Updating the Cholesky Factorization

Igor Kohanovsky

Prague, Czech Republic

Abstract

The problem of updating Cholesky LL^T factorization was treated in [1] and [2]. Based on these works algorithms are described that compute the factorization $\tilde{A} = \tilde{L}\tilde{L}^T$ where $\tilde{A}$ is the matrix $A = LL^T$ after it has a row and the symmetric column added or deleted. This is achieved by updating the factor L.

Let's $A \in R^{n \times n}$ be symmetric positive definite matrix. Applying Cholesky's method to A yields the factorization

$$A = LL^T$$

where L is lower triangular matrix with positive diagonal elements.

Suppose $\tilde{A}$ is a positive definite matrix created by adding a row and the symmetric column to A:

$$\tilde{A} = \begin{pmatrix} A & d \\ d^T & \gamma \end{pmatrix} , \qquad d^T \in R^n$$

Then its Cholesky factorization is (see [1]):

$$\tilde{L} = \begin{pmatrix} L & \\ e^T & \alpha \end{pmatrix} ,$$

where

$$e = L^{-1} d ,$$

and

$$\alpha = \sqrt{\tau} , \quad \tau = \gamma - e^T e, \quad \tau > 0.$$

If $\tau \leq 0$ then $\tilde{A}$ is not positive definite.

Now assume $\tilde{A}$ is obtained from A by deleting row and column r of A. Matrix $\tilde{A}$ is the positive definite matrix (as any principal square submatrix of the positive definite matrix). It is shown in [2] how to get $\tilde{L}$ from L in such case. Let us partition A and L along row and column r as follows:

$$A = \left(\begin{array}{ccc} A_{11} & a_{1r} & A_{12} \\ a_{1r}^T & a_{rr} & a_{2r}^T \\ A_{21} & a_{2r} & A_{22} \end{array} \right), \qquad L = \left(\begin{array}{ccc} L_{11} & & \\ l_{1r}^T & l_{rr} & \\ L_{21} & l_{2r} & L_{22} \end{array} \right).$$

Then $\tilde{A}$ can be written as

$$\tilde{A} = \left(\begin{array}{cc} A_{11} & A_{12} \\ A_{21} & A_{22} \end{array} \right).$$

By deleting row r from L we obtain the matrix

$$H = \left(\begin{array}{ccc} L_{11} & & \\ L_{21} & l_{2r} & L_{22} \end{array} \right)$$

with the property that $HH^T = \tilde{A}$. Factor $\tilde{L}$ can be obtained from H by applying Givens rotations to the matrix elements $h_{r,r+1}, h_{r+1,r+2}, \ldots, h_{n-1,n}$ and deleting the last column of the obtained matrix $\tilde{H}$:

$$\tilde{H} = HR = \left(\begin{array}{ccc} L_{11} & & \\ L_{21} & \tilde{L}_{22} & 0 \end{array} \right) = \left(\tilde{L} \; 0 \right),$$

where $R = R_0 R_1 \ldots R_{n-1-r}$ is the matrix of Givens rotations, L_{11} and $\tilde{L}_{22}$ — lower triangle matrices, and

$$\begin{aligned} &\tilde{L} = \left(\begin{array}{cc} L_{11} & \\ L_{21} & \tilde{L}_{22} \end{array} \right), \\ &\tilde{H}\tilde{H}^T = HRR^T H^T = HH^T = \tilde{A}, \\ &\tilde{H}\tilde{H}^T = \tilde{L}\tilde{L}^T, \quad \tilde{L}\tilde{L}^T = \tilde{A}. \end{aligned}$$

Orthogonal matrix R_t $(0 \le t \le n-1-r), R_t \in R^{n\times n}$ which defines the Givens rotation annulating $(r+t, r+t+1)$th element of the $H_k R_0 \ldots R_{t-1}$ matrix is defined as

$$R_t = \left(\begin{array}{cccccc} 1 & \ldots & 0 & 0 & \ldots & 0 \\ \vdots & \ldots & \ldots & \ldots & \ldots & \vdots \\ 0 & \ldots & c & s & \ldots & 0 \\ 0 & \ldots & -s & c & \ldots & 0 \\ \vdots & \ldots & \ldots & \ldots & \ldots & \vdots \\ 0 & \ldots & 0 & 0 & \ldots & 1 \end{array} \right).$$

where entries $(r+t, r+t)$ and $(r+t+1, r+t+1)$ equal c, $(r+t, r+t+1)$ entry equals s, and $(r+t+1, r+t)$ one equals $-s$, where $c^2+s^2=1$. Let's $\tilde{l}_{ij}^{t-1}$ — coefficients of matrix $H_k R_0 \dots R_{t-1}$ ($\tilde{l}_{ij}^{-1}$ — coefficients of H_k). and

$$c = \frac{\tilde{l}_{r+t,r+t}^{t-1}}{\sqrt{(\tilde{l}_{r+t,r+t}^{t-1})^2 + (\tilde{l}_{r+t,r+t+1}^{t-1})^2}} \ , \quad s = \frac{\tilde{l}_{r+t,r+t+1}^{t-1}}{\sqrt{(\tilde{l}_{r+t,r+t}^{t-1})^2 + (\tilde{l}_{r+t,r+t+1}^{t-1})^2}} \ ,$$

Then matrix $H_k R_0 \dots R_t$ will differ from $H_k R_0 \dots R_{t-1}$ with entries of $(r+t)$ $(r+t+1)$ columns only, thereby

$$\tilde{l}_{i,r+t}^{t} = \tilde{l}_{i,r+t}^{t-1} = 0 \, , \ \tilde{l}_{i,r+t+1}^{t} = \tilde{l}_{i,r+t+1}^{t-1} = 0 \, , \quad 1 \le i \le r+t-1 \, ,$$

$$\begin{aligned} \tilde{l}_{i,r+t}^{t} &= c\tilde{l}_{i,r+t}^{t-1} + s\tilde{l}_{i,r+t+1}^{t-1} \, , \\ \tilde{l}_{i,r+t+1}^{t} &= -s\tilde{l}_{i,r+t}^{t-1} + c\tilde{l}_{i,r+t+1}^{t-1} \quad r+t \le i \le n-1 \, . \end{aligned}$$

Where

$$\tilde{l}_{r+t,r+t}^{t} = \sqrt{(\tilde{l}_{r+t,r+t}^{t-1})^2 + (\tilde{l}_{r+t,r+t+1}^{t-1})^2} \, , \qquad \tilde{l}_{r+t,r+t+1}^{t} = 0 \, .$$

Also, coefficient $\tilde{l}_{r+t,r+t}^{t}$ is a nonnegative one. In order to avoid unnecessary overflow or underflow during computation of c and s, it was recommended (see [3]) to calculate value of the square root $w = \sqrt{x^2+y^2}$ as folows :

$$\begin{aligned} v &= \max\{|x|, |y|\} \ , \qquad u = \min\{|x|, |y|\} \ , \\ w &= \begin{cases} v\sqrt{1 + \left(\frac{u}{v}\right)^2} \ , & v \ne 0 \\ 0 \ , \quad v = 0 \end{cases} \ . \end{aligned}$$

This formula is for machine that use normalize base 2 arithmetic.

Applying the updating technique to Cholesky factorization allows significally reduce the complexity of calculations. So, in case of adding a row and the symmetric column to the original matrix it will be necessary to carry out about n^2 flops instead of about $\frac{(n+1)^3}{3}$ flops for the direct calculation of the new Cholesky factor. In the case of deleting a row and the symmetric column from the original matrix, the new Cholesky factor can be obtained with about $3(n-r)^2$ flops (the worst case requires about $3(n-1)^2$ operations) instead of about $\frac{(n-1)^3}{3}$ flops required for its direct calculation.

References

[1] J. A. George and J. W-H. Liu. *Computer Solution of Large Sparse Positive Definite Systems*, Prentice-Hall, 1981

[2] T. F. Coleman, L. A. Hulbert. A direct active set algorithm for large sparse quadratic programs with simple bounds. // *Mathematical Programming*, **45** (1989), 373–406.

[3] C. L. Lawson, R. J. Hanson. *Solving least squares problems*, SIAM, 1995