CHAPTER 4

THE CHOLESKY DECOMPOSITION

In Chapter 8 we will discuss a general procedure for generating correlated standard normal random variables. The key machinery there is the *Cholesky decomposition* of a positive definite (PD) or positive semi-definite (PSD) matrix. Here we prove the existence of the relevant decomposition. The mechanics of its calculation are left for Chapter 8.

§1. The Cholesky Decomposition

Theorem 1 (Cholesky Decomposition) Let $\mathbf{V} = (v_{ij})_{n\times n}$ be a symmetric PSD matrix. Then there is a lower triangular matrix $\mathbf{L}$ with non-negative diagonal entries such that $\mathbf{V} = \mathbf{L}\mathbf{L}^T$. If $\mathbf{V}$ is PD, there is a lower triangular matrix $\mathbf{L}$ with strictly positive diagonal entries such that $\mathbf{V} = \mathbf{L}\mathbf{L}^T$.

Before proving this we develop a few preliminaries. We assume the reader is familiar with Gaussian elimination, the properties of matrix multiplication, the notion of a matrix's rank, and a matrix's inverse. Using Gaussian elimination, we plan to *diagonalize* the matrix $\mathbf{V}$. For $1 \le i, j \le n$ with $i \ne j$, let $\mathbf{E}_{i,j;c}$ denote the $n \times n$ elementary matrix with ones on the diagonal and the number c at row i, column j, and zeros elsewhere:

$$\mathbf{E}_{i,j;c} = \begin{array}{cc} & \begin{matrix} & & j & & & \end{matrix} \\ \begin{matrix} \\ \\ \\ i \\ \\ \\ \end{matrix} & \begin{bmatrix} 1 & & \vdots & & & \\ & 1 & \vdots & & & \\ & & \ddots & & & \\ \cdots & \cdots & c & \ddots & \cdots & \cdots \\ & & \vdots & & 1 & \\ & & \vdots & & & 1 \end{bmatrix} \end{array} .$$

If $\mathbf{A}$ is any $n \times n$ matrix, the matrix product $\mathbf{E}_{i,j;c}\mathbf{A}$ is identical to $\mathbf{A}$ except that the i^{th} row of $\mathbf{E}_{i,j;c}\mathbf{A}$ is the i^{th} row of $\mathbf{A}$ plus c times the j^{th} row of $\mathbf{A}$. We call this "pre-multiplication" by $\mathbf{E}_{i,j;c}$. Furthermore, the matrix product $\mathbf{A}\mathbf{E}_{i,j;c}^T$ is the identical to $\mathbf{A}$ except that the i^{th} column of $\mathbf{A}\mathbf{E}_{i,j;c}^T$ is the i^{th} column of $\mathbf{A}$

plus c times the j^{th} column of $\mathbf{A}$. We call this "post-multiplication" by $\mathbf{E}_{i,j;c}^T$. If $1 \le j < i \le n$, then $\mathbf{E}_{i,j;c}$ is of the form

$$\text{lower triangular with all ones on the diagonal.} \tag{1}$$

The product of any two matrices of the form (1) is also of this form; any matrix of this form is invertible since its determinant is one (the product of the diagonal entries).

As we have seen, there are two other kinds of elementary matrices that we will not use here — one switches two specified rows and the other multiplies a specified row by a non-zero constant. Any reference to an "LTE matrix" in the following refers to a lower triangular elementary matrix $\mathbf{E}_{i,j;c}$. Here we allow $c = 0$ to include the $n \times n$ identity matrix $\mathbf{I}$.

In the following two lemmas, $\mathbf{A} = (a_{ij})$ is a symmetric $n \times n$ matrix.

Fact 1. Let $\mathbf{L}$ be of form (1). If $\mathbf{A}$ is PD then $\mathbf{LAL}^T$ is symmetric and PD. If $\mathbf{A}$ is PSD then $\mathbf{LAL}^T$ is symmetric and PSD.

Proof. It is easy to see that, in either case, $\mathbf{LAL}^T$ is symmetric:

$$(\mathbf{LAL}^T)^T \;=\; \mathbf{L}^{T^T}\mathbf{A}^T\mathbf{L}^T \;=\; \mathbf{LAL}^T.$$

Now suppose $\mathbf{A}$ is PD and let $\mathbf{x}$ be any non-zero column vector. Then

$$\mathbf{x}^T(\mathbf{LAL}^T)\mathbf{x} \;=\; (\mathbf{L}^T\mathbf{x})^T\mathbf{A}(\mathbf{L}^T\mathbf{x}) \;>\; 0,$$

so $\mathbf{LAL}^T$ is PD. (Note that $\mathbf{L}^T\mathbf{x}$ is non-zero since $\mathbf{L}^T$ is invertible and $\mathbf{x}$ is non-zero.) If $\mathbf{A}$ is PSD replace the "$>$" with "$\ge$". $\square$

Fact 2. If $\mathbf{A}$ is PD, then all its diagonal entries are positive. If $\mathbf{A}$ is PSD, then all its diagonal entries are non-negative. If the PSD matrix $\mathbf{A}$ has $a_{ii} = 0$ then both the i^{th} row and i^{th} column of $\mathbf{A}$ are entirely 0.

Proof. Let $\mathbf{e}_i = (0, \ldots, 0, 1, 0, \ldots, 0)^T$, where the 1 is in the i^{th} position. Then $a_{ii} = \mathbf{e}_i^T\mathbf{A}\mathbf{e}_i$. If $\mathbf{A}$ is PD, this must be positive. If $\mathbf{A}$ is PSD, this must be non-negative. Suppose $\mathbf{A}$ is PSD and $a_{ii} = 0$ for some row/column i. Fix any $j \ne i$ and let $\mathbf{x} = (0, \ldots, 0, h, 0, \ldots, 0, 1, 0, \ldots, 0)^T$, where the h is in the i^{th} position and the 1 is in the j^{th} position. (The illustrated vector shows this for $j > i$, but we could also have $j < i$.) Using that $a_{ii} = 0$ and $a_{ji} = a_{ij}$, we have

$$\begin{aligned}
\mathbf{x}^T\mathbf{A}\mathbf{x} \;&=\; \sum_{s=1}^{n}\sum_{t=1}^{n} x_s a_{st} x_t \\
&=\; x_i a_{ii} x_i \;+\; x_i a_{ij} x_j \;+\; x_j a_{ji} x_i \;+\; x_j a_{jj} x_j \\
&=\; 2h a_{ij} + a_{jj}.
\end{aligned}$$

If $a_{ij} \neq 0$ we can make $\mathbf{x}^T\mathbf{A}\mathbf{x} = -1$ by taking $h = (-1 - a_{jj})/(2a_{ij})$. Hence we must have $a_{ij} = 0$. This proves the entire i^{th} row is 0. Since $\mathbf{A}$ is symmetric, the entire i^{th} column is also 0. □

Proof of Theorem 1. Suppose $\mathbf{V}$ is PSD. We wish to convert $\mathbf{V}$ to a diagonal matrix via elementary row and column operations. By Fact 2, $v_{11} \geq 0$.

Case 1: $v_{11} > 0$. Using Gaussian elimination, we may make the first column below v_{11} all zeros. We may accomplish this by a sequence of pre-multiplications by LTE matrices, i.e, for some LTE matrices $\mathbf{E}_1, \ldots, \mathbf{E}_{n-1}$, we have

$$\mathbf{E}_{n-1}\mathbf{E}_{n-2}\cdots\mathbf{E}_2\mathbf{E}_1\mathbf{V} \;=\; \mathbf{V}_2,$$

where $\mathbf{V}_2$ is of the form:

$$\mathbf{V}_2 \;=\; \left[\begin{array}{c|ccc} d_1 & * & \cdots & * \\ \hline 0 & & & \\ \vdots & & & \\ 0 & & & \end{array}\right],$$

where $d_1 = v_{11}$. Since $\mathbf{V}$ is symmetric and the first row of $\mathbf{V}$ and $\mathbf{V}_2$ are identical, the same sequence of operations performed on the *columns* of $\mathbf{V}_2$ will make the entries to the right of d_1 all zero. In other words, $\mathbf{V}_2\mathbf{E}_1^T\cdots\mathbf{E}_{n-1}^T = \mathbf{V}_3$, where $\mathbf{V}_3$ is of the form

$$\mathbf{V}_3 \;=\; \left[\begin{array}{c|ccc} d_1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & & \\ 0 & & & \end{array}\right].$$

Now

$$\mathbf{V}_3 \;=\; \mathbf{E}_{n-1}\cdots\mathbf{E}_1\,\mathbf{V}\,\mathbf{E}_1^T\cdots\mathbf{E}_{n-1}^T \;=\; \mathbf{L}_1\mathbf{V}\mathbf{L}_1^T,$$

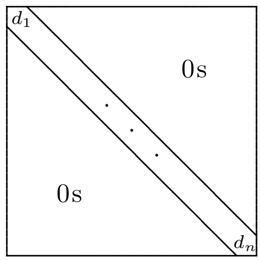
where $\mathbf{L}_1 = \mathbf{E}_{n-1}\cdots\mathbf{E}_1$ is of form (1). By Fact 1, $\mathbf{V}_3$ is still symmetric and PSD.

Case 2: $v_{11} = 0$. By Fact 2, the entire first row and column of $\mathbf{V}$ are already zeros, and we may take $\mathbf{V}_3 = \mathbf{L}_1\mathbf{V}\mathbf{L}_1^T$ where $\mathbf{L}_1 = \mathbf{I}$, the $n \times n$ identity matrix, and $d_1 = 0$. In this case also, $\mathbf{L}_1$ is a matrix of the form (1) and $\mathbf{V}_3$ (which in this case is just $\mathbf{V}$) is symmetric and PSD. We may continue row-and-column by row-and-column in this fashion until we produce

$$\mathbf{L}_n \cdots \mathbf{L}_1 \, \mathbf{V} \, \mathbf{L}_1^T \cdots \mathbf{L}_n^T \;=\; \mathbf{D},$$

where $\mathbf{D}$ is of the form

$\mathbf{D}$ =



,

and each $\mathbf{L}_i$ is a product of LTE matrices. The product $\mathbf{M} = \mathbf{L}_n \cdots \mathbf{L}_1$ is therefore also a product of LTE matrices and hence is of form (1). Now $\mathbf{D} = \mathbf{M}\mathbf{V}\mathbf{M}^T$ so, by Fact 1, $\mathbf{D}$ is PSD and therefore by Fact 2 has non-negative diagonal entries (strictly positive if $\mathbf{V}$ is PD).

Let $\mathbf{N} = \mathbf{M}^{-1}$. Note that $\mathbf{N}$ is also a product of LTE matrices since $\mathbf{M}$ is such a product and the inverse of an LTE matrix is an LTE matrix (the reader will verify that $\mathbf{E}_{i,j;c}^{-1} = \mathbf{E}_{i,j;-c}$). It follows that $\mathbf{N}$ is also of form (1). Since $(\mathbf{M}^T)^{-1} = (\mathbf{M}^{-1})^T$, we have $\mathbf{V} = \mathbf{N}\mathbf{D}\mathbf{N}^T$, where $\mathbf{N}$ is of form (1).

Letting $\sqrt{\mathbf{D}}$ denote the diagonal matrix whose ii^{th} entry is $\sqrt{d_i}$, we have that $\mathbf{V} = \mathbf{N}\mathbf{D}\mathbf{N}^T = \mathbf{N}\sqrt{\mathbf{D}}\sqrt{\mathbf{D}}\mathbf{N}^T = \mathbf{N}\sqrt{\mathbf{D}}(\mathbf{N}\sqrt{\mathbf{D}})^T$. (Using that $(\sqrt{\mathbf{D}})^T = \sqrt{\mathbf{D}}$.) Taking $\mathbf{L} = \mathbf{N}\sqrt{\mathbf{D}}$ establishes the decomposition. If $\mathbf{V}$ is PD, then all the diagonal entries of $\mathbf{D}$, and hence $\sqrt{\mathbf{D}}$ and $\mathbf{L}$, are strictly positive. □