Perturbation Analyses for the Cholesky Downdating Problem

Xiao-Wen Chang   Christopher C. Paige

Technical Report No. SOCS-96.2
May 1996

PERTURBATION ANALYSES FOR THE CHOLESKY DOWNDATING PROBLEM *

XIAO-WEN CHANG† AND CHRISTOPHER. C. PAIGE†

Abstract. New perturbation analyses are presented for the block Cholesky downdating problem $U^TU = R^TR - X^TX$. These show how changes in R and X alter the Cholesky factor U. There are two main cases for the perturbation matrix ΔR in R: (1) ΔR is a general matrix; (2)ΔR is an upper triangular matrix. For both cases, first order perturbation bounds for the downdated Cholesky factor U are given using two approaches — a detailed "matrix-vector equation" analysis which provides tight bounds and resulting true condition numbers, which unfortunately are costly to compute, and a simpler "matrix-equation" analysis which provides results that are weaker but easier to compute or estimate. The analyses more accurately reflect the sensitivity of the problem than previous results. As $X \to 0$, the asymptotic values of the new condition numbers for case (1) have bounds that are independent of $\kappa_2(R)$ if R was found using the standard pivoting strategy in the Cholesky factorization, and the asymptotic values of the new condition numbers for case (2) are unity. Simple reasoning shows this last result must be true for the sensitivity of the problem, but previous condition numbers did not exhibit this.

Key Words. perturbation analysis, sensitivity, condition, asymptotic condition, Cholesky factorization, downdating

AMS Subject Classifications: 15A23, 65F35

1. Introduction. Let $A \in \mathcal{R}^{n\times n}$ be a symmetric positive definite matrix. Then there exists a unique upper triangular matrix $R \in \mathcal{R}^{n\times n}$ with positive diagonal elements such that $A = R^TR$. This factorization is called the Cholesky factorization, and R is called the Cholesky factor of A.

In this paper we give perturbation analyses of the following problem: given an upper triangular matrix $R \in \mathcal{R}^{n\times n}$ and a matrix $X \in \mathcal{R}^{k\times n}$ such that $R^TR - X^TX$ is positive definite, find an upper triangular matrix $U \in \mathcal{R}^{n\times n}$ with positive diagonal elements such that

$$U^TU = R^TR - X^TX. \tag{1.1}$$

This problem is called the block Cholesky downdating problem, and the matrix U is referred to as the downdated Cholesky factor. The block Cholesky downdating problem has many important applications, and the case for k=1 has been extensively studied in the literature (see [1, 2, 3, 8, 11, 12, 16, 17, 18]).

Perturbation results for the single Cholesky downdating problem were presented by Stewart [18]. Eldén and Park [10] made an analysis for block downdating. But these two papers just considered the case that only R or X is perturbed. More complete analyses, with both R and X being perturbed, were given by Pan [15] and Sun [20]. Pan [15] gave first order perturbation bounds for single downdating. Sun [20] gave strict, also first order perturbation bounds for single downdating and first order perturbation bounds for block downdating.

The main purpose of this paper is to establish new first order perturbation results and present new condition numbers which more closely reflect the true sensitivity of

* This research was partially supported by NSERC of Canada Grant OGP0009236.

† School of Computer Science, McGill University, Montreal, Quebec, Canada, H3A 2A7, (chang@cs.mcgill.ca), (chris@cs.mcgill.ca).

the problem. In Section 2 we will give the key result of Sun [20], and a new result using the approach of these earlier papers. In Section 3 we present new perturbation results, first by the straightforward matrix equation approach, then by the more detailed and tighter matrix–vector equation approach. The basic ideas behind these two approaches were discussed in Chang, Paige and Stewart [6, 7]. We give numerical results and suggest practical condition estimators in Section 4.

Previous papers implied the change ΔR in R was upper triangular, and Sun [20] said this, but neither he nor the others made use of this. In fact a backward stable algorithm for computing U given R and X would produce the exact result $U_c = U+\Delta U$ for nearby data $R+\Delta R$ and $X+\Delta X$, where it is not clear that ΔR would be upper triangular — the form of the equivalent backward error ΔR would depend on the algorithm, and if it were upper triangular, it would require a rounding error analysis to show this. Thus for completeness it seems necessary to consider two separate cases — upper triangular ΔR and general ΔR. We do this throughout Sections 3–4, and get stronger results for upper triangular ΔR than in the general case.

In any perturbation analysis it is important to examine how good the results are. In Section 3.2 we produce provably tight bounds, leading to the true condition numbers (for the norms chosen). The numerical example in Section 4 indicates how much better the results of this new analysis can be compared with some earlier ones, but a theoretical understanding is also desirable. By considering the asymptotic case as $X \to 0$, the results simplify, and are easily understandable. We show the new results have the correct properties as $X \to 0$, in contrast to earlier results.

Before proceeding, let us introduce some notation. Let $B = (b_{ij}) \in \mathcal{R}^{n\times n}$, then $\mathrm{up}(B)$, $\mathrm{sut}(B)$, $\mathrm{slt}(B)$ and $\mathrm{diag}(B)$ are defined by

$$\mathrm{up}(B) \equiv \begin{pmatrix} \frac{1}{2}b_{11} & b_{12} & \cdot & b_{1n} \\ 0 & \frac{1}{2}b_{22} & \cdot & b_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \frac{1}{2}b_{nn} \end{pmatrix}, \tag{1.2}$$

$$\mathrm{sut}(B) \equiv \begin{pmatrix} 0 & b_{12} & b_{13} & \cdot & b_{1n} \\ 0 & 0 & b_{23} & \cdot & b_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & 0 & b_{n-1,n} \\ 0 & \cdot & \cdot & \cdot & 0 \end{pmatrix}, \qquad \mathrm{slt}(B) \equiv \mathrm{sut}(B^T)^T, \tag{1.3}$$

and

$$\mathrm{diag}(B) = \mathrm{diag}(b_{11}, b_{22}, \ldots, b_{nn}). \tag{1.4}$$

2. Previous perturbation results, and an improvement. The condition number for downdating presented by Pan [15] involves the square of the condition number of R, $\kappa_2^2(R)$, where $\kappa_2(R) \equiv \|R\|_2\|R^{-1}\|_2$. Sun [20] proposed new condition numbers which are simple and proportional to $\kappa_2(R)$. The condition number for the block Cholesky downdating problem proposed in [20] is

$$\beta_1 \equiv \sqrt{2}\frac{\kappa_2(U)}{\sigma_n^2(\Gamma)},$$

where Γ satisfies $\Gamma^T\Gamma = I_n - R^{-T}X^TXR^{-1}$, and $\sigma_n(\Gamma)$ is the smallest singular value of Γ. Notice that for fixed R, $\Gamma^T\Gamma \to I_n$ as $X \to 0$, so $\sigma_n(\Gamma) \to 1$. Now we use a similar approach to derive a new bound, from which Sun's bound follows.

First we derive some relationships among U, R, X and Γ.

1) From (1.1) obviously we have

$$\|U\|_2 \le \|R\|_2, \quad \|X\|_2 \le \|R\|_2. \tag{2.1}$$

2) From (1.1) it follows that

$$RU^{-1}U^{-T}R^T = (I_n - R^{-T}X^TXR^{-1})^{-1},$$

so that taking the 2-norm gives

$$\|RU^{-1}\|_2 = \frac{1}{\sigma_n(\Gamma)}. \tag{2.2}$$

3) From (1.1) we have

$$U^{-T}X^TXU^{-1} = U^{-T}R^TRU^{-1} - I_n,$$

which, combined with (2.2), gives

$$\|XU^{-1}\|_2 = \sqrt{\|RU^{-1}\|_2^2 - 1} = \frac{\sqrt{1-\sigma_n^2(\Gamma)}}{\sigma_n(\Gamma)} = \sqrt{1-\sigma_n^2(\Gamma)}\|RU^{-1}\|_2. \tag{2.3}$$

4) From (1.1) we have

$$R^{-T}X^TXR^{-1} = I_n - R^{-T}U^TUR^{-1},$$

which, combined with (2.2), gives

$$\|XR^{-1}\|_2 = \sqrt{1-\sigma_{\min}^2(UR^{-1})} = \sqrt{1 - \frac{1}{\|RU^{-1}\|_2^2}} = \sqrt{1-\sigma_n^2(\Gamma)}. \tag{2.4}$$

5) By (2.2) we have

$$\|R\|_2 = \|RU^{-1}U\|_2 \le \|RU^{-1}\|_2\|U\|_2 = \frac{\|U\|_2}{\sigma_n(\Gamma)}. \tag{2.5}$$

6) Finally from (2.4) we see

$$\frac{\|X\|_2}{\|R\|_2} \le \|XR^{-1}\|_2 = \sqrt{1-\sigma_n^2(\Gamma)}. \tag{2.6}$$

To derive first order perturbation results we consider the perturbed version of (1.1):

$$(U+\Delta U)^T(U+\Delta U) = (R+\Delta R)^T(R+\Delta R) - (X+\Delta X)^T(X+\Delta X), \tag{2.7}$$

where U, $U+\Delta U$ and R are upper triangular matrices with positive diagonal elements. Clearly when ΔR and ΔX are sufficiently small, (2.7) has a unique solution ΔU.

Multiplying out the two sides of (2.7) and ignoring second order terms, we obtain a linear matrix equation for the first order approximation $\widehat{\Delta U}$ to ΔU:

$$U^T \widehat{\Delta U} + \widehat{\Delta U}^T U = R^T \Delta R + \Delta R^T R - X^T \Delta X - \Delta X^T X. \tag{2.8}$$

In fact it is straightforward to show $\widehat{\Delta U} = \dot{U}(0)$, the rate of change of $U(\tau)$ at $\tau = 0$, where $U(\tau)^T U(\tau) \equiv (R + \tau\Delta R)^T (R + \tau\Delta R) - (X + \tau\Delta X)^T (X + \tau\Delta X)$, $0 \le \tau \le 1$, so $\widehat{\Delta U}$ also has a precise meaning. From (2.8), we have

$$\widehat{\Delta U} U^{-1} + (\widehat{\Delta U} U^{-1})^T = U^{-T}(R^T \Delta R + \Delta R^T R - X^T \Delta X - \Delta X^T X)U^{-1}.$$

Notice since $\widehat{\Delta U} U^{-1}$ is upper triangular, it follows, with (1.2), that

$$\widehat{\Delta U} = \mathrm{up}[U^{-T}(R^T \Delta R + \Delta R^T R - X^T \Delta X - \Delta X^T X)U^{-1}]U. \tag{2.9}$$

But for any symmetric matrix B,

$$2\|\mathrm{up}(B)\|_F^2 = \|B\|_F^2 - \frac{1}{2}(b_{11}^2 + b_{22}^2 + \cdots + b_{nn}^2) \le \|B\|_F^2.$$

Thus from (2.9) we have

$$\begin{aligned}\|\widehat{\Delta U}\|_F &\le \frac{1}{\sqrt{2}} \|U^{-T}(R^T \Delta R + \Delta R^T R - X^T \Delta X - \Delta X^T X)U^{-1}\|_F \|U\|_2 \\ &\le \sqrt{2}\|U\|_2 \|U^{-1}\|_2 (\|RU^{-1}\|_2 \|\Delta R\|_F + \|XU^{-1}\|_2 \|\Delta X\|_F),\end{aligned}$$

which, combined with (2.2) and (2.3), gives

$$\|\widehat{\Delta U}\|_F \le \frac{\sqrt{2}\|U\|_2 \|U^{-1}\|_2}{\sigma_n(\Gamma)} (\|\Delta R\|_F + \sqrt{1 - \sigma_n^2(\Gamma)} \|\Delta X\|_F),$$

resulting in the new perturbation bound for relative changes

$$\frac{\|\widehat{\Delta U}\|_F}{\|U\|_2} \le \sqrt{2} \frac{\|U^{-1}\|_2 \|R\|_2}{\sigma_n(\Gamma)} \frac{\|\Delta R\|_F}{\|R\|_2} + \sqrt{2} \frac{\sqrt{1 - \sigma_n^2(\Gamma)} \|U^{-1}\|_2 \|X\|_2}{\sigma_n(\Gamma)} \frac{\|\Delta X\|_F}{\|X\|_2}, \tag{2.10}$$

which leads to the condition numbers for the Cholesky downdating problem:

$$\beta_R \equiv \sqrt{2} \frac{\|U^{-1}\|_2 \|R\|_2}{\sigma_n(\Gamma)}, \quad \beta_X \equiv \sqrt{2} \frac{\sqrt{1 - \sigma_n^2(\Gamma)} \|U^{-1}\|_2 \|X\|_2}{\sigma_n(\Gamma)}$$

for U with respect to relative changes in R and X, respectively. Notice from (2.1) $\beta_R > \beta_X$. So we can define a new overall condition number

$$\beta_2 \equiv \beta_R = \sqrt{2} \frac{\|U^{-1}\|_2 \|R\|_2}{\sigma_n(\Gamma)}. \tag{2.11}$$

Rewriting (2.10) as

$$\frac{\|\widehat{\Delta U}\|_F}{\|U\|_2} \le \sqrt{2} \frac{\|U^{-1}\|_2 \|R\|_2}{\sigma_n(\Gamma)} \left(\frac{\|\Delta R\|_F}{\|R\|_2} + \sqrt{1 - \sigma_n^2(\Gamma)} \frac{\|X\|_2}{\|R\|_2} \frac{\|\Delta X\|_F}{\|X\|_2} \right),$$

and combining it with (2.5) and (2.6), gives Sun's bound

$$\frac{\|\widehat{\Delta U}\|_F}{\|U\|_2} \le \sqrt{2}\frac{\kappa_2(U)}{\sigma_n^2(\Gamma)}\left(\frac{\|\Delta R\|_F}{\|R\|_2} + (1-\sigma_n^2(\Gamma))\frac{\|\Delta X\|_F}{\|X\|_2}\right). \tag{2.12}$$

We have seen the right hand side of (2.10) is never worse than that of (2.12), so

$$\beta_2 \le \beta_1. \tag{2.13}$$

Although β_2 is a minor improvement on β_1, it is still not what we want. We can see this from the asymptotic behavior of these "condition numbers". The Cholesky factorization is unique, so as $X \to 0$, $U \to R$, and $X^T\Delta X \to 0$ in (2.8). Now for any upper triangular perturbation ΔR in R, $\Delta U \to \Delta R$, so the true condition number should approach unity. Here $\beta_1, \beta_2 \to \sqrt{2}\kappa_2(R)$. The next section shows how we can overcome this inadequacy.

3. New perturbation results. In Section 2 we saw the key to deriving first order perturbation bounds for U in the block Cholesky downdating problem is the equation (2.8). We will now analyze it in two new approaches. The two approaches have been used in the perturbation analyses of the Cholesky factorization, the QR factorization (see Chang, Paige and Stewart [6, 7]), and LU factorization (see Chang [4] and Stewart [19]). The first approach, the refined matrix equation approach, gives a clear improvement on the previous results, while the second, the matrix-vector equation approach, gives a further improvement still, which leads to the true condition numbers for the block Cholesky downdating problem.

3.1. Refined matrix equation analysis. In the last section we used (2.8) to produce the matrix equation (2.9), and derived the bounds directly from this. We now look at this approach more closely.

Let $\mathcal{D}_n$ be the set of all $n \times n$ real positive definite diagonal matrices. For any $D = \mathrm{diag}(\delta_1, \ldots, \delta_n) \in \mathcal{D}_n$, let $U = D\bar{U}$. Note that for any matrix B we have $\mathrm{up}(BD^{-1}) = \mathrm{up}(B)D^{-1}$ and $\mathrm{up}(D^{-1}B) = D^{-1}\mathrm{up}(B)$.

First *with no restriction on* ΔR we have from (2.9)

$$\begin{aligned}\widehat{\Delta U} &= \{\mathrm{up}(U^{-T}R^T\Delta R\bar{U}^{-1}) + D^{-1}\mathrm{up}(\bar{U}^{-T}\Delta R^T RU^{-1})D\}\bar{U} \\ &\quad - \{\mathrm{up}(U^{-T}X^T\Delta X\bar{U}^{-1}) + D^{-1}\mathrm{up}(\bar{U}^{-T}\Delta X^T XU^{-1})D\}\bar{U},\end{aligned}$$

so taking the F-norm gives

$$\begin{aligned}\|\widehat{\Delta U}\|_F &\le \|\mathrm{up}(U^{-T}R^T\Delta R\bar{U}^{-1}) + D^{-1}\mathrm{up}(\bar{U}^{-T}\Delta R^T RU^{-1})D\|_F\|\bar{U}\|_2 \\ &\quad + \|\mathrm{up}(U^{-T}X^T\Delta X\bar{U}^{-1}) + D^{-1}\mathrm{up}(\bar{U}^{-T}\Delta X^T XU^{-1})D\|_F\|\bar{U}\|_2.\end{aligned} \tag{3.1}$$

It is easy to show for any $B \in \mathcal{R}^{n\times n}$ (see Lemma 5.1 in [7])

$$\|\mathrm{up}(B) + D^{-1}\mathrm{up}(B^T)D\|_F \le \sqrt{1+\zeta_D^2}\|B\|_F, \tag{3.2}$$

where $\zeta_D = \max_{1\le i<j\le n}\{\delta_j/\delta_i\}$. Thus from (3.1) we have

$$\begin{aligned}\|\widehat{\Delta U}\|_F &\le \sqrt{1+\zeta_D^2}(\|U^{-T}R^T\Delta R\bar{U}^{-1}\|_F + \|U^{-T}X^T\Delta X\bar{U}^{-1}\|_2)\|\bar{U}\|_2 \\ &\le \sqrt{1+\zeta_D^2}\kappa_2(\bar{U})(\|RU^{-1}\|_2\|\Delta R\|_F + \|XU^{-1}\|_2\|\Delta X\|_F). \\ &= \sqrt{1+\zeta_D^2}\frac{\kappa_2(\bar{U})}{\sigma_n(\Gamma)}(\|\Delta R\|_F + \sqrt{1-\sigma_n^2(\Gamma)}\|\Delta X\|_F), \quad \text{(using (2.2), (2.3))}\end{aligned}$$

which is an elegant result in the changes alone. It leads to the following perturbation bound in terms of relative changes

$$
\begin{aligned}
(3.3) \qquad \frac{\|\widehat{\Delta U}\|_F}{\|U\|_2} \leq & \sqrt{1+\zeta_D^2}\frac{\kappa_2(D^{-1}U)}{\sigma_n(\Gamma)}\frac{\|R\|_2}{\|U\|_2}\frac{\|\Delta R\|_F}{\|R\|_2} \\
& + \sqrt{1+\zeta_D^2}\sqrt{1-\sigma_n^2(\Gamma)}\frac{\kappa_2(D^{-1}U)}{\sigma_n(\Gamma)}\frac{\|X\|_2}{\|U\|_2}\frac{\|\Delta X\|_F}{\|X\|_2}.
\end{aligned}
$$

Although here it would be simpler to just define an overall condition number, for later comparisons it is necessary for us to define the following two quantities as condition numbers for U with respect to relative changes in R and X, respectively (here subscript $_G$ refers to *general* ΔR, and later the subscript $_T$ will refer to *upper triangular* ΔR):

$$
(3.4) \qquad c_{RG}(R,X) \equiv \inf_{D\in\mathcal{D}_n} c_{RG}(R,X,D), \quad c_X(R,X) \equiv \inf_{D\in\mathcal{D}_n} c_X(R,X,D),
$$

where

$$
(3.5) \qquad c_{RG}(R,X,D) \equiv \sqrt{1+\zeta_D^2}\frac{\kappa_2(D^{-1}U)}{\sigma_n(\Gamma)}\frac{\|R\|_2}{\|U\|_2},
$$

$$
(3.6) \qquad c_X(R,X,D) \equiv \sqrt{1+\zeta_D^2}\sqrt{1-\sigma_n^2(\Gamma)}\frac{\kappa_2(D^{-1}U)}{\sigma_n(\Gamma)}\frac{\|X\|_2}{\|U\|_2}.
$$

Then an overall condition number can be defined as

$$
(3.7) \qquad c_G(R,X) \equiv \inf_{D\in\mathcal{D}_n} c_G(R,X,D),
$$

where

$$
c_G(R,X,D) \equiv \max\{c_{RG}(R,X,D), c_X(R,X,D)\} = c_{RG}(R,X,D).
$$

Obviously we have

$$
(3.8) \qquad c_G(R,X) = c_{RG}(R,X) \geq c_X(R,X).
$$

Thus with these, we have from (3.3) that

$$
\begin{aligned}
(3.9) \qquad \frac{\|\widehat{\Delta U}\|_F}{\|U\|_2} &\leq c_{RG}(R,X)\frac{\|\Delta R\|_2}{\|R\|_2} + c_X(R,X)\frac{\|\Delta X\|_F}{\|X\|_2} \\
&\leq c_G(R,X)(\frac{\|\Delta R\|_F}{\|R\|_2} + \frac{\|\Delta X\|_F}{\|X\|_2}).
\end{aligned}
$$

Clearly if we take $D = I_n$, (3.3) will become (2.10), and

$$
(3.10) \qquad c_{RG}(R,X) \leq c_{RG}(R,X,I_n) = \beta_R, \quad c_X(R,X) \leq c_X(R,X,I_n) = \beta_X,
$$

$$
(3.11) \qquad c_G(R,X) \leq c_G(R,X,I_n) = \beta_2.
$$

It is not difficult to give an example to show β_2 can be arbitrarily larger than $c_G(R,X)$, as can be seen from the following asymptotic behaviour.

If $X \to 0$ we saw $U \to R$ and $\sigma_n(\Gamma) \to 1$, so

$$
c_G(R,X,D) \to \sqrt{1+\zeta_D^2}\kappa(D^{-1}R).
$$

It is shown in [7, §5.1, (5.14)] that with an appropriate choice of D, $\sqrt{1+\zeta_D^2}\kappa(D^{-1}R)$ has a bound which is a function of n only, if R was found using the standard pivoting strategy in the Cholesky factorization, and in this case, we see the condition number $c_G(R,X)$ of the problem here is bounded independently of $\kappa_2(R)$ as $X \to 0$, for *general* ΔR. At the end of this section we give an even stronger result when $X \to 0$ for the case of upper triangular ΔR. Note in the case here that β_2 in (2.11) can be made as large as we like, and thus arbitrarily larger than $c_G(R,X)$.

In the case where ΔR is *upper triangular*, we can refine the analysis further. From (2.9) we have

$$\begin{aligned}\widehat{\Delta U} = [&\mathrm{up}(U^{-T}R^T\Delta RU^{-1} + U^{-T}\Delta R^T RU^{-1}) \\ &- \mathrm{up}(U^{-T}X^T\Delta XU^{-1} + U^{-T}\Delta X^T XU^{-1})]U.\end{aligned} \tag{3.12}$$

Notice with (1.3) and (1.4)

$$\begin{aligned}&U^{-T}R^T\Delta RU^{-1} + U^{-T}\Delta R^T RU^{-1} \\ &= [\mathrm{slt}(U^{-T}R^T) + \mathrm{diag}(U^{-T}R^T)]\Delta RU^{-1} \\ &\quad + U^{-T}\Delta R^T[\mathrm{sut}(RU^{-1}) + \mathrm{diag}(RU^{-1})] \\ &= \mathrm{diag}(U^{-T}R^T)\cdot\Delta RU^{-1} + U^{-T}\Delta R^T\cdot\mathrm{diag}(RU^{-1}) \\ &\quad + \mathrm{slt}(U^{-T}R^T)\cdot\Delta RU^{-1} + U^{-T}\Delta R^T\cdot\mathrm{sut}(RU^{-1}).\end{aligned} \tag{3.13}$$

But for any *upper* triangular matrix T we have

$$\mathrm{up}(T) + \mathrm{up}(T^T) = T,$$

so that if we define $T \equiv \mathrm{diag}(U^{-T}R^T)\cdot\Delta RU^{-1}$, then

$$\begin{aligned}&\mathrm{up}[\mathrm{diag}(U^{-T}R^T)\cdot\Delta RU^{-1} + U^{-T}\Delta R^T\cdot\mathrm{diag}(RU^{-1})] \\ &= \mathrm{diag}(U^{-T}R^T)\cdot\Delta RU^{-1}.\end{aligned} \tag{3.14}$$

Thus from (3.12), (3.13) and (3.14) we obtain

$$\begin{aligned}&\widehat{\Delta U} \\ &= \mathrm{diag}(U^{-T}R^T)\cdot\Delta R + \{\mathrm{up}[\mathrm{slt}(U^{-T}R^T)\cdot\Delta RU^{-1} + U^{-T}\Delta R^T\cdot\mathrm{sut}(RU^{-1})] \\ &\quad - \mathrm{up}(U^{-T}X^T\Delta XU^{-1} + U^{-T}\Delta X^T XU^{-1})\}U.\end{aligned} \tag{3.15}$$

As before, let $U = D\bar{U}$, where $D = \mathrm{diag}(\delta_1,\ldots,\delta_n) \in \mathcal{D}_n$. From (3.15) it follows that

$$\begin{aligned}\|\widehat{\Delta U}\|_F \le\, &\|\mathrm{diag}(U^{-T}R^T)\|_2\|\Delta R\|_F \\ &+ \|\mathrm{up}[\mathrm{slt}(U^{-T}R^T)\cdot\Delta R\bar{U}^{-1}] + D^{-1}\mathrm{up}[\bar{U}^{-T}\Delta R^T\cdot\mathrm{sut}(RU^{-1})]D\|_F\|\bar{U}\|_2 \\ &+ \|\mathrm{up}(U^{-T}X^T\Delta X\bar{U}^{-1}) + D^{-1}\mathrm{up}(\bar{U}^{-T}\Delta X^T XU^{-1})D\|_F\|\bar{U}\|_2.\end{aligned}$$

Then, applying (3.2) to this, we get the following perturbation bound

$$\begin{aligned}\frac{\|\widehat{\Delta U}\|_F}{\|U\|_2} \le\, &\Big(\|\mathrm{diag}(RU^{-1})\|_2 + \sqrt{1+\zeta_D^2}\kappa_2(D^{-1}U)\|\mathrm{sut}(RU^{-1})\|_2\Big)\frac{\|R\|_2}{\|U\|_2}\frac{\|\Delta R\|_F}{\|R\|_2} \\ &+ \sqrt{1+\zeta_D^2}\kappa_2(D^{-1}U)\|XU^{-1}\|_2\frac{\|X\|_2}{\|U\|_2}\frac{\|\Delta X\|_F}{\|X\|_2}.\end{aligned} \tag{3.16}$$

Comparing (3.16) with (3.3) and noticing (2.3), we see the sensitivity of U with respect to changes in X does not change, so $c_X(R,X)$ defined in (3.4) can still be regarded as a condition number for U with respect to changes in X. But we now need to define a new condition number for U with respect to *upper triangular* changes in R, that is (subscript $_T$ indicates upper triangular ΔR)

$$c_{RT}(R,X) \equiv \inf_{D\in\mathcal{D}_n} c_{RT}(R,X,D),$$

where

$$(3.17)\quad c_{RT}(R,X,D) \equiv (\|\mathrm{diag}(RU^{-1})\|_2 + \sqrt{1+\zeta_D^2}\kappa_2(D^{-1}U)\|\mathrm{sut}(RU^{-1})\|_2)\frac{\|R\|_2}{\|U\|_2}.$$

Thus an overall condition number can be defined as

$$c_T(R,X) = \inf_{D\in\mathcal{D}_n} c_T(R,X,D),$$

where

$$c_T(R,X,D) = \max\{c_{RT}(R,X,D), c_X(R,X,D)\}.$$

Obviously we have

$$(3.18)\qquad c_T(R,X) = \max\{c_{RT}(R,X), c_X(R,X)\}.$$

With these, we have from (3.16) that

$$(3.19)\qquad \begin{aligned}\frac{\|\widehat{\Delta U}\|_F}{\|U\|_2} &\le c_{RT}(R,X)\frac{\|\Delta R\|_F}{\|R\|_2} + c_X(R,X)\frac{\|\Delta X\|_F}{\|X\|_2}\\ &\le c_T(R,X)(\frac{\|\Delta R\|_F}{\|R\|_2} + \frac{\|\Delta X\|_F}{\|X\|_2}).\end{aligned}$$

What is the relationship between $c_T(R,X)$ and $c_G(R,X) = c_{RG}(R,X)$? For any $n\times n$ upper triangular matrix $T=(t_{ij})$, observe the following two facts:

1) t_{ii}, $i=1,2,\ldots,n$ are the eigenvalues of T, so that

$$|t_{ii}| \le \|T\|_2,$$

which gives

$$\|\mathrm{diag}(T)\|_2 \le \|T\|_2.$$

2)

$$\|\mathrm{sut}(T)\|_2 \le \|\mathrm{sut}(T)\|_F \le \|T\|_F \le \sqrt{n}\|T\|_2.$$

(Note: In fact we can prove a slightly sharper inequality $\|\mathrm{sut}(T)\|_2 \le \sqrt{n-1}\|T\|_2$).

Therefore

$$\begin{aligned} c_{RT}(R,X,D) &= \left(\|\mathrm{diag}(RU^{-1})\|_2 + \sqrt{1+\zeta_D^2}\kappa_2(D^{-1}U)\|\mathrm{sut}(RU^{-1})\|_2\right)\frac{\|R\|_2}{\|U\|_2}\\ &\le \left(\|RU^{-1}\|_2 + \sqrt{n}\sqrt{1+\zeta_D^2}\kappa_2(D^{-1}U)\|RU^{-1}\|_2\right)\frac{\|R\|_2}{\|U\|_2}\\ &< (1+\sqrt{n})\sqrt{1+\zeta_D^2}\kappa_2(D^{-1}U)\|RU^{-1}\|_2\frac{\|R\|_2}{\|U\|_2}\\ &= (1+\sqrt{n})\sqrt{1+\zeta_D^2}\frac{\kappa_2(D^{-1}U)}{\sigma_n(\Gamma)}\frac{\|R\|_2}{\|U\|_2} \quad (\text{ using (2.2)})\\ &= (1+\sqrt{n})c_{RG}(R,X,D),\end{aligned}$$

so that

$$c_{RT}(R,X) \leq (1+\sqrt{n})c_{RG}(R,X). \tag{3.20}$$

Thus we have from (3.8) and (3.18)

$$c_T(R,X) \leq (1+\sqrt{n})c_G(R,X). \tag{3.21}$$

On the other hand, $c_T(R,X)$ can be arbitrarily smaller than $c_G(R,X)$. This can be seen from the asymptotic behaviour, which is important in its own right. As $X \to 0$, since $U \to R$, $\sigma_n(\Gamma) \to 1$ and $RU^{-1} \to I_n$, we have

$$c_T(R,X,I_n) \to 1,$$

so for upper triangular changes in R, whether pivoting was used in finding R or not,

$$c_T(R,X) \to 1.$$

Thus when $X \to 0$, the bound in (3.19) reflects the true sensitivity of the problem. For the case of general ΔR, if we do not use pivoting it is straightforward to make $c_G(R,X)$ in (3.7) arbitrarily large even with $X=0$, see (3.5).

3.2. Matrix-vector equation analysis. In the last subsection, based on the structure of ΔR, we gave two perturbation bounds using the so called refined matrix equation approach. Also based on the structure of ΔR, we can now obtain provably sharp, but less intuitive results by viewing the matrix equation (2.8) as a large matrix-vector equation. For any matrix $C \equiv (c_{ij}) \equiv [c_1, \ldots, c_n] \in \mathcal{R}^{n\times n}$, denote by $c_j^{(i)}$ the vector of the first i elements of c_j. With this, we define ("u" denotes "upper")

$$\mathrm{uvec}(C) \equiv \begin{pmatrix} c_1^{(1)} \\ c_2^{(2)} \\ \cdot \\ c_n^{(n)} \end{pmatrix}.$$

It is the vector formed by stacking the columns of the upper triangular part of C into one long vector.

First assume ΔR is a *general* real $n\times n$ matrix. It is easy to show (2.8) can be rewritten into the following matrix-vector form (cf [7])

$$W_U \mathrm{uvec}(\widehat{\Delta U}) = Z_R \mathrm{vec}(\Delta R) - Y_X \mathrm{vec}(\Delta X), \tag{3.22}$$

where $W_U \in \mathcal{R}^{\frac{n(n+1)}{2}\times\frac{n(n+1)}{2}}$ is

$$\left[\begin{array}{cccccccccccc}
u_{11} &&&&&&&&&&& \\
u_{12} & u_{11} &&&&&&&&&& \\
& u_{12} & u_{22} &&&&&&&&& \\
u_{13} &&& u_{11} &&&&&&&& \\
& u_{13} & u_{23} & u_{12} & u_{22} &&&&&&& \\
&&& u_{13} & u_{23} & u_{33} &&&&&& \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot &&&&& \\
u_{1n} &&&&&&& u_{11} &&&& \\
& u_{1n} & u_{2n} &&&&& u_{12} & u_{22} &&& \\
&&& u_{1n} & u_{2n} & u_{3n} && u_{13} & u_{23} & u_{33} && \\
&&&&&& \cdot & \cdot & \cdot & \cdot & \cdot & \\
&&&&&&& u_{1n} & u_{2n} & u_{3n} & \cdot & u_{nn}
\end{array}\right],$$

$Z_R \in \mathcal{R}^{\frac{n(n+1)}{2} \times n^2}$ is

$$\left[\begin{array}{cccc|cccc|c|cccc}
r_{11} & & & & & & & & & & & & \\
\hline
r_{12} & r_{22} & & & r_{11} & & & & & & & & \\
 & & & & r_{12} & r_{22} & & & & & & & \\
\hline
\cdot & \cdot & \cdot & & \cdot & \cdot & \cdot & & \cdot & & & & \\
\hline
r_{1n} & r_{2n} & \cdot & r_{nn} & & & & & & r_{11} & & & \\
 & & & & r_{1n} & r_{2n} & \cdot & r_{nn} & & r_{12} & r_{22} & & \\
 & & & & & & & & \cdot & \cdot & \cdot & \cdot & \\
 & & & & & & & & & r_{1n} & r_{2n} & \cdot & r_{nn}
\end{array}\right],$$

and $Y_X \in \mathcal{R}^{\frac{n(n+1)}{2} \times kn}$ is

$$\left[\begin{array}{cccc|cccc|c|cccc}
x_{11} & x_{21} & \cdot & x_{k1} & & & & & & & & & \\
\hline
x_{12} & x_{22} & \cdot & x_{k2} & x_{11} & x_{21} & \cdot & x_{k1} & & & & & \\
 & & & & x_{12} & x_{22} & \cdot & x_{k2} & & & & & \\
\hline
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & & & & \\
\hline
x_{1n} & x_{2n} & \cdot & x_{kn} & & & & & & x_{11} & x_{21} & \cdot & x_{k1} \\
 & & & & x_{1n} & x_{2n} & \cdot & x_{kn} & & x_{12} & x_{22} & \cdot & x_{kn} \\
 & & & & & & & & \cdot & \cdot & \cdot & \cdot & \cdot \\
 & & & & & & & & & x_{1n} & x_{2n} & \cdot & x_{nn}
\end{array}\right].$$

Since U is nonsingular, W_U is also, and from (3.22)

$$\text{uvec}(\widehat{\Delta U}) = W_U^{-1} Z_R \text{vec}(\Delta R) - W_U^{-1} Y_X \text{vec}(\Delta X), \tag{3.23}$$

so taking the 2-norm gives

$$\|\widehat{\Delta U}\|_F \leq \|W_U^{-1} Z_R\|_2 \|\Delta R\|_F + \|W_U^{-1} Y_X\|_2 \|\Delta X\|_F,$$

resulting in the following perturbation bound

$$\begin{aligned}
\frac{\|\widehat{\Delta U}\|_F}{\|U\|_2} &\leq \kappa_{RG}(R,X) \frac{\|\Delta R\|_F}{\|R\|_2} + \kappa_X(R,X) \frac{\|\Delta X\|_F}{\|X\|_2} \\
&\leq \kappa_{CDG}(R,X) \left(\frac{\|\Delta R\|_F}{\|R\|_2} + \frac{\|\Delta X\|_F}{\|X\|_2}\right),
\end{aligned} \tag{3.24}$$

where

$$\kappa_{RG}(R,X) \equiv \frac{\|W_U^{-1} Z_R\|_2 \|R\|_2}{\|U\|_2}, \quad \kappa_X(R,X) \equiv \frac{\|W_U^{-1} Y_X\|_2 \|X\|_2}{\|U\|_2}, \tag{3.25}$$

$$\kappa_{CDG}(R,X) \equiv \max\{\kappa_{RG}(R,X), \kappa_X(R,X)\}. \tag{3.26}$$

Now we would like to show

$$\kappa_{RG}(R,X) \leq c_{RG}(R,X), \quad \kappa_X(R,X) \leq c_X(R,X). \tag{3.27}$$

Before showing this, we will prove a more general result. Suppose from (2.8) we are able to obtain a perturbation bound of the form

$$\frac{\|\widehat{\Delta U}\|_F}{\|U\|_2} \leq \alpha_R \frac{\|\Delta R\|_F}{\|R\|_2} + \alpha_X \frac{\|\Delta X\|_F}{\|X\|_2}, \tag{3.28}$$

where α_R and α_X, two functions of R and X, are other measures of the sensitivity of the Cholesky downdating problem with respect to changes in R and X. Let $\Delta X = 0$. Then from (3.23) and (3.28) we have

$$\frac{\|W_U^{-1} Z_R \mathrm{vec}(\Delta R)\|_2}{\|U\|_2} \leq \alpha_R \frac{\|\Delta R\|_F}{\|R\|_2}.$$

Notice ΔR can be any (sufficiently small) $n \times n$ real matrix, so we must have

$$\|W_U^{-1} Z_R\|_2 \leq \alpha_R \frac{\|U\|_2}{\|R\|_2},$$

which gives

$$\kappa_{RG}(R, X) \leq \alpha_R.$$

Similarly, we can show

$$\kappa_X(R, X) \leq \alpha_X.$$

Notice since (3.9) is a particular case of (3.28), (3.27) follows. Thus we have from (3.8) and (3.26)

$$\kappa_{CDG}(R, X) \leq c_G(R, X). \tag{3.29}$$

The above analysis shows for general ΔR, $\kappa_{RG}(R, X)$ and $\kappa_X(R, X)$ are optimal measures of the sensitivity of U with respect to changes in R and X, respectively, and thus the bound (3.24) is optimal. So we propose $\kappa_{RG}(R, X)$ and $\kappa_X(R, X)$ as the *true* condition numbers for U with respect to general changes in R and X, respectively, and $\kappa_{CDG}(R, X)$ as the *true* overall condition number of the problem in this case.

It is easy to observe that if $X \to 0$, $\kappa_{CDG}(R, X) \to \|W_R^{-1} Z_R\|_2$, where W_R is just W_U with each entry u_{ij} replaced by r_{ij}. If R was found using the standard pivoting strategy in the Cholesky factorization, then $\|W_R^{-1} Z_R\|_2$ has a bound which is a function of n alone (see [5] for a proof). So in this case our condition number $\kappa_{CDG}(R, X)$ also has a bound which is a function of n alone as $X \to 0$.

Remark 1: Our numerical experiments suggest $c_G(R, X)$ is usually a good approximation to $\kappa_{CDG}(R, X)$. But the following example shows $c_G(R, X)$ can sometimes be arbitrarily larger than $\kappa_{CDG}(R, X)$.

$$R = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \epsilon^3 & 0 \\ 0 & 0 & 0 & \epsilon^2 \end{pmatrix}, \quad X = \begin{pmatrix} \sqrt{3} & 2/\sqrt{3} & 0 & 0 \\ 0 & \sqrt{2/3 - \epsilon^2} & 0 & 0 \end{pmatrix}, \quad U = \mathrm{diag}(1, \epsilon, \epsilon^3, \epsilon^2),$$

where ϵ is a small positive number. It is not difficult to show

$$c_G(R, X) = O(\frac{1}{\epsilon^2}), \quad \kappa_{CDG}(R, X) = O(\frac{1}{\epsilon}).$$

But $c_G(R, X)$ has an advantage over $\kappa_{CDG}(R, X)$ — it can be quite easy to estimate — all we need do is choose a suitable D in $c_G(R, X, D)$. We consider how to do this in the next section. In contrast $\kappa_{CDG}(R, X)$ is, as far as we can see, unreasonably expensive to compute or estimate.

Now we consider the case where ΔR is *upper triangular.* (2.8) can now be rewritten as the following matrix-vector form

$$W_U \mathrm{uvec}(\widehat{\Delta U}) = W_R \mathrm{uvec}(\Delta R) - Y_X \mathrm{vec}(\Delta X), \tag{3.30}$$

where $W_U \in \mathcal{R}^{\frac{n(n+1)}{2} \times \frac{n(n+1)}{2}}$ and $Y_X \in \mathcal{R}^{\frac{n(n+1)}{2} \times kn}$ are defined as before, and $W_R \in \mathcal{R}^{\frac{n(n+1)}{2} \times \frac{n(n+1)}{2}}$ is just W_U with each entry u_{ij} replaced by r_{ij}. Since U is nonsingular, W_U is also, and from (3.30)

$$\mathrm{uvec}(\widehat{\Delta U}) = W_U^{-1} W_R \mathrm{uvec}(\Delta R) - W_U^{-1} Y_X \mathrm{vec}(\Delta X), \tag{3.31}$$

so taking the 2-norm gives

$$\|\widehat{\Delta U}\|_F \leq \|W_U^{-1} W_R\|_2 \|\Delta R\|_F + \|W_U^{-1} Y_X\|_2 \|\Delta X\|_F,$$

which leads to the following perturbation bound

$$\begin{aligned} \frac{\|\widehat{\Delta U}\|_F}{\|U\|_2} &\leq \kappa_{RT}(R, X) \frac{\|\Delta R\|_F}{\|R\|_2} + \kappa_X(R, X) \frac{\|\Delta X\|_F}{\|X\|_2} \\ &\leq \kappa_{CDT}(R, X)\left(\frac{\|\Delta R\|_F}{\|R\|_2} + \frac{\|\Delta X\|_F}{\|X\|_2}\right), \end{aligned} \tag{3.32}$$

where

$$\begin{aligned} \kappa_{RT}(R, X) \equiv \frac{\|W_U^{-1} W_R\|_2 \|R\|_2}{\|U\|_2}, &\quad \kappa_X(R, X) \equiv \frac{\|W_U^{-1} Y_X\|_2 \|X\|_2}{\|U\|_2}, \\ \kappa_{CDT}(R, X) &\equiv \max\{\kappa_{RT}(R, X), \kappa_X(R, X)\}. \end{aligned} \tag{3.33}$$

Note $\kappa_X(R, X)$ is the same as that defined in (3.25).

As before, we can show that for the case where ΔR is upper triangular, $\kappa_{RT}(R, X)$ and $\kappa_X(R, X)$ are optimal measures of the sensitivity of U with respect to changes in R and X, respectively, and thus the bound (3.32) is optimal. In particular, we have

$$\kappa_{RT}(R, X) \leq c_{RG}(R, X), c_{RT}(R, X), \kappa_{RG}(R, X); \quad \kappa_X(R, X) \leq c_X(R, X). \tag{3.34}$$

In fact $\kappa_{RT}(R, X) \leq \kappa_{RG}(R, X)$ can also be proved directly by the fact that the columns of W_R form a proper subset of the columns of Z_R, and the second inequality has been proved before. Thus we have from (3.8), (3.18), (3.26) and (3.33)

$$\kappa_{CDT}(R, X) \leq c_G(R, X), c_T(R, X), \kappa_{CDG}(R, X). \tag{3.35}$$

By the above analysis, we propose $\kappa_{RT}(R, X)$ and $\kappa_X(R, X)$ as the *true* condition numbers for U with respect to changes in R and X, respectively, and $\kappa_{CDT}(R, X)$ as the *true* overall condition number, in the case that ΔR is upper triangular.

If as well $X \to 0$, then since $U \to R$, $W_U^{-1} W_R \to I_{\frac{n(n+1)}{2}}$, and $\kappa_{CDT}(R, X) \to 1$. So in this case the Cholesky downdating problem becomes very well conditioned no matter how ill conditioned R or U is.

Remark 2: Numerical experiments also suggest $c_T(R, X)$ is usually a good approximation to $\kappa_{CDT}(R, X)$. But sometimes $c_T(R, X)$ can be arbitrarily larger than

$\kappa_{CDT}(R, X)$. This can also be seen from the example in Remark 1. In fact, it is not difficult to obtain

$$c_T(R, X) = O(\frac{1}{\epsilon^2}), \quad \kappa_{CDT}(R, X) = O(\frac{1}{\epsilon}).$$

Like $\kappa_{CDG}(R, X)$, $\kappa_{CDT}(R, X)$ is difficult to compute or estimate. But $c_T(R, X)$ is easy to estimate, which is discussed in the next section.

4. Numerical tests and condition estimators. In Section 3 we presented new first order perturbation bounds for the the downdated Cholesky factor U using first the refined matrix equation approach, and then the matrix vector equation approach. We defined $\kappa_{CDG}(R, X)$ for general ΔR, and $\kappa_{CDT}(R, X)$ for upper triangular ΔR, as the true overall condition numbers of the problem. Also we gave two corresponding practical but weaker condition numbers $c_G(R, X)$ and $c_T(R, X)$ for the two ΔR cases.

We would like to choose D such that $c_G(R, X, D)$ and $c_T(R, X, D)$ are good approximations to $c_G(R, X)$ and $c_T(R, X)$, respectively. We see from (3.5), (3.6) and (3.17) that we want to find D such that $\sqrt{1+\zeta_D^2}\kappa_2(D^{-1}U)$ approximates its infimum. By a well known result of van der Sluis [21], $\kappa_2(D^{-1}U)$ will be nearly minimal when the rows of $D^{-1}U$ are equilibrated. But this could lead to a large ζ_D. So a reasonable compromise is to choose D to equilibrate U as far as possible while keeping $\zeta_D \leq 1$. Specifically, take $\zeta_1 = \sqrt{\sum_{j=1}^n u_{1j}^2}$, $\zeta_i = \sqrt{\sum_{j=i}^n u_{ij}^2}$ if $\sqrt{\sum_{j=i}^n u_{ij}^2} \leq \zeta_{i-1}$ otherwise $\zeta_i = \zeta_{i-1}$, for $i = 2, \ldots, n$. Then we use a standard condition estimator to estimate $\kappa_2(D^{-1}U)$ in $O(n^2)$ operations.

Notice from (2.4) we have $\sigma_n(\Gamma) = \sqrt{1 - \|XR^{-1}\|_2^2}$. Usually k, the number of rows of X, is much smaller than n, so $\sigma_n(\Gamma)$ can be computed in $O(n^2)$. If k is not much smaller than n, then we use a standard norm estimator to estimate $\|XR^{-1}\|_2$ in $O(n^2)$. Similarly $\|U\|_2$ and $\|R\|_2$ can be estimated in $O(n^2)$. So finally $c_G(R, X, D)$ can be estimated in $O(n^2)$. Estimating $c_T(R, X, D)$ is not as easy as estimating $c_G(R, X, D)$. The part $\|\mathrm{diag}(RU^{-1})\|_2$ in $c_{RT}(R, X, D)$ can easily be computed in $O(n)$, since $\mathrm{diag}(RU^{-1}) = \mathrm{diag}(r_{11}/u_{11}, \ldots, r_{nn}/u_{nn})$. The part $\|\mathrm{sut}(RU^{-1})\|_2$ in $c_{RT}(R, X, D)$ can roughly be estimated in $O(n^2)$, based on

$$\frac{1}{\sqrt{n-1}}\|\mathrm{sut}(RU^{-1})\|_F \leq \|\mathrm{sut}(RU^{-1})\|_2 \leq \|\mathrm{sut}(RU^{-1})\|_F,$$
$$\|\mathrm{sut}(RU^{-1})\|_F = \sqrt{\|RU^{-1}\|_F^2 - \|\mathrm{diag}(RU^{-1})\|_F^2}$$

and the fact that $\|RU^{-1}\|_F$ can be estimated by a standard norm estimator in $O(n^2)$. The value of $\|XU^{-1}\|_2$ in $c_X(R, X, D)$ can be calculated (if $k << n$) or estimated by a standard estimator in $O(n^2)$. All the remaining values $\|R\|_2$, $\|X\|_2$ and $\|U\|_2$ can also be estimated by a standard norm estimator in $O(n^2)$. Hence $c_{RT}(R, X, D)$, $c_X(R, X, D)$, and thus $c_T(R, X, D)$ can be estimated in $O(n^2)$. For standard condition estimators and norm estimators, see Chapter 14 of [14].

The relationships among the various overall condition numbers for the Cholesky downdating problem presented in Section 2 and Section 3 are as follows.

$$\beta_1 \geq \beta_2 \geq c_G(R, X) \geq \kappa_{CDG}(R, X) \geq \kappa_{CDT}(R, X),$$
$$(1+\sqrt{n})c_G(R, X) \geq c_T(R, X) \geq \kappa_{CDT}(R, X).$$

Now we give one numerical example to illustrate these. The example, quoted from Sun [20], is as follows.

$$R = \mathrm{diag}(1, s, s^2, s^3, s^4) \begin{pmatrix} 1 & -c & -c & -c & -c \\ 0 & 1 & -c & -c & -c \\ 0 & 0 & 1 & -c & -c \\ 0 & 0 & 0 & 1 & -c \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \quad X^T = \tau \begin{pmatrix} 0.240 \\ -0.899 \\ 0.899 \\ 1.560 \\ 2.390 \end{pmatrix},$$

where $c = 0.95$, $s = \sqrt{1 - c^2}$. The results obtained using MATLAB are shown in Table 4.1 for various values of τ:

$$\tau_1 = 1.004015006005433e - 2, \quad \tau_2 = 1.003021021209640e - 2,$$
$$\tau_3 = 9.036225416303058e - 3,$$

and $\tau_4 = \tau_3 \cdot e - 01$, $\tau_5 = \tau_3 \cdot e - 03$, $\tau_6 = \tau_3 \cdot e - 5$.

Table 4.1

τ	τ_1	τ_2	τ_3	τ_4	τ_5	τ_6
$\Vert XR^{-1}\Vert_2$	0.99999	0.999	0.9	0.09	0.0009	0.000009
β_1	2.25e+10	2.25e+07	2.60e+04	2.72e+03	2.69e+03	2.69e+03
β_2	1.01e+08	1.01e+06	1.14e+04	2.71e+03	2.69e+03	2.69e+03
$c_G(R,X,D)$	3.60e+03	3.61e+02	3.79e+01	1.79e+01	1.78e+01	1.78e+01
$\kappa_{CDG}(R,X)$	1.66e+03	1.66e+02	1.71e+01	8.42e+00	8.41e+00	8.41e+00
$c_T(R,X,D)$	2.12e+03	2.12e+02	1.79e+01	1.07e+00	1.00e+00	1.00e+00
$\kappa_{CDT}(R,X)$	2.43e+02	2.43e+01	2.44e+00	1.01e+00	1.00e+00	1.00e+00

Note in Table 4.1 how β_1 and β_2 can be far worse than the true condition numbers $\kappa_{CDG}(R,X)$ and $\kappa_{CDT}(R,X)$, although β_2 is not as bad as β_1. Also we observe that $c_G(R,X,D)$ and $c_T(R,X,D)$ are very good approximations to $\kappa_{CDG}(R,X)$ and $\kappa_{CDT}(R,X)$, respectively. When X become small, all of the condition numbers decrease. The asymptotic behavior of $c_G(R,X,D)$, $c_T(R,X,D)$, $\kappa_{CDG}(R,X)$ and $\kappa_{CDT}(R,X)$ coincides with our theoretical results — when $X \to 0$, $c_G(R,X)$ and $\kappa_{CDG}(R,X)$ will be bounded in terms of n since here R corresponds to the Cholesky factor of a correctly pivoted A, and $c_T(R,X), \kappa_{CDT}(R,X) \to 1$.

Acknowledgement. We would like to thank Ji-guang Sun for suggesting to us that the approach describe in (a draft version of) [6] might apply to the Cholesky downdating problem.

REFERENCES

[1] S. T. Alexander, C.-T. Pan and R. J. Plemmons, *Analysis of a recursive least squares hyperbolic rotation algorithm for signal processing*, Linear Algebra Appl., 98 (1988), pp. 3–40.

[2] A. Björck, H. Park and L. Eldén, *Accurate downdating of least squares solutions*, SIAM J. Matrix Anal. Appl., 15 (1994), pp. 549–568.

[3] A. W. Bojanczyk, R. P. Brent, P. Van Dooren, and F. R. de Hoog, *A note on downdating the Cholesky factorization*, SIAM J. Sci. Stat. Comput., 8 (1987), pp. 210–221.

[4] X.-W. Chang, PhD Thesis, in preparation, McGill University, School of Computer Science, 1996.

[5] X.-W. Chang and C. C. Paige, A perturbation analysis for R in the QR factorization, Technical Report, **SOCS-95.7**, McGill University, School of Computer Science, 1995.

[6] X.-W. Chang, C. C. Paige and G. W. Stewart, *New perturbation analyses for the Cholesky factorization*, 1995. To appear in IMA, J. Numer. Anal.

[7] X.-W. Chang, C. C. Paige and G. W. Stewart, *Perturbation analyses for the QR factorization*, 1996, submitted for publication.

[8] J. J. Dongarra, J. R. Bunch, C. B. Moler and G. W. Stewart, *LINPACK User's Guide*, SIAM, Philadelphia, 1979.

[9] L. Eldén and H. Park, *Block downdating of least squares solutions*, SIAM J. Matrix Anal. Appl., 15 (1994), pp. 1018–1034.

[10] L. Eldén and H. Park, *Perturbation analysis for block downdating of a Cholesky decomposition*, Numerische Mathematik, 68 (1994), pp. 457–467.

[11] G. H. Golub and G. P. Styan, *Numerical computations for univariate linear models*, J. Stat. Comp. Simul., 2 (1973), pp. 253–272.

[12] P. E. Gill, G. H. Golub, W. Murray and M. A. Saunders, *Methods for modifying matrix factorizations*, Math. Comp., 28 (1974), pp. 505–535.

[13] G. H. Golub and C. F. Van Loan, *Matrix computations*, Second Edition, The Johns Hopkins University Press, Baltimore, Maryland, 1989.

[14] N. J. Higham, *Accuracy and Stability of Numerical Algorithms*, Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 1996.

[15] C.-T. Pan, *A perturbation analysis of the problem of downdating a Cholesky factorization*, Linear Algebra Appl., 183 (1993), pp. 103–116.

[16] C.-T. Pan and R. Plemmons, *Least squares modification with inverse factorizations: parallel implications*, J. Comp. Appl. Math., 27 (1989), pp. 109–127.

[17] M. A. Saunders, *Large-scale linear programming using the Cholesky factorization*, Technical Report CS252, Computer Science Department, Stanford University, 1972.

[18] G. W. Stewart, *The effects of rounding error on an algorithm for downdating a Cholesky factorization*, J. Inst. Math. Appl., 23 (1979), pp. 203–213.

[19] G. W. Stewart, *On the Perturbation of LU and Cholesky Factors*, Technical Report, **CS-TR-3535 UMIACS-TR-95-93**, University of Maryland, Department of Computer Science, 1995.

[20] J.-G. Sun, *Perturbation analysis of the Cholesky downdating and QR updating problems*, SIAM J. Matrix Anal. Appl., 16 (1995), pp. 760–775.

[21] A. van der Sluis, *Condition numbers and equilibration of matrices*, Numerische Mathematik, 14 (1969), pp. 14–23.