Available at www.ElsevierMathematics.com
POWERED BY SCIENCE DIRECT®

Applied Mathematics and Computation 147 (2004) 601–606

APPLIED MATHEMATICS AND COMPUTATION

www.elsevier.com/locate/amc

Perturbation analysis for the generalized Cholesky factorization

Wei-guo Wang [a], Jin-xi Zhao [b,*]

[a] *Department of Mathematics, Ocean University of Qingdao, Shandong 266071, PR China*
[b] *State Key Laboratory for Novel Software Technology, Nanjing University, Nanjing 210093, PR China*

Abstract

Let K be a symmetric indefinite matrix. Suppose that $K = LJL^{\mathrm{T}}$ is the generalized Cholesky factorization of K. In this paper we present perturbation analysis for the generalized Cholesky factorization. We obtain the first-order bound on the norm of the perturbation in the generalized Cholesky factor. Also, we give rigorous perturbation bounds.

Keywords: Generalized Cholesky factorization; Perturbation bounds; Augmented matrix

1. Introduction

Consider the problem of solving the structured linear system

$$\begin{pmatrix} A & B^{\mathrm{T}} \\ B & -C \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} b \\ d \end{pmatrix}, \tag{1}$$

for x and y, where $A \in R^{m\times m}$ is symmetric positive definite matrix, $B \in R^{m\times n}$, x, $b \in R^m$, and $y, d \in R^n$, $C \in R^{n\times n}$. This system is called an augmented system, or an equilibrium system. The system (1) has been investigated by many authors for numerical algorithms. (See, [1,2,4,7,11,12])

* Corresponding author. Address: Department of Computer Science and Technology, Nanjing University, Nanjing 210008, China.
E-mail address: jxzhao@nju.edu.cn (J. Zhao).

0096-3003/$ - see front matter
doi:10.1016/S0096-3003(02)00798-1

In [1], the generalized Cholesky factorization is presented and this method inherits the advantage of Cholesky factorization with small storage and low computation costs. We first derive the generalized Cholesky factorization theorem.

Theorem 1.1 [1]. *Given any symmetric indefinite matrix*

$$K = \begin{pmatrix} A & B^{\mathrm{T}} \\ B & -C \end{pmatrix}, \tag{2}$$

where A, B *and* C *are the same as that defined in* (1). *Then we have*

$$K = LJL^{\mathrm{T}}, \tag{3}$$

$$L = \begin{pmatrix} L_{11} & \\ L_{21} & L_{22} \end{pmatrix}, \quad J = \begin{pmatrix} I_m & \\ & -I_n \end{pmatrix}, \tag{4}$$

where $L_{11} \in R^{m\times m}$ *and* $L_{22} \in R^{n\times n}$ *are lower triangular*, $L_{21} \in R^{n\times m}, I_m$ *and* I_n *are identity matrices.*

Let $\widetilde{K} = K + \Delta K$ be a perturbation of K in which ΔK is symmetric. If ΔK is sufficiently small, then $\widetilde{K}$ also has a generalized Cholesky factorization:

$$K + \Delta K = (L + \Delta L)J(L + \Delta L)^{\mathrm{T}}. \tag{5}$$

There have been several results dealing with the perturbation analysis for the Cholesky factor (see [3,8,9]). They obtained the first-order perturbation result. The result is sharpened in [5,6].

About (5), K. Veslić [10] has given some eigenvalue perturbation results. In this paper we derive the first-order perturbation bound and the rigorous perturbation bound for the generalized Cholesky factorization.

2. Perturbation theorems for the generalized Cholesky factorization

In the section, two perturbation theorems on the generalized Cholesky factor will be given. The symbols $\|\cdot\|_2$ and $\|\cdot\|_{\mathrm{F}}$ will be used for the spectral norm and the Frobenius norm, respectively.

We need a lemma controlling the triangular indefinite decomposition of $J + N$ for small N.

Lemma 2.1 [10]. *Let* N *be a Hermitian matrix with* $\|N\| < 1/2$, *then there exists a unique lower triangular matrix such that*

$$J + N = (I + \Gamma)J(I + \Gamma^*),$$

where

$$\|\Gamma\|_{\mathrm{F}} \leqslant \frac{\sqrt{2}}{1+\sqrt{1-2\|N\|_{\mathrm{F}}}}.$$

Theorem 2.2. *Let K be a symmetric indefinite matrix and $K = LJL^{\mathrm{T}}$ its generalized Cholesky factorization, let $G \in R^{(m+n)\times(m+n)}$ be symmetric matrix, and let $\Delta K = \varepsilon G$, for some $\varepsilon \geqslant 0$. If*

$$\rho((LL^{\mathrm{T}})^{-1}\Delta K) < \frac{1}{2}, \tag{6}$$

then $K + \Delta K$ has the unique generalized Cholesky factorization,

$$K + \Delta K = (L + \Delta L)J(L + \Delta L)^{\mathrm{T}}, \tag{7}$$

with ΔL satisfying

$$\Delta L = \varepsilon \dot{L}(0) + \mathrm{O}(\varepsilon^2), \tag{8}$$

where $\dot{L}(0)$ is defined by the unique generalized Cholesky factorization

$$K + tG = L(t)JL^{\mathrm{T}}(t), \quad |t| \leqslant \varepsilon, \tag{9}$$

and so satisfies the equations

$$LJ\dot{L}^{\mathrm{T}}(0) + \dot{L}(0)JL^{\mathrm{T}} = G, \tag{10}$$

$$\dot{L}^{\mathrm{T}}(0) = J\ \mathrm{up}(L^{-1}GL^{-\mathrm{T}}L^{\mathrm{T}}), \tag{11}$$

where the 'up'and 'low' notation is defined by

$$\mathrm{up}(X) = \begin{pmatrix} \frac{1}{2}x_{11} & x_{12} & \cdot & x_{1n} \\ 0 & \frac{1}{2}x_{22} & \cdot & x_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & x_{nn} \end{pmatrix}, \tag{12}$$

$$\mathrm{low}(X) = X - \mathrm{up}(X) = \mathrm{up}(X^{\mathrm{T}})^{\mathrm{T}}. \tag{13}$$

Proof. If (6) holds, then for all $|t| \leqslant \varepsilon$ the spectral radius of $tL^{-1}GL^{-\mathrm{T}}$ satisfies

$$\rho(tL^{-1}GL^{-\mathrm{T}}) = \rho(tL^{-\mathrm{T}}L^{-1}G) = \rho(t(LL^{\mathrm{T}})^{-1}G) < \frac{1}{2}.$$

Therefore for all $|t| \leqslant \varepsilon$, $K + tG = L(J + tL^{-1}GL^{-\mathrm{T}})L^{\mathrm{T}}$ is symmetric nonsingular matrix, and $J + tL^{-1}GL^{-\mathrm{T}}$ has m positive eigenvalues and n negative eigenvalues. From the previous Lemma 2.1., so $K + tG$ has the generalized Cholesky factorization (9). Notice that $R(0) = R$ and $R(\varepsilon) = R + \Delta R$, so (7) holds.

If we differentiate (9) and set $t = 0$ in the result we obtain the linear equation (10). From upper triangular $\dot{L}^{\mathrm{T}}(0)L^{-\mathrm{T}}$ in

$$(J\dot{L}^{\mathrm{T}}(0)L^{-\mathrm{T}})^{\mathrm{T}} + J\dot{L}^{\mathrm{T}}(0)L^{-\mathrm{T}} = L^{-1}GL^{-\mathrm{T}},$$

we see with the 'up' notation in (12) that (11) holds. Finally the Taylor expansion for $R(t)$ about $t = 0$ gives (8) at $t = \varepsilon$. $\square$

Using Theorem 2.2 we can now easily obtain the first-order perturbation bound by a different approach.

Theorem 2.3. *Let K be symmetric indefinite matrix and $K = LJL^{\mathrm{T}}$ be generalized Cholesky factorization, and let ΔK be a real symmetric matrix satisfying $\|\Delta K\|_{\mathrm{F}} \leqslant \varepsilon \|K\|_2$. If*

$$\varepsilon \|L^{-1}\|_2^2 \|K\|_2 < \frac{1}{2}, \tag{14}$$

then $K + \Delta K$ has the generalized Cholesky factorization

$$K + \Delta K = (L + \Delta L)J(L + \Delta L)^{\mathrm{T}}, \tag{15}$$

where

$$\frac{\|\Delta L\|_{\mathrm{F}}}{\|L\|_2} \leqslant \frac{1}{\sqrt{2}} \|L^{-1}\|_2^2 \|K\|_2 \varepsilon + \mathrm{O}(\varepsilon^2). \tag{16}$$

Proof. Let $G \equiv \Delta K/\varepsilon$ (if $\varepsilon = 0$, the theorem is trivial). Then

$$\|G\|_{\mathrm{F}} \leqslant \|K\|_2, \tag{17}$$

since

$$\rho((LL^{\mathrm{T}})^{-1}\Delta K) \leqslant \|(LL^{\mathrm{T}})^{-1}\Delta K\|_2 \leqslant \|L^{-1}\|_2^2 \varepsilon \|K\|_2 < \frac{1}{2}.$$

So the conclusion of Theorem 2.2 hold here. Because

$$2\|\mathrm{up}(X)\|_{\mathrm{F}}^2 = 2\|\mathrm{low}(X)\|_{\mathrm{F}}^2 = \|X\|_{\mathrm{F}}^2 - \frac{1}{2}(x_{11}^2 + x_{22}^2 + \cdots + x_{nn}^2) \leqslant \|X\|_{\mathrm{F}}^2,$$

i.e. $\|\mathrm{up}(X)\|_{\mathrm{F}} \leqslant (1/\sqrt{2})\|X\|_{\mathrm{F}}$ for any symmetric X.

We have from (11) that

$$\begin{aligned} \|\dot{L}^{\mathrm{T}}(0)\|_{\mathrm{F}} &= \|J\ \mathrm{up}(L^{-1}GL^{-\mathrm{T}})L^{\mathrm{T}}\|_{\mathrm{F}} \\ &= \|\mathrm{up}(L^{-1}GL^{-\mathrm{T}})L^{\mathrm{T}}\|_{\mathrm{F}} \leqslant \frac{1}{\sqrt{2}}\|L^{-1}GL^{-\mathrm{T}}\|_{\mathrm{F}}\|L^{\mathrm{T}}\|_2 \\ &\leqslant \frac{1}{\sqrt{2}}\|L^{-\mathrm{T}}\|_2^2\|L^{\mathrm{T}}\|_2\|G\|_{\mathrm{F}}, \end{aligned} \tag{18}$$

which, with (17), gives

$$\frac{\|\dot{L}^{\mathrm{T}}(0)\|_{\mathrm{F}}}{\|L^{\mathrm{T}}\|_2} \leqslant \frac{1}{\sqrt{2}}\|L^{-\mathrm{T}}\|_2^2\|K\|_2. \tag{19}$$

Then from the Taylor expansion, (16) follows immediately. □

Clearly from (16) we see $(1/\sqrt{2})\|L^{-1}\|_2^2\|K\|_2$ can be regarded as a measure of the sensitivity of the generalized Cholesky factorization.

3. New perturbation bound

Multiplying out the right-hand side of (5) and ignoring higher-order terms, we obtain a linear matrix equation for first-order approximation $\widetilde{\Delta L}$ to ΔL [5]

$$L\widetilde{\Delta L}^{\mathrm{T}} + \widetilde{\Delta L}L^{\mathrm{T}} = \Delta K. \tag{20}$$

About this equation, Our basic result is the following

$$\widetilde{\Delta L} = L\ \mathrm{low}(L^{-1}\Delta K(JL^{\mathrm{T}})^{-1}) \text{ and } \widetilde{\Delta L}^{\mathrm{T}} = \mathrm{up}(L^{-1}\Delta K(JL^{\mathrm{T}})^{-1})(JL^{\mathrm{T}}).$$

To see this, write

$$\begin{aligned} &L(\mathrm{up}(L^{-1}\Delta K(JL^{\mathrm{T}})^{-1})(JL^{\mathrm{T}})) + L\ (\mathrm{low}(L^{-1}\Delta K(JL^{\mathrm{T}})^{-1})(JL^{\mathrm{T}})) \\ &\quad = L(L^{-1}\Delta K(JL^{\mathrm{T}})^{-1})J\Delta L^{\mathrm{T}} = \Delta K. \end{aligned}$$

We can take norms in the expressions $\widetilde{\Delta L}^{\mathrm{T}}$ and $\widetilde{\Delta L}$ to get first-order perturbation bounds for the generalized Cholesky factorization, but it is possible to introduce degrees of freedom in the expressions that can later used to reduce the bounds. Specifically, for any nonsingular diagonal matrix D_L, we have

$$\begin{aligned} \widetilde{\Delta L} &= L\ \mathrm{low}(L^{-1}\Delta K(JL^{\mathrm{T}})^{-1}) = LD_L\mathrm{low}(D_L^{-1}L^{-1}\Delta K(JL^{\mathrm{T}})^{-1}) \\ &= \widehat{L}\ \mathrm{low}(\widehat{L}^{-1}\Delta K(JL^{\mathrm{T}})^{-1}), \end{aligned}$$

consequently

$$\begin{aligned} \|\widetilde{\Delta L}\|_{\mathrm{F}} &= \|\widehat{L}\ \mathrm{low}(\widehat{L}^{-1}\Delta K(JL^{\mathrm{T}})^{-1})\|_{\mathrm{F}} \leqslant \|\widehat{L}\|_2\|\widehat{L}^{-1}\|_2\|\Delta K\|_{\mathrm{F}}\|J\|_2\|L^{-\mathrm{T}}\|_2 \\ &= \|\widehat{L}\|_2\|\widehat{L}^{-1}\|_2\|L^{-\mathrm{T}}\|_2\|\Delta K\|_{\mathrm{F}}, \end{aligned} \tag{21}$$

or

$$\frac{\|\widetilde{\Delta L}\|_{\mathrm{F}}}{\|L\|_2} \leqslant \kappa(\widehat{L})\kappa(L)\frac{\|\Delta K\|_{\mathrm{F}}}{\|L\|_2^2},$$

where $\kappa(L) = \|L\|_2\|L^{-1}\|_2$.

Since $\|K\|_2 \leqslant \|L\|_2^2$, we have

$$\frac{\|\widetilde{\Delta L}\|_{\mathrm{F}}}{\|L\|_2} \leqslant \kappa(\widehat{L})\kappa(L)\frac{\|\Delta K\|_{\mathrm{F}}}{\|K\|_2}.$$

If $K(\widehat{L}) = 1$(it cannot be less), then the bound (21) reduces to

$$\|\widetilde{\Delta L}\|_{\mathrm{F}} \leqslant \|L^{-1}\|_2\|\Delta K\|_{\mathrm{F}}. \tag{22}$$

References

[1] J. Zhao, The generalized Cholesky factorization method for solving saddle point problems, Appl. Math. Comput. 92 (1998) 49–58.

[2] G.H. Golub, C.F. Vanloan, Matrix Computation, third ed., The Johns Hopkins University Press, Baltimore, MD, 1996.

[3] J. Sun, Perturbation bounds for the Cholesky and QR factorizations, BIT 31 (1991) 341–352.

[4] P.E. Gill, M.A. Saunders, J.R. Shinnerl, On the stability of Cholesky factorization for symmetric quasidefinite systems, SIAM J. Matrix Anal. Appl. 17 (1) (1996) 35–46.

[5] G.W. Stewart, On the Perturbation of LU and Cholesky factors, IMA J. Numer. Anal. 17 (1997) 1–6.

[6] X. Chang, C.C. Paige, G.W. Stewart, New perturbation analyses for the Cholesky factorization, IMA J. Numer. Anal. 16 (1996) 457–484.

[7] A. Forsgren, P.E. Gill, J.R. Shinnerl, Stability of symmetric ill-conditioned systems arising in interior methods for constrained optimization, SIAM J. Matrix Anal. 17 (1) (1996) 187–211.

[8] J. Stoer, R. Bulirsch, Introduction to Numerical Analysis, second ed., Springer-Verlag, New York, 1993.

[9] G.W. Stewart, Perturbation bounds for the QR factorization of matrix, SIAM J. Numer. Anal. 14 (1977) 509–518.

[10] K. Veslić, Perturbation theory for the eigenvalues of factorised symmetric matrices, Linear Algebra Appl. 309 (2000) 85–102.

[11] Å. Björck, C.C. Paige, Solution of augmented linear systems using orthogonal factorizations, BIT 34 (1994) 1–24.

[12] Å. Björck, Numerical stability of methods for solving augmented systems, Contemp. Math. 204 (1997) 51–60.