Linear transformations and Gaussian random vectors.

We fix a positive integer n.

1. Symmetric matrices; psd matrices.

When we write $\mathbf{x} \in \mathbb{R}^n$ we mean that

$$\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

Let

$$\mathbf{Sym}(n)$$

be the vector space of n by n symmetric matrices. We say the n by n matrix B is **positive definite symmetric (psd)** if B is symmetric and

(1) $$\mathbf{x}^T B\mathbf{x} > 0 \quad \text{whenever } \mathbf{x} \in \mathbb{R}^n.$$

If the n by n matrix B is symmetric then (1) is equivalent to the statement that the eigenvalues of B are positive.Remember, n-vectors are the same as $n \times 1$ matrices.

Suppose $B \in \mathbf{Sym}(n)$. Then there are n by n matrices P, D such that P is orthogonal (which means P is invertible and $P^{-1} = P^T$), $D = \mathbf{diag}\,(\lambda_1, \ldots, \lambda_n)$ and

$$B = PDP^{-1}.$$

This amounts to saying that, for each $j = 1, \ldots, n$,

$$BP_j = \lambda_j P_j$$

where P_j is the jth column of B; in other words, P_j is a (nonzero) eigenvector of B with eigenvalue λ_j.

Theorem 1.1. Suppose $f : \mathbb{R} \to \mathbb{R}$. There is one and only one function, also denoted by

$$f : \mathbf{Sym}(n) \to \mathbf{Sym}(n)$$

(ambiguously denoted by f!) which preserves the matrix operations and such that

$$f(\mathbf{diag}\,(\lambda_1, \ldots, \lambda_n) = \mathbf{diag}\,(f(\lambda_1), \ldots, f(\lambda_n)).$$

Moreover, if $g : \mathbb{R} \to \mathbb{R}$ we have

$$(fg)(B) = f(B)g(B) \quad \text{for } B \in \mathbf{Sym}(n).$$

Proof. We let

$$f(B) = P\mathbf{diag}\,(f(\lambda_1), \ldots, f(\lambda_n))P^{-1}$$

where P, D and λ_i, $i = 1, \ldots, n$ are as above. Note that P, D and λ_i, $i = 1, \ldots, n$ are *not* unique; we leave it to the reader that $f(B)$ is, nonetheless, well defined. We also leave to the reader the simple task of verifying the other asserted properties of the mapping $B \to f(B)$. $\square$

2. Gaussian random vectors.

Let $\mathbf{X}$ a random n-vector. Recall the definitions of

$$\mathrm{E}(\mathbf{X}), \quad \mathrm{Cov}(\mathbf{X}, \mathbf{Y});$$

these are, respectively, a vector in $\mathbb{R}^n$ and an n by n symmetric matrix.

Theorem 2.1. Suppose $\mathbf{Y}$ is a random vector, B is an n by n psd matrix and $\mathbf{m} \in \mathbb{R}^n$. The following are equivalent:

(I) $\mathbf{Y} = \sqrt{B}\mathbf{X} + \mathbf{m}$ for some standard normal $\mathbf{X}$.

(II) $\mathbf{Y}$ is continuous and

$$f_{\mathbf{Y}} = (2\pi)^{-n/2}\sqrt{\mathbf{det}\, B}e^{-(\mathbf{y}-\mathbf{m})^T B^{-1}(\mathbf{y}-\mathbf{m})/2} \quad \text{for } \mathbf{y} \in \mathbb{R}^n.$$

If these conditions hold then

$$\mathrm{E}(Y) = \mathbf{m} \quad \text{and} \quad \mathrm{Cov}(Y, Y) = B.$$

Proof. Suppose (I) holds. Then $\mathbf{Y}$ is continuous and, if $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ are such that $\mathbf{y} = \sqrt{B}\mathbf{x} + \mathbf{m}$ then $\mathbf{x} = \sqrt{B}^{-1}(\mathbf{y} - \mathbf{m})$ so

$$|\mathbf{x}|^2 = \mathbf{x}^T\mathbf{x} = (\sqrt{B}^{-1}(\mathbf{y} - \mathbf{m}))^T\sqrt{B}^{-1}(\mathbf{y} - \mathbf{m}) = (\mathbf{y} - \mathbf{m}))^T B^{-1}(\mathbf{y} - \mathbf{m});$$

thus, by Change of Variables Formula for Random Vectors,

$$\begin{aligned} f_{\mathbf{Y}}(\mathbf{y}) &= f_{\mathbf{X}}(\mathbf{x})\frac{1}{\mathbf{det}\left(\sqrt{B}^{-1}\right)} \\ &= (2\pi)^{-n/2}\sqrt{\mathbf{det}\, B}e^{-|\mathbf{x}|^2/2} \\ &= (2\pi)^{-n/2}\sqrt{\mathbf{det}\, B}e^{-(\mathbf{y}-\mathbf{m})^T B^{-1}(\mathbf{y}-\mathbf{m})/2}. \end{aligned}$$

Thus (II) holds.

Suppose (II) holds. Let $\mathbf{X} = \sqrt{B}^{-1}(\mathbf{Y} - \mathbf{m})$. Then $\mathbf{X}$ is continuous and, if $\mathbf{x}, \mathbf{y}$ are as above,

$$\begin{aligned} f_{\mathbf{X}}(\mathbf{x}) &= f_{\mathbf{Y}}(\mathbf{y})\frac{1}{\mathbf{det}\,\sqrt{B}^{-1}} \\ &= (2\pi)^{-n/2}\sqrt{\mathbf{det}\, B}e^{-(\mathbf{y}-\mathbf{m})^T B^{-1}(\mathbf{y}-\mathbf{m})/2}\frac{1}{\mathbf{det}\,\sqrt{B}^{-1}} \\ &= (2\pi)^{-n/2}e^{-|\mathbf{x}|^2/2} \end{aligned}$$

by th Change of Variables Formula for Random Vectors. So $\mathbf{X}$ is standard normal.

The assertion about the mean and covariance follow from straightforward homework exercises. □

Theorem 2.2. Suppose Y is Gaussian, A is a nonsingular n by n matrix and $\mathbf{b} \in \mathbb{R}^n$. Then $A\mathbf{Y} + \mathbf{b}$ is Gaussian.

Proof. Exercise for the reader; just turn the crank using the Change of Variables Formula for Random Vectors. □