

MATH 545: INTRODUCTION TO TIME SERIES ANALYSIS

WINTER 2025

THE MULTIVARIATE NORMAL DISTRIBUTION

The multivariate Normal (or Gaussian) distribution arises as a linear transform of a vector of iid standard Normal components: let $\mathbf{Z} = (Z_1, \dots, Z_n)^\top$ be a vector of independent rvs with $Z_i \sim \text{Normal}(0, 1)$. Consider the transform

$$\mathbf{X} = \boldsymbol{\mu} + \mathbf{V}\mathbf{Z}$$

where $\boldsymbol{\mu}$ is $n \times 1$ and $\mathbf{V}$ is $n \times n$ and non-singular. Then $\mathbf{X} \sim \text{Normal}_n(\boldsymbol{\mu}, \Sigma)$, with $\Sigma = \mathbf{V}\mathbf{V}^\top$. To see this we may use the multivariate transformation theorem: we have that

$$f_{\mathbf{Z}}(\mathbf{z}) = \prod_{i=1}^n \left(\frac{1}{2\pi} \right)^{1/2} \exp \left\{ -\frac{1}{2} z_i^2 \right\} = \left(\frac{1}{2\pi} \right)^{n/2} \exp \left\{ -\frac{1}{2} \mathbf{z}^\top \mathbf{z} \right\}$$

and hence by the transformation theorem

$$f_{\mathbf{X}}(\mathbf{x}) = f_{\mathbf{Z}}(\mathbf{V}^{-1}(\mathbf{x} - \boldsymbol{\mu})) |\mathbf{J}|$$

where $|\mathbf{J}|$ is the absolute value of the determinant of the transformation. For this linear transformation, basic linear algebra results allow us to conclude that $|\mathbf{J}| = |\mathbf{V}|^{-1}$, that is, the reciprocal of the determinant of $\mathbf{V}$. Hence if $\Sigma = \mathbf{V}\mathbf{V}^\top$, we have that

$$f_{\mathbf{X}}(\mathbf{x}) = \left(\frac{1}{2\pi} \right)^{n/2} \frac{1}{|\Sigma|^{1/2}} \exp \left\{ -\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})^\top \Sigma^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right\}$$

where $|\Sigma| = |\mathbf{V}|^2$, and the $\mathbf{z}^\top \mathbf{z}$ term becomes

$$\{\mathbf{V}^{-1}(\mathbf{x} - \boldsymbol{\mu})\}^\top \{\mathbf{V}^{-1}(\mathbf{x} - \boldsymbol{\mu})\} = (\mathbf{x} - \boldsymbol{\mu})^\top \{\mathbf{V}^{-1}\}^\top \{\mathbf{V}^{-1}\}(\mathbf{x} - \boldsymbol{\mu}) = (\mathbf{x} - \boldsymbol{\mu})^\top \{\mathbf{V}\mathbf{V}^\top\}^{-1}(\mathbf{x} - \boldsymbol{\mu}).$$

Notes:

- If Σ has a block diagonal structure

$$\Sigma = \begin{bmatrix} \Sigma_{11} & \mathbf{0} \\ \mathbf{0} & \Sigma_{22} \end{bmatrix} \quad \Sigma^{-1} = \begin{bmatrix} \Sigma_{11}^{-1} & \mathbf{0} \\ \mathbf{0} & \Sigma_{22}^{-1} \end{bmatrix}$$

where Σ_{11} is $k_1 \times k_1$ and Σ_{22} is $k_2 \times k_2$, and $|\Sigma| = |\Sigma_{11}||\Sigma_{22}|$. Hence if we consider the partition $\mathbf{X} = (\mathbf{X}_1^\top, \mathbf{X}_2^\top)^\top$ where $\mathbf{X}_1$ is an $k_1 \times 1$ vector, then

$$(\mathbf{x} - \boldsymbol{\mu})^\top \Sigma^{-1} (\mathbf{x} - \boldsymbol{\mu}) = (\mathbf{x}_1 - \boldsymbol{\mu}_1)^\top \Sigma_{11}^{-1} (\mathbf{x}_1 - \boldsymbol{\mu}_1) + (\mathbf{x}_2 - \boldsymbol{\mu}_2)^\top \Sigma_{22}^{-1} (\mathbf{x}_2 - \boldsymbol{\mu}_2)$$

where $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$ are the relevant components of $\boldsymbol{\mu}$. Therefore in this block diagonal case, we deduce that

$$f_{\mathbf{X}}(\mathbf{x}) = f_{\mathbf{X}_1}(\mathbf{x}_1) f_{\mathbf{X}_2}(\mathbf{x}_2)$$

where

$$f_{\mathbf{X}_1}(\mathbf{x}_1) = \left(\frac{1}{2\pi} \right)^{k_1/2} \frac{1}{|\Sigma_{11}|^{1/2}} \exp \left\{ -\frac{1}{2} (\mathbf{x}_1 - \boldsymbol{\mu}_1)^\top \Sigma_{11}^{-1} (\mathbf{x}_1 - \boldsymbol{\mu}_1) \right\}$$

$$f_{\mathbf{X}_2}(\mathbf{x}_2) = \left(\frac{1}{2\pi} \right)^{k_2/2} \frac{1}{|\Sigma_{22}|^{1/2}} \exp \left\{ -\frac{1}{2} (\mathbf{x}_2 - \boldsymbol{\mu}_2)^\top \Sigma_{22}^{-1} (\mathbf{x}_2 - \boldsymbol{\mu}_2) \right\}$$

and hence $\mathbf{X}_1$ and $\mathbf{X}_2$ are independent. The zero blocks in the variance-covariance matrix Σ indicate independence of the components $\mathbf{X}_1$ and $\mathbf{X}_2$.

- **Marginal and Conditional Distributions:** We attempt to deduce

- (a) the **marginal** distribution of $\mathbf{X}_1$, and
- (b) the **conditional** distribution of $\mathbf{X}_2$ **given** that $\mathbf{X}_1 = \mathbf{x}_1$.

and for simplicity initially assume $\boldsymbol{\mu} = \mathbf{0}$.

First, write

$$\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \quad \Sigma^{-1} = \begin{bmatrix} \mathbf{V}_{11} & \mathbf{V}_{12} \\ \mathbf{V}_{21} & \mathbf{V}_{22} \end{bmatrix}$$

so that

$$\begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \begin{bmatrix} \mathbf{V}_{11} & \mathbf{V}_{12} \\ \mathbf{V}_{21} & \mathbf{V}_{22} \end{bmatrix} = \begin{bmatrix} \mathbf{I}_{k_1} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{k_2} \end{bmatrix}$$

where Σ_{11} is $k_1 \times k_1$, Σ_{22} is $k_2 \times k_2$, $\Sigma_{21} = \Sigma_{12}^\top$, where ($\mathbf{I}_r$ is the $r \times r$ identity matrix) where $\mathbf{0}$ represents the zero matrix of appropriate dimension. More specifically,

$$\Sigma_{11} \mathbf{V}_{11} + \Sigma_{12} \mathbf{V}_{21} = \mathbf{I}_{k_1} \quad (1)$$

$$\Sigma_{11} \mathbf{V}_{12} + \Sigma_{12} \mathbf{V}_{22} = \mathbf{0} \quad (2)$$

$$\Sigma_{21} \mathbf{V}_{11} + \Sigma_{22} \mathbf{V}_{21} = \mathbf{0} \quad (3)$$

$$\Sigma_{21} \mathbf{V}_{12} + \Sigma_{22} \mathbf{V}_{22} = \mathbf{I}_{k_2}. \quad (4)$$

We can re-express the term in the exponent as

$$\mathbf{x}^\top \Sigma^{-1} \mathbf{x} = \mathbf{x}_1^\top \mathbf{V}_{11} \mathbf{x}_1 + \mathbf{x}_1^\top \mathbf{V}_{12} \mathbf{x}_2 + \mathbf{x}_2^\top \mathbf{V}_{21} \mathbf{x}_1 + \mathbf{x}_2^\top \mathbf{V}_{22} \mathbf{x}_2. \quad (5)$$

We complete the square in $\mathbf{x}_2$ in this expression and write

$$\mathbf{x}^\top \Sigma^{-1} \mathbf{x} = (\mathbf{x}_2 - \mathbf{m})^\top M (\mathbf{x}_2 - \mathbf{m}) + \mathbf{c} \quad (6)$$

and by comparing with equation (5) we can deduce that, for quadratic terms in $\mathbf{x}_2$,

$$\mathbf{x}_2^\top \mathbf{V}_{22} \mathbf{x}_2 = \mathbf{x}_2^\top \mathbf{M} \mathbf{x}_2 \quad \therefore \quad \mathbf{M} = \mathbf{V}_{22} \quad (7)$$

for linear terms

$$\mathbf{x}_2^\top \mathbf{V}_{21} \mathbf{x}_1 = -\mathbf{x}_2^\top \mathbf{M} \mathbf{m} \quad \therefore \quad \mathbf{m} = -\mathbf{V}_{22}^{-1} \mathbf{V}_{21} \mathbf{x}_1 \quad (8)$$

and for constant terms

$$\mathbf{x}_1^\top \mathbf{V}_{11} \mathbf{x}_1 = \mathbf{c} + \mathbf{m}^\top M \mathbf{m} \quad \therefore \quad \mathbf{c} = \mathbf{x}_1^\top (\mathbf{V}_{11} - \mathbf{V}_{21}^\top \mathbf{V}_{22}^{-1} \mathbf{V}_{21}) \mathbf{x}_1 \quad (9)$$

thus yielding all the terms required for equation (6), that is

$$\mathbf{x}^\top \Sigma^{-1} \mathbf{x} = (\mathbf{x}_2 + \mathbf{V}_{22}^{-1} \mathbf{V}_{21} \mathbf{x}_1)^\top \mathbf{V}_{22} (\mathbf{x}_2 + \mathbf{V}_{22}^{-1} \mathbf{V}_{21} \mathbf{x}_1) + \mathbf{x}_1^\top (\mathbf{V}_{11} - \mathbf{V}_{21}^\top \mathbf{V}_{22}^{-1} \mathbf{V}_{21}) \mathbf{x}_1, \quad (10)$$

which, crucially, is a sum of two terms, where the first can be interpreted as a function of $\mathbf{x}_2$, given $\mathbf{x}_1$, and the second is a function of $\mathbf{x}_1$ only. Hence we have an immediate factorization of the full joint pdf

$$f_{\mathbf{X}}(\mathbf{x}) = f_{\mathbf{X}_2|\mathbf{X}_1}(\mathbf{x}_2|\mathbf{x}_1) f_{\mathbf{X}_1}(\mathbf{x}_1) \quad (11)$$

where

$$f_{\mathbf{X}_2|\mathbf{X}_1}(\mathbf{x}_2|\mathbf{x}_1) \propto \exp \left\{ -\frac{1}{2} (\mathbf{x}_2 + \mathbf{V}_{22}^{-1} \mathbf{V}_{21} \mathbf{x}_1)^\top \mathbf{V}_{22} (\mathbf{x}_2 + \mathbf{V}_{22}^{-1} \mathbf{V}_{21} \mathbf{x}_1) \right\} \quad (12)$$

giving that

$$\mathbf{X}_2|\mathbf{X}_1 = \mathbf{x}_1 \sim \text{Normal}_{k_2}(-\mathbf{V}_{22}^{-1} \mathbf{V}_{21} \mathbf{x}_1, \mathbf{V}_{22}^{-1}) \quad (13)$$

and

$$f_{\mathbf{X}_1}(\mathbf{x}_1) \propto \exp \left\{ -\frac{1}{2} \mathbf{x}_1^\top (\mathbf{V}_{11} - \mathbf{V}_{21}^\top \mathbf{V}_{22}^{-1} \mathbf{V}_{21}) \mathbf{x}_1 \right\} \quad (14)$$

giving that

$$\mathbf{X}_1 \sim \text{Normal}_{k_1} \left(\mathbf{0}, (\mathbf{V}_{11} - \mathbf{V}_{21}^\top \mathbf{V}_{22}^{-1} \mathbf{V}_{21})^{-1} \right). \quad (15)$$

But, from equation (2), $\Sigma_{12} = -\Sigma_{11} \mathbf{V}_{12} \mathbf{V}_{22}^{-1}$, and then from equation (1), substituting in Σ_{12} ,

$$\Sigma_{11} \mathbf{V}_{11} - \Sigma_{11} \mathbf{V}_{12} \mathbf{V}_{22}^{-1} \mathbf{V}_{21} = \mathbf{I}_{k_1} \quad \therefore \quad \Sigma_{11} = (\mathbf{V}_{11} - \mathbf{V}_{12} \mathbf{V}_{22}^{-1} \mathbf{V}_{21})^{-1} = (\mathbf{V}_{11} - \mathbf{V}_{21}^\top \mathbf{V}_{22}^{-1} \mathbf{V}_{21})^{-1}.$$

Hence, by inspection of equation (15), we conclude that

$$\boxed{\mathbf{X}_1 \sim \text{Normal}_{k_1}(\mathbf{0}, \Sigma_{11})}, \quad (16)$$

that is, we can extract the Σ_{11} block of Σ to define the marginal sigma matrix of $\mathbf{X}_1$.

Using similar arguments, we can define the conditional distribution from equation (13) more precisely. First, from equation (2), $\mathbf{V}_{12} = -\Sigma_{11}^{-1} \Sigma_{12} \mathbf{V}_{22}$, and then from equation (4), substituting in $\mathbf{V}_{12}$

$$-\Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12} \mathbf{V}_{22} + \Sigma_{22} \mathbf{V}_{22} = \mathbf{I}_{k-k_1} \quad \therefore \quad \mathbf{V}_{22}^{-1} = \Sigma_{22} - \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12} = \Sigma_{22} - \Sigma_{12}^\top \Sigma_{11}^{-1} \Sigma_{12}.$$

Finally, from equation (2), taking transposes on both sides, we have that $\mathbf{V}_{21} \Sigma_{11} + \mathbf{V}_{22} \Sigma_{21} = \mathbf{0}$. Then pre-multiplying by $\mathbf{V}_{22}^{-1}$, and post-multiplying by Σ_{11}^{-1} , we have

$$\mathbf{V}_{22}^{-1} \mathbf{V}_{21} + \Sigma_{21} \Sigma_{11}^{-1} = \mathbf{0} \quad \therefore \quad \mathbf{V}_{22}^{-1} \mathbf{V}_{21} = -\Sigma_{21} \Sigma_{11}^{-1},$$

so we have, substituting into equation (13), that

$$\boxed{\mathbf{X}_2 | \mathbf{X}_1 = \mathbf{x}_1 \sim \text{Normal}_{k_2}(\Sigma_{21} \Sigma_{11}^{-1} \mathbf{x}_1, \Sigma_{22} - \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12})}. \quad (17)$$

Thus any marginal, and any conditional distribution of a multivariate normal joint distribution is also multivariate normal, as the choices of $\mathbf{X}_1$ and $\mathbf{X}_2$ are arbitrary. When $\boldsymbol{\mu} \neq \mathbf{0}$

$$\mathbf{X}_1 \sim \text{Normal}_{k_1}(\boldsymbol{\mu}_1, \Sigma_{11})$$

$$\mathbf{X}_2 | \mathbf{X}_1 = \mathbf{x}_1 \sim \text{Normal}_{k_2}(\boldsymbol{\mu}_2 + \Sigma_{21} \Sigma_{11}^{-1}(\mathbf{x}_1 - \boldsymbol{\mu}_1), \Sigma_{22} - \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12}).$$