

Result 4.6. Let $\mathbf{X} = \begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{bmatrix}$ be distributed as $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ with $\boldsymbol{\mu} = \begin{bmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{bmatrix}$, $\boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{bmatrix}$, and $|\boldsymbol{\Sigma}_{22}| > 0$. Then the conditional distribution of \mathbf{X}_1 , given that $\mathbf{X}_2 = \mathbf{x}_2$, is normal and has

$$\text{Mean} = \boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2)$$

and

$$\text{Covariance} = \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}$$

Note that the covariance does not depend on the value \mathbf{x}_2 of the conditioning variable.

Proof. We shall give an indirect proof. (See Exercise 4.13, which uses the densities directly.) Take

$$\mathbf{A} = \begin{bmatrix} \mathbf{I} & -\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}$$

$\begin{matrix} (q \times q) & q \times (p-q) \\ (p-q) \times q & (p-q) \times (p-q) \end{matrix}$

so

$$\mathbf{A}(\mathbf{X} - \boldsymbol{\mu}) = \mathbf{A} \begin{bmatrix} \mathbf{X}_1 - \boldsymbol{\mu}_1 \\ \mathbf{X}_2 - \boldsymbol{\mu}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{X}_1 - \boldsymbol{\mu}_1 - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{X}_2 - \boldsymbol{\mu}_2) \\ \mathbf{X}_2 - \boldsymbol{\mu}_2 \end{bmatrix}$$

is jointly normal with covariance matrix $\mathbf{A}\boldsymbol{\Sigma}\mathbf{A}'$ given by

$$\begin{bmatrix} \mathbf{I} & -\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{0}' \\ (-\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1})' & \mathbf{I} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21} & \mathbf{0}' \\ \mathbf{0} & \boldsymbol{\Sigma}_{22} \end{bmatrix}$$

Since $\mathbf{X}_1 - \boldsymbol{\mu}_1 - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{X}_2 - \boldsymbol{\mu}_2)$ and $\mathbf{X}_2 - \boldsymbol{\mu}_2$ have zero covariance, they are independent. Moreover, the quantity $\mathbf{X}_1 - \boldsymbol{\mu}_1 - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{X}_2 - \boldsymbol{\mu}_2)$ has distribution $N_q(\mathbf{0}, \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21})$. Given that $\mathbf{X}_2 = \mathbf{x}_2$, $\boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2)$ is a constant. Because $\mathbf{X}_1 - \boldsymbol{\mu}_1 - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{X}_2 - \boldsymbol{\mu}_2)$ and $\mathbf{X}_2 - \boldsymbol{\mu}_2$ are independent, the conditional distribution of $\mathbf{X}_1 - \boldsymbol{\mu}_1 - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2)$ is the same as the unconditional distribution of $\mathbf{X}_1 - \boldsymbol{\mu}_1 - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{X}_2 - \boldsymbol{\mu}_2)$. Since $\mathbf{X}_1 - \boldsymbol{\mu}_1 - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{X}_2 - \boldsymbol{\mu}_2)$ is $N_q(\mathbf{0}, \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21})$, so is the random vector $\mathbf{X}_1 - \boldsymbol{\mu}_1 - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2)$ when \mathbf{X}_2 has the particular value \mathbf{x}_2 . Equivalently, given that $\mathbf{X}_2 = \mathbf{x}_2$, \mathbf{X}_1 is distributed as $N_q(\boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2), \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21})$. ■