BIT ? (????), ???–???.

ON THE SENSITIVITY OF THE LU FACTORIZATION *

XIAO-WEN CHANG[1] and CHRISTOPHER C. PAIGE[2] †

[1] *Department of Computer Science, University of British Columbia, Vancouver, B. C. Canada V6T 1Z4. email: chang@cs.ubc.ca*

[2] *School of Computer Science, McGill University, Montreal, Quebec Canada H3A 2A7. email: chris@cs.mcgill.ca*

Abstract.

This paper gives sensitivity analyses by two approaches for L and U in the factorization $A = LU$ for general perturbations in A which are sufficiently small in norm. By the matrix-vector equation approach, we derive the condition numbers for the L and U factors. By the matrix equation approach we derive corresponding condition estimates. We show how partial pivoting and complete pivoting affect the sensitivity of the LU factorization.

AMS subject classification: 15A23, 65F35

Key words: LU factorization, Sensitivity, Condition number, Condition estimate, Pivoting.

1 Introduction

The LU factorization is a basic and effective tool in numerical linear algebra: given a real $n \times n$ matrix A whose first $n-1$ leading principal submatrices are all nonsingular, there exists a unique unit lower triangular matrix L and unique upper triangular matrix U such that

$$A = LU.$$

Notice here we require the diagonal elements of L to be 1. L and U are referred to as the LU factors.

Let ΔA be a sufficiently small $n \times n$ matrix such that the first $n-1$ leading principal submatrices of $A + \Delta A$ are still all nonsingular; then $A + \Delta A$ has a unique LU factorization

$$A + \Delta A = (L + \Delta L)(U + \Delta U).$$

*Received August 1997.

†The material presented here is a part of the first author's PhD Thesis [2] under the supervision of the second author. This research was supported by NSERC of Canada Grant OGP0009236.

The goal of the sensitivity analysis for the LU factorization is to determine a bound on $\|\Delta L\|$ (or $|\Delta L|$) and a bound on $\|\Delta U\|$ (or $|\Delta U|$) in terms of (a bound on) $\|\Delta A\|$ (or $|\Delta A|$).

The sensitivity analysis of the LU factorization has been considered by other authors. For the case when a bound is given on $\|\Delta A\|$, the first rigorous perturbation bounds on $\|\Delta L\|$ and $\|\Delta U\|$ were presented by Barrlund [1]. Using a different approach, Stewart [7] gave first-order perturbation bounds, which were recently improved by Stewart [8]. In [8], L was not assumed to be *unit* lower triangular, and a parameter p was used to control how much of the perturbation is attached to the diagonals of L and U. For the case when a bound is given on $|\Delta A|$, rigorous perturbation bounds on $|\Delta L|$ and $|\Delta U|$ were given by Sun [9].

The main purpose of this paper is to establish new first-order bounds, derive condition numbers, give new condition estimates, and shed light on the effect of partial pivoting and complete pivoting on the sensitivity of the LU factorization problem. We deal with the case when a bound is given on $\|\Delta A\|$ (norm-bounded perturbations). Our perturbation bounds and condition estimates for this case give improvements on those in [7] and [8], and probably this is the first time the actual condition numbers have been delineated in the literature.

The rest of this paper is organized as follows. In Section 2 we obtain expressions for $\dot{L}(0)$ and $\dot{U}(0)$ in the LU factorization $A+tG = L(t)U(t)$. These basic sensitivity expressions will be used to obtain our new perturbation bounds in Sections 3. In Section 3 we derive perturbation results, first by the so called matrix-vector equation approach, which leads to sharp bounds, then by the so called matrix equation approach, which leads to weaker but practical bounds. The basic ideas behind these two approaches were discussed in Chang, Paige and Stewart [4, 5], see also Chang [2]. This paper is essentially a rewrite of Chapter 4 of [2].

Throughout the paper, for a nonsingular matrix A, we use the notation $\kappa_p(A) \equiv \|A^{-1}\|_p\|A\|_p$ and $\mathrm{cond}_p(A) \equiv \|\,|A^{-1}|\cdot|A|\,\|_p$ for a consistent matrix norm $\|\cdot\|_p$.

2 Rate of change of L and U

To simplify the presentation, for any $n\times n$ matrix $X=(x_{ij})$, we define the *strictly lower triangular* matrix and *upper triangular* matrix

$$(2.1)\quad \mathrm{slt}(X)\equiv(s_{ij}),\quad s_{ij}\equiv\begin{cases} x_{ij} & \text{if } i>j \\ 0 & \text{otherwise}\end{cases},\qquad \mathrm{ut}(X)\equiv X-\mathrm{slt}(X).$$

Here we derive, for later use, the basic results on how L and U change as A changes. If A is singular then so is U. To handle this case we introduce a nonsingular $\bar{U}$ in the theorem.

THEOREM 2.1. *Let $A\in\mathbf{R}^{n\times n}$ have nonsingular leading $k\times k$ principal submatrices for $k=1,\ldots,n-1$, and the LU factorization $A=LU$, and let $\Delta A\in\mathbf{R}^{n\times n}$ satisfy $\Delta A=\epsilon G$. If ϵ is small enough such that the first $n-1$ leading principal submatrices of $A+tG$ are nonsingular for all $|t|\le\epsilon$, then $A+tG$ has the unique*

LU factorization

$$A + tG = L(t)U(t), \qquad |t| \le \epsilon, \tag{2.2}$$

which leads to

$$L\dot{U}(0) + \dot{L}(0)U = G, \tag{2.3}$$

$$\dot{L}(0) = L\,\mathrm{slt}(L^{-1}G\bar{U}^{-1}), \tag{2.4}$$

$$\dot{U}(0) = \mathrm{ut}(L^{-1}G\bar{U}^{-1})\bar{U}, \tag{2.5}$$

with $\bar{U} = U + (\alpha - u_{nn})e_n e_n^T$ *for some* $\alpha \neq 0$. *In particular,* $A + \Delta A$ *has the LU factorization*

$$A + \Delta A = (L + \Delta L)(U + \Delta U), \tag{2.6}$$

where ΔL *and* ΔU *satisfy*

$$\Delta L = \epsilon\, \dot{L}(0) + O(\epsilon^2), \tag{2.7}$$

$$\Delta U = \epsilon\, \dot{U}(0) + O(\epsilon^2). \tag{2.8}$$

Proof. Since the first $n-1$ leading principal submatrices of $A + tG$ are nonsingular for all $|t| \le \epsilon$, $A + tG$ has the unique LU factorization (2.2). Note that $L(0) = L$, $L(\epsilon) = L + \Delta L$, $U(0) = U$ and $U(\epsilon) = U + \Delta U$. When $t = \epsilon$, (2.2) becomes (2.6). It is easy to observe that $L(t)$ and $U(t)$ are continuously differentiable for $|t| \le \epsilon$ from a standard algorithm for the LU factorization. If we differentiate (2.2) and set $t = 0$ in the result, we obtain (2.3) which we will see is a linear equation *uniquely* defining the elements of strictly lower triangular $\dot{L}(0)$ and upper triangular $\dot{U}(0)$ in terms of the elements of G. Since $\dot{L}(0)e_n = 0$, (2.3) may be rewritten as $L\dot{U}(0) + \dot{L}(0)\bar{U} \equiv G$, giving

$$L^{-1}\dot{L}(0) + \dot{U}(0)\bar{U}^{-1} = L^{-1}G\bar{U}^{-1}.$$

Note that $L^{-1}\dot{L}(0)$ is strictly lower triangular and $\dot{U}(0)\bar{U}^{-1}$ is upper triangular, thus we have

$$L^{-1}\dot{L}(0) = \mathrm{slt}(L^{-1}G\bar{U}^{-1}), \qquad \dot{U}(0)\bar{U}^{-1} = \mathrm{ut}(L^{-1}G\bar{U}^{-1}),$$

which give (2.4) and (2.5). Finally the Taylor expansions for $L(t)$ and $U(t)$ about $t = 0$ give (2.7) and (2.8). If A, and so U, is nonsingular, we can replace $\bar{U}$ above by U. □

3 Main results

The basis for deriving first-order perturbation bounds is the equation (2.3), or (2.4) and (2.5). There are two ways to proceed. The matrix-vector equation approach (which here is based on (2.3)) will be used to provide sharp bounds, resulting in the condition numbers of the problem, while the matrix equation approach (based on (2.4) and (2.5)) can be used to obtain more practical bounds, resulting in easily computable upper bounds on the condition numbers.

Throughout this section we invoke all the assumptions in Theorem 2.1, so we can use its conclusions. Also we assume

$$\|\Delta A\|_F \le \epsilon \|A\|_F, \tag{3.1}$$

so using $\Delta A = \epsilon G$ as in Theorem 2.1 we see

$$\|G\|_F \le \|A\|_F \tag{3.2}$$

(if $\epsilon = 0$, all results we will present are obviously true). The one exception to (3.1) occurs in Remark 3.3 where we assume $\|\Delta A\|_{1,\infty} \le \epsilon \|A\|_{1,\infty}$.

3.1 Matrix-vector equation analysis and condition numbers

For any matrix $C \equiv (c_{ij}) \equiv [c_1, \ldots, c_n] \in \mathbf{R}^{n\times n}$, denote by $c_j^{(i)}$ the vector of the first i elements of c_j, and by $c_j^{\underline{(i)}}$ the vector of the last i elements of c_j. With these, we define ("u" denotes "upper", "sl" denotes "strictly lower")

$$\mathrm{vec}(C) \equiv \begin{bmatrix} c_1 \\ c_2 \\ \cdot \\ c_n \end{bmatrix}, \qquad \mathrm{uvec}(C) \equiv \begin{bmatrix} c_1^{(1)} \\ c_2^{(2)} \\ \cdot \\ c_n^{(n)} \end{bmatrix}, \qquad \mathrm{slvec}(C) \equiv \begin{bmatrix} c_1^{\underline{(n-1)}} \\ c_2^{\underline{(n-2)}} \\ \cdot \\ c_{n-1}^{\underline{(1)}} \end{bmatrix}.$$

These last two are the vectors formed by stacking the columns of the upper triangular part of C into one long vector and by stacking the columns of the strictly lower triangular part of C into one long vector, respectively.

Using this notation we will rewrite (2.3) in the form (3.3) below—the "matrix-vector equation" form. With obvious notation ($\dot{u}_j$ for $\dot{u}_j(0)$ etc) the jth column of $L\dot{U}(0) + \dot{L}(0)U = G$, which is (2.3), is

$$\begin{aligned} g_j &= [l_1, l_2, \ldots, l_n]\dot{u}_j + [\dot{l}_1, \dot{l}_2, \ldots, \dot{l}_n]u_j = [l_1, \ldots, l_j]\dot{u}_j^{(j)} + [\dot{l}_1, \ldots, \dot{l}_j]u_j^{(j)} \\ &= [l_1, \ldots, l_j]\dot{u}_j^{(j)} + \begin{bmatrix} 0 \\ u_{1j}I \end{bmatrix} \dot{l}_1^{\underline{(n-1)}} + \begin{bmatrix} 0 \\ 0 \\ u_{2j}I \end{bmatrix} \dot{l}_2^{\underline{(n-2)}} + \ldots + \begin{bmatrix} 0 \\ \cdot \\ 0 \\ u_{jj}I \end{bmatrix} \dot{l}_j^{\underline{(n-j)}}. \end{aligned}$$

It follows that the matrix equation (2.3) can be rearranged as

$$W \begin{bmatrix} \mathrm{uvec}(\dot{U}(0)) \\ \mathrm{slvec}(\dot{L}(0)) \end{bmatrix} = \mathrm{vec}(G), \tag{3.3}$$

where $W \equiv [W_L, W_U]$, with $W_L \in \mathbf{R}^{n^2 \times \frac{n(n+1)}{2}}$ being the n-block by n-block

$$
\left[\begin{array}{cccc}
\boxed{\begin{array}{c} 1 \\ l_{21} \\ \vdots \\ l_{n1} \end{array}} & & & \\
 & \boxed{\begin{array}{cc} 1 & \\ l_{21} & 1 \\ \vdots & \vdots \\ l_{n1} & l_{n2} \end{array}} & & \\
 & & \ddots & \\
 & & & \boxed{\begin{array}{ccccc} 1 & & & & \\ l_{21} & 1 & & & \\ \vdots & \vdots & \ddots & & \\ l_{n-1,1} & l_{n-1,2} & \cdots & 1 & \\ l_{n1} & l_{n2} & \cdots & l_{n,n-1} & 1 \end{array}}
\end{array}\right],
$$

and $W_U \in \mathbf{R}^{n^2 \times \frac{n(n-1)}{2}}$ being the n-block by $(n-1)$-block

$$
\left[\begin{array}{cccc|ccc|c|c}
0 & & & & & & & & \\
u_{11} & & & & & & & & \\
 & u_{11} & & & & & & & \\
 & & \ddots & & & & & & \\
 & & & u_{11} & & & & & \\
\hline
0 & & & & 0 & & & & \\
u_{12} & & & & 0 & & & & \\
 & u_{12} & & & u_{22} & & & & \\
 & & \ddots & & & \ddots & & & \\
 & & & u_{12} & & & u_{22} & & \\
\hline
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdots & \\
\hline
0 & & & & 0 & & & & \\
u_{1,n-1} & & & & 0 & & & & \\
 & u_{1,n-1} & & & u_{2,n-1} & & & & \\
 & & \ddots & & & \ddots & & \cdots & \\
 & & & u_{1,n-1} & & & u_{2,n-1} & & u_{n-1,n-1} \\
\hline
0 & & & & 0 & & & & \\
u_{1n} & & & & 0 & & & & \\
 & u_{1n} & & & u_{2n} & & & & \\
 & & \ddots & & & \ddots & & \cdots & \\
 & & & u_{1n} & & & u_{2n} & & u_{n-1,n}
\end{array}\right].
$$

It is easy to observe that after appropriate column permutations, $[W_L, W_U]$ will become lower triangular with diagonal elements

$$\underbrace{1, u_{11}, u_{11}, \cdots, u_{11}}_{n}, \underbrace{1, 1, u_{22}, \cdots, u_{22}}_{n}, \cdots\cdots, \underbrace{1, 1, \cdots, 1, u_{n-1,n-1}}_{n}, \underbrace{1, 1, \cdots, 1}_{n}.$$

Since the leading $(n-1)\times(n-1)$ block of U is nonsingular, W is also, and from (3.3) we have

$$\begin{bmatrix} \mathrm{uvec}(\dot{U}(0)) \\ \mathrm{slvec}(\dot{L}(0)) \end{bmatrix} = W^{-1}\mathrm{vec}(G). \tag{3.4}$$

If we partition W^{-1} into two blocks, we have from (3.4) that

$$W^{-1} \equiv \begin{bmatrix} Y_U \\ Y_L \end{bmatrix}, \tag{3.5}$$

$$\mathrm{slvec}(\dot{L}(0)) = Y_L\mathrm{vec}(G), \qquad \mathrm{uvec}(\dot{U}(0)) = Y_U\mathrm{vec}(G). \tag{3.6}$$

We see $\|\mathrm{slvec}(\dot{L}(0))\|_2 = \|\dot{L}(0)\|_F$ and $\|\mathrm{uvec}(\dot{U}(0))\|_2 = \|\dot{U}(0)\|_F$, and by using these norms we will be able to obtain tight bounds. Thus taking the 2-norm and using $\|G\|_F \le \|A\|_F$ we have

$$\|\dot{L}(0)\|_F \le \|Y_L\|_2\|G\|_F \le \|Y_L\|_2\|A\|_F, \tag{3.7}$$

$$\|\dot{U}(0)\|_F \le \|Y_U\|_2\|G\|_F \le \|Y_U\|_2\|A\|_F, \tag{3.8}$$

where individual equalities can be obtained by choosing G such that $\mathrm{vec}(G)$ lies in the space spanned by the right singular vectors corresponding to the largest singular value of either Y_L or Y_U, respectively, and $\|G\|_F = \|A\|_F$. Thus from the Taylor expansions (2.7) and (2.8), we obtain the following individually attainable first-order bounds,

$$\frac{\|\Delta L\|_F}{\|L\|_F} \le \frac{\|Y_L\|_2\|A\|_F}{\|L\|_F}\epsilon + O(\epsilon^2), \tag{3.9}$$

$$\frac{\|\Delta U\|_F}{\|U\|_F} \le \frac{\|Y_U\|_2\|A\|_F}{\|U\|_F}\epsilon + O(\epsilon^2). \tag{3.10}$$

These two sharp bounds imply that the condition numbers for the L factor and the U factor defined respectively by

$$\kappa_L(A) \equiv \lim_{\epsilon\to 0}\sup_{\Delta A}\{\frac{\|\Delta L\|_F}{\epsilon\,\|L\|_F} : A + \Delta A = (L+\Delta L)(U+\Delta U),\, \|\Delta A\|_F \le \epsilon\,\|A\|_F\},$$

$$\kappa_U(A) \equiv \lim_{\epsilon\to 0}\sup_{\Delta A}\{\frac{\|\Delta U\|_F}{\epsilon\,\|U\|_F} : A + \Delta A = (L+\Delta L)(U+\Delta U),\, \|\Delta A\|_F \le \epsilon\,\|A\|_F\}$$

are given by

$$\kappa_L(A) = \frac{\|Y_L\|_2\|A\|_F}{\|L\|_F}, \qquad \kappa_U(A) = \frac{\|Y_U\|_2\|A\|_F}{\|U\|_F}. \tag{3.11}$$

3.2 Matrix equation analysis and condition estimates

Given a bound on the norm of the change in A, in Section 3.1 we derived sharp perturbation bounds for the L factor and U factor, and presented the corresponding condition numbers. But it is difficult to estimate these condition numbers by using the usual approach. Now we add the matrix equation approach to this to derive more practical perturbation bounds.

First we derive a perturbation bound for the L factor. Let U_{n-1} denote the leading $(n-1)\times(n-1)$ block of U. If we write $U = \begin{bmatrix} U_{n-1} & u \\ 0 & u_{nn} \end{bmatrix}$, then from (2.4)

$$\dot{L}(0) = L\,\mathrm{slt}\Big(L^{-1}G\begin{bmatrix} U_{n-1}^{-1} & -U_{n-1}^{-1}u/\alpha \\ 0 & 1/\alpha \end{bmatrix}\Big) = L\,\mathrm{slt}\Big(L^{-1}G\begin{bmatrix} U_{n-1}^{-1} & 0 \\ 0 & 0 \end{bmatrix}\Big). \tag{3.12}$$

Denote by $\mathbf{D}_n$ the set of all $n\times n$ real positive definite diagonal matrices. Let $D = \mathrm{diag}(\delta_1,\ldots,\delta_n) \in \mathbf{D}_n$. Note that for any $n\times n$ matrix B we have $D\,\mathrm{slt}(B) = \mathrm{slt}(DB)$, so from (3.12) we obtain

$$\dot{L}(0) = LD^{-1}\,\mathrm{slt}\Big(DL^{-1}G\begin{bmatrix} U_{n-1}^{-1} & 0 \\ 0 & 0 \end{bmatrix}\Big). \tag{3.13}$$

Noting $\|\mathrm{slt}(B)\|_F \le \|B\|_F$ for any $B\in\mathbf{R}^{n\times n}$, we have

$$\|\dot{L}(0)\|_F \le \|LD^{-1}\|_2\|DL^{-1}\|_2\|U_{n-1}^{-1}\|_2\|G\|_F,$$

which with $\|G\|_F \le \|A\|_F$ in (3.2) gives

$$\frac{\|\dot{L}(0)\|_F}{\|L\|_F} \le \kappa_2(LD^{-1})\frac{\|U_{n-1}^{-1}\|_2\|A\|_F}{\|L\|_F}.$$

Since this is true for all $D\in\mathbf{D}_n$, by the Taylor expansion (2.7) we have

$$\frac{\|\Delta L\|_F}{\|L\|_F} \le \kappa_L'(A)\epsilon + O(\epsilon^2) \tag{3.14}$$

where

$$\kappa_L'(A) = \inf_{D\in\mathbf{D}_n}\kappa_L'(A,D), \tag{3.15}$$

$$\kappa_L'(A,D) \equiv \kappa_2(LD^{-1})\frac{\|U_{n-1}^{-1}\|_2\|A\|_F}{\|L\|_F}. \tag{3.16}$$

Since $\kappa_L(A)$ given by (3.11) is the condition number for the L factor, we certainly have

$$\kappa_L(A) \le \kappa_L'(A), \tag{3.17}$$

and as it is reasonably easy to find good approximations to $\kappa_L'(A)$, we can use this as a condition estimate.

Also we can obtain a lower bound on $\kappa_L(A)$. Let nonzero $v \in \mathbf{R}^{n-1}$ be such that $\|U_{n-1}^{-T}v\|_2 = \|U_{n-1}^{-1}\|_2\|v\|_2$. In (3.12), take $G = [e_n v^T, 0]$, where $e_n = (0,\ldots,0,1)^T \in \mathbf{R}^n$. Then it is easy to verify that

$$\dot{L}(0) = e_n[v^T U_{n-1}^{-1}, 0],$$

so

$$\|\dot{L}(0)\|_F = \|v^T U_{n-1}^{-1}\|_2 = \|U_{n-1}^{-1}\|_2\|v\|_2 = \|U_{n-1}^{-1}\|_2\|G\|_F.$$

Combining this with the first equality of (3.6), we have for this special G that

$$\|U_{n-1}^{-1}\|_2\|G\|_F = \|\dot{L}(0)\|_F = \|\mathrm{slvec}(\dot{L}(0))\|_2 = \|Y_L \mathrm{vec}(G)\|_2 \le \|Y_L\|_2\|G\|_F,$$

which gives the general result

$$\|Y_L\|_2 \ge \|U_{n-1}^{-1}\|_2, \tag{3.18}$$

or with (3.11)

$$\kappa_L(A) \equiv \frac{\|Y_L\|_2\|A\|_F}{\|L\|_F} \ge \frac{\|U_{n-1}^{-1}\|_2\|A\|_F}{\|L\|_F} \ge \sqrt{1-1/n}. \tag{3.19}$$

This last inequality follows from $A\,\mathrm{diag}(U_{n-1}^{-1}, 0) = [\bar{L}, 0]$, where $\bar{L}$ is the first $n-1$ columns of L, so $\|\bar{L}\|_F \le \|U_{n-1}^{-1}\|_2\,\|A\|_F$ and $\|L\|_F^2 = \|\bar{L}\|_F^2 + 1 \ge n$. Taking $A = \mathrm{diag}(I_{n-1}, 0)$ makes W in (3.3) a permutation matrix, giving $\kappa_L(A) = \sqrt{1-1/n}$ in (3.11), so the overall lower bound is attainable. We would like to point out that (3.18) can also be derived directly from the structure of W in (3.3).

Now we derive a practical perturbation bound for the U factor. In what follows we can replace $\bar{U}$ (see Theorem 2.1) by U when U is nonsingular. Let $D \in \mathbf{D}_n$. Notice for any $n \times n$ matrix B we have $\mathrm{ut}(BD) = \mathrm{ut}(B)D$, so from (2.5) we obtain

$$\dot{U}(0) = \mathrm{ut}(L^{-1}G\bar{U}^{-1}D)D^{-1}\bar{U}.$$

Thus

$$\|\dot{U}(0)\|_F \le \|L^{-1}\|_2\|\bar{U}^{-1}D\|_2\|D^{-1}\bar{U}\|_2\|G\|_F, \tag{3.20}$$

which with $\|G\|_F \le \|A\|_F$ gives

$$\frac{\|\dot{U}(0)\|_F}{\|U\|_F} \le \kappa_2(D^{-1}\bar{U})\frac{\|L^{-1}\|_2\|A\|_F}{\|U\|_F}.$$

Since this is true for all $D \in \mathbf{D}_n$, by the Taylor expansion (2.8) we have

$$\frac{\|\Delta U\|_F}{\|U\|_F} \le \kappa_U'(A)\epsilon + O(\epsilon^2), \tag{3.21}$$

where

$$\kappa_U'(A) = \inf_{D\in\mathbf{D}_n} \kappa_U'(A, D), \tag{3.22}$$

$$\kappa'_U(A,D) \equiv \kappa_2(D^{-1}\bar{U})\frac{\|L^{-1}\|_2\|A\|_F}{\|U\|_F}. \tag{3.23}$$

Note the freedom α in $\bar{U}$ does not affect the result of the optimization.

Since $\kappa_U(A)$ given by (3.11) is the condition number for the U factor, certainly from (3.21) we have

$$\kappa_U(A) \le \kappa'_U(A). \tag{3.24}$$

Also we can get a lower bound on $\kappa_U(A)$. Let nonzero $v \in \mathbf{R}^n$ be such that $\|L^{-1}v\|_2 = \|L^{-1}\|_2\|v\|_2$, and take $G = ve_n^T$ in (2.5) to give

$$\dot{U}(0) = \mathrm{ut}(L^{-1}ve_n^T\bar{U}^{-1})\bar{U} = L^{-1}ve_n^T.$$

Combining this with the second equality in (3.6) gives here

$$\begin{aligned} \|\dot{U}(0)\|_F &= \|L^{-1}\|_2\|v\|_2 = \|L^{-1}\|_2\|G\|_F \\ &= \|Y_U \mathrm{vec}(G)\|_F \le \|Y_U\|_2\|G\|_F \end{aligned}$$

and so in general

$$\|Y_U\|_2 \ge \|L^{-1}\|_2, \tag{3.25}$$

that is with (3.11)

$$\kappa_U(A) \ge \frac{\|L^{-1}\|_2\|A\|_F}{\|U\|_F} \ge 1. \tag{3.26}$$

Taking $A = I$ gives $\kappa_U(A) = 1$, so again this overall lower bound is attainable. Like (3.18), (3.25) can also be shown directly from the structure of W in (3.3).

These results, and the analysis in Section 3.1, can be summarized in our first-order perturbation theorem:

THEOREM 3.1. *Suppose all the assumptions of Theorem* 2.1 *hold and let* $\|\Delta A\|_F \le \epsilon\|A\|_F$*, then* $A + \Delta A$ *has the unique LU factorization*

$$A + \Delta A = (L + \Delta L)(U + \Delta U),$$

where with $\kappa_L(A)$ *and* $\kappa_U(A)$ *defined in* (3.11)*, we have for the L factor,*

$$\frac{\|\Delta L\|_F}{\|L\|_F} \le \kappa_L(A)\epsilon + O(\epsilon^2), \tag{3.27}$$

with bounds on $\kappa_L(A)$ *of:*

$$\sqrt{1-1/n} \le \frac{\|U_{n-1}^{-1}\|_2\|A\|_F}{\|L\|_F} \le \kappa_L(A) \le \kappa'_L(A) \equiv \inf_{D\in\mathbf{D}_n} \kappa_2(LD^{-1})\frac{\|U_{n-1}^{-1}\|_2\|A\|_F}{\|L\|_F}, \tag{3.28}$$

and for the U factor

$$\frac{\|\Delta U\|_F}{\|U\|_F} \le \kappa_U(A)\epsilon + O(\epsilon^2), \tag{3.29}$$

with bounds on $\kappa_U(A)$ *of:*

$$1 \le \frac{\|L^{-1}\|_2\|A\|_F}{\|U\|_F} \le \kappa_U(A) \le \kappa'_U(A) \equiv \inf_{D\in\mathbf{D}_n} \kappa_2(D^{-1}\bar{U})\frac{\|L^{-1}\|_2\|A\|_F}{\|U\|_F}. \tag{3.30}$$

REMARK 3.1. We might consider simplifying $\kappa'_L(A,D)$ and $\kappa'_U(A,D)$. If we use

$$\|A\|_F \leq \|L\|_F\|U\|_2,\ \|L\|_2\|U\|_F, \tag{3.31}$$

then we have from (3.16) and (3.23)

$$\kappa'_L(A,D) \leq \kappa_2(LD^{-1})\|U\|_2\|U_{n-1}^{-1}\|_2, \tag{3.32}$$

$$\kappa'_U(A,D) \leq \kappa_2(L)\kappa_2(D^{-1}\bar{U}). \tag{3.33}$$

But both of the right hand sides of these inequalities can be arbitrarily larger than the corresponding left hand sides due to the inequality (3.31). Note there can be large cancellation in the product $LU = A$.

REMARK 3.2. If we assume A is nonsingular, take $\bar{U} = U$ and $D = I$ in both $\kappa'_L(A,D)$ and $\kappa'_U(A,D)$, and use $\|L\|_2 \leq \|L\|_F$, $\|U\|_2 \leq \|U\|_F$ and $\|U_{n-1}^{-1}\|_2 \leq \|U^{-1}\|_2$, then from (3.16) and (3.23) we have

$$\kappa'_L(A) \leq \kappa'_L(A,I) \leq \kappa_2(L)\|U_{n-1}^{-1}\|_2\,\|A\|_F/\|L\|_F \leq \|L^{-1}\|_2\,\|U^{-1}\|_2\,\|A\|_F, \tag{3.34}$$

$$\kappa'_U(A) \leq \kappa'_U(A,I) \leq \kappa_2(U)\|L^{-1}\|_2\,\|A\|_F/\|U\|_F \leq \|L^{-1}\|_2\,\|U^{-1}\|_2\,\|A\|_F. \tag{3.35}$$

Thus it follows from Theorem 3.1 that

$$\frac{\|\Delta L\|_F}{\|L\|_F} \lesssim \|L^{-1}\|_2\|U^{-1}\|_2\|A\|_F\epsilon, \tag{3.36}$$

$$\frac{\|\Delta U\|_F}{\|U\|_F} \lesssim \|L^{-1}\|_2\|U^{-1}\|_2\|A\|_F\epsilon. \tag{3.37}$$

These perturbation bounds were obtained by Stewart [7]. They are simple, but can overestimate the true sensitivity of the problem. By using the scaling technology, Stewart [8] obtained significant improvements on the above results. In [8], the diagonal elements of L were not assumed to be 1's, and the diagonal elements of ΔL may not be 0's, and a parameter p was used to control how much of the perturbation is attached to the diagonals of L and U. The perturbation bounds given in [8] are equivalent to

$$\frac{\|\Delta L\|_F}{\|L\|_F} \lesssim \kappa_2(LD^{-1})\kappa_2(U)\epsilon, \tag{3.38}$$

$$\frac{\|\Delta U\|_F}{\|U\|_F} \lesssim \kappa_2(L)\kappa_2(D^{-1}U)\epsilon. \tag{3.39}$$

Note $\kappa_2(L)\kappa_2(D^{-1}U)$ in the second bound is equivalent to the right hand side of (3.33) with nonsingular U, while $\kappa_2(LD^{-1})k_2(U)$ in the first bound can be arbitrarily weaker than the right hand side of (3.32). But we have already seen those two can be arbitrarily weaker than (3.23) and (3.16), so while (3.38) and (3.39) are often useful, they can be unnecessarily weak in some circumstances.

REMARK 3.3. If we use the 1- and ∞-norms, we can get perturbation bounds without involving the scaling matrix D. In fact, suppose $\|\Delta A\|_p \leq \epsilon\|A\|_p$,

$p = 1, \infty$ instead of (3.1), then since $G = \Delta A/\epsilon$ we have $\|G\|_p \leq \|A\|_p$. From (3.12) we have

$$|\dot{L}(0)| \leq |L||L^{-1}||G| \begin{bmatrix} |U_{n-1}^{-1}| & 0 \\ 0 & 0 \end{bmatrix},$$

so taking the p-norm ($p = 1, \infty$) gives

$$\|\dot{L}(0)\|_p \leq \mathrm{cond}_p(L^{-1})\|U_{n-1}\|_p \, \|G\|_p \leq \mathrm{cond}_p(L^{-1})\|U_{n-1}\|_p \, \|A\|_p.$$

Then we have from the Taylor expansion (2.7)

$$\begin{aligned} \frac{\|\Delta L\|_p}{\|L\|_p} &\leq \mathrm{cond}_p(L^{-1})\frac{\|U_{n-1}^{-1}\|_p \, \|A\|_p}{\|L\|_p}\,\epsilon + O(\epsilon^2) \\ &\leq \mathrm{cond}_p(L^{-1})\|U\|_p \, \|U_{n-1}^{-1}\|_p \, \epsilon + O(\epsilon^2). \end{aligned} \tag{3.40}$$

Similarly from (2.5), (where if U is nonsingular $\mathrm{cond}_p(\bar{U}) = \mathrm{cond}_p(U)$)

$$\begin{aligned} \frac{\|\Delta U\|_p}{\|U\|_p} &\leq \mathrm{cond}_p(\bar{U})\frac{\|L^{-1}\|_p \, \|A\|_p}{\|U\|_p}\,\epsilon + O(\epsilon^2), \\ &\leq \kappa_p(L)\mathrm{cond}_p(\bar{U})\,\epsilon + O(\epsilon^2). \end{aligned} \tag{3.41}$$

Note $\mathrm{cond}_p(L^{-1})$ is invariant under any column scaling of L and $\mathrm{cond}_p(\bar{U})$ is invariant under any row scaling of $\bar{U}$ (so it is independent of the freedom α in $\bar{U}$).

As far as we know, it is expensive to estimate the condition numbers $\kappa_L(A)$ and $\kappa_U(A)$ in (3.11) directly by the usual approach except when A has some special structure (for example, A is tridiagonal, see Chang and Paige [3]). Fortunately we can estimate the condition estimates $\kappa'_L(A)$ and $\kappa'_U(A)$ reasonably efficiently. By a well-known result of van der Sluis [10], we have

$$\kappa_2(LD_L^{-1}) \leq \sqrt{n} \inf_{D \in \mathbf{D}_n} \kappa_2(LD^{-1}),$$

where $D_L \equiv \mathrm{diag}(\|L(:,j)\|_2)$. This is to say $\kappa_2(LD^{-1})$ will be near its infimum when each column of LD^{-1} has unit 2-norm. Therefore we have

$$\kappa'_L(A) \leq \kappa'_L(A, D_L) \leq \sqrt{n}\,\kappa'_L(A). \tag{3.42}$$

So in practice we choose $D = D_L$, then use a standard condition estimator and a norm estimator to estimate $\kappa'_L(A, D_L)$, which costs $O(n^2)$ flops. Similarly, we have

$$\kappa'_U(A) \leq \kappa'_U(A, D_U) \leq \sqrt{n}\,\kappa'_U(A), \tag{3.43}$$

where $D_U \equiv \mathrm{diag}(\|\bar{U}(i,:)\|_2)$, i.e., each row of $D_U^{-1}\bar{U}$ has unit 2-norm. So in practice we choose $D = D_U$, and use a standard condition estimator and a norm estimator to estimate $\kappa'_U(A, D_U)$, which costs $O(n^2)$ flops. Numerical experiments (see Section 5) show that $\kappa'_L(A, D_L)$ and $\kappa'_U(A, D_U)$ are good approximations to $\kappa_L(A)$ and $\kappa_U(A)$, respectively.

4 Effects of pivoting on sensitivity

It is well known that the standard algorithms for LU factorization without pivoting are not numerically stable. In order to repair this shortcoming, partial pivoting or complete pivoting should be incorporated in the computation. In this section we examine the effects these two pivoting strategies have on the sensitivity of the factorization. In other words we want to see how different the sensitivity of the LU factorization of A is from that of the LU factorization of the new matrix, A with its rows (and columns) permuted.

Suppose partial pivoting or complete pivoting is used in the LU factorization $PAQ = LU$, where P and Q are permutation matrices ($Q = I$ if partial pivoting is used) and $|l_{ij}| \leq 1$ for $i > j$. It is easy to show $|(L^{-1})_{ij}| \leq 2^{i-j-1}$ for $i > j$ (see Higham [6, p156]). Thus

$$1 \leq \kappa_2(L) \leq \|L\|_F \|L^{-1}\|_F \leq \sqrt{2n(n+1)(4^n + 6n - 1)}/6. \tag{4.1}$$

Then from (3.28) with $D = I$ we have

$$\begin{aligned}
\frac{\|U_{n-1}^{-1}\|_2 \|A\|_F}{\|L\|_F} &\leq \kappa_L(PAQ) \leq \kappa'_L(PAQ) \\
&\leq \frac{\sqrt{2n(n+1)(4^n + 6n - 1)}}{6} \frac{\|U_{n-1}^{-1}\|_2 \|A\|_F}{\|L\|_F},
\end{aligned} \tag{4.2}$$

so standard partial pivoting or complete pivoting keeps the condition number of L within a factor, only involving n, of its lower bound. But it is possible that such pivoting may cause $\|U_{n-1}^{-1}\|_2/\|L\|_F$ to become larger (in fact the crucial factor is $\|U_{n-1}^{-1}\|_2$ as $\sqrt{n} \leq \|L\|_F \leq \sqrt{n(n+1)/2}$). We cannot say the condition number $\kappa_L(PAQ)$ is larger or smaller than $\kappa_L(A)$.

For the U factor, when partial pivoting is used, noticing

$$1 \leq \|L^{-1}\|_2 \|A\|_F / \|U\|_F \leq \kappa_2(L) \leq \sqrt{2n(n+1)(4^n + 6n - 1)}/6,$$

we have from (3.30) that

$$1 \leq \kappa_U(PA) \leq \kappa'_U(PA) \leq \frac{\sqrt{2n(n+1)(4^n + 6n - 1)}}{6} \inf_{D \in \mathbf{D}_n} \kappa_2(D^{-1}\bar{U}). \tag{4.3}$$

Again from this bound we cannot say whether $\kappa_U(PA)$ is larger or smaller than $\kappa_U(A)$. But there is an essential difference between these upper bounds on $\kappa_U(PA)$ and $\kappa_L(PA)$. The former has a choice of D, which may make $\inf_{D \in \mathbf{D}_n} \kappa_2(D^{-1}\bar{U})$ not increase much. Furthermore if the ill-conditioning (with respect to inversion) of U is mostly due to the bad scaling of its rows, then $\inf_{D \in \mathbf{D}_n} \kappa_2(D^{-1}\bar{U})$ will be close to 1. In this case the U factor is not sensitive. When complete pivoting is used, the elements of U satisfy $|u_{ii}| \geq |u_{ij}|$ for all $j > i$. Then choosing $D = D_0 = \mathrm{diag}(u_{11}, \ldots, u_{n-1,n-1}, \alpha)$ and setting $\hat{U} \equiv D_0^{-1}\bar{U}$, we have $1 = |\hat{u}_{ii}| \geq |\hat{u}_{ij}|$ for all $j > i$. Thus just as in (4.1), we have

$$\kappa_2(D_0^{-1}\bar{U}) = \kappa_2(\hat{U}) \leq \sqrt{2n(n+1)(4^n + 6n - 1)}/6,$$

so from (4.3) we obtain the much better result

$$1 \leq \kappa_U(PAQ) \leq \kappa'_U(PAQ) \leq n(n+1)(4^n + 6n - 1)/18.$$

Note the upper bound is only a function of n, which suggests complete pivoting can give a significant improvement in $\kappa_U(PAQ)$ over $\kappa_U(A)$.

5 Numerical experiments

In Section 3 we presented first-order sharp perturbation bounds for the LU factors, obtained the corresponding condition numbers $\kappa_L(A)$ and $\kappa_U(A)$, and suggested $\kappa_L(A)$ and $\kappa_U(A)$ could be estimated in practice by $\kappa'_L(A, D_L)$ and $\kappa'_U(A, D_U)$ with $D_L = \mathrm{diag}(\|L(:,j)\|_2)$ and $D_U = \mathrm{diag}(\|\bar{U}(i,:)\|_2)$. The condition numbers and condition estimates satisfy the following inequalities (see (3.28), (3.30), (3.34), (3.35), (3.42), and (3.43)):

$$\|U_{n-1}^{-1}\|_2 \, \|A\|_F/\|L\|_F \leq \kappa_L(A) \leq \kappa'_L(A) \leq \|L^{-1}\|_2 \, \|U^{-1}\|_2 \, \|A\|_F, \tag{5.1}$$

$$\kappa'_L(A) \leq \kappa'_L(A, D_L) \leq \sqrt{n}\, \kappa'_L(A), \tag{5.2}$$

$$\|L^{-1}\|_2 \, \|A\|_F/\|U\|_F \leq \kappa_U(A) \leq \kappa'_U(A) \leq \|L^{-1}\|_2 \, \|U^{-1}\|_2 \, \|A\|_F, \tag{5.3}$$

$$\kappa'_U(A) \leq \kappa'_U(A, D_U) \leq \sqrt{n}\, \kappa'_U(A). \tag{5.4}$$

where for the last inequalities in (5.1) and (5.3) we assume A is nonsingular. In Section 4 we discussed the effect of partial pivoting and complete pivoting on the sensitivity of the LU factorization.

Now we give some numerical tests to illustrate our theoretical analyses. The matrices have the form $A = D_1 B D_2$, where $D_1 = \mathrm{diag}(1, d_1, \ldots, d_1^{n-1})$, $D_2 = \mathrm{diag}(1, d_2, \ldots, d_2^{n-1})$ and B is an $n \times n$ random matrix (produced by the MATLAB function `randn`). The results for $n = 10$; d_1, $d_2 \in \{0.2, 1, 2\}$; and the same matrix B, are shown in Table 5.1 without pivoting, in Table 5.2 with partial pivoting, and in Table 5.3 with complete pivoting, where $\beta_L \equiv \|U_{n-1}^{-1}\|_2 \, \|A\|_F/\|L\|_F$, $\beta_U \equiv \|L^{-1}\|_2 \, \|A\|_F/\|U\|_F$ are the lower bounds, and $\beta \equiv \|L^{-1}\|_2 \, \|U^{-1}\|_2 \, \|A\|_F$ is the upper bound in (5.1) and (5.3).

Table 5.1: Results without pivoting

d_1	d_2	β_L	$\kappa_L(A)$	$\kappa'_L(A, D_L)$	β_U	$\kappa_U(A)$	$\kappa'_U(A, D_U)$	β
0.2	0.2	2.4e+09	2.4e+09	1.7e+10	3.8e+00	8.7e+00	1.8e+01	1.5e+12
0.2	1	1.4e+05	1.7e+05	9.7e+05	2.8e+00	2.0e+02	7.3e+02	1.7e+07
0.2	2	5.3e+06	7.2e+06	3.8e+07	1.2e+00	2.7e+04	1.2e+05	7.9e+08
1	0.2	2.8e+03	3.4e+04	4.2e+05	4.9e+01	1.1e+02	2.3e+02	1.1e+07
1	1	4.5e+00	2.2e+02	6.7e+02	2.8e+00	1.9e+02	7.3e+02	4.1e+03
1	2	7.9e+02	3.7e+04	1.2e+05	1.5e+00	3.2e+04	1.4e+05	7.2e+05
2	0.2	2.6e+01	3.4e+04	1.0e+06	3.1e+05	3.2e+05	1.4e+06	2.6e+08
2	1	3.2e+00	3.2e+04	1.3e+05	3.2e+02	3.6e+03	8.2e+04	3.2e+07
2	2	2.7e+02	2.7e+06	1.0e+07	6.7e+01	1.6e+05	6.4e+06	2.7e+09

Table 5.2: Results with partial pivoting, $\tilde{A} \equiv PA$

d_1	d_2	β_L	$\kappa_L(\tilde{A})$	$\kappa'_L(\tilde{A}, D_L)$	β_U	$\kappa_U(\tilde{A})$	$\kappa'_U(\tilde{A}, D_U)$	β
0.2	0.2	3.5e+09	3.5e+09	8.5e+09	1.6e+00	1.7e+00	3.1e+00	6.6e+11
0.2	1	2.0e+05	2.4e+05	4.8e+05	1.6e+00	2.2e+01	8.8e+01	7.5e+06
0.2	2	7.7e+06	1.0e+07	1.9e+07	1.6e+00	3.5e+03	2.1e+04	3.4e+08
1	0.2	1.1e+05	2.1e+05	6.2e+05	4.7e+00	4.7e+00	6.7e+00	3.9e+06
1	1	7.1e+00	1.3e+01	4.0e+01	2.5e+00	1.2e+01	4.3e+01	8.5e+01
1	2	8.0e+02	1.4e+03	4.5e+03	1.6e+00	1.5e+03	6.0e+03	9.8e+03
2	0.2	8.2e+06	1.2e+07	2.7e+07	2.1e+00	2.1e+00	3.2e+00	2.6e+08
2	1	4.9e+02	6.3e+02	1.6e+03	1.7e+00	1.8e+01	7.8e+01	5.0e+03
2	2	2.2e+04	2.7e+04	7.4e+04	1.7e+00	2.7e+03	1.4e+04	1.7e+05

Table 5.3: Results with complete pivoting, $\hat{A} \equiv PAQ$

d_1	d_2	β_L	$\kappa_L(\hat{A})$	$\kappa'_L(\hat{A}, D_L)$	β_U	$\kappa_U(\hat{A})$	$\kappa'_U(\hat{A}, D_U)$	β
0.2	0.2	3.5e+09	3.5e+09	8.5e+09	1.6e+00	1.7e+00	3.1e+00	6.6e+11
0.2	1	1.4e+05	1.4e+05	2.0e+05	1.2e+00	2.5e+00	6.6e+00	5.6e+06
0.2	2	2.6e+06	2.6e+06	5.8e+06	1.4e+00	1.5e+00	4.4e+00	2.9e+08
1	0.2	1.1e+05	2.1e+05	6.2e+05	4.7e+00	4.7e+00	6.7e+00	3.9e+06
1	1	2.8e+00	5.0e+00	1.9e+01	3.4e+00	4.9e+00	1.4e+01	6.7e+01
1	2	1.4e+02	3.3e+02	1.2e+03	5.0e+00	7.0e+00	1.6e+01	5.8e+03
2	0.2	8.2e+06	1.2e+07	2.7e+07	2.1e+00	2.1e+00	3.2e+00	2.6e+08
2	1	3.1e+02	3.4e+02	9.6e+02	1.8e+00	3.0e+00	1.3e+01	3.9e+03
2	2	1.1e+04	1.2e+04	4.4e+04	2.2e+00	3.0e+00	7.5e+00	1.3e+05

We give some comments on the results.

- The results confirm that $\beta = \|L^{-1}\|_2 \|U^{-1}\|_2 \|A\|_F$ can be much larger than $\kappa_L(A)$ and $\kappa_U(A)$, especially for the latter, so the first-order bounds (3.36) and (3.37) can significantly overestimate the true sensitivities of the L and U factors.
- $\kappa'_L(A, D_L)$ and $\kappa'_U(A, D_U)$ are good approximations here of $\kappa_L(A)$ and $\kappa_U(A)$, respectively, no matter whether pivoting is used or not. This was also the case in our other numerical experiments.
- Both $\kappa_L(PA)$ and $\kappa_L(PAQ)$ can be much larger or smaller than $\kappa_L(A)$. So partial pivoting and complete pivoting can make the L factor more sensitive or less sensitive. But from Tables 5.1–5.2 we see that partial pivoting can give a significant improvement on the condition of the U factor. In fact here $\kappa_U(PA) \leq \kappa_U(A)$ for all cases. From Table 5.3 we see that complete pivoting can give a more significant improvement. This strongly suggests we use complete pivoting when we want an accurately computed U using finite precision.
- It can be seen that for most cases the L factor is more sensitive than the U factor no matter whether pivoting is used or not.

- When partial pivoting is used, we see that κ_L is close to its lower bound β_L. When complete pivoting is used, we see that both κ_L and κ_U are close to their lower bounds β_L and β_U, respectively.

6 Summary and future work

The first-order perturbation analyses presented here show what the sensitivity of each of L and U is in the LU factorization of A, and in so doing provide their condition numbers $\kappa_L(A)$ and $\kappa_U(A)$ (with respect to the measures used, and for sufficiently small ΔA), as well as efficient ways of approximating them.

As we know, $\kappa_2(L)$ is usually (much) smaller than $\kappa_2(U)$, especially in practice when we use partial pivoting in computing the LU factorization. So we can expect that the computed solution of the linear system $Lx = b$ will usually be more accurate than that of the linear system $Uy = b$. However our analysis and numerical experiments suggest that usually the L factor is more sensitive than the U factor in the LU factorization, so we expect U to be more accurate than L. This is an interesting phenomenon. Also we see the effect of partial pivoting and complete pivoting on the sensitivity of L is uncertain—both $\kappa_L(PA)$ and $\kappa_L(PAQ)$ can be much larger or smaller than $\kappa_L(A)$. But partial pivoting can usually make U less sensitive, and complete pivoting can give significant improvement.

In the future we would like to investigate the ratios (see (3.28) and (3.30)) $\kappa_L(A)/\kappa'_L(A)$ and $\kappa_U(A)/\kappa'_U(A)$, and extend our analysis to the case where $|\Delta A| \leq \epsilon|A|$ and to the case where ΔA has the equivalent form of backward errors resulting from standard algorithms for the LU factorization.

Acknowledgement.

We would like to thank Jim Varah for helpful discussions on this and for carefully reading the manuscript.

REFERENCES

1. A. Barrlund, *Perturbation bounds for the LDL^H and the LU factorizations*, BIT, 31 (1991), pp. 358–363.

2. X.-W. Chang, *Perturbation Analysis of Some Matrix Factorizations*, PhD thesis, Computer Science, McGill University, Montreal, Canada, February 1997.

3. X.-W. Chang and C. C. Paige, *Sensitivity Analyses for Factorizations of Sparse or Structured Matrices*, Linear Algebra Appl.. 20 pages, submitted in September 1997.

4. X.-W. Chang, C. C. Paige, and G. W. Stewart, *New perturbation analyses for the cholesky factorization*, IMA J. Numer. Anal., 16 (1996), pp. 457–484.

5. X.-W. CHANG, C. C. PAIGE, AND G. W. STEWART, *Perturbation analyses for the QR factorization*, SIAM J. Matrix Anal. Appl., 18 (1997), pp. 775–791.

6. N. J. HIGHAM, *Accuracy and Stability of Numerical Algorithms*, SIAM, Philadelphia, 1996.

7. G. W. STEWART, *On the perturbation of LU, Cholesky, and QR factorizations*, SIAM J. Matrix Anal. Appl., 14 (1993), pp. 1141–1145.

8. ———, *On the perturbation of LU and Cholesky factors*, IMA J. Numer. Anal., 17 (1997), pp. 1–6.

9. J.-G. SUN, *Componentwise perturbation bounds for some matrix decompositions*, BIT, 32 (1992), pp. 702–714.

10. A. VAN DER SLUIS, *Condition numbers and equilibration of matrices*, Numer. Math., 14 (1969), pp. 14–23.