Section 5.8. Random Vectors and Matrices

Note. In this section we define random vectors and matrices, and define the variance-covariance matrix of a random vector. We introduce the multivariate normal distribution and its density function. By convention, we consider a vector as a column vector; that is, we treat an n-dimensional vector as a $n \times 1$ matrix.

Definition. A *random vector* or a *random matrix* is one whose elements are random variables. For random vector $\mathbf{Y} = [Y_i]$ (where Y_i is are the random variable entries of the random vector), the *expected value* is the vector $\mathbf{E}\{\mathbf{Y}\} = [E\{Y_i\}]$. Similarly, the *expected value* of random matrix $\mathbf{Y} = [Y_{ij}]$ is the matrix $\mathbf{E}\{\mathbf{Y}\} = [E\{Y_{ij}\}]$.

Note. For an example of a random vector related to regression, let $\varepsilon_1, \varepsilon_2, \varepsilon_3$ be the error terms in three regression applications. Then the expectation of each is zero, $E\{\varepsilon_1\} = E\{\varepsilon_2\} = E\{\varepsilon_3\} = 0$. So for random vector

$$\boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{bmatrix}$$

we have the expectation $\mathbf{E}\{\boldsymbol{\varepsilon}\} = \mathbf{0}$.

Definition. Let $\mathbf{Y}$ be a random vector with three entries, Y_1, Y_2, Y_3. We use the variances of the three random variables $\sigma^2\{Y_i\}$ and the covariances between two of the random variables $\sigma\{Y_i, Y_j\}$, to define the *variance-covariance matrix* of $\mathbf{Y}$ as

$$\boldsymbol{\sigma}^2\{\mathbf{Y}\} = \begin{bmatrix} \sigma^2\{Y_1\} & \sigma\{Y_1, Y_2\} & \sigma\{Y_1, Y_3\} \\ \sigma\{Y_2, Y_1\} & \sigma^2\{Y_2\} & \sigma\{Y_2, Y_3\} \\ \sigma\{Y_3, Y_1\} & \sigma\{Y_3, Y_2\} & \sigma^2\{Y_3\} \end{bmatrix}.$$

Note 5.8.A. Since the covariance satisfies $\sigma\{Y_i, Y_j\} = \sigma\{Y_j, Y_i\}$, for $i \neq j$, then the variance-covariance matrix $\boldsymbol{\sigma}^2\{\mathbf{Y}\}$ is symmetric. Since $E\{(Y_i - E\{Y_i\})^2\} = \sigma^2\{Y_i\}$ and $E\{(Y_i - E\{Y_i\})(Y_j - E\{Y_j\})\} = \sigma\{Y_i, Y_j\}$ for $i \neq j$, then (since the expectation operator is linear; see my online notes for Mathematical Statistics 1 [STAT 4057/5057] on Section 1.8. Expectation of a Random Variable and Theorem 1.8.2) the variance covariance matrix above can be written as the matrix product

$$\boldsymbol{\sigma}^2\{\mathbf{Y}\} = \mathbf{E}\left\{ \begin{bmatrix} Y_1 - E\{Y_1\} \\ Y_2 - E\{Y_2\} \\ Y_3 - E\{Y_3\} \end{bmatrix} [Y_1 - E\{Y_1\} \quad Y_2 - E\{Y_2\} \quad Y_3 - E\{Y_3\}] \right\}.$$

We can extend the variance-covariance matrix to a random vector with n entries, as follows. Notice that the factorization given here also extends to the resulting $n \times n$ matrix $\boldsymbol{\sigma}^2\{\mathbf{Y}\}$.

Definition. Let $\mathbf{Y}$ be a random vector with n entries, $Y_1, Y_2, \ldots, Y_n$. We use the variances of the n random variables $\sigma^2\{Y_i\}$ and the covariances between two of the random variables $\sigma\{Y_i, Y_j\}$, to define the *variance-covariance matrix* of $\mathbf{Y}$ as

$$\boldsymbol{\sigma}^2\{\mathbf{Y}\} = \begin{bmatrix} \sigma^2\{Y_1\} & \sigma\{Y_1, Y_2\} & \cdots & \sigma\{Y_1, Y_n\} \\ \sigma\{Y_2, Y_1\} & \sigma^2\{Y_2\} & \cdots & \sigma\{Y_2, Y_n\} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma\{Y_n, Y_1\} & \sigma\{Y_n, Y_2\} & \cdots & \sigma^2\{Y_n\} \end{bmatrix}.$$

Theorem 5.8.A. For $\mathbf{Y}$ a random vector and $\mathbf{A}$ a matrix of scalars (i.e., constants), we can define the new random variable $\mathbf{W} = \mathbf{AY}$. Then:

(5.44) Expectation of a Constant Matrix: $\mathbf{E}\{\mathbf{A}\} = \mathbf{A}$.

(5.45) Expectation of a Constant Matrix Times a Random Vector:

$$\mathbf{E}\{\mathbf{W}\} = \mathbf{E}\{\mathbf{AY}\} = \mathbf{A}\mathbf{E}\{\mathbf{Y}\}.$$

(5.46) Variance of a Constant Matrix Times a Random Vector:

$$\boldsymbol{\sigma}^2\{\mathbf{W}\} = \boldsymbol{\sigma}^2\{\mathbf{AY}\} = \mathbf{A}\boldsymbol{\sigma}^2\{\mathbf{Y}\}\mathbf{A}'.$$

Note. With $\mathbf{Y}$ a (column) vector of p random variables, $Y_1, Y_2, \ldots, Y_p$, we denote the mean vector $\mathbf{E}\{\mathbf{Y}\}$ and $\boldsymbol{\mu}$. With $E\{Y_i\} = \mu_i$ for $1 \le i \le p$, we then have $\boldsymbol{\mu} = [\mu_i]$. If we denote the variance of Y_i as σ_i^2 and the covariance of Y_i and Y_j as σ_{ij}, then we can abbreviate the variance-covariance matrix $\boldsymbol{\sigma}^2\{\mathbf{Y}\}$, which we also denote $\mathbf{\Sigma}$, as

$$\mathbf{\Sigma} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1} & \sigma_{p2} & \cdots & \sigma_p^2 \end{bmatrix}.$$

This equips us with what we need to define the density function for the multivariate normal distribution.

Definition. The *multivariate normal distribution* is the distribution with the density function

$$f(\mathbf{Y}) = \frac{1}{(2\pi)^{p/2}|\mathbf{\Sigma}|^{1/2}} \exp\left[-\frac{1}{2}(\mathbf{Y} - \boldsymbol{\mu})'\mathbf{\Sigma}^{-1}(\mathbf{Y} - \boldsymbol{\mu})\right].$$

Note. In the previous definition, $|\mathbf{\Sigma}|$ denotes the determinant of $\mathbf{\Sigma}$, so we are concerned that this is nonnegative since the square root is taken. We also need $\mathbf{\Sigma}^{-1}$, so we need $\mathbf{\Sigma}$ to be invertible and hence we need $|\mathbf{\Sigma}| > 0$. It is and $\mathbf{\Sigma}$ is a "positive definite" matrix. We also have concerns that the function given really is a density function and that it integrates to 1. We content ourselves with omitting these details, since they are covered in the undergraduate/graduate class Mathematical Statistics 1 (STAT 4047/5047). See my online notes for this class on Section 3.5. The Multivariate Normal Distribution, where *many* details and rigor are given leading up to the definition of the multivariate normal distribution density function (see Note 3.5.F) and the multivariate normal moment generating function (see Definition 3.5.1 and the comments preceding it that explains details on how the matrices and exponential functions interact).

Revised: 9/16/2022