Linear Algebra and its Applications 436 (2012) 659-681



On the infinitesimal limits of the Schur complements of tridiagonal matrices

Naveen Somasunderam^{a,*} Shivkumar Chandrasekaran^b

^a Dept. of Mechanical Engineering, University of California, Santa Barbara, United States ^b Dept. of Electrical Engineering, University of California, Santa Barbara, United States

ARTICLEINFO

Article history: Received 15 February 2011 Accepted 14 July 2011 Available online 6 September 2011

Submitted by R.A. Brualdi

AMS classification: 15A23 15A99 65F05 65Q10

Keywords: Asymptotic Analysis Difference equations LU factorization Schur complements Tridiagonal matrices

ABSTRACT

In this paper we consider diagonally dominant tridiagonal matrices whose diagonals come from smooth functions. It is shown that the Schur complements or pivots that arise from Gaussian elimination of these matrices can be given point-wise limits on a grid as the matrices grow in size to infinity. Numerical results are presented to verify the theory.

© 2011 Elsevier Inc. All rights reserved.

1. Introduction

Linear systems of the form Ax = b are ubiquitous in applications. A direct solution to such systems requires the *LU* factorization of the matrix *A*. Performing direct Gaussian elimination would require an $O(n^3)$ algorithm for an $n \times n$ system [3], unless some special structure of the matrix *A* could be exploited. Therefore, it is of considerable interest to look for special structures either in *A* or its *LU* factors.

The article [2] considered block tridiagonal matrices that come from the discretization of constant coefficient elliptic PDEs on the unit cube. It was shown that the final schur complement of such matrices converged to a known fixed point as the grid sizes grew to infinity. The same result for the constant

* Corresponding author. E-mail addresses: naveen@umail.ucsb.edu (N. Somasunderam), shiv@ece.ucsb.edu (S. Chandrasekaran).

0024-3795/\$ - see front matter © 2011 Elsevier Inc. All rights reserved. doi:10.1016/j.laa.2011.07.037

scalar case has been known for some time in the dynamical systems literature [6]. In this paper we shall look at the *LU* factorization of diagonally dominant tridiagonal matrices, whose diagonals come from the samples of smooth functions on a uniform grid. Such matrices can arise in many applications such as the discretization of differential equations [4]. In particular, we prove that the Schur complements of these matrices have point-wise limits on the grid as the discretization size goes to zero.

This local behavior of the Schur complements raises many interesting questions. For example, one can consider the possibility of interpolating the *LU* factors of the operator from coarser to finer grids (as compared to multigrid methods that interpolate the solution [1]). These ideas may also be used for constructing approximate inverses. These issues are under investigation and shall be published in a future paper. The aim of this paper is to lay a theoretical foundation to show that such limits are possible.

The matrices we wish to analyze would look like

$$A = \begin{pmatrix} a_1 & b_1 & & & \\ c_1 & a_2 & b_2 & & & \\ & c_2 & a_3 & b_3 & & \\ & \ddots & \ddots & \ddots & \\ & & c_{n-2} & a_{n-1} & b_{n-1} \\ & & & c_{n-1} & a_n \end{pmatrix}_{n \times n}$$

One may think of the a_k 's, b_k 's, and c_k 's as coming from the samples of smooth functions on the interval [0, 1]. That is $a(k) = a\left(\frac{k}{n+1}\right)$ for k = 1, ..., n, and so too with b_k and c_k . This would reflect many naturally occurring systems, for example the discretization of differential equations.

If we were to perform Gaussian elimination on the matrix A, we would first have to zero out c_1 in the (2, 1) position of the matrix. This can be done by adding $-c_1/a_1$ times the first row to the second row. Therefore, our first Schur complement would just be

$$s_1 = a_1,$$

and the second Schur complement in the (2, 2) position would be

$$s_2 = a_2 - \frac{c_1 b_1}{s_1}.$$

Now we have to use the second row to zero out the entry in the (3, 2) position. This would produce the third Schur complement in the (3, 3) position

$$s_3 = a_3 - \frac{c_2 b_2}{s_2}$$

Continuing this recursive process, any (k + 1)th Schur complement would be

$$s_{k+1} = a_{k+1} - \frac{c_k b_k}{s_k}.$$

The question we wish to answer is as the size of *A* grows, that is as *n* tends to infinity, does the above recursion have a limit. An intuitive argument would be as follows. Assuming a_k , b_k , and c_k are all constant, the above recursion would be

$$s_{k+1}=a-\frac{cb}{s_k}.$$

Presume that s_k is converging to some point *s*. Then the above recursion becomes a quadratic in the limit since

 $s_{k+1}s_k = as_k - cb,$

and with $s_k \approx s$ we get

$$s^2 - as + cb = 0.$$



Fig. 1. Point-wise limits of the Schur complements.

The two roots of this equation are then

$$s=\frac{a+\sqrt{a^2-4bc}}{2},$$

and

$$s = \frac{a - \sqrt{a^2 - 4bc}}{2}$$

We shall show that the above recursion does converge to the more positive root under some appropriate assumptions on the matrix A such that its LU factorization exists.

In particular, we claim that the Schur complements have point-wise limits as the grid becomes finer and finer. That is, consider Fig. 1 shown below. Then, if you look at a fixed discretization point x' on the grid, the Schur complements corresponding to that point have a local limit falling onto a know curve s(x). We need to make one more relevant point. We claim that the Schur complements are converging to the positive quadratic root. However, if we look at the matrix A the very first Schur complement shall always be fixed at a_1 , where as the positive root is not equal to a_1 . Therefore, this introduces a natural discontinuity at the origin. So the convergence we look for will not be uniform.

This paper will follow the theory laid out in [8]. The organization of the rest of this paper is as follows. We shall start by looking at a simple example in Section 2, the discrete equivalent of the second derivative operator. This should set the stage for the more general case of tridiagonal matrices with constant diagonals in Section 3, followed by variable tridiagonal matrices in Section 4. We shall present various numerical results to support the theory in section 5. Finally, we shall conclude with a summary and further work in Section 6.

Before proceeding, we need to raise two simple lemmas that are nevertheless useful. We will also make the following notational simplification through out this paper. Suppose that α_1 , α_2 , to α_n are an arbitrary set of *n* numbers. Then we will denote their product as

$$\prod_{i=1}^{n} \alpha_i = \alpha_{1..n}$$

Lemma 1.1. Let B be a bidiagonal matrix

$$B = \begin{pmatrix} a_1 & & & \\ c_1 & a_2 & & \\ & c_2 & \ddots & \\ & & \ddots & \ddots & \\ & & & c_{n-1} & a_n \end{pmatrix},$$

then its inverse is given by

$$(B_{in\nu})_{kj} = \begin{cases} \frac{(-1)^{|k-j|}c_{j.k-1}}{a_{j.k}} & \text{for } k > j, \\ \frac{1}{a_k} & \text{for } k = j, \\ 0 & \text{else,} \end{cases}$$

for k = 1, ..., n.

Proof. Consider the multiplication of any *k*th row of B_{inv} by a *j*th column of *B*, for k > j, then

$$(B_{inv})_{k*} (B)_{*j} = \frac{(-1)^{|k-j|} c_{j..k-1}}{a_{j..k}} a_j + \frac{(-1)^{|k-(j+1)|} c_{j+1..k-1}}{a_{j+1..k}} c_j,$$

= $\frac{(-1)^{|k-j|} c_{j..k-1}}{a_{j+1..k}} - \frac{(-1)^{|k-j|} c_{j..k-1}}{a_{j+1..k}},$
= 0.

Assume k = j then,

$$(B_{inv})_{j*} (B)_{*j} = \frac{1}{a_k} a_k,$$

= 1.

If k < j then it is obvious that

$$(B_{inv})_{k*} (B)_{*i} = 0.$$

This completes the proof of Lemma 1.1. \Box

The next Lemma relates the final Schur complement of a matrix A to its inverse.

Lemma 1.2. Suppose that X is the last Schur complement in the LU factorization of a matrix A. Then the bottom right-most entry of A^{-1} is equal to X^{-1} .

Proof. Note that since *X* is the last Schur complement of *A*, it must be the last diagonal entry of the upper factor *U*. Since $A^{-1} = U^{-1}L^{-1}$, the last entry of A^{-1} must be X^{-1} . \Box

2. A simple example

In this section we analyze the second order operator with Dirichlet boundary conditions,

$$-\frac{d^2}{dx^2}(u) = f,$$

with $u(0) = u_0$ and $u(1) = u_1$ on [0, 1].

The discrete equivalent of this operator with a second order finite difference scheme is

$$A = \begin{pmatrix} 2 & -1 & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & -1 \\ & & & -1 & 2 \end{pmatrix}_{n \times n}$$
(1)

We ignore the scaling factor of $\frac{1}{h^2}$, where *h* is the discretization step size. This matrix is symmetric positive definite and therefore has a Cholesky factorization (see [7]). We will analyze the Cholesky factorization of this matrix by first proving the following lemma.

Lemma 2.1. The Schur complement at each discretization point for the matrix in Eq. 1 converges to one.

Proof. Note that

$$s_1 = 2,$$

 $s_2 = 2 - \frac{1}{s_1},$
 $= 1 + \frac{1}{2}.$

So, if the *m*th Schur complement is $s_m = 1 + \frac{1}{m}$, then

$$s_{m+1} = 2 - \frac{1}{s_m},$$

= $2 - \frac{m}{m+1}$
= $1 + \frac{1}{m+1}$

Now for any discretization point $x \in (0, 1]$, the Schur complement corresponding to that point is $s(x) = s(kh) = 1 + \frac{1}{k}$, where *h* is the discretization step size and *k* is a positive integer such that x = kh. Since *kh* is constant as *h* goes to zero, we see that s(x) converges to 1. This completes the proof of Lemma 2.1. \Box

A natural question that arises is to what matrix this limiting Schur complements correspond to? We will call this Cholesky factor as L_{∞}

$$L_{\infty} = \begin{pmatrix} 1 & & \\ -1 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & \ddots & \\ & & & -1 & 1 \end{pmatrix}.$$

Multiplying this matrix out we get

$$L_{\infty}L_{\infty}^{T} = \begin{pmatrix} 1 & -1 & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & -1 \\ & & & -1 & 2 \end{pmatrix}$$

Note that the resulting matrix is almost similar to A. In fact,

$$A = L_{\infty}L_{\infty}^{T} + uu^{T},$$

= $L_{\infty} \left(I + L_{\infty}^{-1}uu^{T}L_{\infty}^{-T} \right) L_{\infty}^{T},$

where $u = (1 \ 0 \ \cdots \ 0)^T$. It is easy to verify that the inverse of L_{∞} is

$$L_{\infty}^{-1} = \operatorname{tril}(\mathbf{11}^{\mathrm{T}})$$

where tril(*) indicates the lower triangular part of *, and $\mathbf{1}$ is a vector of all ones. Therefore, we can write the following decomposition of A

$$A = L_{\infty} \left(I + v v^T \right) L_{\infty}^T,$$

where v = 1. The matrix in the middle is an identity plus rank-one matrix. In the next few sections we will show that this type of factorization could be extended to more general tridiagonal matrices.

3. Constant coefficient case

We start off our analysis of the general case by considering the constant coefficient tridiagonal matrix, where we assume the diagonal is *a* and the sub and super-diagonals are *b*. We assume without loss of generality that *a* is positive. The matrix looks like

$$A = \begin{pmatrix} a & b & & & \\ b & a & b & & & \\ b & a & b & & & \\ & \ddots & \ddots & \ddots & & \\ & & b & a & b \\ & & & & b & a \end{pmatrix}_{n \times n}$$
(2)

First we make the following assumption on A.

Assumption 3.1. Let a and b be such that $a \ge 2|b|$.

It follows from Assumption 3.1 that the term $a^2 - 4b^2$ is non-negative. The Schur complements of *A* are given by

 $s_1 = a$,

and

$$s_{k+1} = a - \frac{b^2}{s_k}.$$

The above non-linear recursion has two fixed points

$$X_p = \frac{a + \sqrt{a^2 - 4b^2}}{2},$$
$$X_n = \frac{a - \sqrt{a^2 - 4b^2}}{2}.$$

We make the following claim.

Theorem 3.1. The Schur complements of the matrix A in Eq. 2 converge point-wise to X_p , in the limit as the matrix size n tends to infinity.

The rest of this section shall be devoted to the Proof of Theorem 3.1. Consider the matrix

$$L_{\infty} = \begin{pmatrix} X_p^{1/2} & & \\ bX_p^{-1/2} & X_p^{1/2} & & \\ & bX_p^{-1/2} & \ddots & \\ & & \ddots & \ddots \\ & & & bX_p^{-1/2} & X_p^{1/2} \end{pmatrix}.$$

Looking at the product $L_{\infty}L_{\infty}^{T}$ we see that

$$L_{\infty}L_{\infty}^{T} = \begin{pmatrix} X_{p} & b & & \\ b & X_{p} + b^{2}X_{p}^{-1} & b & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & X_{p} + b^{2}X_{p}^{-1} & b & \\ & & & b & X_{p} + b^{2}X_{p}^{-1} \end{pmatrix},$$

$$= \begin{pmatrix} X_{p} & b & & & \\ b & X_{p} + X_{n} & b & & \\ & \ddots & \ddots & \ddots & & \\ & & & \ddots & X_{p} + X_{n} & b & \\ & & & & b & X_{p} + X_{n} \end{pmatrix},$$

$$= \begin{pmatrix} X_{p} & b & & & \\ b & a & b & & \\ & \ddots & \ddots & \ddots & & \\ & & \ddots & a & b & \\ & & & & b & a \end{pmatrix}.$$

The matrix A can now be written as

$$A=L_{\infty}SL_{\infty}^{T},$$

where

$$S = I + vv^{T},$$

$$v = L_{\infty}^{-1}u,$$

$$u = \left(X_{n}^{1/2} \ 0 \ \cdots \ 0\right)^{T}.$$

Now L_{∞} is a bidiagonal matrix whose inverse is given by Lemma 1.1. So, we can write out L_{∞}^{-1} as

,

N. Somasunderam, S. Chandrasekaran / Linear Algebra and its Applications 436 (2012) 659-681

From this we get the expression for *v* to be

$$v = L_{\infty}^{-1} u$$

= $\left(1 - bX_p^{-1} b^2 X_p^{-2} - b^3 X_p^{-3} \cdots\right)^T$. (3)

We now take a look at the Schur complements of $S = I + vv^T$. First we note that the inverse of *S* is given by the Sherman–Morrison formula (see [7]),

$$S^{-1} = \left(I + vv^T\right)^{-1},$$

= $I - \frac{vv^T}{1 + v^Tv}.$

Let us denote the components of *v* as

$$\boldsymbol{v} = \begin{pmatrix} v_1 & v_2 & \cdots & v_n \end{pmatrix}^T.$$

We can now make use of Lemma 1.2, and see that the inverse of the last Schur complement of *S* is the last entry in the inverse of *S*. Therefore, we have the following expression for the inverse of the last Schur complement \tilde{s}_n of *S*

$$\tilde{s}_n^{-1} = 1 - \frac{v_n^2}{1 + \sum_{m=1}^n v_m^2},$$
$$= \frac{1 + \sum_{m=1}^{n-1} v_m^2}{1 + \sum_{m=1}^n v_m^2}.$$

We can now write the last Schur complement of S as

$$\tilde{s}_n = \frac{1 + \sum_{m=1}^n v_m^2}{1 + \sum_{m=1}^{n-1} v_m^2},$$

= $1 + \frac{v_n^2}{1 + \sum_{m=1}^{n-1} v_m^2}.$

Note that by using a similar argument and considering the $k \times k$ principal block of *S*, we can write any *k*th Schur complement of *S* to be

$$\tilde{s}_k = 1 + \frac{v_k^2}{1 + \sum_{m=1}^{k-1} v_m^2}.$$

By taking a look at Eq. 3, we can write down v_{k+1}^2 to be

$$v_{k+1}^2 = b^{2k} X_p^{-2k},$$

= $\left(b^{2k} X_p^{-k}\right) X_p^{-k},$
= $\gamma^k,$

666

where we define γ as

$$\gamma = \frac{X_n}{X_p}.$$

We make one more point. It is easily verifiable that *S* is a symmetric positive definite matrix whose eigenvalues are just 1 and $1 + v^T v$. Therefore, *S* has a Cholesky factorization which we denote by $\hat{L}\hat{L}^T$. We can now give a proof of Theorem 3.1.

Proof of Theorem 3.1. We can write down the Cholesky factorization of the matrix A as

$$A = L_{\infty} \hat{L} \hat{L}^T L_{\infty}^T$$

where \hat{L} is the Cholesky factor of S. Since the kth diagonal entry of the product of two upper triangular matrices is just the product of the kth diagonal entries of the two matrices, we can now write the (k + 1)th Schur complement of A as

$$s_{k+1} = \tilde{s}_{k+1} X_p,$$

= $\left(1 + \frac{v_{k+1}^2}{1 + \sum_{m=1}^k v_m^2}\right) X_p,$
= $\left(1 + \frac{\gamma^k}{1 + \sum_{m=1}^{k-1} \gamma^m}\right) X_p.$

Now notice that with Assumption 3.1, we have $\gamma \leq 1$. Therefore, as $k \to \infty$, the first expression on the right approaches 1. So, s_{k+1} approaches X_p . This completes the proof. \Box

We can now extend Theorem 3.1 to non-symmetric tridiagonal matrices of the form

$$A = \begin{pmatrix} a & b & & \\ c & a & b & & \\ c & a & b & & \\ & \ddots & \ddots & \ddots & \\ & & c & a & b \\ & & & c & a \end{pmatrix},$$
(4)

where we assume that *A* is sign-symmetric.

Assumption 3.2. Let b and c have the same sign.

Assumption 3.3. Let *a*, *b* and *c* be such that $a \ge 2\sqrt{bc}$.

Corollary 3.1. Consider the matrix A as given by Eq. 4. Then the Schur complements of A converge pointwise to

$$X_p = \frac{a + \sqrt{a^2 - 4bc}}{2}$$

Proof. First we claim that for every sign symmetric matrix as A, there exists a diagonal matrix D such that DAD^{-1} is symmetric. For consider the diagonal matrix

$$D = \begin{pmatrix} 1 & \begin{pmatrix} \frac{b}{c} \end{pmatrix}^{\frac{1}{2}} & & \\ & \begin{pmatrix} \frac{b}{c} \end{pmatrix}^{\frac{3}{2}} & & \\ & & \ddots & \\ & & & \begin{pmatrix} \frac{b}{c} \end{pmatrix}^{\frac{n-1}{2}} \end{pmatrix}.$$

Now it is easy to verify that DAD^{-1} is a symmetric matrix of the form

$$A_{sym} = DAD^{-1},$$

$$= \begin{pmatrix} a & \sqrt{bc} & & \\ \sqrt{bc} & a & \sqrt{bc} & & \\ & \ddots & \ddots & \ddots & \\ & & \sqrt{bc} & a & \sqrt{bc} \\ & & & \sqrt{bc} & a \end{pmatrix}.$$

For if we consider any 2×2 block interaction of the above multiplication at any given *k*th level, we find that

$$\begin{pmatrix} \left(\frac{b}{c}\right)^{(k-1)/2} \\ \left(\frac{b}{c}\right)^{k/2} \end{pmatrix} \begin{pmatrix} a & b \\ c & a \end{pmatrix} \begin{pmatrix} \left(\frac{c}{b}\right)^{(k-1)/2} \\ \left(\frac{c}{b}\right)^{k/2} \end{pmatrix} \\ = \begin{pmatrix} a & \left(\frac{b}{c}\right)^{(k-1)/2} & b & \left(\frac{b}{c}\right)^{(k-1)/2} \\ c & \left(\frac{b}{c}\right)^{k/2} & a & \left(\frac{b}{c}\right)^{k/2} \end{pmatrix} \begin{pmatrix} \left(\frac{c}{b}\right)^{(k-1)/2} \\ \left(\frac{c}{b}\right)^{k/2} \\ \left(\frac{c}{b}\right)^{k/2} \end{pmatrix}, \\ = \begin{pmatrix} a & \sqrt{bc} \\ \sqrt{bc} & a \end{pmatrix}.$$

Now by applying Theorem 3.1 to the matrix A_{sym} , we find that its Schur complements converge to

$$X_p = \frac{a + \sqrt{a^2 - 4bc}}{2}$$

Suppose now that $A_{sym} = LU$ is an LU factorization of A_{sym} . Then we can write

$$A = D^{-1}LUD.$$

It is apparent that the above diagonal transformation does not affect the Schur complements of A since we can write A as

$$A = D^{-1}LUD,$$

= $(D^{-1}LD)(D^{-1}UD)$

Now notice that $D^{-1}LD$ is a lower triangular matrix with a unit diagonal since *L* has a unit diagonal. Then $D^{-1}UD$ is the unique upper triangular factor in the *LU* factorization of *A*. We point out here that the *LU* factorization of a matrix is only uniquely determined up to the diagonal entries of the factors. And here we are concerned with the factorization such that the lower triangular part has a unit diagonal. Moreover, it has the same diagonal as *U*. Therefore, the Schur complements of *A* converge to X_p . This finishes the proof of Corollary 3.1. \Box

668

We now move onto the next section, where we show that for a variable tridiagonal matrix, under some suitable assumptions on the diagonals, we can still establish point-wise limits on the Schur complements.

4. Variable tridiagonal matrix

We consider first the Cholesky factorization of symmetric positive definite tridiagonal matrices. It is assumed that the diagonal entries of the matrix are generated by some underlying smooth functions. For example, let a(x), b(x) be smooth functions on [0, 1], and we assume without loss of generality that a is a positive function. Then consider the matrix

$$A = \begin{pmatrix} a_{1} & b_{1} & & & \\ b_{1} & a_{2} & b_{2} & & & \\ b_{2} & a_{3} & b_{3} & & \\ & \ddots & \ddots & \ddots & \\ & & b_{n-2} & a_{n-1} & b_{n-1} \\ & & & b_{n-1} & a_{n} \end{pmatrix},$$
(5)

where $a_k = a\left(\frac{k}{n+1}\right)$ and $b_k = b\left(\frac{k}{n+1}\right)$. The Schur complements of this matrix are given by the recursion

$$s_0 = a_1$$

$$s_{k+1} = a_{k+1} - \frac{b_k}{s_k}.$$

Let us first make our assumptions on A explicit.

Assumption 4.1. *a and b are continuous functions with bounded derivatives on* [0, 1].

Assumption 4.2. There exists a constant δ such that

$$a-2|b| \ge \delta > 0.$$

With these assumptions, we make the following claim.

Theorem 4.1. The Schur complement at a point converges in the limit to

$$s(x) = \frac{a(x) + \sqrt{a^2(x) - 4b^2(x)}}{2}.$$

The rest of this section shall be devoted to the Proof of Theorem 4.1 and its extension to the non-symmetric case. Suppose that *A* is an $n \times n$ matrix. Then we define

$$\begin{split} X_{p_k} &= \frac{a\left(\frac{k}{n+1}\right) + \sqrt{a^2\left(\frac{k}{n+1}\right) - 4b^2\left(\frac{k}{n+1}\right)}}{2},\\ X_{n_k} &= \frac{a\left(\frac{k}{n+1}\right) - \sqrt{a^2\left(\frac{k}{n+1}\right) - 4b^2\left(\frac{k}{n+1}\right)}}{2}, \end{split}$$

and

$$\gamma_k = \frac{X_{nk}}{X_{pk}},$$

for k = 1, ..., n. Now consider the matrix

$$L_{\infty} = \begin{pmatrix} X_{p_1}^{1/2} & & & \\ \frac{b_1}{X_{p_1}^{1/2}} & X_{p_2}^{1/2} & & & \\ & \frac{b_2}{X_{p_2}^{1/2}} & X_{p_3}^{1/2} & & \\ & & \frac{b_3}{X_{p_3}^{1/2}} & X_{p_4}^{1/2} & & \\ & & \ddots & \ddots & \\ & & & \frac{b_{n-1}}{X_{p_n}^{1/2}} & X_{p_n}^{1/2} \end{pmatrix}$$

then the product $L_{\infty}L_{\infty}^{T}$ looks like

We can write the diagonal terms as

$$\begin{aligned} X_{p_k} + \frac{b_{k-1}^2}{X_{p_{k-1}}} &= X_{p_k} + X_{n_{k-1}}, \\ &= \frac{a_k + \sqrt{a_k^2 - 4b_k^2}}{a_k + \epsilon_k^2} + \frac{a_{k-1} - \sqrt{a_{k-1}^2 - 4b_{k-1}^2}}{2}, \end{aligned}$$

where ϵ_k depends on the local continuity of *a*. In fact,

$$\epsilon_k = X_{p_k} + X_{nk-1} - a_k$$

= - (X_{nk} - X_{nk-1})
= - \Delta X_{nk}.

By Assumptions 4.1 and 4.2, the derivative of X_n is well defined since

$$X'_{n} = \left(\frac{a - \sqrt{a^{2} - 4b^{2}}}{2}\right)',$$
$$= \frac{2a' \left(a^{2} - 4b^{2}\right)^{1/2} - (2aa' - 8bb')}{4 \left(a^{2} - 4b^{2}\right)^{1/2}}.$$

Therefore, we have the estimate

$$\begin{aligned} |\epsilon_k| &= |\Delta X_{nk}|, \\ &\leqslant \|X'_n\|_{\infty} h. \end{aligned}$$

where *h* is the discretization step size. So ϵ_k is of the order of O(h) for all *k*. Let us define D_{ϵ} to be a diagonal matrix of ϵ_k 's. We can now proceed to write *A* in terms of the factorization

$$A = L_{\infty}L_{\infty}^{T} + uu^{T} + D_{\epsilon},$$

= $L_{\infty} \left(I + vv^{T} + L_{\infty}^{-1}D_{\epsilon}L_{\infty}^{-T} \right) L_{\infty}^{T},$
= $L_{\infty} \left(\tilde{S} + \Delta \tilde{S} \right) L_{\infty}^{T},$

where

$$u = \begin{pmatrix} X_{n1}^{1/2} & 0 & \cdots & 0 \end{pmatrix}^T,$$
$$v = L_{\infty}^{-1} u,$$
$$\tilde{S} = I + v v^T,$$
$$\Delta \tilde{S} = L_{\infty}^{-1} D_{\epsilon} L_{\infty}^{-T}.$$

The proof of Theorem 4.1 is similar to that of Theorem 3.1 in Section 3, but we now have to analyze the effect of the $\Delta \tilde{S}$ perturbation to \tilde{S} . Let us look at the expression for L_{∞}^{-1} .

$$L_{\infty} = \begin{pmatrix} 1 & & & \\ \frac{b_1}{X_{p_1}} & 1 & & \\ & \frac{b_2}{X_{p_2}} & 1 & & \\ & & \ddots & \ddots & \\ & & & \frac{b_{n-1}}{X_{p_{n-1}}} & 1 \end{pmatrix} \begin{pmatrix} X_{p_1}^{1/2} & & & \\ & X_{p_2}^{1/2} & & \\ & & X_{p_3}^{1/2} & \\ & & & X_{p_n}^{1/2} \end{pmatrix},$$
$$= L_{\mathcal{B}}L_{\mathcal{D}},$$

where L_B denotes the bidiagonal part and L_D denotes the diagonal part. The inverse of the bidiagonal part L_B of the above matrix, using Lemma 1.1, is

$$L_{Bkj}^{-1} = \begin{cases} \frac{(-1)^{|k-j|} b_{j..k}}{X_{p_{j..k}}} & \text{for } k > j \\ 1 & \text{for } k = j \\ 0 & \text{else.} \end{cases}$$

This expression can be reduced to

$$L_{B_{kj}}^{-1} = \begin{cases} \frac{(-1)^{|k-j|}}{\gamma_{j,k}^{1/2}} & \text{for } k > j \\ 1 & \text{for } k = j \\ 0 & \text{else.} \end{cases}$$

That is

$$L_{\infty}^{-1} = L_D^{-1} L_B^{-1},$$

where L_B^{-1} looks like

$$\begin{pmatrix} 1 & & & \\ -\gamma_1^{1/2} & 1 & & \\ \gamma_{1..2}^{1/2} & -\gamma_2^{1/2} & 1 & & \\ -\gamma_{1..3}^{1/2} & \gamma_{2..3}^{1/2} & -\gamma_3^{1/2} & 1 & \\ & \ddots & \ddots & \ddots & \ddots & \\ & & & -\gamma_{n-1}^{1/2} & 1 \end{pmatrix}.$$

Note that the $(-1)^{|k-j|}$ term in the above expression needs to be replaced by 1 if *b* is a negative function. But, as this does not alter the analysis we proceed assuming the above expression. In particular, we get *v* to be

$$v = L_D^{-1} \begin{pmatrix} 1 \\ -\gamma_1^{1/2} \\ \gamma_{1..2}^{1/2} \\ -\gamma_{1..3}^{1/2} \\ \vdots \end{pmatrix} X_{n_1}^{1/2}.$$

/

The expression for $v^T v$ is

$$v^{T}v = \gamma_{1} + X_{n_{1}}\gamma_{1}X_{p_{2}}^{-1} + X_{n_{1}}\gamma_{1}\gamma_{2}X_{p_{3}}^{-1} + X_{n_{1}}\gamma_{1}\gamma_{2}\gamma_{3}X_{p_{4}}^{-1} + \cdots,$$

= $\gamma_{1} + \sum_{k=2}^{n} X_{n_{1}}\gamma_{1}...\gamma_{k-1}X_{p_{k}}^{-1}.$

Using Assumption 4.2, we see that there exists an α < 1 such that

 $\sup \gamma \leqslant \alpha.$

We can then produce an upper bound on $v^T v$ as follows

$$v^T v \leq \beta \sum_{k=0}^{\infty} \alpha^{k+1} < \infty,$$

where

$$\beta = \frac{\max_{x} X_n}{\min_{x} X_p}.$$

Note that β is well defined from our assumptions. Moreover, the square of the (k + 1)th entry of v is bounded by

$$v_{k+1}^2 = X_{n_1} \gamma_1 \dots \gamma_{k-1} X_{p_k}^{-1}$$

 $\leq \beta \alpha^k.$

Therefore, v_k^2 goes to zero as k tends to infinity. Now the kth Schur complement of \tilde{S} is given by,

$$\tilde{s}_k = 1 + \frac{v_k^2}{1 + \sum_{m=1}^{k-1} v_m^2}.$$

Therefore, \tilde{s}_k converges to 1 as k tends to infinity. We now have to look at the effect of the perturbation $\Delta \tilde{S}$ on \tilde{S} . We know that for $\gamma < 1$, the identity plus rank-one matrix has a Schur complement converging to one, so what is the effect of the perturbation? The perturbation bounds of LU and Cholesky factorizations have been studied by Stewart [9] and Sun [10]. We will use a non-trivial result of Sun (see [5]) on the Cholesky factors of a perturbed symmetric positive definite matrix.

Theorem 4.2. Suppose A be symmetric positive definite and R its Cholesky factor. If ΔA is a symmetric perturbation to A with Cholesky factor $R + \Delta R$, and $||A^{-1}\Delta A||_2 < 1$ then,

$$\|\Delta R\|_F \leqslant \frac{1}{\sqrt{2}} \frac{\|A^{-1}\|_2 \|\Delta A\|_F}{1 - \|A^{-1}\|_2 \|\Delta A\|_F} \|R\|_2.$$

With $A = I + vv^T$, note that $||A||_2 = \sqrt{1 + v^T v}$ and $||A^{-1}||_2 = 1$. Also, $||R||_2 \leq ||A||_2 = \sqrt{1 + v^T v}$. So, we just need to make an estimate on $||\Delta A||_F$. With $\Delta A = L_{\infty}^{-1} D_{\epsilon} L_{\infty}^{-T}$, first note that,

$$\begin{split} \|\Delta \tilde{S}\| &= \|L_{\infty}^{-1} D_{\epsilon} L_{\infty}^{-T}\|_{F} \\ &\leq \| |L_{\infty}^{-1}| |D_{\epsilon} L_{\infty}^{-T}| \|_{F}, \\ &\leq \epsilon \| |L_{\infty}^{-1}| |L_{\infty}^{-T}| \|_{F}, \end{split}$$

where ϵ is the maximal entry in D_{ϵ} . Now

$$L_{\infty}^{-1} = L_D^{-1} L_B^{-1},$$

and since $||DA||_F \leq m ||A||_F$, for any matrix A and diagonal matrix D, with $m = \max_i |D_i|$, by taking m to be

$$m=\frac{1}{\min_{x}X_{p}},$$

we have

$$\begin{split} \|\Delta \tilde{S}\|_{F} &\leqslant m \epsilon \, \||L_{B}^{-1}||L_{B}^{-T}||_{F}, \\ &\leqslant m \epsilon \, \left\| \begin{pmatrix} 1 & & & \\ \alpha & 1 & & \\ \alpha^{2} & \alpha & 1 & & \\ \alpha^{3} & \alpha^{2} & \alpha & 1 & \\ \ddots & \ddots & \ddots & \ddots & \\ & & \alpha & 1 \end{pmatrix} \begin{pmatrix} 1 & & & & \\ \alpha^{2} & \alpha & 1 & & \\ \alpha^{3} & \alpha^{2} & \alpha & 1 & \\ \ddots & \ddots & \ddots & \ddots & \\ & & & \alpha & 1 \end{pmatrix}^{T} \right\|_{F}, \\ &= m \epsilon \, \left\| \begin{pmatrix} 1 & \alpha & \alpha^{2} & \alpha^{3} & & \\ \alpha & 1 + \alpha^{2} & \alpha(1 + \alpha^{2}) & & \\ \alpha^{2} & \alpha(1 + \alpha^{2}) & 1 + \alpha^{2} + \alpha^{4} & & \\ \alpha^{3} & \alpha^{2}(1 + \alpha^{2}) & \alpha(1 + \alpha^{2} + \alpha^{4}) & \ddots & \\ & \ddots & \ddots & \ddots & \ddots & \\ & & & 1 + \alpha^{2} + \ldots + \alpha^{n} \end{pmatrix} \right\|_{F} \end{split}$$

,

with α < 1. Now suppose

$$r=\sum_k \alpha^{2k},$$

then

$$\|\Delta \tilde{S}\|_{F} \leqslant \epsilon m \left\| \begin{pmatrix} r & r\alpha & r\alpha^{2} & r\alpha^{3} & \\ r\alpha & r & r\alpha & & \\ r\alpha^{2} & r\alpha & r & & \\ r\alpha^{3} & r\alpha^{2} & r\alpha & \ddots & \\ & \ddots & \ddots & \ddots & \ddots & \\ & & & r\alpha & r \end{pmatrix} \right\|_{F}$$
$$\leqslant \epsilon m \sqrt{2nr^{3}}.$$



Fig. 2. Schur complements with $a(x) = 5 + x^2$ and $b(x) = c(x) = 1 + e^{-x}$ for grid sizes of n = 10, 100, 200 and 1000.

Since $\epsilon = O(1/n)$, $\|\triangle A\|_F$ goes to zero as *n* tends to infinity. As $\|\triangle A\|_F$ goes to zero, we have that $\|\triangle R\|_F$ goes to zero, therefore the Schur complements of the perturbed matrix must converge to one. We are now ready to give a proof of Theorem 4.1.

Proof of Theorem 4.1. The matrix A in Eq. 5 could be written as

 $A = L_{\infty} \left(\tilde{S} + \Delta \tilde{S} \right) L_{\infty}^{T},$

where it has been shown that the Schur complements of the middle term goes to one as n tends to infinity. Since the kth Schur complement of A is just the product of the kth Schur complement of middle term times X_{p_k} , then the Schur complement corresponding to a point x = kh on the grid converges to s(x). This completes the proof of Theorem 4.1.

We can now extend Theorem 4.1 to non-symmetric tridiagonal matrices of the form

$$A = \begin{pmatrix} a_1 & b_1 & & & \\ c_1 & a_2 & b_2 & & & \\ & c_2 & a_2 & b_3 & & & \\ & & \ddots & \ddots & \ddots & & \\ & & c_{n-2} & a_{n-1} & b_{n-1} \\ & & & c_{n-1} & a_n \end{pmatrix},$$
(7)

where we assume that A is sign-symmetric. We make the assumptions explicit.

Assumption 4.3. *a,b, c are continuous functions with bounded derivatives on* [0, 1].

Assumption 4.4. b, c are bounded away from zero and have the same sign.

Assumption 4.5. There exists a constant δ such that

$$a - 2\sqrt{bc} \ge \delta > 0.$$

Corollary 4.1. Consider the matrix A as given by Eq. 7. Then the Schur complements of A converge pointwise to

$$s(x) = \frac{a(x) + \sqrt{a^2(x) - 4b(x)c(x)}}{2}$$



Fig. 3. Log error of the Schur complements to the limiting function s(x) with $a(x) = 5 + x^2$ and $b(x) = c(x) = 1 + e^{-x}$ for grid sizes of n = 100, 1000, 10,000 and 100,000.



Fig. 4. Schur complements with $a(x) = 3 + \cos(10\pi x)$ and $b(x) = c(x) = e^{-|\sin(\pi x)|}$ for grid sizes of n = 10, 100, 200 and 1000.

Proof of Corollary 4.1. First, we claim that for every sign-symmetric matrix of the form given in Eq. 7 there exists a diagonal matrix D such that DAD^{-1} is symmetric. In fact, consider the following diagonal matrix

$$D = \begin{pmatrix} 1 & \left(\frac{b_1}{c_1}\right)^{\frac{1}{2}} & & \\ & \left(\frac{b_1 b_2}{c_1 c_2}\right)^{\frac{1}{2}} & & \\ & & \ddots & \\ & & & \left(\frac{b_1 .. b_{n-1}}{c_1 .. c_{n-1}}\right)^{\frac{1}{2}} \end{pmatrix}.$$

Then, DAD^{-1} is a symmetric matrix of the form



Fig. 5. Log error of the Schur complements to the limiting function s(x) with $a(x) = 3 + \cos(10\pi x)$ and $b(x) = c(x) = e^{-|\sin(\pi x)|}$ for grid sizes of n = 100, 1000, 10,000 and 100,000.



Fig. 6. Schur complements with $a(x) = 2\sqrt{2 + |x - 0.5|}$, $b(x) = 1 + e^{-0.5x^2}$, and $c(x) = |\sin(2\pi x)|$ for grid sizes of n = 10, 100, 200 and 1000.

$$A_{sym} = DAD^{-1},$$

$$= \begin{pmatrix} a_1 & \sqrt{b_1c_1} & & \\ \sqrt{b_1c_1} & a_2 & \sqrt{b_2c_2} & & \\ & \ddots & \ddots & \ddots & \\ & & \sqrt{b_{n-2}c_{n-2}} & a_{n-1} & \sqrt{b_{n-1}c_{n-1}} \\ & & & \sqrt{b_{n-1}c_{n-1}} & a_n \end{pmatrix}.$$

This is easy to see by considering any 2 × 2 block interaction of the above multiplication at any given kth level,

$$\begin{pmatrix} \left(\frac{b_{1..(k-2)}}{c_{1..(k-2)}}\right)^{1/2} \\ \left(\frac{b_{1..k-1}}{c_{1..k-1}}\right)^{1/2} \end{pmatrix} \begin{pmatrix} a_{k-1} & b_{k-1} \\ c_{k-1} & a_k \end{pmatrix} \begin{pmatrix} \left(\frac{c_{1..k-2}}{b_{1..k-2}}\right)^{1/2} \\ \left(\frac{c_{1..k-1}}{b_{1..k-1}}\right)^{1/2} \end{pmatrix}$$



Fig. 7. Log error of the Schur complements to the limiting function s(x) with $a(x) = 2\sqrt{2 + |x - 0.5|}$, $b(x) = 1 + e^{-0.5x^2}$, and $c(x) = |\sin(2\pi x)|$ for grid sizes of n = 100, 1000, 10,000 and 100,000.



Fig. 8. Schur complements with a(x) = 2 + sign(x - 0.5), b(x) = 0.5(1 - x), and $c(x) = |x - 0.75|^3$ for grid sizes of n = 10, 100, 200 and 1000.

$$= \begin{pmatrix} a_{k-1} \left(\frac{b_{1..(k-2)}}{c_{1..(k-2)}}\right)^{1/2} b_{k-1} \left(\frac{b_{1..(k-2)}}{c_{1..(k-2)}}\right)^{1/2} \\ c_{k-1} \left(\frac{b_{1..(k-1)}}{c_{1..(k-1)}}\right)^{1/2} a_{k} \left(\frac{b_{1..(k-1)}}{c_{1..(k-1)}}\right)^{1/2} \end{pmatrix} \begin{pmatrix} \left(\frac{c_{1..(k-2)}}{b_{1..(k-2)}}\right)^{1/2} \\ \left(\frac{c_{1..(k-1)}}{b_{1..(k-1)}}\right)^{1/2} \end{pmatrix},$$
$$= \begin{pmatrix} a_{k-1} & \sqrt{b_{k-1}c_{k-1}} \\ \sqrt{b_{k-1}c_{k-1}} & a_{k} \end{pmatrix}.$$

Now we can apply Theorem 4.1 to the matrix A_{sym} . Therefore, the Schur complements of A_{sym} converge to s(x). Since $A = D^{-1}A_{sym}D$, and as we saw in Section 3 such a diagonal transformation does not change the diagonal elements of the upper factor of A from that of A_{sym} , we conclude that the Schur complements of A converge to s(x). This finishes the proof of Corollary 4.1.



Fig. 9. Log error of the Schur complements to the limiting function s(x) with a(x) = 2 + sign(x - 0.5), b(x) = 0.5(1 - x), and $c(x) = |x - 0.75|^3$ for grid sizes of n = 100, 1000, 10,000 and 100,000.



Fig. 10. Schur complements with $p(x) = e^x$ for grid sizes of n = 100, 200 and 1000.

5. Numerical results

In this section we look at some numerical plots of the Schur complements of matrices generated by a variety of functions. Each case is accompanied by a figure that shows the distribution of the Schur complements as dashed lines, for different grid sizes. The solid-line curve denotes the limiting function s(x). There is also a figure of the \log_{10} error between the Schur complements at each point and the curve s(x).

5.1. Diagonally dominant matrices

Figs. 2 and 3 show the Schur complements of a matrix *A* with $a(x) = 5 + x^2$ and $b(x) = c(x) = 1 + e^{-x}$. Figs. 4 and 5 show the Schur complements of a matrix *A* with $a(x) = 3 + \cos(10\pi x)$ and $b(x) = c(x) = e^{-|\sin(\pi x)|}$, which makes the entries of *A* highly oscillatory.

Figs. 6 and 7 show the Schur complements of a matrix A with $a(x) = 2\sqrt{2 + |x - 0.5|}$, $b(x) = 1 + e^{-0.5x^2}$, and $c(x) = |\sin(2\pi x)|$. This is an interesting case in that both a(x) and b(x) have singularities



Fig. 11. Log error of the Schur complements with $p(x) = e^x$ for grid sizes of n = 100, 1000, 10,000 and 100,000.



Fig. 12. Schur complements with $p(x) = \sqrt{2 + sign(x - 0.5)}$ for grid sizes of n = 100, 200 and 1000.

in their derivatives. In fact, the derivative of *b* has a jump discontinuity at x = 0.5 and the derivative of *a* is blowing up at the same point.

Figs. 8 and 9 show the Schur complements of a matrix A with a(x) = 2 + sign(x - 0.5), b(x) = 0.5(1 - x), and $c(x) = |x - 0.75|^3$. Note that in this case, there is a jump discontinuity in the diagonal at x = 0.5.

5.2. Second order variable coefficient operator

Here we also include examples of the matrix equivalent of the equation

$$-\frac{d}{dx}\left(p(x)\frac{d}{dx}(u)\right) = f,$$



Fig. 13. Log error of the Schur complements with $p(x) = \sqrt{2 + sign(x - 0.5)}$ for grid sizes of n = 100, 1000, 10,000 and 100,000.

with $u(0) = u_0$ and $u(1) = u_1$ on [0, 1]. A second order discretization of this operator leads to a matrix of the form

$$A = \begin{pmatrix} p_{3/2} + p_{1/2} & -p_{3/2} \\ -p_{3/2} & p_{5/2} + p_{3/2} & -p_{5/2} \\ & \ddots & \ddots & \ddots \\ & & & & -p_{n-1/2} \\ & & & & -p_{n-1/2} & p_{n-1/2} + p_{n+1/2} \end{pmatrix}$$

where $p_{k+1/2} = p((k + 1/2)/(n + 1))$ (See [4]). Note that Assumption 4.2 is not valid for this case, and therefore the Proof of Theorem 4.1 does not hold. However, we can always gain the validity of Assumption 4.2 by adding an ϵ to the diagonal of *A*. For example, consider $A + \sqrt{\epsilon_{mach}}I$ where ϵ_{mach} is the machine precision.

6. Conclusions

In this paper we proved that it was possible for certain matrices to have Schur complements that exhibit limiting behavior as the discretization sizes go to zero. This opens up the possibility of interpolating the *LU* factors of the underlying operator. This property is currently being explored further for other classes of matrices, such as diagonal plus semiseparable matrices [11] which include the inverse of tridiagonal matrices. These matters will be published in future papers.

References

- [1] W.L. Briggs, V.E. Henson, S.F. McCormick, A Multigrid Tutorial, Society of Industrial and Applied Mathematics, 2000.
- [2] S. Chandrasekaran, P. Dewilde, M. Gu, N. Somasunderam, On the numerical ranks of the off-diagonal blocks of the Schur complements of discretized elliptic PDEs, SIAM J. Matrix Anal. Appl. 31 (5) (2010) 2261–2290.
- [3] G.H. Golub, C.F. Van Loan, Matrix Computations, third ed., John Hopkins University Press, 1996.
- [4] G.H. Golub, J.M. Ortega, Scientific Computing and Differential Equations, Academic Press, 1992.
- [5] N.J. Higham, Accuracy and Stability of Numerical Algorithmsm, second ed., SIAM, 2002.
- [6] M.R.S. Kulenovic, G. Ladas, Dynamics of Second Order Rational Difference Equations, with Open Problems and Conjectures, Chapman and Hall, 2002.
- [7] Carl D. Meyer, Matrix Analysis and Applied Linear Algebra, SIAM, 2000.

- [8] N. Somasunderam, LU factorizations in the limit of infinitesimal grids, Ph.D. Dissertation, University of California, Santa Barbara, 2010.
- [9] G.W. Stewart, On the Perturbation of LU, Cholesky, and QR Factorizations, SIAM J. Matrix Anal. 14 (4) (1993) 1141-1145.
- [10] J. Sun, Perturbation bounds for the Cholesky and QR factorizations, BIT Numer. Math. 31 (2) (1991) 341–352.
- [11] R. Vanderbril, M. Van Barel, N. Mastronardi, Matrix computations and semiseparable matrices, John Hopkins University Press, 2008.