Cholesky decomposition

Canonical name	CholeskyDecomposition
Date of creation	2013-03-22 12:07:38
Last modified on	2013-03-22 12:07:38
Owner	gufotta (12050)
Last modified by	gufotta (12050)
Numerical id	16
Author	gufotta (12050)
Entry type	Definition
Classification	msc 62J05
Classification	msc 65-00
Classification	msc 15-00
Synonym	Cholesky factorization
Synonym	matrix square root
Related topic	SquareRootOfPositiveDefiniteMatrix
Defines	Cholesky triangle

1 Cholesky Decomposition

A symmetric and positive definite matrix can be efficiently decomposed into a lower and upper triangular matrix. For a matrix of any type, this is achieved by the LU decomposition which factorizes $A = LU$. If A satisfies the above criteria, one can decompose more efficiently into $A = LL^T$ where L is a lower triangular matrix with positive diagonal elements. L is called the *Cholesky triangle*.

To solve $Ax = b$, one solves first $Ly = b$ for y, and then $L^T x = y$ for x.

A variant of the Cholesky decomposition is the form $A = R^T R$, where R is upper triangular.

Cholesky decomposition is often used to solve the normal equations in linear least squares problems; they give $A^T Ax = A^T b$, in which $A^T A$ is symmetric and positive definite.

To derive $A = LL^T$, we simply equate coefficients on both sides of the equation:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ a_{31} & a_{32} & \cdots & a_{3n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} = \begin{bmatrix} l_{11} & 0 & \cdots & 0 \\ l_{21} & l_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ l_{n1} & l_{n2} & \cdots & l_{nn} \end{bmatrix} \begin{bmatrix} l_{11} & l_{21} & \cdots & l_{n1} \\ 0 & l_{22} & \cdots & l_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & l_{nn} \end{bmatrix}$$

Solving for the unknowns (the nonzero l_{ji}s), for $i = 1, \cdots, n$ and $j = i - 1, \ldots, n$, we get:

$$\begin{aligned} l_{ii} &= \sqrt{\left(a_{ii} - \sum_{k=1}^{i-1} l_{ik}^2 \right)} \\ l_{ji} &= \left(a_{ji} - \sum_{k=1}^{i-1} l_{jk} l_{ik} \right) / l_{ii} \end{aligned}$$

Because A is symmetric and positive definite, the expression under the square root is always positive, and all l_{ij} are real.

References

[1] Originally from The Data Analysis Briefbook (`http://rkb.home.cern.ch/rkb/titleA.html`http://rkb.home.cern.ch/rkb/titleA.html)