2.10 Multivariate Normal Distribution

Definition: An n dimensional random vector $\mathbf{Y}$ is said to have a multivariate normal (MVN) or Gaussian (G) distribution if

$$\mathbf{Y} = \boldsymbol{\mu} + \mathbf{B}\mathbf{Z}$$

where

- $\boldsymbol{\mu}$ is a $n \times 1$ vector
- $\mathbf{B}$ is a $n \times m$ matrix
- $\mathbf{Z}$ is a vector of $m \leq n$ independent normal random variables

By the independence of the elements of $\mathbf{Z}$ and their univariate normality we have the following forms for the density and moment generating function of $\mathbf{Z}$:

$$\begin{aligned} f_{\mathbf{Z}}(\mathbf{z}) &= (2\pi)^{-\frac{m}{2}} \exp\left\{-\frac{1}{2}\mathbf{z}^T\mathbf{z}\right\} \\ M_{\mathbf{Z}}(\mathbf{t}) &= \exp\left\{\frac{1}{2}\mathbf{t}^T\mathbf{t}\right\} \end{aligned}$$

Hence the joint moment generating function of $\mathbf{Y}$ is

$$M_{\mathbf{Y}}(\mathbf{t}) = \exp\left\{\mathbf{t}^T\boldsymbol{\mu} + \frac{1}{2}\mathbf{t}^T\boldsymbol{\Sigma}\mathbf{t}\right\}$$

where $\mathbf{B}\mathbf{B}^T = \boldsymbol{\Sigma}$. Note that $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ represent the mean vector and covariance matrix of $\mathbf{Y}$. Since the moment generating function depends only on these two parameters it follows that the parameters $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ completely characterize the MVN distribution.

The following results about MVN random variables follow from the definition:

- If $\mathbf{Y}$ is MVN with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$ denoted

  $$\mathbf{Y} \sim \text{MVN}\left(\boldsymbol{\mu}, \boldsymbol{\Sigma}\right)$$

  then $\mathbf{X} = \mathbf{c} + \mathbf{D}\mathbf{Y}$ where $\mathbf{c}, \mathbf{D}$ are known $p \times 1$ and $p \times n$ matrices, respectively, is MVN with mean $\mathbf{c} + \mathbf{D}\boldsymbol{\mu}$ and covariance matrix $\mathbf{D}\boldsymbol{\Sigma}\mathbf{D}^T$. Note that the mean and covaiance expressions for $\mathbf{X}$ follow from the general moment results. To show that $\mathbf{X}$ is MVN, write $\mathbf{Y} = \boldsymbol{\mu} + \mathbf{B}\mathbf{Z}$ so that

  $$\mathbf{X} = (\mathbf{c} + \mathbf{D}\boldsymbol{\mu}) + (\mathbf{D}\mathbf{B})\mathbf{Z}$$

- If $\mathbf{Y} \sim \text{MVN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ then any subset of coordinates of $\mathbf{Y}$ is also MVN with mean and covariance matrix being the appropriate sub matrices of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$. To show this, write

$$\mathbf{Y} = \begin{bmatrix} \mathbf{Y}_1 \\ \mathbf{Y}_2 \end{bmatrix}$$

where $\mathbf{Y}_1$ is $p \times 1$, $\mathbf{Y}_2$ is $n - p \times 1$ and express $\mathbf{Y}_1$ as the following linear combination of $\mathbf{Y}$

$$\mathbf{Y}_1 = \begin{bmatrix} \mathbf{I}_p & \mathbf{0} \end{bmatrix} \mathbf{Y}$$

Then from the result above $\mathbf{Y}_1$ is MVN with mean $\boldsymbol{\mu}_1$ and covariance matrix $\boldsymbol{\Sigma}_{11}$ where

$$\begin{bmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{bmatrix}$$

represent the appropriate partitions of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$.

- If $\mathbf{Y} \sim \text{MVN}\ (\boldsymbol{\mu}, \boldsymbol{\Sigma})$ and

$$\mathbf{Y} = \begin{bmatrix} \mathbf{Y}_1 \\ \mathbf{Y}_2 \end{bmatrix}$$

where $\mathbf{Y}_1$ is $p \times 1$, $\mathbf{Y}_2$ is $(n-p) \times 1$ then $\mathbf{Y}_1$ and $\mathbf{Y}_2$ are statistically independent if and only if $\boldsymbol{\Sigma}_{12} = \text{cov}\ (\mathbf{Y}_1, \mathbf{Y}_2) = \mathbf{0}$. This result follows from the ability to factor the moment generating function of $\mathbf{Y}$ if and only if $\boldsymbol{\Sigma}_{12} = \mathbf{0}$.

- It also follows that if subsets of a MVN variable are pairwise independent, then they are mutually independent as well.

Note that the density of a MVN variable has not yet been described. This is because unless rank $(\mathbf{B}) = n$ (in the definition of the MVN), the mass of $\mathbf{Y}$ $(n \times 1)$ is concentrated on a subspace of $\mathbf{R}^n$. In fact, by definition, $\mathbf{Y}$ lies in the space spanned by the columns of $\mathbf{B}$ with probability one. Thus, if rank $(\mathbf{B}) < n$ the density of $\mathbf{Y}$ with respect to Lebesgue measure in $\mathbf{R}^n$ does not exist. As an example, consider $n = 2$ where the correlation between Y_1 and Y_2 is unity. Then all the mass is concentrated on the subspace consisting of the line through the origin with slope given by

$$\frac{\text{var}\,(Y_1)}{\text{var}\,(Y_2)}$$

If rank $(\mathbf{B}) = n$, then the density in $\mathbf{R}^n$ with respect to n dimensional Lesbesque measure exists and has the form

$$f_{\mathbf{Y}}(\mathbf{y}) = (2\pi)^{-\frac{n}{2}}(\det(\mathbf{\Sigma}))^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(\mathbf{y}-\boldsymbol{\mu})^T\mathbf{\Sigma}^{-1}(\mathbf{y}-\boldsymbol{\mu})\right\}$$

Result: If $\mathbf{Y}$ is MVN with mean $\boldsymbol{\mu}$ and covariance matrix $\mathbf{\Sigma}$ where

$$\boldsymbol{\mu} = \begin{bmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{bmatrix} \text{ and } \mathbf{\Sigma} = \begin{bmatrix} \mathbf{\Sigma}_{11} & \mathbf{\Sigma}_{12} \\ \mathbf{\Sigma}_{21} & \mathbf{\Sigma}_{22} \end{bmatrix}$$

Then the conditional distribution of $\mathbf{Y}_2$ given $\mathbf{Y}_1 = \mathbf{y}_2$ is also MVN with mean $\boldsymbol{\mu}^*$ and covariance matrix $\mathbf{\Sigma}^*$ where

$$\boldsymbol{\mu}^* = \boldsymbol{\mu}_2 + \mathbf{\Sigma}_{21}\mathbf{\Sigma}_{11}^{-1}(\mathbf{y}_1 - \boldsymbol{\mu}_1) \text{ and } \mathbf{\Sigma}^* = \mathbf{\Sigma}_{22} - \mathbf{\Sigma}_{21}\mathbf{\Sigma}_{11}^{-1}\mathbf{\Sigma}_{12}$$