ELSEVIER Linear Algebra and its Applications 287 (1999) 181–189

LINEAR ALGEBRA AND ITS APPLICATIONS

Stability of block $\mathrm{LDL}^{\mathrm{T}}$ factorization of a symmetric tridiagonal matrix

Nicholas J. Higham [1]

Department of Mathematics, University of Manchester, Manchester M13 9PL, UK

Received 30 September 1997; accepted 20 April 1998

Submitted by D.D. Olesky

Dedicated to Ludwig Elsner on the occasion of his 60th birthday

Abstract

For symmetric indefinite tridiagonal matrices, block $\mathrm{LDL}^{\mathrm{T}}$ factorization without interchanges is shown to have excellent numerical stability when a pivoting strategy of Bunch is used to choose the dimension (1 or 2) of the pivots. © 1999 Elsevier Science Inc. All rights reserved.

AMS classification: 65F05; 65G05

Keywords: Tridiagonal matrix; Symmetric indefinite matrix; Diagonal pivoting method; $\mathrm{LDL}^{\mathrm{T}}$ factorization; Growth factor; Numerical stability; Rounding error analysis; LAPACK; LINPACK

1. Introduction

Linear systems involving symmetric indefinite tridiagonal matrices arise in a number of situations. For example, Aasen's method with partial pivoting [1] produces a factorization $\Pi A \Pi^{\mathrm{T}} = LTL^{\mathrm{T}}$ of a symmetric matrix A, where Π is a permutation matrix, L is unit lower triangular, and T is tridiagonal. To solve a linear system $Ax = b$ using Aasen's method it is necessary to solve a system with coefficient matrix T. A recent application that produces linear systems with symmetric tridiagonal coefficient matrices is a Lanczos-based trust region method for unconstrained optimization of Gould et al. [8].

[1] E-mail: higham@ma.man.ac.uk.

PII: S0024-3795(98)10074-5

Symmetric tridiagonal linear systems are most commonly solved by Gaussian elimination with partial pivoting (GEPP) or by $\mathrm{LDL}^{\mathrm{T}}$ factorization without pivoting. Neither method is completely satisfactory. GEPP destroys the symmetry, and therefore cannot be used to determine the inertia, while an $\mathrm{LDL}^{\mathrm{T}}$ factorization yields the inertia of A directly from the diagonal of D, but can fail to exist and its computation can be numerically unstable when it does exist.

A method that promises to combine the benefits of GEPP and $\mathrm{LDL}^{\mathrm{T}}$ factorization was proposed by Bunch [3], but has received little attention in the literature. Bunch's idea is to compute a *block* $\mathrm{LDL}^{\mathrm{T}}$ factorization without interchanges, with a particular strategy for choosing the pivot size (1 or 2) at each stage of the factorization. Bunch's method requires less storage but slightly more computation than GEPP (see [3] for the details).

The purpose of this work is to examine the numerical stability of block $\mathrm{LDL}^{\mathrm{T}}$ factorization with Bunch's pivoting strategy. In Section 2 we define the pivoting strategy and explain how Bunch's derivation of it yields a bound of order 1 for the growth factor. In Section 3 we show that $\|L\|/\|A\|$ can be arbitrarily large and explain why numerical stability is therefore not a consequence of error analysis for general block LU factorization. We prove normwise backward stability of the method in Section 3, making use of results of Higham [10] on the stability of general block $\mathrm{LDL}^{\mathrm{T}}$ factorization.

2. Block $\mathrm{LDL}^{\mathrm{T}}$ factorization and the choice of pivot

Consider the computation of a block $\mathrm{LDL}^{\mathrm{T}}$ factorization without interchanges of a symmetric tridiagonal matrix $A \in \mathbb{R}^{n\times n}$. In the first stage of the factorization we choose an integer $s = 1$ or 2 and partition

$$A = \begin{array}{c c} & \begin{array}{cc} s & n-s \end{array} \\ \begin{array}{c} s \\ n-s \end{array} & \begin{bmatrix} E & C^{\mathrm{T}} \\ C & B \end{bmatrix} \end{array}. \tag{2.1}$$

If E is singular for both choices of s then $a_{11} = a_{21} = 0$, but $a_{21} = 0$ means that the first row and column is already in diagonal form and we can skip to the next stage of the factorization. Therefore, we can assume that E is nonsingular. Then we can factorize

$$A = \begin{bmatrix} I_s & 0 \\ CE^{-1} & I_{n-s} \end{bmatrix} \begin{bmatrix} E & 0 \\ 0 & B - CE^{-1}C^{\mathrm{T}} \end{bmatrix} \begin{bmatrix} I_s & E^{-1}C^{\mathrm{T}} \\ 0 & I_{n-s} \end{bmatrix}. \tag{2.2}$$

This process can be repeated recursively on the $(n-s)\times(n-s)$ Schur complement

$$S = B - CE^{-1}C^{\mathrm{T}}.$$

The result is a factorization

$$A = LDL^{\mathrm{T}}, \tag{2.3}$$

where L is unit lower triangular and D is block diagonal with each diagonal block having dimension 1 or 2. While the factorization always exists, whether it can be computed in a numerically stable way depends on the choice of pivots.

Bunch's strategy [3] for choosing the pivot size s at each stage of the factorization is fully defined by describing the choice of the first pivot.

Algorithm 1 (Bunch's pivoting strategy). This algorithm determines the pivot size, s, for the first stage of block $\mathrm{LDL}^{\mathrm{T}}$ factorization applied to a symmetric tridiagonal matrix $A \in \mathbb{R}^{n\times n}$.

$\sigma := \max\{|a_{ij}|: i,j = 1:n\}$ (compute once, at the start of the factorization)
$\alpha := (\sqrt{5}-1)/2 \approx 0.62$
if $\sigma|a_{11}| \geqslant \alpha a_{21}^2$
  $s = 1$
else
  $s = 2$
end

Bunch excludes a_{11} from the maximization defining σ; we find it more natural to include it, because it increases the probability that a 1×1 pivot will be chosen, while having no effect on the analysis below.

Bunch's choice of pivot can be explained by considering element growth in the factorization [3]. Since A is tridiagonal, the matrix C in (2.1) has the form $C = a_{s+1,s}e_1e_s^{\mathrm{T}}$, for unit vectors $e_1 \in \mathbb{R}^{n-s}$, $e_s \in \mathbb{R}^s$. Hence the Schur complement

$$S = B - a_{s+1,s}^2(e_s^{\mathrm{T}}E^{-1}e_s)e_1e_1^{\mathrm{T}}, \tag{2.4}$$

which shows that only the (1,1) element of the Schur complement differs from the corresponding element of A. We now examine the possible element growth in this position.

Consider first the case $s = 1$. We have

$$s_{11} = a_{22} - a_{21}^2/a_{11}.$$

Hence, from the conditions in Algorithm 1,

$$|s_{11}| \leqslant \sigma + \frac{\sigma}{\alpha}.$$

The choice $s = 2$ is made when

$$\sigma|a_{11}| < \alpha a_{21}^2. \tag{2.5}$$

For $s = 2$ we therefore have

$$\begin{aligned}\det(E) &= a_{11}a_{22} - a_{21}^2 \leqslant |a_{11}a_{22}| - a_{21}^2 \leqslant \alpha a_{21}^2|a_{22}|/\sigma - a_{21}^2 \\ &\leqslant (\alpha - 1)a_{21}^2 < 0,\end{aligned} \tag{2.6}$$

since $\alpha < 1$. Hence E is indefinite. We have

$$E^{-1} = \frac{1}{\det(E)}\begin{bmatrix} a_{22} & -a_{21} \\ -a_{21} & a_{11} \end{bmatrix} \tag{2.7}$$

and so, from (2.4), $s_{11} = a_{33} - a_{32}^2 a_{11}/\det(E)$. Hence, using (2.5),

$$|s_{11}| \leqslant |a_{33}| + \frac{a_{32}^2|a_{11}|}{(1-\alpha)a_{21}^2} \leqslant \sigma + \frac{\sigma^2\alpha}{(1-\alpha)\sigma} = \frac{\sigma}{1-\alpha}.$$

We have obtained bounds depending only on α and σ for the size of the (1,1) element of the Schur complement. This element is not subsequently modified and becomes a diagonal element of D. It follows that growth in any particular element takes place over a single stage of the factorization and is not cumulative. The value of α can therefore be determined by equating the maximal element growth for an $s = 1$ step with that for an $s = 2$ step. Hence we set

$$\sigma + \frac{\sigma}{\alpha} = \frac{\sigma}{1-\alpha},$$

which is a quadratic in α having the positive root $\alpha := (\sqrt{5} - 1)/2$. With α so chosen, the growth factor ρ_n for the factorization satisfies

$$\rho_n := \frac{\max_{i,j}|d_{ij}|}{\max_{i,j}|a_{ij}|} \leqslant \frac{1}{2}(\sqrt{5} + 3) \approx 2.62.$$

3. Error analysis

That the growth factor is nicely bounded does not, by itself, imply that computation of the block LDL^T factorization is a numerically stable process; see [10] for a discussion in the case of block LDL^T factorization of general symmetric matrices. From results on block LU factorization [6], numerical stability could be deduced if we could show that $\|L\|/\|A\|$ is suitably bounded. We therefore examine the size of the block CE^{-1} of L in (2.2). For $s = 1$ we have

$$\|CE^{-1}\|_\infty = \frac{|a_{21}|}{|a_{11}|} \leqslant \frac{\sigma}{\alpha|a_{21}|},$$

and the bound is sharp. It follows that $\|L\|/\|A\|$ can be arbitrarily large. A parametrized example is given by

$$A = \begin{bmatrix} \epsilon & \epsilon^{1/2} \\ \epsilon^{1/2} & 2 \end{bmatrix}, \qquad 0 \leqslant \epsilon \ll 1,$$

for which the first pivot is 1×1 and

$$L = \begin{bmatrix} 1 & 0 \\ \epsilon^{-1/2} & 1 \end{bmatrix}, \qquad D = \begin{bmatrix} \epsilon & 0 \\ 0 & 1 \end{bmatrix}, \quad \|L\|_\infty / \|A\|_\infty \approx \epsilon^{-1/2}/2.$$

For $s = 2$,

$$\begin{aligned} \|CE^{-1}\|_\infty &= |a_{32}| \|e_2^{\mathrm{T}} E^{-1}\|_1 \\ &\leqslant |a_{32}| \frac{(|a_{21}| + |a_{11}|)}{(1-\alpha)a_{21}^2} \\ &\leqslant \frac{|a_{32}|}{1-\alpha} \left(\frac{1}{|a_{21}|} + \frac{\alpha}{\sigma} \right), \end{aligned}$$

using (2.5) again. This bound is sharp and again it easy to construct a parametrized example in which $\|CE^{-1}\|_\infty / \|A\|_\infty$ can be arbitrarily large. We conclude that numerical stability does not follow from results on general block LU factorization.

Higham [10] proves the following general result. We employ the usual model of floating point arithmetic

$$fl(x \text{ op } y) = (x \text{ op } y)(1 + \delta), \qquad |\delta| \leqslant u, \quad \text{op} = +, *, /,$$

where u is the unit roundoff. Absolute values of matrices and inequalities between matrices are to be interpreted componentwise.

Theorem 3.1. *Let block* $\mathrm{LDL}^{\mathrm{T}}$ *factorization with any pivoting strategy be applied to a symmetric matrix* $A \in \mathbb{R}^{n \times n}$ *to yield the computed factorization* $\Pi A \Pi^{\mathrm{T}} \approx \hat{L}\hat{D}\hat{L}^{\mathrm{T}}$, *where* Π *is a permutation matrix and* D *has diagonal blocks of dimension* 1 *or* 2. *Let* $\hat{x}$ *be the computed solution to* $Ax = b$ *obtained using the factorization. Assume that for all linear systems* $Ey = f$ *involving* 2×2 *pivots* E *the computed solution* $\hat{x}$ *satisfies*

$$(E + \Delta E)\hat{y} = f, \qquad |\Delta E| \leqslant (cu + O(u^2))|E|, \tag{3.1}$$

where c *is a constant. Then,*

$$\Pi(A + \Delta A_1)\Pi^{\mathrm{T}} = \hat{L}\hat{D}\hat{L}^{\mathrm{T}}, \qquad (A + \Delta A_2)\hat{x} = b,$$

where

$$|\Delta A_i| \leqslant p(n)u\Big(|A| + \Pi^{\mathrm{T}}|\hat{L}||\hat{D}||\hat{L}^{\mathrm{T}}|\Pi\Big) + O(u^2), \qquad i = 1{:}2, \tag{3.2}$$

with p *a linear polynomial.*

For our tridiagonal A we can set the polynomial p in Theorem 3.1 to be of zero degree, and we have $\Pi = I$. However, to verify that the theorem is applicable, we have to check condition (3.1). It suffices to consider the first stage of the factorization. Suppose, first, that GEPP is used to solve $Ey = f$. For a 2×2 pivot E to be selected we must have

$$\sigma|a_{11}| < \alpha a_{21}^2 \leqslant \alpha|a_{21}|\sigma,$$

which implies

$$|a_{11}| < \alpha|a_{21}| < |a_{21}|.$$

Hence GEPP interchanges rows 1 and 2 of E and factorizes

$$PE = \begin{bmatrix} a_{21} & a_{22} \\ a_{11} & a_{21} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \frac{a_{11}}{a_{21}} & 1 \end{bmatrix} \begin{bmatrix} a_{21} & a_{22} \\ 0 & a_{21} - \frac{a_{11}a_{22}}{a_{21}} \end{bmatrix} = LU.$$

From ([9], Theorem. 9.4), we have the backward error result

$$(PE + \Delta E)\hat{y} = Pf, \qquad |\Delta E| \leqslant (6u + O(u^2))|\hat{L}||\hat{U}|.$$

Now, using (2.5),

$$|L||U| \leqslant \begin{bmatrix} |a_{21}| & |a_{22}| \\ |a_{11}| & \left|\frac{a_{11}a_{22}}{a_{21}}\right| + \left|a_{21} - \frac{a_{11}a_{22}}{a_{21}}\right| \end{bmatrix} \leqslant \begin{bmatrix} |a_{21}| & |a_{22}| \\ |a_{11}| & (2\alpha + 1)|a_{21}| \end{bmatrix}.$$

Hence $|L||U| \leqslant \sqrt{5}P|E|$. It follows that (3.1) holds with $c = 6\sqrt{5}$.

Another way to solve the linear systems $Ey = f$ is by the use of the explicit inverse, as is done in LINPACK [7] and LAPACK [2] in their implementations of block $\mathrm{LDL^T}$factorization with the pivoting strategyof Bunch and Kaufman [4] for general symmetric matrices. The formula used in LINPACK and LAPACK is suitable here too

$$y = \frac{1}{a_{21}\left(\frac{a_{11}}{a_{21}} \cdot \frac{a_{22}}{a_{21}} - 1\right)} \begin{bmatrix} \frac{a_{22}}{a_{21}} & -1 \\ -1 & \frac{a_{11}}{a_{21}} \end{bmatrix} f. \tag{3.3}$$

It is not hard to show that condition (3.1) holds when the formula (3.3) is used; the proof is very similar to that in [10] for the pivoting strategy of Bunch and Kaufman.

We have now established that Theorem 3.1 is applicable. To deduce stability of the factorization we have to show that $|L||D||L^T|$ is suitably bounded in norm (we have replaced the computed L and D by their exact counterparts, which affects only the second order term of (3.2)).

We write

$$
\begin{aligned}
|L||D||L^{\mathrm{T}}| &= \begin{bmatrix} I & \\ |L_{21}| & |L_S| \end{bmatrix} \begin{bmatrix} |E| & \\ & |D_S| \end{bmatrix} \begin{bmatrix} I & |L_{21}^{\mathrm{T}}| \\ & |L_S^{\mathrm{T}}| \end{bmatrix} \\
&= \begin{bmatrix} |E| & |E||L_{21}^{\mathrm{T}}| \\ |L_{21}||E| & |L_{21}||E||L_{21}^{\mathrm{T}}| + |L_S||D_S||L_S^{\mathrm{T}}| \end{bmatrix},
\end{aligned} \tag{3.4}
$$

where L_{21} and E are from the first stage of the factorization. We first bound

$$F := |L_{21}||E| \leq |C||E^{-1}||E|.$$

For $s = 1$ we have, trivially, $\|F\|_{\infty} = |a_{21}| \leqslant \sigma$. For $s = 2$, using (2.5)–(2.7),

$$
\begin{aligned}
|E^{-1}||E| &\leqslant \frac{1}{(1-\alpha)a_{21}^2} \begin{bmatrix} |a_{22}| & |a_{21}| \\ |a_{21}| & |a_{11}| \end{bmatrix} \begin{bmatrix} |a_{11}| & |a_{21}| \\ |a_{21}| & |a_{22}| \end{bmatrix} \\
&= \frac{1}{(1-\alpha)a_{21}^2} \begin{bmatrix} |a_{22}||a_{11}| + a_{21}^2 & 2|a_{22}||a_{21}| \\ 2|a_{21}||a_{11}| & a_{21}^2 + |a_{11}||a_{22}| \end{bmatrix} \\
&\leqslant \frac{1}{1-\alpha} \begin{bmatrix} \frac{|a_{22}|\alpha}{\sigma} + 1 & 2\frac{|a_{22}|}{|a_{21}|} \\ 2\frac{|a_{11}|}{|a_{21}|} & 1 + \frac{\alpha|a_{22}|}{\sigma} \end{bmatrix} \\
&\leqslant \frac{1}{1-\alpha} \begin{bmatrix} 1+\alpha & 2\frac{|a_{22}|}{|a_{21}|} \\ 2\frac{|a_{11}|}{|a_{21}|} & 1+\alpha \end{bmatrix}.
\end{aligned} \tag{3.5}
$$

Hence

$$
\begin{aligned}
\|F\|_{\infty} &\leqslant |a_{32}|\|e_1 e_2^{\mathrm{T}}|E^{-1}||E|\|_{\infty} \leqslant \frac{|a_{32}|}{1-\alpha}\left(2\frac{|a_{11}|}{|a_{21}|} + 1 + \alpha\right) \\
&\leqslant \frac{\sigma}{1-\alpha}\left(2\frac{\alpha|a_{21}|}{\sigma} + 1 + \alpha\right) \leqslant \frac{(3\alpha+1)\sigma}{1-\alpha} < 8\sigma.
\end{aligned} \tag{3.6}
$$

Now we bound $G := |L_{21}||E||L_{21}^{\mathrm{T}}|$. For $s = 1$,

$$\|G\|_{\infty} = \frac{a_{21}^2}{|a_{11}|} \leqslant \frac{\sigma}{\alpha} < 2\sigma.$$

For $s = 2$,

$$G \leqslant |C||E^{-1}||E||E^{-1}|C^{\mathrm{T}}| = a_{32}^2 e_2^{\mathrm{T}}(|E^{-1}||E||E^{-1}|)e_2 e_1 e_1^{\mathrm{T}}.$$

We bound the (2,2) element of $|E^{-1}||E||E^{-1}|$ starting with (3.5) and find that

$$\|G\|_{\infty} \leqslant \frac{a_{32}^2}{(1-\alpha)^2 a_{21}^2}(3+\alpha)|a_{11}| \leqslant \frac{(3+\alpha)a_{32}^2}{(1-\alpha)^2}\frac{\alpha}{\sigma} \leqslant \frac{(3+\alpha)\alpha}{(1-\alpha)^2}\sigma < 16\sigma. \tag{3.7}$$

We have now bounded all the terms in (3.4) except the term $|L_S||D_S||L_S^T|$. But L_S and D_S are block $\mathrm{LDL^T}$ factors of the Schur complement of D in A, and every Schur complement S satisfies

$$\|S\|_M \leqslant \rho_n \|A\|_M \leqslant 2.62\|A\|_M, \tag{3.8}$$

where

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

From the bounds (3.6)–(3.8) and the structure of the (2,2) block in (3.4) we deduce that

$$\||L||D||L^T|\|_M \leqslant 16 \times 2.62\|A\|_M < 42\|A\|_M.$$

The following result summarizes the stability of block $\mathrm{LDL^T}$ factorization with Bunch's pivoting strategy.

Theorem 3.2. *Let block* $\mathrm{LDL^T}$ *factorization with the pivoting strategy of Algorithm* 1 *be applied to a symmetric tridiagonal matrix* $A \in \mathbb{R}^{n\times n}$ *to yield the computed factorization* $A \approx \hat{L}\hat{D}\hat{L}^T$, *and let* $\hat{x}$ *be the computed solution to* $Ax = b$ *obtained using the factorization. Assume that all linear systems* $Ey = f$ *involving* 2×2 *pivots* E *are solved by GEPP or by using the explicit inverse formula* (3.3). *Then*

$$A + \Delta A_1 = \hat{L}\hat{D}\hat{L}^T, \qquad (A + \Delta A_2)\hat{x} = b,$$

where

$$\|\Delta A_i\|_M \leqslant cu\|A\|_M + O(u^2), \qquad i = 1{:}2, \tag{3.9}$$

with c *a constant.*

4. Conclusions

Theorem 3.2 shows that block $\mathrm{LDL^T}$ factorization with the pivoting strategy of Algorithm 1 is a normwise backward stable way to factorize a symmetric tridiagonal matrix A and to solve a linear system $Ax = b$. Block $\mathrm{LDL^T}$ factorization therefore provides an attractive alternative to GEPP for solving such linear systems.

Since the inertia of A is the same as that of the block diagonal factor D, the factorization also provides a normwise backward stable way to compute the inertia. However, for computing inertias of symmetric tridiagonal matrices standard $\mathrm{LDL^T}$ factorization without pivoting has the stronger componentwise relative form of backward stability ([5], Lemma 5.3), and so is preferable in the bisection method for computing eigenvalues, for example.

Acknowledgements

I thank Nick Gould for pointing out the open question of the numerical stability of block LDL^T factorization with Bunch's pivoting strategy.

References

[1] J.O. Aasen, On the reduction of a symmetric matrix to tridiagonal form, BIT 11 (1971) 233–242.

[2] E. Anderson, Z. Bai, C.H. Bischof, J.W. Demmel, J.J. Dongarra, J.J. Du Croz, A. Greenbaum, S.J. Hammarling, A. McKenney, S. Ostrouchov, D.C. Sorensen, LAPACK Users' Guide, Release 2.0. 2nd ed., Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 1995, xix + 325 pp. ISBN 089871345-5.

[3] J.R. Bunch, Partial pivoting strategies for symmetric matrices, SIAM J. Numer. Anal. 11 (3) (1974) 521–528.

[4] J.R. Bunch, L. Kaufman, Some stable methods for calculating inertia and solving symmetric linear systems, Math. Comp. 31 (137) (1977) 163–179.

[5] J.W. Demmel, Applied numerical linear algebra, Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 1997. xi+419 pp. ISBN 0-89871-389-7.

[6] J.W. Demmel, N.J. Higham, R.S. Schreiber, Stability of block *LU* factorization, Numerical Linear Algebra with Applications 2 (2) (1995) 173–190.

[7] J.J. Dongarra, J.R. Bunch, C.B. Moler, G.W. Stewart, LINPACK users' guide, Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 1979. ISBN 0-89871-172-X.

[8] N.I.M. Gould, S. Lucidi, M. Roma, P.L. Toint, Solving the trust-region subproblem using the Lanczos method, Report RAL-97-028, Atlas Centre, Rutherford Appleton Laboratory, Didcot, Oxon, UK, 1997, p. 28.

[9] N.J. Higham, Accuracy and stability of numerical algorithms, Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 1996, xxviii+688 pp. ISBN 0-89871-355-2.

[10] N.J. Higham, Stability of the diagonal pivoting method with partial pivoting, SIAM J. Matrix Anal. Appl. 18 (1) (1997) 52–65.