DOI 10.1007/s11253-017-1340-8
Ukrainian Mathematical Journal, Vol. 68, No. 12, May, 2017 (Ukrainian Original Vol. 68, No. 12, December, 2016)

PERTURBATION AND ERROR ANALYSES OF THE PARTITIONED *LU* FACTORIZATION FOR BLOCK TRIDIAGONAL LINEAR SYSTEMS

C.-Y. Wu[1] **and T.-Zh. Huang**[2]

UDC 519.62

We present the perturbation and backward error analyses of the partitioned LU factorization for block tridiagonal matrices. In addition, we consider the bounds of perturbations for the partitioned LU factorization for block-tridiagonal linear systems. Finally, numerical examples are given to verify the obtained results.

1. Introduction

We consider a linear system $Ax = b$, where A is a nonsingular block tridiagonal matrix of the form:

$$A = \begin{pmatrix} A_1 & C_1 & & & \\ B_2 & A_2 & C_2 & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & C_{s-1} \\ & & & B_s & A_s \end{pmatrix}, \tag{1.1}$$

where

$$A_i \in \mathbb{R}^{k_i \times k_i}, \qquad B_i \in \mathbb{R}^{k_i \times k_{i-1}}, \qquad \text{and} \qquad C_i \in \mathbb{R}^{k_i \times k_{i+1}}$$

for all $1 \leq i \leq s$.

Our aim is to solve the linear system $Ax = b$ efficiently and accurately for the indicated nonsingular block tridiagonal matrix. Applying the partitioned LU factorization to the general matrix, we get the following representation of the partitioned LU factorization for nonsingular block tridiagonal matrices:

$$A = \begin{pmatrix} L_{11} & & & \\ B_2 U_{11}^{-1} & I_2 & & \\ & & \ddots & \\ & & & I_s \end{pmatrix} \begin{pmatrix} I_1 & \\ & S_1 \end{pmatrix} \begin{pmatrix} U_{11} & L_{11}^{-1} C_1 & & \\ & I_2 & & \\ & & \ddots & \\ & & & I_s \end{pmatrix} = L_1 D_1 U_1,$$

[1] Jinan University and School of Mathematical Sciences, University of Electronic Science and Technology, Guangzhou, China.
[2] School of Mathematical Sciences, University of Electronic Science and Technology, Chengdu, China.

Published in Ukrains'kyi Matematychnyi Zhurnal, Vol. 68, No. 12, pp. 1683–1695, December, 2016. Original article submitted April 25, 2013.

0041-5995/17/6812–1949

where

$$S_1 = \begin{pmatrix} A_2 - B_2 U_{11}^{-1} L_{11}^{-1} C_1 & C_2 & & & \\ B_3 & A_3 & C_3 & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & C_{s-1} \\ & & & B_s & A_s \end{pmatrix}.$$

If $A_2 - B_2 U_{11}^{-1} L_{11}^{-1}$ can be factorized as follows:

$$A_2 - B_2 U_{11}^{-1} L_{11}^{-1} = L_{22} U_{22},$$

then D_1 satisfies the relation

$$D_1 = L_2 D_2 U_2 = \begin{pmatrix} I_1 & & & & \\ & L_{22} & & & \\ & B_3 U_{22}^{-1} & I_3 & & \\ & & & \ddots & \\ & & & & I_s \end{pmatrix}$$

$$\times \begin{pmatrix} I_1 & & \\ & I_2 & \\ & & S_2 \end{pmatrix} \begin{pmatrix} I_1 & & & & \\ & U_{22} & L_{11}^{-1} C_2 & & \\ & & I_3 & & \\ & & & \ddots & \\ & & & & I_s \end{pmatrix},$$

where the form of S_2 is similar to the form of S_1 and, hence, it can be ignored. For a given i, if the first block of S_i can be factorized, then the partitioned LU factorization may run up to the $(i+1)$st step. Otherwise, the factorization must terminate at the ith step. Suppose that the factorization may run up to the completion. As a result, we obtain

$$A = L_1 \ldots L_{s-1} L_s U_s U_{s-1} \ldots U_1,$$

where

$$D_{s-1} = L_s U_s.$$

Note that the form and content of the partitioned LU factorization and the general block LU factorization are different because every step in the process of the former factorization needs one more LU factorization comparing to the latter and the factors L_i and U_i in the former are triangular forms that are not satisfied for the latter.

In the literature, there are numerous papers dealing with the perturbation bounds for usual or pointwise LU, Cholesky, or QR factorizations. The references dealing with this problem include Barrlund [1], Stewart [2, 3, 4], Chang and Paige [5], Dopico and Molera [6], etc. The first-order perturbation bounds are frequently used, see Chang, Paige, and Stewart [7] and Stewart [2, 3]. Dopico and Molera [6] presented expressions for the terms of any order in the series expressions for the perturbed LU and Cholesky factors. When the above factorization for the original matrix A in (1.1) runs to completion, the question is whether the perturbed matrix $A+E$ exists for the partitioned LU factorization. If E satisfies the inequality

$$|E| \le \epsilon |A|,$$

where ϵ is sufficiently small and $|A|$ stands for a matrix of the absolute values of the entries of A, then the partitioned LU factorization for the perturbed matrix $A+E$ exists. The relationships between

$$S_{ij}^{(k)}(A+E), \quad D_{ij}^{(k)}(A+E), \qquad \text{and} \qquad S_{ij}^{(k)}(A), \quad D_{ij}^{(k)}(A),$$

respectively, are considered, where $S_{ij}^{(k)}(A)$ and $D_{ij}^{(k)}(A)$ stand for the blocks (i,j) of S_k and D_k, respectively. Moreover, the perturbation bounds for the factors are also established.

The error analysis is one of the most powerful tools for studying the accuracy and stability of numerical algorithms. The references relevant to this problem include Higham [8, 9, 10], Amodio and Mazzia [11], Demmel, Higham, and Shreiber [12], Zhao, Wang, and Ren [13], Mattheij [14], Forsgren, Gill, and Shinnerl [15], Bueno and Dopico [16], etc. In the present paper, we apply the special property that the factors L_i and U_i are triangular forms. Then some assumptions on the BLAS3 that cannot be used in the error analysis of the general block LU factorization can be used in the error analysis of the partitioned LU factorization. Hence, the error analysis of the partitioned LU factorization for block tridiagonal linear systems can be considered. Comparing the results of Higham [8], Demmel and Higham [17] with the results obtained in the present paper, we conclude that the difference between the former and the latter is well visible. In view of the assumptions, the latter conditions are weaker than the former. Finally, two numerical examples are considered to illustrate the results of our theory for the indicated matrices generated by the discretization of the partial differential equation $-\Delta u = f$ and the random block tridiagonal matrices from the MATLAB 6.5, respectively.

2. Perturbation Theory

In this section, we perform the perturbation analysis of the factors of the partitioned LU factorization.

2.1. Some Properties.

We first consider the relationship between the first block of $S_k(A+E)$ and the first block of $S_k(A)$.

Theorem 2.1. *Suppose that the partitioned LU factorization for the block tridiagonal matrix A in (1.1) runs up to the completion. Assume that ϵ is sufficiently small such that $|E| \le \epsilon|A|$. Then*

$$S_{11}^{(k)}(A+E) = S_{11}^{(k)}(A) + T_k + O(\epsilon^2),$$

where T_k, $1 \le k \le s$, satisfy

$$T_1 = E_{11}, \qquad T_k = \begin{pmatrix} B_k \left(S_{11}^{(k-1)}\right)^{-1} & I_k \end{pmatrix} \begin{pmatrix} -T_{k-1} & -E_{k-1,k} \\ -E_{k,k-1} & E_{k,k} \end{pmatrix} \begin{pmatrix} \left(S_{11}^{(k-1)}\right)^{-1} C_k \\ I_k \end{pmatrix}.$$

Proof. To save clutter, we omit "$+O(\epsilon^2)$." The proof is essentially inductive. For $k = 1$, we find

$$S_{11}^{(1)}(A+E) = (A_2 + E_{22}) - (B_2 + E_{21})U_{11}^{-1}(A+E)L_{11}^{-1}(A+E)(C_1 + E_{12}).$$

Since $A^{-1} = U^{-1}L^{-1}$ and $|E| \le \epsilon|A|$, we get

$$\begin{aligned} S_{11}^{(1)}(A+E) &= (A_2 + E_{22}) - (B_2 + E_{21})(A_1^{-1} + A_1^{-1}E_{11}A_1^{-1})(C_1 + E_{12}) \\ &= A_2 - B_2A_1^{-1}C_1 + E_{22} - E_{21}A_1^{-1}C_1 - B_2A_1^{-1}E_{11}A_1^{-1}C_1 - B_2A_1^{-1}E_{12} \\ &= S_{11}^{(1)}(A) + \begin{pmatrix} B_2A_1^{-1} & I_2 \end{pmatrix} \begin{pmatrix} -E_{11} & -E_{12} \\ -E_{21} & E_{22} \end{pmatrix} \begin{pmatrix} A_1^{-1}C_1 \\ I_2 \end{pmatrix}. \end{aligned}$$

For $k = i - 1$, by the assumption, we find

$$S_{11}^{(i-1)}(A+E) = S_{11}^{(i-1)}(A) + T_i,$$

where it follows from the structure of T_i that $T_i = O(\epsilon)$. For $k = i$, we get

$$\begin{aligned} S_{11}^{(i)}(A+E) &= (A_{i+1} + E_{i+1,i+1}) - (B_{i+1} + E_{i+1,i})\left(S_{11}^{(i-1)} + T_i\right)^{-1}(C_i + E_{i,i+1}) \\ &= A_{i+1} - B_{i+1}U_{ii}^{-1}C_i + E_{i+1,i+1} - E_{i+1,i}\left(S_{11}^{(i-1)}\right)^{-1}C_i - B_{i+1}\left(S_{11}^{(i-1)}\right)^{-1}T_i\left(S_{11}^{(i-1)}\right)^{-1}C_i \\ &= S_{11}^{(i)}(A) + \left(B_{i+1}\left(S_{11}^{(i-1)}\right)^{-1} \quad I_{i+1}\right)\begin{pmatrix} -T_i & -E_{i,i+1} \\ -E_{i+1,i} & E_{i+1,i+1} \end{pmatrix}\begin{pmatrix} \left(S_{11}^{(i-1)}\right)^{-1}C_i \\ I_{i+1} \end{pmatrix}. \end{aligned}$$

Theorem 2.1 is proved.

As above, this result implies the following theorem:

Theorem 2.2. *Suppose that the partitioned LU factorization for the block tridiagonal matrix A in (1.1) runs to the completion. Assume that ϵ is sufficiently small so that $|E| \le \epsilon|A|$. Then*

$$S_{ij}^{(k)}(A+E) = S_{ij}^{(k)}(A) + \alpha_{ij}(T_k + O(\epsilon^2)) + (1 - \alpha_{ij})E_{ij},$$

$$D_{ij}^{(k)}(A+E) = D_{ij}^{(k)}(A) + \beta_i\big(\alpha_{ij}(T_k + O(\epsilon^2)) + (1 - \alpha_{ij})E_{ij}\big),$$

where

$$\beta_i = \begin{cases} 1, & k \le i \le s-1, \\ 0, & 1 \le i < k, \end{cases} \qquad \alpha_{ij} = \begin{cases} 1, & i = j = 1, \\ 0, & \textit{otherwise}. \end{cases}$$

Proof. By the partitioned LU factorization, we obtain

$$S_{ij}^{(k)}(A+E) = S_{ij}^{(k)}(A) + E_{ij}, \quad i,j \neq 1. \tag{2.1}$$

Combining (2.1) with Theorem 2.1, we find

$$S_{ij}^{(k)}(A+E) = S_{ij}^{(k)}(A) + \alpha_{ij}(T_k + O(\epsilon^2)) + (1-\alpha_{ij})E_{ij}, \tag{2.2}$$

where

$$\alpha_{ij} = \begin{cases} 1, & i=j=1, \\ 0, & \text{otherwise.} \end{cases}$$

In view of the form of D_k, we conclude that

$$D_{ij}^{(k)}(A+E) = D_{ij}^{(k)}(A), \quad 1 \le i < k. \tag{2.3}$$

It follows from (2.2) that

$$D_{ij}^{(k)}(A+E) = D_{ij}^{(k)}(A) + \alpha_{ij}(T_k + O(\epsilon^2)) + (1-\alpha_{ij})E_{ij}, \quad k \le i \le s-1. \tag{2.4}$$

In view of (2.3) and (2.4), we can write

$$D_{ij}^{(k)}(A+E) = D_{ij}^{(k)}(A) + \beta_i \left(\alpha_{ij}(T_k + O(\epsilon^2)) + (1-\alpha_{ij})E_{ij}\right),$$

where

$$\beta_i = \begin{cases} 1, & k \le i \le s-1, \\ 0, & 1 \le i < k. \end{cases}$$

Theorem 2.2 is proved.

Corollary 2.1. *Let the partitioned LU factorization for the block tridiagonal matrix A in (1.1) run to completion. Assume that ϵ is sufficiently small so that $|E| \le \epsilon|A|$. Then*

$$S_{ij}^{(k)}(A+E) = S_{ij}^{(k)}(A) + O(\epsilon), \qquad D_{ij}^{(k)}(A+E) = D_{ij}^{(k)}(A) + O(\epsilon).$$

Proof. From the proof of Theorem 2.1 and the form of T_k, it follows that $T_k = O(\epsilon)$. Then

$$T_k + E_{ij} + O(\epsilon^2) = O(\epsilon).$$

Therefore,

$$S_{ij}^{(k)}(A+E) = S_{ij}^{(k)}(A) + O(\epsilon), \qquad D_{ij}^{(k)}(A+E) = D_{ij}^{(k)}(A) + O(\epsilon).$$

In view of the inequality $|E| \leq \epsilon|A|$, if ϵ is sufficiently small, then, for the spectral radius, we get

$$\rho(L^{-1}EU^{-1}) < 1.$$

Therefore, it has the unique block LU factorization (see Theorem 12.1 in [8] for details). In this case, the question is whether the matrices $S_{11}^{(k)}$, $1 \leq k \leq s-1$, admit the LU factorization, i.e., whether the perturbed matrix $A+E$ admits the partitioned LU factorization. By Theorem 2.1, we conclude that

$$S_{11}^{(k)}(A+E) = S_{11}^{(k)}(A) + T_k + O(\epsilon^2).$$

Under the assumption of Theorem 2.1, we find

$$\begin{aligned} S_{11}^{(k)}(A+E) &= L_{k+1,k+1}U_{k+1,k+1} + T_k + O(\epsilon^2) \\ &= L_{k+1,k+1}\left(I_{k+1} + L_{k+1,k+1}^{-1}(T_k + O(\epsilon^2))U_{k+1,k+1}^{-1}\right)U_{k+1,k+1}. \end{aligned}$$

Since

$$T_k + O(\epsilon^2) = O(\epsilon), \qquad \rho\left(L^{-1}EU^{-1}\right) < 1,$$

we get

$$\rho\Big(L_{k+1,k+1}^{-1}(T_k + O(\epsilon^2))U_{k+1,k+1}^{-1}\Big) < 1,$$

i.e., $I_{k+1} + L_{k+1,k+1}^{-1}(T_k + O(\epsilon^2))U_{k+1,k+1}^{-1}$ admits the LU factorization. Thus, $S_{11}^{(k)}(A+E)$ $(1 \leq k \leq s-1)$ have the LU factorization. Hence, the perturbed matrix $A+E$ admits the partitioned LU factorization. These results enable us to formulate the following theorem:

Theorem 2.3. *Suppose that the partitioned LU factorization for the block tridiagonal matrix A in (1.1) runs to completion. Assume that ϵ is sufficiently small and such that $|E| \leq \epsilon|A|$. Then the perturbed matrix $A+E$ admits the partitioned LU factorization.*

2.2. Perturbation Bounds for the Factors. In this section, we present the bounds for the factors. First, we consider the bound for $S_{ij}^{(k)}$. Obviously, we can easily get the following componentwise perturbation bound for $S_{11}^{(1)}$ by applying Theorem 2.1:

$$\left|S_{11}^{(1)}(A+E) - S_{11}^{(1)}(A)\right| \leq \epsilon|A_1|.$$

Unless otherwise stated, in this section, we assume that the norm without subscripts $\|.\|$ is an arbitrary subordinate and monotone matrix norm. For $S_{11}^{(k)}$, we get the following theorem:

Theorem 2.4. *Assume that the partitioned LU factorization for the block tridiagonal matrix A in (1.1) runs to completion and that*

$$\chi = \max_k \left\{ \left\| \left(S_{11}^{(k-1)}\right)^{-1} C_k \right\| \|A_{k,k-1}\| + \left\| B_k \left(S_{11}^{(k-1)}\right)^{-1} \right\| \|A_{k-1,k}\| + \|A_{kk}\| \right\},$$

$$\omega = \max_k \left\{ \left\| B_k \left(S_{11}^{(k-1)} \right)^{-1} \right\| \left\| \left(S_{11}^{(k-1)} \right)^{-1} C_k \right\| \right\}$$

with

$$\left\| B_k \left(S_{11}^{(k-1)} \right)^{-1} \right\| \left\| \left(S_{11}^{(k-1)} \right)^{-1} C_k \right\| \neq 1.$$

Assume that ϵ *is sufficiently small so that* $|E| \le \epsilon |A|$. *Then*

$$\left\| S_{11}^{(k)}(A+E) - S_{11}^{(k)}(A) \right\| \le \omega^{k-1} \left(\|A_1\| + \frac{\chi(\omega^{k-1}-1)}{\omega^{k-1}(\omega-1)} \right) \epsilon + O(\epsilon^2).$$

Proof. We first consider the bound for T_k. It follows from Theorem 2.1 that

$$\begin{aligned} T_k &= \left(B_k \left(S_{11}^{(k-1)} \right)^{-1} \;\; I_k \right) \begin{pmatrix} -T_{k-1} & -E_{k-1,k} \\ -E_{k,k-1} & E_{k,k} \end{pmatrix} \begin{pmatrix} \left(S_{11}^{(k-1)} \right)^{-1} C_k \\ I_k \end{pmatrix} \\ &= -B_k \left(S_{11}^{(k-1)} \right)^{-1} T_{k-1} \left(S_{11}^{(k-1)} \right)^{-1} C_k \\ &\qquad - E_{k,k-1} \left(S_{11}^{(k-1)} \right)^{-1} C_k - B_k \left(S_{11}^{(k-1)} \right)^{-1} E_{k-1,k} + E_{k,k}. \end{aligned}$$

Taking the monotone norm on both sides, we conclude that

$$\begin{aligned} \|T_k\| \le{}& \left\| B_k \left(S_{11}^{(k-1)} \right)^{-1} \right\| \left\| \left(S_{11}^{(k-1)} \right)^{-1} C_k \right\| \|T_{k-1}\| \\ &+ \epsilon \left\| \left(S_{11}^{(k-1)} \right)^{-1} C_k \right\| \|A_{k,k-1}\| + \epsilon \left\| B_k \left(S_{11}^{(k-1)} \right)^{-1} \right\| \|A_{k-1,k}\| + \epsilon \|A_{kk}\|. \end{aligned}$$

Rearranging this inequality, we find

$$\begin{aligned} \|T_k\| + \frac{\chi\epsilon}{\omega-1} &\le \omega \left(\|T_{k-1}\| + \frac{\chi\epsilon}{\omega-1} \right) \\ &\le \omega^{k-1} \left(\|T_1\| + \frac{\chi\epsilon}{\omega-1} \right) \\ &\le \omega^{k-1} \left(\|A_1\| + \frac{\chi}{\omega-1} \right) \epsilon. \end{aligned}$$

Then

$$\|T_k\| \le \omega^{k-1} \left(\|A_1\| + \frac{\chi(\omega^{k-1}-1)}{\omega^{k-1}(\omega-1)} \right) \epsilon.$$

Theorem 2.4 is proved.

By using Theorems 2.2 and 2.4, we can easily establish the perturbation bounds for $S_{ij}^k(A)$ and $D_{ij}^k(A)$, i.e.,

$$\left\|S_{ij}^{(k)}(A+E) - S_{ij}^{(k)}(A)\right\| \leq \omega^{k-1}\left(\|A_1\| + \frac{\chi(\omega^{k-1}-1)}{\omega^{k-1}(\omega-1)}\right)\epsilon + 2\|A_{ij}\|\epsilon + O(\epsilon^2),$$

$$\left\|D_{ij}^{(k)}(A+E) - D_{ij}^{(k)}(A)\right\| \leq \omega^{k-1}\left(\|A_1\| + \frac{\chi(\omega^{k-1}-1)}{\omega^{k-1}(\omega-1)}\right)\epsilon + 2\|A_{ij}\|\epsilon + O(\epsilon^2).$$

3. Error Analysis

Throughout this section, we use the conventional error model of floating-point arithmetic. The evaluation of an expression in the floating-point arithmetic is denoted by $fl(\cdot)$ with

$$fl(a\ o\ b) = (a\ o\ b)(1+\delta), \qquad |\delta| \leq u, \quad o = +, -, *, /$$

(see, e.g., [8]). Here, u is the unit roundoff of the applied machine.

Unless otherwise stated, in this section, the norm without subscripts denotes

$$\|A\| := \max_{i,j}|a_{ij}|.$$

Note that, for this norm, the best inequality is

$$\|AB\| \leq n\|A\|\|B\|,$$

where $A \in \mathbb{R}^{m\times n}$ and $B \in \mathbb{R}^{n\times p}$. It is well known that this norm is inconsistent but, for sparse matrices, it is a simple and proper choice.

Based on the techniques of fast matrix multiplication, the use of BLAS3 affects the stability only by increasing constant terms in the normwise backward error bounds (see [17] for details). We have the following assumptions concerning the underlying level-3 BLAS:

(a) The computed approximation $\hat{C}$ to $C = AB$, where $A \in \mathbb{R}^{m\times n}$ and $B \in \mathbb{R}^{n\times p}$, satisfies the relations

$$\hat{C} = AB + \Delta C, \qquad \|\Delta C\| \leq c_1(m,n,p)u\|A\|\|B\| + O(u^2),$$

where $c_1(m,n,p)$ is a constant depending on m, n, and p.

(b) If $T \in \mathbb{R}^{m\times m}$ and $B \in \mathbb{R}^{m\times p}$, then the computed solution $\hat{X}$ of the triangular systems $TX = B$ satisfies

$$T\hat{X} = B + \Delta B, \qquad \|\Delta B\| \leq c_2(m,p)u\|T\|\|\hat{X}\| + O(u^2),$$

where $c_2(m,p)$ denotes a constant depending on m and p.

Assumption (b) in the BLAS3 cannot be applied in the error analysis of the general block LU factorization because the factor U is not a triangular form. In view of this consideration, the partitioned LU is presented for block tridiagonal matrices because the factors L and U are triangular forms, i.e., the errors accumulated in the process of partitioned LU factorization and substitution can be represented by using the assumption (b). We first recall the error analyses of the partitioned LU factorization for a general partitioned matrix $A \in \mathbb{R}^{n\times n}$ and of the corresponding computed solution of $Ax = b$.

Lemma 3.1 [17]**.** *Under the assumptions (a), (b), and*

$$\hat{L}_{11}\hat{U}_{11} = A_1 + \Delta A_1, \qquad \|\Delta A_1\| \le c_3(k_1)u\|\hat{L}_{11}\|\|\hat{U}_{11}\| + O(u^2),$$

the LU factors of $A \in \mathbb{R}^{n\times n}$ computed by using the partitioned outer product form of the LU factorization with block size k_1 satisfy relation

$$\hat{L}\hat{U} = A + \Delta A,$$

where

$$\|\Delta A\| \le u\Big(\delta(n,k_1)\|A\| + \theta(n,k_1)\|\hat{L}\|\|\hat{U}\|\Big) + O(u^2),$$

and, in addition,

$$\delta(n,k_1) = 1 + \delta(n-k_1,k_1), \qquad \delta(k_1,k_1) = 0,$$

$$\theta(n,k_1) = \max\big\{c_3(k_1),\ c_2(k_1,n-k_1),\ 1 + c_1(n-k_1,k_1,n-k_1) + \delta(n-k_1,k_1) + \theta(n-k_1,k_1)\big\},$$

$$\theta(k_1,k_1) = 0.$$

The following lemma is known as Problem 12.6 in [8].

Lemma 3.2 [8]**.** *Under the conditions of Lemma 3.1 the computed solution of the equation $Ax = b$ satisfies*

$$(A + \delta A)\hat{x} = b, \qquad \|\delta A\| \le c_n u(\|A\| + \|\hat{L}\|\|\hat{U}\|) + O(u^2),$$

where c_n is a constant depending on n and the block size.

The corresponding error analyses of the block tridiagonal matrix A in (1.1) and its linear systems are presented in what follows:

Theorem 3.1. *Assume that the partitioned LU factorization for the block tridiagonal matrix A in (1.1) runs up to the completion. Then, under the assumptions (a) and (b),*

$$A + \Delta A = \hat{L}\hat{U}, \qquad \|\Delta A\| \le (\xi_{ij}\|A\| + \zeta_{ij}\|L\|\|U\|)u + O(u^2),$$

where

$$\xi_{ij} = \begin{cases} 0, & i = j = 1, \\ 1, & i = j \ne 1, \\ c_2(k_i,k_i)k_i\kappa(L_{ii}), & i = j-1, \\ c_2(k_i,k_i)k_i\kappa(U_{ii}), & i = j+1, \end{cases} \qquad \zeta_{ij} = \begin{cases} c_2(k_1,k_1), & i = j = 1, \\ c_i, & i = j \ne 1, \\ 0, & \textit{others}, \end{cases}$$

$$c_i = \max\big\{1 + c_1(k_{i-1},k_{i-1},k_i), c_2(k_i,k_i)\big\}.$$

Proof. To save the clutter, we omit "$+O(u^2)$." By the assumption (b), we get

$$\hat{\hat{U}}_{ii}^{-1}\hat{U}_{ii} = I_i + \Delta I_i, \qquad \|\Delta I_i\| \leq c_2(k_i,k_i)u\|\hat{\hat{U}}_{ii}^{-1}\|\|\hat{U}_{ii}\|, \tag{3.1}$$

where $\hat{\hat{U}}_{ii}^{-1}$ are the computed quantities used to invert $\hat{U}_{ii}$. Thus,

$$\hat{\hat{U}}_{ii}^{-1} = (I_i + \Delta I_i)\hat{U}_{ii}^{-1}.$$

Therefore,

$$\|\hat{\hat{U}}_{ii}^{-1}\| \leq \|\hat{U}_{ii}^{-1}\| + O(u).$$

By virtue of representation (3.1), we conclude that

$$\|\Delta I_i\| \leq c_2(k_i,k_i)u\|\hat{U}_{ii}^{-1}\|\|\hat{U}_{ii}\|.$$

Similarly, we get

$$\hat{\hat{U}}_{ii}^{-1}\hat{U}_{ii} = I_i + \Delta I_i, \qquad \|\Delta I_i\| \leq c_2(k_i,k_i)u\|U_{ii}^{-1}\|\|U_{ii}\| = c_2(k_i,k_i)\kappa(U_{ii})u. \tag{3.2}$$

A similar assertion also holds for the following case:

$$\hat{L}_{ii}\hat{\hat{L}}_{ii}^{-1} = I_i + \Delta I_i, \qquad \|\Delta I_i\| \leq c_2(k_i,k_i)u\|L_{ii}\|\|L_{ii}^{-1}\| = c_2(k_i,k_i)\kappa(L_{ii})u. \tag{3.3}$$

The process of partitioned factorization gives

$$\begin{aligned} B_2\hat{\hat{U}}_{11}^{-1}\hat{U}_{11} &= B_2 + \Delta B_2, \\ \hat{L}_{11}\hat{\hat{L}}_{11}^{-1}C_1 &= C_1 + \Delta C_1. \end{aligned} \tag{3.4}$$

It follows from representations (2.2)–(2.4) that

$$\|\Delta B_2\| \leq c_2(k_1,k_1)k_1\kappa(U_{11})\|B_{22}\|u,$$

$$\|\Delta C_1\| \leq c_2(k_1,k_1)k_1\kappa(L_{11})\|C_1\|u.$$

By the assumption (b), we get the following bound for ΔA_i :

$$\|\Delta A_1\| \leq c_2(k_1,k_1)u\|L_{11}\|\|U_{11}\|.$$

As for ΔB_3, ΔA_2, and ΔC_2, due to the errors acquired in the processes of multiplication and subtraction of the matrices and the LU factorization for $A_2 - B\hat{\hat{U}}_{11}^{-1}\hat{\hat{L}}^{-1}C_1$, they are different from ΔB_2, ΔA_1 and ΔC_1, respectively. Let

$$L_{21}U_{12} = B_2U_{11}^{-1}L_{11}^{-1}C_1 = H.$$

Then the computed approximation $\hat{H}$ satisfies the relations

$$B_2\hat{\hat{U}}_{11}^{-1}\hat{\hat{L}}_{11}^{-1}C_1 + \Delta H = \hat{H},$$

$$\|\Delta H\| \le c_1(k_1, k_1, k_2)u\|B_2U_{11}^{-1}\|\|L_{11}^{-1}C_1\|.$$

Let

$$A_2 - H = G.$$

Then the computed approximation $\hat{G}$ to G satisfies

$$A_2 - \hat{H} + \Delta G' = \hat{G}, \qquad \|\Delta G'\| \le u(\|A_2\| + \|\hat{H}\|),$$

i.e.,

$$\begin{gathered} A_2 - B_2\hat{\hat{U}}_{11}^{-1}\hat{\hat{L}}_{11}^{-1}C_1 + \Delta G = \hat{G}, \\ \|\Delta G\| \le u(\|A_2\| + (1 + c_1(k_1, k_1, k_2))\|B_2U_{11}^{-1}\|\|L_{11}^{-1}C_1\|). \end{gathered} \tag{3.5}$$

Applying the LU factorization to $\hat{G}$, we find

$$\hat{G} + \Delta G'' = \hat{L}_{22}\hat{U}_{22}, \qquad \|\Delta G''\| \le c_2(k_2, k_2)u\|L_{22}\|\|U_{22}\|.$$

Combining (3.1) with (3.2), we find

$$A_2 + \Delta A_2 = \hat{L}_{22}\hat{U}_{22} + B_2\hat{\hat{U}}_{11}^{-1}\hat{\hat{L}}_{11}^{-1}C_1,$$

$$\begin{aligned} \|\Delta A_2\| \le u\Big(&\|A_2\| + (1 + c_1(k_1, k_1, k_2))\|B_2U_{11}^{-1}\|\|L_{11}^{-1}C_1\| \\ &+ c_2(k_2, k_2)\|L_{22}\|\|U_{22}\|\Big). \end{aligned}$$

From the factors $\hat{L}_{11}$, $\hat{U}_{11}$, $\hat{L}_{22}$, and $\hat{U}_{22}$ obtained in the process of factorization, it is clear that $\hat{L}_{11}$ and $\hat{U}_{11}$ are different from $\hat{L}_{22}$ and $\hat{U}_{22}$, respectively, because the latter contain the errors caused by the multiplication and subtraction in addition to the process of factorization. For ΔA_i, $3 \le i \le s$, we get the following similar results:

$$\begin{aligned} \|\Delta A_i\| &\le u\Big(\|A_i\| + (1 + c_1(k_{i-1}, k_{i-1}, k_i))\big\|B_iU_{i-1,i-1}^{-1}\big\|\big\|L_{i-1,i-1}^{-1}C_{i-1}\big\| \\ &\qquad + c_2(k_i, k_i)\|L_{ii}\|\|U_{ii}\|\Big) \\ &\le u\big(\|A_i\| + c(\|L_{i,i-1}\|\|U_{i-1,i-1}\| + \|L_{ii}\|\|U_{ii}\|)\big), \end{aligned}$$

where

$$c_i = \max\left\{1 + c_1(k_{i-1}, k_{i-1}, k_i), c_2(k_i, k_i)\right\}.$$

For ΔB_{i+1} and ΔC_i, for all $2 \le i \le s-1$, we get

$$\|\Delta B_{i+1}\| \le c_2(k_i, k_i) k_i \kappa(U_{ii}) \|B_{i+1,i+1}\| u,$$

$$\|\Delta C_i\| \le c_2(k_i, k_i) k_i \kappa(L_{ii}) \|C_i\| u.$$

Therefore,

$$\|\Delta A\| \le (\xi_{ij} \|A\| + \zeta_{ij} \|L\| \|U\|) u,$$

where

$$\xi_{ij} = \begin{cases} 0, & i = j = 1, \\ 1, & i = j \neq 1, \\ c_2(k_i, k_i) k_i \kappa(L_{ii}), & i = j-1, \\ c_2(k_i, k_i) k_i \kappa(U_{ii}), & i = j+1, \end{cases} \qquad \zeta_{ij} = \begin{cases} c_2(k_1, k_1), & i = j = 1, \\ c, & i = j \neq 1, \\ 0, & \text{others.} \end{cases}$$

Theorem 3.1 is proved.

Remark 3.1. Comparing Lemma 3.1 with Theorem 3.1, we can formulate the following remarks:

(1) The assumption of Lemma 3.1

$$\hat{L}_{11} \hat{U}_{11} = A_1 + \Delta A_1, \qquad \|\Delta A_1\| \le c_3(k_1) u \|\hat{L}_{11}\| \|\hat{U}_{11}\| + O(u^2),$$

is omitted in Theorem 3.1. This means that the assumptions of Theorem 3.1 are weaker than the assumptions of Lemma 3.1.

(2) It is clear that the proof of the theorem differs from the proof of the lemma.

(3) In the result of the lemma, we use the computed approximate $\hat{L}$ and $\hat{U}$. At the same time, the exact quantities L and U are used in the result of the theorem.

From Theorem 3.1, we get the following assertion for the block tridiagonal linear systems (note that ΔL and ΔU are obtained in solving the equations $Ly = b$ and $Ux = y$, respectively):

Theorem 3.2. *Let A be as in (1.1). Suppose that the partitioned LU factorization gives an approximate solution $\hat{x}$ of the system $Ax = b$, where $\hat{x}$ is the exact solution of the system*

$$(A + \delta A)\hat{x} = b.$$

Then

$$\|\delta A\| \le (\xi_{ij} \|A\| + \delta_{ij} \|L\| \|U\|) u + O(u^2),$$

$$\frac{\|\hat{x} - x\|}{\|\hat{x}\|} \le n \left(\left(\xi_{ij} \kappa(A) + \frac{\gamma_{ns}}{u} \kappa(U) \right) + \left(\zeta_{ij} + \frac{n\gamma_{ns}}{u} \right) \|L\| \|U\| \|A^{-1}\| \right) u + O(u^2),$$

where $\delta_{ij} = \zeta_{ij} + \gamma_{2ns}/u$.

Proof. By the assumption, we find

$$(\hat{L}+\Delta L)(\hat{U}+\Delta U)\hat{x}=b.$$

Then

$$\delta A=\Delta A+\Delta L\hat{U}+\hat{L}\Delta U+\Delta L\Delta U. \tag{3.6}$$

In the subsequent proof, we need the bounds for ΔL and ΔU. Applying the factorization and the result of [18], we find

$$(\hat{U}_1+\Delta U_1)x=y^{(1)}, \qquad |\Delta U_1|\le \frac{nu}{1-nu}|\hat{U}_1|.$$

For a given i, we obtain

$$(\hat{U}_i+\Delta U_i)y^{(i-1)}=y^{(i)}, \qquad |\Delta U_i|\le \frac{nu}{1-nu}|\hat{U}_i|.$$

Thus,

$$(\hat{U}_s+\Delta U_s)\dots(\hat{U}_1+\Delta U_1)x=y,$$

$$|\Delta U|\le \frac{nsu}{1-nu}|\hat{U}_s|\dots|\hat{U}_1|\le \gamma_{ns}|\hat{U}|,$$

where $\gamma_{ns}=nsu/(1-nsu)$. By the definition of the norm, we get

$$\|\Delta U\|\le \gamma_{ns}\|\hat{U}\|. \tag{3.7}$$

On the other hand, we can write

$$(\hat{L}_1+\Delta L_1)\dots(\hat{L}_s+\Delta L_s)y=b, \qquad \|\Delta L\|\le \gamma_{ns}\|\hat{L}\|. \tag{3.8}$$

Combining (3.6), (3.7) with (3.8), by Theorem 3.1, we conclude that

$$\begin{aligned}\|\delta A\| &\le (\xi_{ij}\|A\|+\zeta_{ij}\|L\|\|U\|)u+(2\gamma_{ns}+\gamma_{ns}^2)n\|\hat{L}\|\|\hat{U}\|\\ &\le (\xi_{ij}\|A\|+\delta_{ij}\|L\|\|U\|)u,\end{aligned}$$

where $\delta_{ij}=\zeta_{ij}+\gamma_{2ns}n/u$ and $2\gamma_{ns}+\gamma_{ns}^2\le\gamma_{2ns}$ [8, 9]. The remaining part of the proof deals with the relative error. According to Higham [10], we get

$$\|\hat{x}-x\|\le \|A^{-1}(\Delta A+\Delta L\hat{U})+\hat{U}^{-1}\Delta U\|\|\hat{x}\|. \tag{3.9}$$

Applying Theorem 3.1, from (3.7) and (3.8), we find

$$\frac{\|\hat{x}-x\|}{\|\hat{x}\|}\le n\left(\left(\xi_{ij}\kappa(A)+\frac{\gamma_{ns}}{u}\kappa(U)\right)+\left(\zeta_{ij}+\frac{n\gamma_{ns}}{u}\right)\|L\|\|U\|\|A^{-1}\|\right)u.$$

Theorem 3.2 is proved.

Actually, for $k_i = 1,\ 1 \le i \le s,$ there exists a relationship between $\kappa(U)$ and $\kappa(A)$ and $\|L\| \le 1$ holds if a partial pivoting strategy is applied during factorization. Then the relative error mentioned above can be $O(\kappa(A)u)$. On the other hand, the triangular form of the factors L_i and U_i in the partitioned LU factorization used in the present paper gives an advantage of the relative error

$$\|\hat{x} - x\|/\|\hat{x}\|.$$

Remark 3.2. Comparing Lemma 3.2 with Theorem 3.2, we can make the following remarks in addition to the first comment in Remark 3.1:

1. The coefficient c_n in Lemma 3.2 is a faint constant; however, the coefficients in Theorem 3.2 are given exactly.
2. In Theorem 3.2, the relative error of the solution is also considered; however, its form is not analyzed.

4. Numerical Experiments

In this section, we apply MATLAB 6.5 to illustrate the theoretical results on the backward error generated by the partitioned LU factorization for block tridiagonal matrices and on the relative error of the solution to linear systems.

Example 4.1. Assume that the block tridiagonal matrices are generated by the discretization of the partial differential equation

$$-\Delta u = f,$$

where

$$A_i = \text{tridiag}(-1, 4, -1)_{k_i \times k_i}.$$

Some results corresponding to this example are listed in Table 4.1.

Example 4.2. Let A be random block tridiagonal matrices, where A_i, B_i, and C_i are random matrices with approximately

$$0.8 \times k_i \times k_i, \qquad 0.2 \times k_i \times k_{i-1}, \qquad \text{and} \qquad 0.2 \times k_{i-1} \times k_i$$

uniformly distributed nonzero entries, respectively. The results are listed in Table 2.

It follows from the results presented above that the errors

$$\|A - \hat{L} * \hat{U}\| \qquad \text{and} \qquad \|x - \hat{x}\|/\|\hat{x}\|$$

are very small. However, it is impossible to say that the partitioned LU factorization must be stable because the backward error contains $\|L\|$. Thus,

$$A = \begin{pmatrix} \epsilon & 0 & 1 & 0 & & \\ 0 & \epsilon & 0 & 1 & & \\ 1 & 0 & \epsilon & 0 & 1 & 0 \\ 0 & 1 & 0 & \epsilon & 0 & 1 \\ & & 1 & 0 & 1 & 0 \\ & & 0 & 1 & 0 & 1 \end{pmatrix},$$

Table 4.1

Size	$\|A - \hat{L} * \hat{U}\|$	$\|x - \hat{x}\|/\|\hat{x}\|$
900×900	$1.7764e - 015$	$2.2204e - 015$
1600×1600	$2.6645e - 015$	$1.0880e - 014$
3600×3600	$3.5527e - 015$	$1.4655e - 014$

Table 4.2

Size	$\|A - \hat{L} * \hat{U}\|$	$\|x - \hat{x}\|/\|\hat{x}\|$
900×900	$5.6843e - 014$	$3.4195e - 013$
1600×1600	$1.2967e - 013$	$1.2765e - 012$
3600×3600	$8.1712e - 014$	$3.3598e - 012$

where ϵ is sufficiently small. Applying the partitioned LU factorization studied in the present paper, we get

$$L = L_1 L_2 L_3 = \begin{pmatrix} 1 & & & & & \\ & 1 & & & & \\ \frac{1}{\epsilon} & & 1 & & & \\ & \frac{1}{\epsilon} & & 1 & & \\ & & \frac{\epsilon}{\epsilon^2 - 1} & & 1 & \\ & & & \frac{\epsilon}{\epsilon^2 - 1} & & 1 \end{pmatrix}.$$

Hence, $\|L\|$ is boundless if ϵ is sufficiently small.

The present research was supported by the NSFC (No. 11301223), GDNSF (No. S2012040007333), financial engineering and algorithm research center of the Jinan University, high-level university construction and overall planning project of the Jinan University (Shenzhen Tourism College) (No. 88015201), and the teaching reform research project of the Jinan University (No. 55610196).

REFERENCES

1. A. Barrland, "Perturbation bounds for the LDL^H and the LU factorizations," *BIT*, **31**, 358–363 (1991).
2. G. W. Stewart, "Perturbation bounds for the QR factorization of a matrix," *SIAM J. Numer. Anal.*, **14**, 509–518 (1977).
3. G. W. Stewart, "On the perturbation of LU, Cholesky, and QR factorizations," *SIAM J. Numer. Anal.*, **14**, 1141–1145 (1993).
4. G. W. Stewart, "On the perturbation of LU and Cholesky factors," *SIAM J. Numer. Anal.*, **17**, 1–6 (1997).
5. X.-W. Chang and C. C. Paige, "On the sensitivity of the LU factorization," *BIT*, **38**, 486–501 (1998).
6. F. M. Dopico and J. M. Molera, "Perturbation theory of factorizations of LUF type through series expansions," *SIAM J. Matrix Anal. Appl.*, **27**, 561–581 (2005).

7. X.-W. Chang, C. C. Paige, and G. W. Stewart, "Perturbation analyses for the QR factorization," *SIAM J. Matrix Anal. Appl.*, **18**, 775–791 (1997).
8. N. J. Higham, *Accuracy and Stability of Numerical Algorithm*, SIAM, Philadelphia (1996).
9. N. J. Higham, "Accuracy and stability of algorithms in numerical linear algebra," *Manchester Center Comput. Math. Numer. Anal. Rept.*, No. 333 (1998).
10. N. J. Higham, "The accuracy of solutions to triangular systems," *SIAM J. Numer. Anal.*, **26**, 1252–1256 (1989).
11. P. Amodio and F. Mazzia, "A new approach to the backward error analysis in the LU factorization algorithm," *BIT*, **39**, 385–402 (1999).
12. J. W. Demmel, N. J. Higham, and R. S. Shreiber, "Stability of block LU factorization," *Numer. Linear Algebra Appl.*, **2**, 173–190 (1995).
13. J. Zhao, W. Wang, and W. Ren, "Stability of the matrix factorization for solving block tridiagonal symmetric indefinite linear systems," *BIT Numer. Math.*, **44**, 181–188 (2004).
14. R. M. M. Mattheij, "The stability of LU-decompositions of block tridiagonal matrices," *Bull. Aust. Math. Soc.*, **29**, 177–205 (1984).
15. A. Forsgren, P. E. Gill, and J. R. Shinnerl, "Stability of symmetric ill-conditioned systems arising in interior methods for constrained optimization," *SIAM J. Matrix Anal. Appl.*, **17**, 187–211 (1996).
16. M. I. Bueno and F. M. Dopico, "Stability and sensitivity of tridiagonal LU factorization without pivoting," *BIT Numer. Math.*, **44**, 651–673 (2004).
17. J. W. Demmel and N. J. Higham, "Stability of block algorithms with fast level-3 BLAS," *ACM Trans. Math. Software*, **18**, 274–291 (1992).
18. J. Stoer and R. Bulirsch, *Introduction to Numerical Analysis*, 2nd ed, Springer-Verlag, New York (1993).