Cholesky Decomposition

For any symmetric positive definite matrix, Cholesky Decomposition provides a unique solution for the lower triangular matrix.

The method is analytically tractable, and we can see how it works with an example and general algorithm.

Two-Dimensional Example

Let's introduce a two-variate covariance matrix $\mathbf{\Sigma}$

$$\mathbf{\Sigma} = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix}$$

Cholesky Decomposition takes the form the product of two triangular matrices $\mathbf{\Sigma} = \mathbf{L}\mathbf{L}^{\mathbf{T}} = \mathbf{A}\mathbf{A}^{\mathbf{T}}$

$$\mathbf{\Sigma} = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix} = \begin{pmatrix} a_{11} & 0 \\ a_{21} & a_{22} \end{pmatrix} \times \begin{pmatrix} a_{11} & a_{21} \\ 0 & a_{22} \end{pmatrix}$$

with real and positive diagonal elements $a_{11} = \sigma_1^2$ and $a_{22} = \sigma_2^2$.

Two-Dimensional Example (Continued)

Completing multiplication $\mathbf{AA^T}$ we obtain

$$\begin{pmatrix} a_{11}^2 & a_{11}a_{21} \\ a_{21}a_{11} & a_{21}^2 + a_{22}^2 \end{pmatrix} = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix}$$

Given that $\sigma_{12} = \sigma_{21} = \rho\sigma_1\sigma_2$, we have 3 equations for 3 unknowns

$$\begin{cases} a_{11}^2 = \sigma_1^2 \\ a_{21}a_{11} = \rho\sigma_1\sigma_2 \\ a_{21}^2 + a_{22}^2 = \sigma_2^2 \end{cases}$$

Solving for a_{ij} in sequence, we obtain the solution for $\mathbf{A}$ as

$$\begin{pmatrix} \sigma_1 & 0 \\ \rho\sigma_2 & \sqrt{1-\rho^2}\sigma_2 \end{pmatrix}$$

Two-Dimensional Example (Application)

Remember that the purpose of obtaining **A** was to impose correlation on an arbitrarily drawn vector of independent standard Normal variables **Z** by

$$\mathbf{X} = \mathbf{AZ}$$

Continuing our example

$$\mathbf{X} = \begin{pmatrix} \sigma_1 & 0 \\ \rho\sigma_2 & \sqrt{1-\rho^2}\sigma_2 \end{pmatrix} \times \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}$$

so that

$$\begin{aligned} x_1 &= \sigma_1 z_1 \\ x_2 &= \rho\sigma_2 z_1 + \sqrt{1-\rho^2}\sigma_2 z_2 \end{aligned}$$

x_i are converted to default times using their own CDF as $\tau_i = F^{-1}(\Phi(x_i))$.

General Algorithm for Cholesky Decomposition

For the d-dimensional symmetric matrix $\boldsymbol{\Sigma}$, we need to solve the system of equations stemming from the product $\boldsymbol{\Sigma} = \mathbf{L}\mathbf{L}^{\mathbf{T}} = \mathbf{A}\mathbf{A}^{\mathbf{T}}$

$$\boldsymbol{\Sigma} = \begin{pmatrix} a_{11} & & & \\ a_{21} & a_{22} & & \\ \vdots & \vdots & \ddots & \\ a_{d1} & a_{d2} & \cdots & a_{dd} \end{pmatrix} \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{d1} \\ & a_{22} & \cdots & a_{d2} \\ & & \ddots & \vdots \\ & & & a_{dd} \end{pmatrix} =$$

$$\begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1d} \\ \vdots & \sigma_{22} & \cdots & \sigma_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{d1} & \cdots & \cdots & \sigma_{dd} \end{pmatrix}$$

where, again, all diagonal elements are real and strictly positive.

General Algorithm (Continued)

Looping over the σ_{ij} by rows i and then columns j gives (starting with the first row of $\mathbf{A}$ by the first column of $\mathbf{A}^{\mathbf{T}}$)

$$
\begin{aligned}
a_{11}^2 &= \sigma_{11} \\
a_{11}a_{21} &= \sigma_{12} \\
&\vdots \\
a_{11}a_{d1} &= \sigma_{1d} \\
a_{21}^2 + a_{22}^2 &= \sigma_{22} \\
&\vdots \\
a_{21}a_{d1} + a_{22}a_{d2} &= \sigma_{2d}
\end{aligned}
$$

Exactly one new entry of the matrix $\mathbf{A}$ appears in each equation, making it possible to solve the system in sequence.

General Algorithm (Completed)

We can express the generalised Cholesky Decomposition in the form

$$\sigma_{ij} = \sum_{k=1}^{i} a_{ik} a_{jk} \quad \forall j \geq i \tag{4}$$

Separate the last term as $\sum_{k=1}^{i-1} a_{ik} a_{jk} + a_{ii} a_{ji}$ and rearranging gives solution for a_{ji}

$$a_{ji} = \frac{1}{a_{ii}} \left(\sigma_{ij} - \sum_{k=1}^{i-1} a_{ik} a_{jk} \right) \quad \forall j \geq i$$

Solution for a_{ii} also comes from (4) by setting $j = i$ and rearranging as

$$a_{ii} = \sqrt{\sigma_{ii} - \sum_{k=1}^{i-1} a_{ik}^2} \quad \forall j = i$$

Spectral Decomposition

The properties of spectral decomposition of a symmetric matrix into its eigenvalues and eigenvectors also allow obtaining matrix $\mathbf{A}$

$$\mathbf{\Sigma} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^T = \mathbf{V}\mathbf{\Lambda}^{\frac{1}{2}}\left(\mathbf{V}\mathbf{\Lambda}^{\frac{1}{2}}\right)^T$$

where $\mathbf{\Lambda}$ is the diagonal matrix of eigenvalues and $\mathbf{V}$ is the vectorised matrix of eigenvectors (in columns). Eigenvectors are orthogonal to each other and represent the directions of zero covariance (among themselves).

We can simply notice by functional similarity that

$$\mathbf{A} = \mathbf{V}\mathbf{\Lambda}^{\frac{1}{2}}$$

Cholesky vs. Spectral: Pros and Cons

It is clear how the Cholesky Decomposition wins a computational advantage: it is a sequential solving through the organised system of equations where only one new unknown comes up at a time.

The restriction is that the Cholesky Decomposition does not work when matrix $\mathbf{\Sigma}$ is only positive semi-definite.

In that case, we carry out the Spectral Decomposition, which is also occasionally useful in terms of Principal Component Analysis in order to reveal the internal structure of correlation or covariance matrix. Imposing correlation by using only several principal components is also a possibility.