CHAPTER 8

GENERATING MULTIVARIATE NORMALS

§1. The Joint Normal Density

The vector of random variables $\mathbf{Y} = (Y_1, \ldots, Y_n)^T$ has a mean-zero n-dimensional multivariate normal distribution if, for some invertible linear transformation $\mathbf{A} = (a_{ij})_{n \times n}$, the vector of random variables $\mathbf{X} = (X_1, \ldots, X_n)^T$ defined by $\mathbf{X} = \mathbf{A}^{-1}\mathbf{Y}$ are iid standard (mean 0, variance 1) normals. Then $\mathbf{Y} = \mathbf{AX}$ and matrix multiplication yields

$$Y_i \;=\; \sum_{j=1}^{n} a_{ij} X_j, \quad 1 \le i \le n.$$

We are interested in calculating the joint probability density of $\mathbf{Y}$. We begin by calculating the joint density of $\mathbf{X}$. This is much easier as the X_is are independent Normal$\,(0,1)$s. With $\mathbf{x} = (x_1, \ldots, x_n)^T$ we have

$$\begin{aligned} f_{\mathbf{X}}(\mathbf{x}) \;&=\; \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} e^{-x_i^2/2} \\ &=\; \frac{1}{(2\pi)^{n/2}} \exp\left(-\sum_{i=1}^{n} x_i^2 \Big/ 2 \right) \;=\; \frac{1}{(2\pi)^{n/2}}\, e^{-(\mathbf{x}^T\mathbf{x})/2}, \end{aligned} \tag{1}$$

where the matrix multiplication $\mathbf{x}^T\mathbf{x}$ gives the inner product (i.e., dot product) of $\mathbf{x}$ with itself.

Now suppose $I \subset \mathbb{R}^n$ is a tiny n-dimensional cube and let $\mathbf{y} = (y_1, \ldots, y_n)^T$ be a point in I. The joint density function $f_{\mathbf{Y}}(\cdot)$ is characterized by the property that

$$P[\mathbf{Y} \in I] \;\approx\; f_{\mathbf{Y}}(\mathbf{y}) \cdot \mathrm{Vol}\,(I),$$

where Vol$\,(\cdot)$ denotes n-dimensional volume. Working on the left side of this and letting $\mathbf{x} = \mathbf{A}^{-1}\mathbf{y} \in \mathbf{A}^{-1}I$, we have

$$\begin{aligned} P[\mathbf{Y} \in I] \;=\; P[\mathbf{AX} \in I] \;&=\; P[\mathbf{X} \in \mathbf{A}^{-1}I] \\ &\approx\; f_{\mathbf{X}}(\mathbf{x}) \cdot \mathrm{Vol}\,(\mathbf{A}^{-1}I) \\ &=\; \frac{1}{(2\pi)^{n/2}}\, e^{-(\mathbf{x}^T\mathbf{x})/2} \cdot |\det \mathbf{A}^{-1}| \cdot \mathrm{Vol}\,(I) \\ &=\; \frac{1}{(2\pi)^{n/2} |\det \mathbf{A}|}\, e^{-(\mathbf{x}^T\mathbf{x})/2} \cdot \mathrm{Vol}\,(I). \end{aligned} \tag{2}$$

See Chapter 2 for a discussion of the determinant function and its relation to volume.

The next step is to restate this result in terms of the covariance matrix $\mathbf{V} = (v_{ij})$ for $\mathbf{Y}$, which is denoted by $\mathrm{Cov}\,(\mathbf{Y})$ and defined by $v_{ij} = \mathrm{Cov}\,(Y_i, Y_j)$. Then we have

$$\begin{aligned}
\mathbf{V}_{ij} &= \mathrm{Cov}\left(\sum_{s=1}^{n} a_{is}X_s, \sum_{t=1}^{n} a_{jt}X_t\right) \\
&= \sum_s \sum_t a_{is}a_{jt}\mathrm{Cov}\,(X_s, X_t) \\
&= \sum_s a_{is}a_{js} \;=\; \sum_s a_{is}a^T_{sj} \;=\; (\mathbf{A}\mathbf{A}^T)_{ij},
\end{aligned}$$

where the third equality holds since $\mathrm{Cov}\,(X_s, X_t) = 0$ if $t \neq s$ and 1 if $t = s$ — the inner summation (in t) has only one non-zero term. We therefore have $\mathrm{Cov}\,(\mathbf{Y}) = \mathbf{V} = \mathbf{A}\mathbf{A}^T$. It follows then that

$$\mathbf{x}^T\mathbf{x} \;=\; (\mathbf{A}^{-1}\mathbf{y})^T\mathbf{A}^{-1}\mathbf{y} \;=\; \mathbf{y}^T\left[(\mathbf{A}^{-1})^T\mathbf{A}^{-1}\right]\mathbf{y} \;=\; \mathbf{y}^T\mathbf{V}^{-1}\mathbf{y},$$

where we have used that $(\mathbf{A}^{-1})^T\mathbf{A}^{-1} = (\mathbf{A}^T)^{-1}\mathbf{A}^{-1} = (\mathbf{A}\mathbf{A}^T)^{-1}$. Furthermore, since $\det\mathbf{A} = \det\mathbf{A}^T$ we have $\det\mathbf{V} = (\det\mathbf{A})^2$ and hence $|\det\mathbf{A}| = \sqrt{\det\mathbf{V}}$. Substituting this back into (2) gives

$$P[\mathbf{Y} \in I] \;\approx\; \frac{1}{(2\pi)^{n/2}\sqrt{\det\mathbf{V}}}\, e^{-(\mathbf{y}^T\mathbf{V}^{-1}\mathbf{y})/2} \cdot \mathrm{Vol}\,(I),$$

and, therefore,

$$f_{\mathbf{Y}}(\mathbf{y}) \;=\; \frac{1}{(2\pi)^{n/2}\sqrt{\det\mathbf{V}}}\, e^{-(\mathbf{y}^T\mathbf{V}^{-1}\mathbf{y})/2}. \tag{3}$$

It is of significance that the distribution of $\mathbf{Y}$ depends only on the correlation matrix $\mathbf{V}$, and not on the particular transformation $\mathbf{A}$. (See Exercises 2 and 3 below where it is shown that many different $\mathbf{A}$s produce the same $\mathbf{V}$). In particular this implies that, for joint normals $\mathbf{Y}$, if they are uncorrelated then they are independent. To see this, suppose that $\mathrm{Cov}\,(Y_i, Y_j) = 0$ for $i \neq j$ and let $\mathrm{Var}\,Y_i = \sigma_i^2$. Then $\mathbf{V}$ is 0 off the diagonal and $V_{ii} = \sigma_i^2$. To simplify calculations, take each $\widetilde{Y}_i = Y_i/\sigma_i$, which is a standard normal, and put $\widetilde{\mathbf{Y}} = (\widetilde{Y}_1, \ldots, \widetilde{Y}_n)^T$. Clearly the Y_i are independent if and only if the $\widetilde{Y}_i$ are independent and $\widetilde{\mathbf{V}} = \mathrm{Cov}\,(\widetilde{\mathbf{Y}}) = \mathbf{I}$, the $n \times n$ identity matrix. But by (3)

$$f_{\widetilde{\mathbf{Y}}}(\mathbf{y}) \;=\; \frac{1}{(2\pi)^{n/2}}\, e^{-(\mathbf{y}^T\mathbf{y})/2},$$

which factors into the marginals — just reverse the steps in (1).

Exercise 1. We saw above that if $\mathrm{Cov}\,(\mathbf{X}) = \mathbf{I}$, then $\mathrm{Cov}\,(\mathbf{AX}) = \mathbf{AA}^T$. Show that, more generally, $\mathrm{Cov}\,(\mathbf{AX}) = \mathbf{A}\mathrm{Cov}\,(\mathbf{X})\mathbf{A}^T$.

Exercise 2. Suppose $\mathbf{X} = (X_1, X_2)^T$ are independent standard normals, $|\rho| < 1$, $\mathbf{A} = \begin{pmatrix} 1 & 0 \\ \rho & \pm\sqrt{1-\rho^2} \end{pmatrix}$, and $\mathbf{Y} = \mathbf{AX}$. Show that $\mathrm{Cov}\,(\mathbf{Y}) = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$.

Exercise 3. With $\mathbf{X}$, ρ, and $\mathbf{Y}$ as in Exercise 2, fix any a and b with $0 < a < 1$ and $a^2 + b^2 = 1$. Now put $c = a\rho \pm b\sqrt{1-\rho^2}$, $d = b\rho \mp a\sqrt{1-\rho^2}$, and $\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Show that $\mathrm{Cov}\,(\mathbf{Y}) = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$.

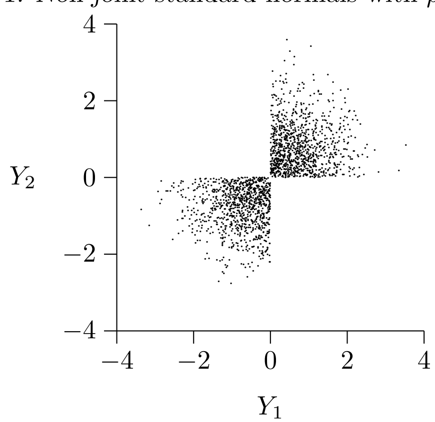
§2. Not all Pairs of Normals are Joint Normals

Consider the following construction. Let X_1 and X_2 be independent standard normals. Now put $Y_1 = X_1$ and let

$$Y_2 = \begin{cases} |X_2| & \text{if } X_1 \geq 0 \\ -|X_2| & \text{if } X_1 < 0 \end{cases}$$

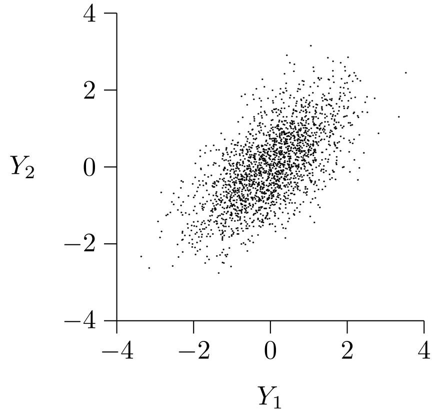
Observe that if Y_1 is positive so is Y_2, and if Y_1 is negative so is Y_2. Thus $f_{\mathbf{Y}}(\cdot)$ is positive only in the first and third quadrants. The scatter plot below (Figure 1) shows 2000 realizations of this construction. It turns out that Y_1 and Y_2 are both standard normals and their correlation is $2/\pi$. However, they are not *joint* normals, as can be seen by the fact that the density function for joint normals as given in (3) is positive *everywhere.*

Figure 1. Non-joint standard normals with $\rho = 2/\pi \approx 0.637$.



To construct *joint* normals with this correlation, we put $Y_1 = X_1$ and $Y_2 = (2/\pi)X_1 + \sqrt{1-(2/\pi)^2}X_2$ (the recipe given in Exercise 2). Figure 2 below shows a scatter plot corresponding to this construction.

Figure 2. Joint standard normals with $\rho = 2/\pi$.



Exercise 4. Show that the construction above for Figure 1 (the non-joint case) produces standard normals Y_1 and Y_2 with correlation $2/\pi$.

§3. Generating Dependent Joint Normals

If the pair of joint normals (X, Y) have mean 0 and variance 1 (they are standard normals), then their joint density function is specified by their correlation ρ, where $|\rho| < 1$:

$$f_\rho(x, y) \;=\; \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left[-\frac{x^2 - 2\rho xy + y^2}{2(1-\rho^2)}\right]. \tag{4}$$

To see this we note that their covariance matrix is $\mathbf{V} = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$, so

$$\det \mathbf{V} = 1 - \rho^2,\;\; \mathbf{V}^{-1} \;=\; \frac{1}{1-\rho^2}\begin{pmatrix} 1 & -\rho \\ -\rho & 1 \end{pmatrix}, \text{ and}$$
$$(x \;\; y)\,\mathbf{V}^{-1}\begin{pmatrix} x \\ y \end{pmatrix} \;=\; (x^2 - 2\rho xy + y^2)/(1-\rho^2);$$

refer back to equation (3). To better understand (4), complete the square in y:

$$x^2 - 2\rho xy + y^2 \;=\; (y - \rho x)^2 + x^2(1-\rho^2).$$

Setting this into (4) yields, after a little algebra, that

$$f_\rho(x, y) \;=\; \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{x^2}{2}\right] \cdot \frac{1}{\sqrt{2\pi(1-\rho^2)}} \exp\left[-\frac{(y-\rho x)^2}{2(1-\rho^2)}\right]. \tag{5}$$

We recognize the first factor above as the density function of a standard normal. In fact, we have factored the joint distribution into the marginal distribution of X evaluated at x times the conditional distribution of Y, given $X = x$, evaluated at y, that is, $f_X(x) \cdot f_{(Y|X=x)}(y)$. From this factorization, we see that, conditioned on $X = x$, $Y \sim \mathrm{Normal}\,(\rho x, 1-\rho^2)$. If we had completed the square for x instead of y, we would have found that Y is a standard normal and, conditioned on $Y = y$, $X \sim \mathrm{Normal}\,(\rho y, 1-\rho^2)$. This follows also from the symmetry of (4) in x and y.

We know how to generate two *independent* standard normals, X and Y. Suppose we put

$$Y_\rho \;=\; \rho X + \sqrt{1-\rho^2}\,Y. \tag{6}$$

Since X and Y are independent with mean 0 and variance 1, $EY_\rho = \rho EX + \sqrt{1-\rho^2}EY = 0$ and $\mathrm{Var}\,Y_\rho = \rho^2 \mathrm{Var}\,X + (1-\rho^2)\mathrm{Var}\,Y = 1$, both as desired. Furthermore, conditioned on $X = x$, we see that

$$\begin{aligned} Y_\rho \;=\; \rho x + \sqrt{1-\rho^2}\,Y \;&\sim\; \rho x + \sqrt{1-\rho^2}\,\mathrm{Normal}\,(0,1) \\ &\sim\; \mathrm{Normal}\,(\rho x, 1-\rho^2). \end{aligned}$$

Here we have used that the conditional distribution of Y does not depend on the particular realization of X since Y is independent of X. Since $X \sim \mathrm{Normal}\,(0,1)$ we see that the pair (X, Y_ρ) has the joint density function given in (5) (or (4)). Figure 3 below shows a scatter plot of 1000 realizations of the pair (X, Y_ρ) produced via (6) for various values of the correlation parameter ρ.

More Than Two Dependent Normals. In this section, we generalize this to more than two joint normals. Suppose we wish to generate n correlated normals $(Y_1, \ldots, Y_n)^T$. Let $v_{ij} = \mathrm{Cov}\,(Y_i, Y_j)$ denote the desired covariance between Y_i and Y_j, so $\mathbf{V} = (v_{ij})_{n\times n}$ is the desired correlation matrix corresponding to $(Y_1, \ldots Y_n)^T$. Not every square matrix may serve as a legitimate covariance matrix. Our first order of business is to characterize those that can:

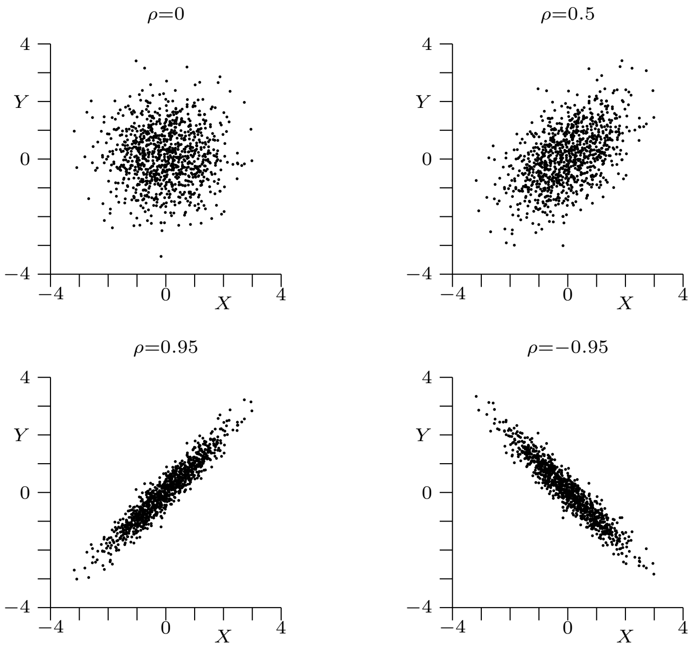
Theorem 1. Let $\mathbf{V} = (v_{ij})$ be an $n \times n$ matrix. Then there are joint random variables $(Y_1, \ldots, Y_n)^T$ whose covariance matrix is $\mathbf{V}$ if and only if $\mathbf{V}$ is symmetric and positive semi-definite.

Proof. Suppose $(Y_1, \ldots, Y_n)^T$ have covariance matrix $\mathbf{V}$, so $\mathrm{Cov}\,(Y_i, Y_j) = v_{ij}$. Then $\mathbf{V}$ is symmetric as $\mathrm{Cov}\,(Y_i, Y_j) = \mathrm{Cov}\,(Y_j, Y_i)$. Let $\mathbf{x} = (x_1, \ldots, x_n)^T$ be any column vector and put $Z = x_1 Y_1 + \cdots + x_n Y_n$. Then

$$\begin{aligned} 0 \;\le\; \mathrm{Var}\,Z \;&=\; \sum_{i=1}^n \sum_{j=1}^n \mathrm{Cov}\,(x_i Y_i, x_j Y_j) \\ &=\; \sum_{i=1}^n \sum_{j=1}^n x_i x_j v_{ij} \;=\; \mathbf{x}^T \mathbf{V} \mathbf{x}. \end{aligned} \tag{7}$$

Hence $\mathbf{V}$ is positive semi-definite. This proves the "only if" part.

Figure 3. Scatter plots for joint standard normals with various correlations.



Now suppose $\mathbf{V}$ is symmetric and positive semi-definite. Then $\mathbf{V}$ factors as $\mathbf{V} = \mathbf{L}\mathbf{L}^T$, where $\mathbf{L} = (\ell_{ij})_{n\times n}$ is a lower triangular matrix with non-negative entries on the diagonal (each $\ell_{ii} \geq 0$). (See Chapter 4 for the Cholesky decomposition.) Now suppose $\mathbf{X} = (X_1, \ldots, X_n)^T$ are independent standard normals and put

$$\begin{aligned} Y_1 &= \ell_{11}X_1 \\ Y_2 &= \ell_{21}X_1 + \ell_{22}X_2 \\ Y_3 &= \ell_{31}X_1 + \ell_{32}X_2 + \ell_{33}X_3 \\ &\cdots \\ Y_n &= \ell_{n1}X_1 + \ell_{n2}X_2 + \ell_{n3}X_3 + \cdots + \ell_{nn}X_n, \end{aligned} \tag{8}$$

so $\mathbf{Y} = \mathbf{L}\mathbf{X}$ where $\mathbf{Y} = (Y_1, \ldots, Y_n)^T$. It follows immediately from §1 that $\mathrm{Cov}\,(\mathbf{Y}) = \mathbf{L}\mathbf{L}^T = \mathbf{V}$. $\square$

Remark. If $\mathbf{V}$ is only PSD, then at least one of the diagonal entries of $\mathbf{L}$ is 0, hence $\det \mathbf{L} = \prod_{i=1}^n \ell_{ii} = 0$ and $\mathbf{L}$ is not invertible. In this case $\mathbf{Y}$ does not follow an n-dimensional joint normal distribution. See §4 for more on this.

We return to the problem of generating n normals $\mathbf{Y} = (Y_1, \ldots, Y_n)^T$ with covariance matrix $\mathbf{V} = (v_{ij})_{n\times n}$. Assume that $\mathbf{V}$ is PD — we'll handle the PSD case in the the next section. Presently we illustrate how to compute the ℓ_{ij}s of (8) that yield the covariance matrix $\mathbf{V}$. We note that Theorem 1 guarantees their existence and, since $\mathbf{V}$ is PD, we may take each $\ell_{ii} > 0$ (see Chapter 4).

First we illustrate with $n = 3$. Since we wish $\ell_{11}^2 = \mathrm{Var}(Y_1) = v_{11}$, we take $\ell_{11} = \sqrt{v_{11}}$. Moving to Y_2, we compute $\mathrm{Cov}(Y_2, Y_1) = \ell_{11}\ell_{21}$. (We will repeatedly use the bi-linearity of covariance and the independence of the X_is.) Since we seek $\mathrm{Cov}(Y_2, Y_1) = v_{21}$, take $\ell_{21} = v_{21}/\ell_{11}$. Now $\mathrm{Var}(Y_2) = \ell_{21}^2 + \ell_{22}^2$ and since we wish $\mathrm{Var}(Y_2) = v_{22}$ we take $\ell_{22} = \sqrt{v_{22} - \ell_{21}^2}$. So far, in terms of the v_{ij}s, we have

$$\begin{aligned} Y_1 &= \sqrt{v_{11}}X_1 \\ Y_2 &= (v_{21}/\sqrt{v_{11}})X_1 + \sqrt{v_{22} - (v_{21}^2/v_{11})}\,X_2, \end{aligned}$$

which generalizes (6) where $v_{11} = v_{22} = 1$ and $v_{21} = \rho$. As for Y_3, begin with $\mathrm{Cov}(Y_3, Y_1) = \ell_{11}\ell_{31}$, which we wish to be v_{31} — so we take $\ell_{31} = v_{31}/\ell_{11}$. Moving across, $\mathrm{Cov}(Y_3, Y_2) = \ell_{21}\ell_{31} + \ell_{22}\ell_{32}$, which we wish to be v_{32}. Now we've already computed ℓ_{21}, ℓ_{31}, and ℓ_{22}, so take $\ell_{32} = (v_{32} - \ell_{21}\ell_{31})/\ell_{22}$. Finally, since $\mathrm{Var}\,Y_3 = \ell_{31}^2 + \ell_{32}^2 + \ell_{33}^2$ (which we wish to be v_{33}), take $\ell_{33} = \sqrt{v_{33} - (\ell_{31}^2 + \ell_{32}^2)}$. (Note that we've already computed ℓ_{31} and ℓ_{32}). This completes the exercise for $n = 3$.

This can be extended to arbitrary n as follows. Suppose we've computed the ℓ_{ij}s up to line $k - 1$. We show how to compute line k, i.e., how to compute the ℓ_{kj} for $1 \le j \le k$. Start with $\mathrm{Cov}(Y_k, Y_1) = \ell_{11}\ell_{k1}$, which we wish to be v_{k1}. We take $\ell_{k1} = v_{k1}/\ell_{11}$. Now, working across row k, assume $\ell_{k1}, \ell_{k2}, \ldots \ell_{k,j-1}$ have been computed, where $j < k$. We may then compute ℓ_{kj} by solving

$$v_{kj} = \mathrm{Cov}(Y_k, Y_j) = \ell_{j1}\ell_{k1} + \ell_{j2}\ell_{k2} + \cdots + \ell_{j,j-1}\ell_{k,j-1} + \ell_{jj}\ell_{kj}. \quad (9)$$

The only ℓ_{**} that is unknown at this step is ℓ_{kj} and we may solve (9) for ℓ_{kj} because we know that $\ell_{jj} > 0$, as $\mathbf{V}$ is PD. Finally, when $\ell_{k1}, \ell_{k2}, \ldots \ell_{k,k-1}$ have been computed, we may obtain ℓ_{kk} from

$$v_{kk} = \mathrm{Var}\,Y_k = \ell_{k1}^2 + \ell_{k2}^2 + \cdots + \ell_{k,k-1}^2 + \ell_{kk}^2,$$

yielding

$$\ell_{kk} = \sqrt{v_{kk} - (\ell_{k1}^2 + \ell_{k2}^2 + \cdots + \ell_{k,k-1}^2)}\,. \quad (10)$$

A key point is that if the wished-for correlation matrix $\mathbf{V}$ is PD, then the number under the radical in (10) will be strictly positive for each k. This is because Theorem 1 guarantees the existence of the requisite ℓ_{ij}s, and the computations of (9) and (10) are forced (there is no choice) line by line beginning with line 1 where we must have $\ell_{11} = \sqrt{v_{11}}$.

Remark. Actually, there appears to be a choice of *sign* to be made in (10). However Theorem 1 guarantees the existence of an $\mathbf{L} = (\ell_{ij})$ with strictly positive diagonal entries. To produce this $\mathbf{L}$, we *must* take the positive solution in (10).

Exercise 5. Suppose you wish to generate 5 standard normals $Y_1, \ldots, Y_5$ so that $\operatorname{Cov}(Y_i, Y_j) = 0.2$ if $i \neq j$ (of course $\operatorname{Cov}(Y_i, Y_i) = 1$). Compute to five decimal places the ℓ_{ij}s that accomplish this. When applying (10), take the positive square root. Verify your computations by executing the program `CholeskyExample1.cpp`.

Exercise 6. Do the same thing, but this time take the *negative* square root when computing ℓ_{22}. What do you observe?

Exercise 7. Starting with six independent standard normals $X_0, X_1, \ldots, X_6$, there is an easier way to generate five Ys with the covariance matrix of Exercise 5. Can you find it? Hint: Display (6).

§4. When V is Only PSD

With $\mathbf{V} = \mathbf{L}\mathbf{L}^T$, so $\det \mathbf{V} = (\det \mathbf{L})^2$, we have

$$\begin{aligned} \mathbf{V} \text{ is only PSD} &\iff \text{some } \ell_{ii} = 0 \\ &\iff \det \mathbf{L} = 0 \\ &\iff \det \mathbf{V} = 0 \\ &\iff \mathbf{V} \text{ is not invertible.} \end{aligned}$$

In this situation, for some non-zero vector $\mathbf{x} = (x_1, \ldots, x_n)^T$ we have $\mathbf{x}^T \mathbf{V} \mathbf{x} = 0$ and the random variable $Z = x_1 Y_1 + \cdots + x_n Y_n$ has $\operatorname{Var} Z = 0$ (see (7)). It follows that $Z = c$ for some constant c. Since our random variables are mean zero, we have $c = 0$. Since $\mathbf{x}$ is non-zero, at least one of the x_is, say x_n, is not zero. This means that

$$Y_n = \left(\frac{-x_1}{x_n}\right) Y_1 + \cdots + \left(\frac{-x_{n-1}}{x_n}\right) Y_{n-1}, \tag{11}$$

i.e., at least one of the Y_is is a linear combination of the others.

Re-order the Y_is, if necessary so that $Y_1, \ldots, Y_k$ are linearly independent (none are linear functions of the others) but each of $Y_{k+1}, \ldots, Y_n$ may be written as a linear function of $Y_1, \ldots, Y_k$ as in (11). Let $\widetilde{\mathbf{V}}_{k\times k}$ denote the covariance matrix for $(Y_1, \ldots, Y_k)^T$. Then $\widetilde{\mathbf{V}}$ is PD and the random variables $Y_{k+1}, \ldots Y_n$ can be expressed as linear functions of $Y_1, \ldots, Y_k$.

Exercise 8. Analyze the case $\mathbf{V} = \begin{pmatrix} 1.0 & 1.4 & 0.2 \\ 1.4 & 5.8 & 2.2 \\ 0.2 & 2.2 & 1.0 \end{pmatrix}$ using (9) and (10). Calculate to five decimal places. Verify your computations by executing the program `CholeskyExample2.cpp`. (Observe that $\mathbf{V}$ is not of full rank, as *row 1* = *row 2* $-$ 2×*row 3*.)

Accompanying Code

`CholeskyExample1.cpp` and `CholeskyExample2.cpp` illustrate Cholesky decomposition for the covariance matrices of Exercises 5 and 8, respectively. The matrix product $\mathbf{LL}^T$ is computed to verify the decomposition.