A Method for Solving Hermitian Pentadiagonal Block Circulant Systems of Linear Equations

Borislav V. Minchev^a, Ivan G. Ivanov^b

^aDepartment of Computer Science, University of Bergen, Thormhlensgate 55, N-5020 Bergen, Norway

^bFaculty of Economics and Business Administration, Sofia University, Sofia 1113, Bulgaria

Abstract

A new effective method and its two modifications for solving Hermitian pentadiagonal block circulant systems of linear equations are proposed. New algorithms based on the proposed method are constructed. Our algorithms are then compared with some classical techniques as far as implementation time is concerned, number of operations and storage. Numerical experiments corroborating the effectiveness of the proposed algorithms are also reported.

Key words: linear system, block circulant matrix, matrix equation, Woodbury's formula 1991 MSC: 65F10

1 Introduction

Linear systems of equations having circulant coefficient matrices appear in many applications. For example, in finite difference approximations to elliptic equations subject to periodic boundary conditions [2,8] and in approximations of periodic functions using splines [1,9]. In case when multidimensional problems are concerned the coefficient matrices of the resulting linear systems are with block circulant structure [7].

Published in LSSC, LNCS 2907, Springer, 481-488, 2004

Email addresses: Borko.Minchev@ii.uib.no (Borislav V. Minchev),

i_ivanov@feb.uni-sofia.bg (Ivan G. Ivanov).

URL: http://www.ii.uib.no/~borko (Borislav V. Minchev).

In this paper we propose a new method and its two modifications for solving Hermitian pentadiagonal block circulant systems of linear equations. It is known that these systems have the form

$$W x = f, (1)$$

where

$$W = \begin{pmatrix} M & N & S & S^* & N^* \\ N^* & M & N & S & S^* \\ S^* & N^* & \ddots & 0 & & \\ S^* & \ddots & \ddots & & & \\ & S^* & \ddots & \ddots & & \\ & 0 & \ddots & \ddots & N & S \\ S & & S^* & N^* & M & N \\ N & S & & S^* & N^* & M \end{pmatrix}$$
(2)

is Hermitian pentadiagonal block circulant matrix with block size n. M, N and S are $m \times m$ matrices, $x = \{x_i\}_{i=1,\dots,n}, f = \{f_i\}_{i=1,\dots,n}$, are column vectors with block size n, x_i and f_i , are blocks with size $m \times 1$.

Our goal is to construct a new effective method for solving (1) and then to compare it with some classical techniques.

The paper is organized as follows: in Section 2 we present the new method and discuss its two modifications based on different applications of the Woodbury's formula [4]; in Section 3 we report some numerical experiments corroborating the effectiveness of the proposed algorithms.

2 A Modification of LU factorizations

Adapting the ideas suggested in [6], we construct a new method for solving linear systems with coefficient matrices of the form (2). Our approach is based on the solution of a special nonlinear matrix equation. One can fine the solution of (1) using the following steps:

Step 1. Solve the parametric linear system

$$T y = f, (3)$$

where

$$T = \begin{pmatrix} X & Y & S & & \\ Y^* & Z & N & & 0 & \\ S^* & N^* & M & . & . & \\ & & \ddots & \ddots & S & \\ 0 & & \ddots & \ddots & N & \\ & & & S^* & N^* & M \end{pmatrix}$$

is pentadiagonal matrix with block size n. It has a block Teoplitz structure except for the north-western corner, $y = \{y_i\}_{i=1,\dots,n}$, and $f = \{f_i\}_{i=1,\dots,n}$ are column vectors with blocks size n, y_i and f_i are blocks with size $m \times 1$.

The matrix T admits the following LU factorization

$$T = LU = \begin{pmatrix} I_m & & \\ Y^*X^{-1} & 0 & \\ S^*X^{-1} & & \\ & \ddots & & \\ 0 & S^*X^{-1} & Y^*X^{-1} & I_m \end{pmatrix} \begin{pmatrix} X & Y & S & 0 \\ & \ddots & & \\ & \ddots & S \\ 0 & & Y \\ & & X \end{pmatrix},$$

where I_m is the identity matrix with size $m \times m$. The above decomposition exists when the parameters $X = X^*$, Y and $Z = Z^*$ satisfy the relations

$$\begin{aligned}
Z &= Y^* X^{-1} Y + X \\
N &= Y^* X^{-1} S + Y \\
M &= S^* X^{-1} S + Z
\end{aligned}$$
(4)

Let us introduce the following notations

$$F = \begin{pmatrix} X & Y \\ Y^* & Z \end{pmatrix}, \quad Q = \begin{pmatrix} S & 0 \\ N & S \end{pmatrix}, \quad R = \begin{pmatrix} M & N \\ N^* & M \end{pmatrix}.$$
(5)

If F is a positive definite solution of the matrix equation

$$F + Q^* F^{-1} Q = R \tag{6}$$

and $X = X^* > 0$, $Z = Z^* > 0$ then the blocks X, Y and Z satisfy the system (4).

Thus, solving the linear system (3) is equivalent to solve two simpler systems

$$L z = f, \quad z = \{z_i\}_{i=1,\dots,n}$$

 $U y = z, \quad y = \{y_i\}_{i=1,\dots,n}.$

Step 2. Solve the pentadiagonal block Teoplitz linear system

$$P \ u = f, \tag{7}$$

where

is Hermitian pentadiagonal block Teoplitz matrix with block size n, u = $\{u_i\}_{i=1,\dots,n}$ and $f = \{f_i\}_{i=1,\dots,n}$ are column vectors with block size n, u_i and f_i are blocks with size $m \times 1$.

The matrices T and P satisfy the relation $P = T + J_2 \hat{V}$, where

$$J_{2} = \begin{pmatrix} I_{m} \ 0 \ 0 \ \dots \ 0 \\ 0 \ I_{m} \ 0 \ \dots \ 0 \end{pmatrix}^{T}, \quad \hat{V} = \begin{pmatrix} M - X \ N - Y \ 0 \ \dots \ 0 \\ N^{*} - Y^{*} \ M - Z \ 0 \ \dots \ 0 \end{pmatrix}$$

are matrices with block size $n \times 2$ and $2 \times n$ respectively. Using the Woodbury's formula we have

$$P^{-1} = T^{-1} - T^{-1} J_2 \left[I_{2m} + \hat{V} T^{-1} J_2 \right]^{-1} \hat{V} T^{-1}, \tag{9}$$

where I_{2m} is the identity matrix with size $2m \times 2m$. Therefore, the solution uof (7) is given by

$$u = P^{-1}f = y - T^{-1}J_2 \left[I_{2m} + \hat{V}T^{-1}J_2 \right]^{-1} \hat{V}y.$$

One can find the matrix $T^{-1}J_2$ by solving 2m linear systems of type (3) with right-hand sides the corresponding two different columns of J_2 . This approach does not take into account the very sparse nonzero structure of J_2 . For real $M, N \text{ and } S \text{ it costs } O(20nm^3) \text{ flops and needs to store } 2nm^2 \text{ real numbers.}$ In order to decrease the number of operations needed to compute $T^{-1}J_2$, we consider a new approach which is motivated by the ideas suggested in [5]. Le

t us denote the block columns vectors of
$$J_2$$
 with E_1 and E_2 respectively i.e

$$E_1 = (I_m \ 0 \ 0 \ \dots \ 0)^T, \quad E_2 = (0 \ I_m \ 0 \ \dots \ 0)^T$$

Put $A = Y^* X^{-1}$ and $B = S^* X^{-1}$.

The matrix T admits the following decomposition

$$T = LDL^*$$
, where $L = \begin{pmatrix} I_m & & \\ A & & 0 \\ B & & \\ & \ddots & \\ 0 & B & A & I_m \end{pmatrix}$

is a square matrix of block size n and $D = \text{diag}(X, \ldots, X)$.

Let $(L^{-1})_{ii}$ be the blocks of the matrix L^{-1} . We have

$$(L^{-1})_{ij} = \begin{cases} 0 & i < j \\ Z_{i-j+1} & i \ge j, \end{cases} \text{ where } \begin{array}{c} Z_1 = I_m, \quad Z_2 = -A, \quad Z_3 = A^2 - B, \\ Z_j = -AZ_{j-1} - BZ_{j-2} & for \ j = 4 \dots n. \end{cases}$$

Obviously $D^{-1} = \text{diag}(X^{-1}, \dots, X^{-1}).$

We propose to compute $T^{-1}J_2$ by consecutive calculations of $T^{-1}E_1$ and $T^{-1}E_2$ using the following algorithm:

Algorithm RP Recursive computations for Pentadiagonal system

• Find the cells $K_i = X^{-1}Z_i$ for i = 1, ..., n by the formulas

$$K_{1} = X^{-1}$$

$$K_{2} = -X^{-1}A$$

$$K_{i} = -K_{i-1}A - K_{i-2}B \qquad for \ i = 3, \dots n.$$

• Compute the blocks $(T^{-1}E_1)_i$ and $(T^{-1}E_2)_i$ by the formulas

$$(T^{-1}E_{1})_{n} = K_{n}$$

$$(T^{-1}E_{1})_{n-1} = K_{n-1} - A^{*}(T^{-1}E_{1})_{n}$$

$$(T^{-1}E_{1})_{i} = K_{i} - A^{*}(T^{-1}E_{1})_{i+1} - B^{*}(T^{-1}E_{1})_{i+2} \quad for \ i = n-2, \dots, 2$$

$$(T^{-1}E_{1})_{1} = X^{-1} - A^{*}(T^{-1}E_{1})_{2} - B^{*}(T^{-1}E_{1})_{3}$$

$$(T^{-1}E_{2})_{n} = K_{n-1}$$

$$(T^{-1}E_{2})_{n-1} = K_{n-2} - A^{*}(T^{-1}E_{2})_{n}$$

$$(T^{-1}E_{2})_{i} = K_{i-1} - A^{*}(T^{-1}E_{2})_{i+1} - B^{*}(T^{-1}E_{2})_{i+2} \quad for \ i = n-2, \dots, 2$$

$$(T^{-1}E_{2})_{1} = -A^{*}(T^{-1}E_{2})_{2} - B^{*}(T^{-1}E_{2})_{3}.$$

If the blocks M, N and S are real the algorithm RP costs $O(12nm^3)$ flops and needs to store $(3n+2)m^2$ real numbers. According to the above algorithm, in the next step, we consider two different approaches for solving (1).

Step 3. Solve the system (1)

3.1 The matrix W satisfies the relation

$$W = P + \tilde{U}\tilde{V},$$

where

$$\tilde{U} = \begin{pmatrix} I_m \ 0 & 0 & 0 \\ 0 & I_m & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & I_m & 0 \\ 0 & 0 & 0 & I_m \end{pmatrix}, \quad \tilde{V} = \begin{pmatrix} 0 & 0 & \dots & S^* & N^* \\ 0 & 0 & \dots & 0 & S^* \\ S & 0 & \dots & 0 & 0 \\ N & S & \dots & 0 & 0 \end{pmatrix}$$

are matrices with block size $n \times 4$ and $4 \times n$ respectively. Using the Woodbury's formula we have

$$W^{-1} = P^{-1} - P^{-1} \tilde{U} \left[I_{4m} + \tilde{V} P^{-1} \tilde{U} \right]^{-1} \tilde{V} P^{-1},$$

where I_{4m} is the identity matrix of size $4m \times 4m$.

The solution x of (1) is obtained from the vector u by the formula

$$x = W^{-1}f = u - P^{-1}\tilde{U}\left[I_{4m} + \tilde{V}P^{-1}\tilde{U}\right]^{-1}\tilde{V}u.$$

Denote the block columns vectors of \tilde{U} by E_1 , E_2 , E_{n-1} and E_n respectively. Thus, the computation of $P^{-1}\tilde{U}$ can be done by consecutive calculations of $P^{-1}E_1$, $P^{-1}E_2$, $P^{-1}E_{n-1}$ and $P^{-1}E_n$ using the formula (9). For i = 1, 2, n-1, n we have

$$P^{-1}E_i = T^{-1}E_i - T^{-1}J_2 \left[I_{2m} + \hat{V}T^{-1}J_2 \right]^{-1} \hat{V}T^{-1}E_i.$$
(10)

Note that the numerical implementation of formulas (10) is very "cheap", since we already know from Step 2 the elements $T^{-1}J_2$ and $\left[I_{2m} + \hat{V}T^{-1}J_2\right]^{-1}$. We recommend formulas (10) instead of solving 4m linear system of the form (7) with right hand side the corresponding column vectors of \tilde{U} . It is easy to observe that the blocks of $T^{-1}E_{n-1}$ and $T^{-1}E_n$ satisfy the relations

$$(T^{-1}E_{n-1})_n = K_2^*,$$

$$(T^{-1}E_{n-1})_i = K_{n-i}^* - K_{n-i+1}^*A \quad for \ i = n-1, \dots, 1,$$

$$(T^{-1}E_n)_i = K_{n+1-i}^* \quad for \ i = n, \dots, 1,$$

where K_i for i = 1, ..., n are the blocks from algorithm RP.

3.2 In order to decrease the size of the inverse matrix in the Woodbury's formula, we propose the following decomposition of the matrix W

$$W = \begin{pmatrix} P & V \\ V^* & R \end{pmatrix},$$

where P is from (8), with block size $n - 2 \times n - 2$, R is from (5) and

$$V^* = \begin{pmatrix} S & 0 & \dots & S^* & N^* \\ N & S & \dots & 0 & S^* \end{pmatrix}$$

is a matrix with block size $2 \times n - 2$. Put

$$\hat{x} = (x_1, \dots, x_{n-2})^T, \quad \tilde{x} = (x_{n-1} x_n)^T, \quad x = \begin{pmatrix} \hat{x} \\ \tilde{x} \end{pmatrix},$$
$$\hat{f} = (f_1, \dots, f_{n-2})^T, \quad \tilde{f} = (f_{n-1} f_n)^T, \quad f = \begin{pmatrix} \hat{f} \\ \tilde{f} \end{pmatrix}.$$

In this notations system (1) can be written in the form

$$\begin{pmatrix} P & V \\ V^* & R \end{pmatrix} \begin{pmatrix} \hat{x} \\ \tilde{x} \end{pmatrix} = \begin{pmatrix} \hat{f} \\ \tilde{f} \end{pmatrix},$$

which is equivalent to

$$\begin{vmatrix} G\hat{x} &= r \\ \tilde{x} &= R^{-1} \left(\tilde{f} - V^* \hat{x} \right), \end{vmatrix}$$

where $G = P - VR^{-1}V^*$, $r = \hat{f} - VR^{-1}\tilde{f}$. By Woodbury's formula we have

$$G^{-1} = P^{-1} + P^{-1}V \left[R - V^* P^{-1}V \right]^{-1} V^* P^{-1}.$$

Therefore,

$$\hat{x} = G^{-1}r = z + P^{-1}V \left[R - V^*P^{-1}V\right]^{-1}V^*z,$$

where $z = P^{-1}r$ can be computed by means of Step 2. Denote the block columns vectors of V with H_1 and H_2 respectively. The computation of $P^{-1}V$ can also be done by consecutive calculations of $P^{-1}H_1$ and $P^{-1}H_2$ using the formula (9). For i = 1, 2

$$P^{-1}H_i = T^{-1}H_i - T^{-1}J_2 \left[I_{2m} + \hat{V}T^{-1}J_2 \right]^{-1} \hat{V}T^{-1}H_i.$$

The numerical implementation of the last formulas is again "cheap", since we already know from Step 2 the elements $T^{-1}J_2$ and $\left[I_{2m} + \hat{V}T^{-1}J_2\right]^{-1}$. The blocks of $T^{-1}H_1$ and $T^{-1}H_2$ satisfy the relations

$$(T^{-1}H_1)_i = (T^{-1}E_1)_i S^* + K^*_{n-2-i}S + K^*_{n-i-1}\tilde{Q} \quad \text{for } i = 1, \dots, n-3$$

$$(T^{-1}H_1)_{n-2} = (T^{-1}E_1)_{n-2}S^* + K^*_1\tilde{Q}$$

$$(T^{-1}H_2)_i = (T^{-1}E_1)_i N^* + (T^{-1}E_2)_i S^* + K^*_{n-1-i}S \quad \text{for } i = 1, \dots, n-2$$

where $\tilde{Q} = N - AS$ and K_i for i = 1, ..., n are the blocks from Algorithm RP.

3 Numerical experiments

In this section we compare our algorithms with some classical techniques for solving (1), with W given as in (2), and the exact solution $x = (1, 1, ..., 1)^T$.

In our numerical experiments, W is Hermitian pentadiagonal block circulant with several block sizes n. The algorithms are compared by means of execution time and accuracy of the solution.

Table 1Execution time (in seconds) and errors for Example 1

m = 3									
Algorithm	n = 4000		n = 6000		n = 8000				
	Err.	time	Err.	time	Err.	time			
LU	1.4482e-015	2.95	1.4482e-015	6.04	1.4482e-015	10.15			
CHOL	2.2888e-015	2.00	2.2888e-015	3.23	2.2888e-015	5.44			
$M_{RP}(4m)$	4.2635 e-014	1.95	4.1064 e-014	3.14	4.7126e-014	5.02			
$M_RP(2m)$	3.3956e-014	1.63	4.1081e-014	2.62	6.0280e-014	4.34			

The codes are written in MATLAB language and the computations are done on an AMD computer. The results of the experiments are given in different tables for each example.

The following notations are used: LU stands for classical LU factorization; CHOL stands for the classical Cholesky factorization; M_RP(4m) stands for algorithm based on the proposed new method using Step 3.1; M_RP(2m) stands for algorithm based on the proposed new method using Step 3.2; $Err. = ||x - \tilde{x}||_{\infty}$, where \tilde{x} is the computed solution.

To solve the system (1) we need to compute a positive definite solution of the matrix equation (6). The sufficient condition for the existence of a positive definite solution is $||R^{-\frac{1}{2}}QR^{-\frac{1}{2}}|| \leq \frac{1}{2}$, (see [3]). In the next two examples the cells of the matrix W, which form the matrices R and Q, are chosen to satisfy this condition.

Example 1 Let

$$M = \begin{pmatrix} 8 & 1-i & 1.5\\ 1+i & 9 & 1\\ 1.5 & 1 & 8 \end{pmatrix}, \quad N = \begin{pmatrix} 0 & 1 & 0\\ 0 & 2 & 0\\ 1-i & 0 & 0 \end{pmatrix},$$
$$S = \begin{pmatrix} 1.2-3i & -0.3-i & 0.1\\ -0.30 & 2.1 & 0.2\\ 0.1 & 0.2 & 0.65+2i \end{pmatrix}.$$

In Table 1, we present the execution time (in seconds) and the error for each algorithm for different values of n.

Example 2 Let

$$M = circ(22, -8, 1, ..., 1, -8), \quad N = circ(-7.2, 1.8, ..., 1.8)$$

are circulant matrices and S = I.

The results from the numerical experiments for this example are given in Table 2.

Algorithm	n = 4000		n = 6000		n = 8000	
	Err.	time	Err.	time	Err.	time
LU	1.7764e-015	3.49	1.7764e-015	6.66	1.7764e-015	11.01
CHOL	1.9984 e-015	2.39	1.9984 e-015	3.90	1.9984 e-015	5.98
$M_{RP}(4m)$	2.6182 e-014	2.14	3.2072e-014	3.40	3.7038e-014	4.88
$M_{RP}(2m)$	2.5537 e-014	1.87	3.1282e-014	2.83	3.6124 e-014	4.47

Table 2Execution time (in seconds) and errors for Example 2

4 Conclusions

The proposed new algorithms $M_RP(2m)$ and $M_RP(4m)$ are faster than the classical LU and CHOL. From the theoretical discussions and the numerical experiments, we can conclude that Algorithm $M_RP(2m)$ is most suitable for implementation. This is due to the size of the inverse matrix in the Woodbury's formula. The inverse matrix in $M_RP(2m)$ is two times smaller than the inverse matrix in $M_RP(4m)$. This leads to a considerable decrease in the execution time.

The complexity of the proposed new algorithms is $O(nm^3)$. For comparison, algorithm based on the Fast Fourier Transform (FFT) has complexity $O(nm \log(n))$, but it can be implemented only when the block size of the matrix W is power of two. Our method does not have these restrictions. The only restriction on the applicability of our method is related with the existence of a solution of the matrix equation (6).

References

- J. H. Ahlberg, E. N. Nilson, J. L. Walsh, The Theory of Splines and Their Applications, Academic Press, New York, 1967.
- [2] R. Chan, T. Chan, Circulant preconditioners for elliptic problems, sJ. Numer. Linear Algebra Appl. 1, (1992), 77-101.
- [3] J. C. Engwerda, On the Existence of a Positive Definite Solution of the Matrix Equation $X + A^T X^{-1} A = I$, Linear Algebra Appl., **194**, (1993), 91-108.
- [4] G. Golub, C. Van Loan, *Matrix Computation*, The John Hopkins University Press, Baltimore, 1989.
- [5] B. Minchev, Some Algorithms for Solving Special Tridiagonal Block Teoplitz Linear Systems, J. Comp. and Appl. Math. vol. 156/1, (2003), 179-200.

- [6] B. Mintchev, I. Ivanov, On the Computer Realization of Algorithms for Solving Special Block Linear Systems, Applications of Mathematics in Engineering and Economics'27, Heron Press, Sofia (2002), 303-313.
- [7] C. Mingkui, On the Solution of circulant linear systems, SIAM J. Numer. Analysis 24, (1987), 668-683.
- [8] W. L. Wood, Periodicity effects on the iterative solution of elliptic difference equations, *SIAM J. Numer. Analysis* 8, (1971), 439-464.
- [9] U. Zavyalov, B. Kvasov, V. Miroshnichenko, Spline Functions Methods (in Russian), Moscow, 1980.