \begin{bmatrix} \beta & -1 & & & \\ -1 & \alpha & -1 & \mathbf{0} & \\ & & & & \\ & \mathbf{0} & & \alpha & -1 \\ & & & -1 & \alpha \end{bmatrix} + \begin{bmatrix} \frac{1}{\beta} & \frac{-1}{\alpha} \\ 0 & 0 \\ \vdots & \vdots \\ \vdots & \vdots \\ 0 & 0 \\ 0 & \frac{-1}{\alpha} \end{bmatrix} \begin{bmatrix} 1 & 0 & \dots & \dots & 0 & 0 \\ 1 & 0 & \dots & \dots & 0 & 1 \end{bmatrix},$$

and

$$\alpha = \beta + \frac{1}{\beta} \quad , \quad \beta = \frac{\alpha}{2} \pm \sqrt{\frac{\alpha^2-4}{4}} \tag{2.7}$$

It should be noticed that both of the vectors $\bar{x}$ and $\bar{b}$ are (n-1)- vectors, that is

$$\bar{x} = [x_1, x_2, \dots, x_{n-1}]^T \text{ and } \bar{b} = [b_1 + b_n/\alpha, b_2, \dots, b_{n-1} + b_n/\alpha]^T. \tag{2.8}$$

The matrix B is of the form

$$B = C + PQ^T \tag{2.9}$$

where

$$C = \begin{bmatrix} \beta & -1 & & & \\ -1 & \alpha & -1 & \mathbf{0} & \\ & & & & \\ & \mathbf{0} & & \alpha & -1 \\ & & & -1 & \alpha \end{bmatrix},$$

$$P = \begin{bmatrix} \frac{1}{\beta} & -\frac{1}{\alpha} \\ \vdots & \vdots \\ \vdots & \vdots \\ 0 & -\frac{1}{\alpha} \end{bmatrix} \text{ and } Q^T = \begin{bmatrix} 1 & 0 & \dots & 0 & 0 \\ 1 & 0 & \dots & 0 & 1 \end{bmatrix}, \tag{2.10}$$

and C, P and Q are (n-1)$\times$(n-1), (n-1)$\times$2 and (n-1)$\times$2 matrices respectively.

Thus applying the Woodbury formula [6],we get

$$B^{-1} = C^{-1} + C^{-1}P\,(I_2 + Q^T C^{-1} P)^{-1}\, Q^T C^{-1} \tag{2.11}$$

and the solution of the system (2.5) is that of the system (2.6), that is

$$\bar{x} = B^{-1}\bar{b} = C^{-1}\bar{b} + C^{-1}P(I_2 + Q^T C^{-1}P)^{-1}\, Q^T C^{-1}\bar{b} \tag{2.12}$$

or

$$\bar{x} = y + DF\,Q^T y \tag{2.13}$$

where

$$y = C^{-1}\bar{b}, D = C^{-1}p \text{ and } F = (I_2 + Q^T C^{-1} P)^{-1} \tag{2.14}$$

and y, D and F are (n-1)$\times$1, (n-1)$\times$2 and 2$\times$2 matrices respectively.

2.2 *Computation of y*

Consider the system

$$C\,y = \bar{b} \tag{2.15}$$

The matrix C is defined in (2.10) and the vector $\bar{b}$ is defined in (2.8). It is easy to see that the matrix C admits the LU factorization, i.e. $C = LU$, where

$$L = \begin{bmatrix} 1 & & & \\ -1/\beta & 1 & \mathbf{0} & \\ & & & \\ & \mathbf{0} & & \\ & & -1/\beta & 1 \end{bmatrix}, U = \begin{bmatrix} \beta & -1 & & \\ & \beta & -1 & \mathbf{0} \\ & & & \\ & \mathbf{0} & & -1 \\ & & & \beta \end{bmatrix} \tag{2.16}$$

and

$$\alpha = \beta + \frac{1}{\beta} \quad , \quad \beta = \frac{\alpha}{2} \pm \sqrt{\frac{\alpha^2-4}{4}}$$

With this factorization, the system (2.15) decomposes into two triangular systems

$$L\, z = \bar{b}\,, \qquad (2.17)$$

$$U\, y = z. \qquad (2.18)$$

where L and U are given in (2.16).

System (2.17) can be solved by forward substitution to obtain

$$z_1 = \bar{b}_1 = b_1 + b_n/\alpha$$

$$z_{i+1} = \bar{b}_{i+1} + z_i/\beta = b_{i+1} + z_i/\beta\,, i=1,2,\ldots,n-3; \qquad (2.19)$$

$$z_{n-1} = \bar{b}_{n-1} + z_{n-2}/\beta = b_{n-1} + b_n/\alpha + z_{n-2}/\beta$$

and system (2.18) by back substitution to get

$$\begin{aligned} y_{n-1} &= z_{n-1}/\beta\,, \\ y_{n-i-1} &= (z_{n-i-1} + y_{n-i})/\beta\,, i=1,2,\ldots,n-2 \end{aligned} \qquad (2.20)$$

From (2.19) and (2.20), the vector $y = C^{-1}\bar{b}$ reveals.

2.3 *Computation of D and F*

To compute the matrices D and F defined in (2.14), we need the first and last column of the matrix C^{-1}. The first column of C^{-1}, can be found by use of the recursion formulae (2.19) and (2.20) with

$$\bar{b}_1 = 1 \text{ and } \bar{b}_2 = \bar{b}_3 = \ldots = \bar{b}_{n-1} = 0.$$

The last column of the matrix C^{-1} can also be computed using the same recursion formulae (2.19) and (2.20) with

$$\bar{b}_1 = \bar{b}_2 = \bar{b}_3 = \ldots = \bar{b}_{n-2} = 0 \text{ and } \bar{b}_{n-1} = 1.$$

Then we get

$$C^{-1} = \begin{bmatrix} \dfrac{\beta^{2(n-1)}-1}{(\beta^2-1)\beta^{2n-3}} & \cdots & \dfrac{1}{\beta^{n-1}} \\ \dfrac{\beta^{2(n-2)}-1}{(\beta^2-1)\beta^{2n-4}} & \cdots & \dfrac{1}{\beta^{n-2}} \\ \vdots & & \\ \dfrac{\beta^{2(n-i)}-1}{(\beta^2-1)\beta^{2(n-1)-i}} & \cdots & \dfrac{1}{\beta^{n-i}} \\ \vdots & & \\ \dfrac{1}{\beta^{n-1}} & \cdots & \dfrac{1}{\beta} \end{bmatrix} \qquad (2.21)$$

Using the equality $\frac{1}{\alpha} = \frac{\beta}{\beta^2+1}$, which comes from (2.7), and from (2.21), we can obtain the 2x2 matrix

$$F = (I_2 + Q^T C^{-1} P)^{-1}.$$

It takes the form

$$F = \frac{1}{(\beta^2-1)(\beta^n-1)} \begin{bmatrix} \beta^{n+2} - 2\beta^n - \beta^2 & \beta^2(\beta^{n-2}+1) \\ -(\beta^2+1)(\beta^{n-2}+1) & (\beta^2+1)(\beta^n+1) \end{bmatrix} \qquad (2.22)$$

The $(n-1)$x2 matrix $D = C^{-1}P$ is, similarly, obtained and takes the form

$$D = \begin{bmatrix} \dfrac{1}{\beta(\beta^2-1)} \dfrac{\beta^{2(n-1)}-1}{\beta^{2n-3}} & \dfrac{-\beta}{\beta^2+1}\left[\dfrac{1}{\beta^{n-1}} + \dfrac{\beta^{2(n-1)}-1}{\beta^{2n-3}(\beta^2-1)}\right] \\ \dfrac{1}{\beta(\beta^2-1)} \dfrac{\beta^{2(n-2)}-1}{\beta^{2n-4}} & \dfrac{-\beta}{\beta^2+1}\left[\dfrac{1}{\beta^{n-2}} + \dfrac{\beta^{2(n-2)}-1}{\beta^{2n-4}(\beta^2-1)}\right] \\ \vdots & \\ \vdots & \\ \dfrac{1}{\beta(\beta^2-1)} \dfrac{\beta^{2(n-i)}-1}{\beta^{2(n-1)-i}} & \dfrac{-\beta}{\beta^2+1}\left[\dfrac{1}{\beta^{n-i}} + \dfrac{\beta^{2(n-i)}-1}{\beta^{2(n-1)-i}(\beta^2-1)}\right] \\ \vdots & \vdots \\ \vdots & \vdots \\ \dfrac{1}{\beta^n} & \dfrac{-\beta}{\beta^2+1}\left[\dfrac{1}{\beta} + \dfrac{1}{\beta^{n-1}}\right] \end{bmatrix} \qquad (2.23)$$

2.4 *Computation of $Q^T y$*

Before writing the final expression for the solution $\bar{x}$ of the system (2.6), we get the vector $Q^T y$ as

$$Q^T y = \begin{bmatrix} y_1 \\ y_1 + y_{n-1} \end{bmatrix} \tag{2.24}$$

3. ALGORITHM FOR SOLVING THE ORIGINAL SYSTEM

Using equations (2.22)-(2.24), we finally get the solution $\bar{x}$ of the system (2.5) as

$$\bar{x}_i = y_i + \frac{\beta^{n-1}}{(\beta^n-1)(\beta^2-1)}\left[\frac{\beta^2-1}{\beta^{n-i}}y_1 + \frac{\beta^{i-2}+\beta^{(n-2)-i}}{\beta^{n-4}}y_{n-1}\right], i=1,2,\ldots,n-1 \tag{3.1}$$

The solution x of the system (2.1) is then

$$\begin{aligned} x_i &= \bar{x}_i \quad, i=1,2,\ldots,n-1 \\ x_n &= (b_n + x_1 + x_{n-1})/\alpha \end{aligned} \tag{3.2}$$

3.1 *Algorithm for the solution of the original system* (1.4)

In the following, an algorithmic procedure to solve the system $Sx=f$ is presented.

Input: $S = circ\ (a_o, a, 0, 0, \ldots, a)$ and f

Procedure:

Step 1: Compute $\alpha = -a_o/a$ and $b_i=-f_i/a$, $i=1,2,\ldots,n$

Step 2: Compute β, where

$$\beta = \frac{\alpha}{2} + \sqrt{\frac{\alpha^2-4}{4}} \quad \text{if } \alpha > 0$$

$$\beta = \frac{\alpha}{2} - \sqrt{\frac{\alpha^2-4}{4}} \quad \text{if } \alpha < 0$$

Step 3: Compute

$z_1=b_1+b_n/\alpha$,

$z_{i+1}=b_{i+1}+z_i/\beta$, $i=1,2,\ldots,n$-3

$z_{n-1}=b_{n-1}+b_n/\alpha+z_{n-2}/\beta$.

Step 4: Compute

$y_{n-1}=z_{n-1}/\beta$,

$y_{n-i-1}=(z_{n-i-1}+y_{n-i})/\beta$.

Step 5: Compute

$$x_i = y_i + \frac{\beta^{n-1}}{(\beta^n-1)(\beta^2-1)}\left[\frac{\beta^2-1}{\beta^{n-i}}y_1 + \frac{\beta^{i-2}+\beta^{(n-2)-i}}{\beta^{n-4}}y_n\right.$$

and

$$x_n = (b_n+x_1+x_{n-1})/\alpha$$

Output

$$x_i\ ,\ i=1,2,\ldots,n$$

3.2 *Stability Discussion*

From (2.19) and (2.20), it is easy to see that

$$\beta = \frac{\alpha}{2} + \sqrt{\frac{\alpha^2-4}{4}} \quad \text{if } \alpha > 0, \text{ and}$$

$$\beta = \frac{\alpha}{2} - \sqrt{\frac{\alpha^2-4}{4}} \quad \text{if } \alpha < 0$$

because we then have $|\beta| > 1$ and numerical stability of the process of solving equation (2.7) follows. Also, we can observe that the power of $1/\beta$ should be computed recursively for increasing values of i to assure numerical stability.

Finally, only a few words of memory are needed for the vectors b, y, and z because each element of these vectors is used once and needs not to be saved.

REFERENCES

[1] R.E. Blahut, *Fast Algorithms for Digital Signal Processing*, Addison-Wesly, Reading, MA., 1984.

[2] J.R. Bunch, "Stability of methods for solving Toeplitz systems of equations", *SIAM J. Sci. Statist. Comput.*, vol. 6, pp. 349-364, 1985.

[3] C. Mingkui, "Modified double sweep method for solving tridiagonal systems of linear equations", *J. Xi'an Jiaotong Univ.*, China, vol. 16, pp. 85-94, 1982.

[4] C. Mingkui, On the solution of circulant linear systems", *SIAM J. Numer. Analysis*, vol. 24, pp. 668-683, 1987.

[5] D. Fischer *et. al.*, "On Fourier-Toeplitz methods for separable elliptic problems", *Math. Comput.*, vol. 28, pp. 349-368, 1974.

[6] A.S. Householder, *The Theory of Matrices in Numerical Analysis*, Blaisdell, N.Y., 1964.

[7] M.A. Shalaby, "Minimal eigenvalue of real symmetric positive definite Toeplitz matrix", *J. Applied Maths. & Computation*, vol. 64, pp. 93-99, 1994.