MAT-269: Sesión 3
Distribuciones matriciales I

Felipe Osorio

fosorios.mat.utfsm.cl

Departamento de Matemática, UTFSM

Definición 1 (Distribución normal multivariada)

Sea $\boldsymbol{x}$ vector aleatorio p-dimensional. Se dice que $\boldsymbol{x}$ tiene distribución normal multivariada con vector de medias $\boldsymbol{\mu} \in \mathbb{R}^p$ y matriz de covarianza $\boldsymbol{\Sigma} \geq 0$ si y sólo si,

$$y = \boldsymbol{t}^\top \boldsymbol{x} \sim \mathsf{N}_1(\mu_Y, \sigma_Y^2), \qquad \text{para todo } \boldsymbol{t} \in \mathbb{R}^p,$$

y anotamos $\boldsymbol{x} \sim \mathsf{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.

Resultado 1 (Función característica)

Si $\boldsymbol{x} \sim \mathsf{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, entonces la función característica de $\boldsymbol{x}$ es dada por

$$\varphi(\boldsymbol{t}) = \exp\left(i\boldsymbol{t}^\top \boldsymbol{\mu} - \tfrac{1}{2}\boldsymbol{t}^\top \boldsymbol{\Sigma} \boldsymbol{t}\right).$$

Resultado 2 (Transformaciones afín)

Suponga que $\boldsymbol{x} \sim \mathsf{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ y considere la transformación $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}$ donde $\boldsymbol{A} \in \mathbb{R}^{m \times p}$ con $\operatorname{rg}(\boldsymbol{A}) = m$. Entonces

$$\boldsymbol{y} \sim \mathsf{N}_m(\boldsymbol{A}\boldsymbol{\mu} + \boldsymbol{b}, \boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{A}^\top).$$

Resultado 3 (Momentos)

Si $\boldsymbol{x} \sim \mathsf{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, entonces

$$\mathsf{E}(\boldsymbol{x}) = \boldsymbol{\mu}, \qquad \mathsf{Cov}(\boldsymbol{x}) = \boldsymbol{\Sigma}.$$

Resultado 3 (Distribución marginal)

Si $\boldsymbol{x} \sim \mathsf{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, entonces la distribución marginal de cualquier subconjunto de k $(< p)$ componentes de $\boldsymbol{x}$ es normal k-variada.

Resultado 4 (Independencia)

Si $\boldsymbol{x} \sim \mathsf{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ y $\boldsymbol{x}$, $\boldsymbol{\mu}$ y $\boldsymbol{\Sigma}$ son particionadas como:

$$\boldsymbol{x} = \begin{pmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \end{pmatrix}, \qquad \boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix}, \qquad \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix},$$

Entonces los vectores $\boldsymbol{x}_1$ y $\boldsymbol{x}_2$ son independientes si y sólo si $\boldsymbol{\Sigma}_{12} = \mathbf{0}$ $(= \boldsymbol{\Sigma}_{21}^\top)$.

Resultado 5 (Función de densidad)

Si $\boldsymbol{x} \sim \mathsf{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ y $\boldsymbol{\Sigma}$ es definida positiva, entonces la densidad de $\boldsymbol{x}$ asume la forma:

$$f(\boldsymbol{x}) = |2\pi\boldsymbol{\Sigma}|^{-1/2} \exp\{-\tfrac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\}, \qquad \boldsymbol{x} \in \mathbb{R}^p.$$

Observación:

Considere $\boldsymbol{\Sigma} = \boldsymbol{B}\boldsymbol{B}^\top$, y sea $\boldsymbol{z} = \boldsymbol{B}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})$. La variable aleatoria[1]

$$(\boldsymbol{x}-\boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu}) = \boldsymbol{z}^\top\boldsymbol{z} = \sum_{i=1}^p z_i^2 \sim \chi_p^2.$$

[1] $D = \{(\boldsymbol{x}-\boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\}^{1/2}$ se conoce como distancia de Mahalanobis de $\boldsymbol{x}$ a $\boldsymbol{\mu}$

Resultado 6 (Distribución condicional)

Sea $\boldsymbol{x} \sim \mathsf{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ y considere la siguiente partición:

$$\boldsymbol{x} = \begin{pmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \end{pmatrix}, \qquad \boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix}, \qquad \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix},$$

Entonces

$$\boldsymbol{x}_1 | \boldsymbol{x}_2 \sim \mathsf{N}_k(\boldsymbol{\mu}_{1\cdot 2}, \boldsymbol{\Sigma}_{11\cdot 2}),$$

donde

$$\boldsymbol{\mu}_{1\cdot 2} = \boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-}(\boldsymbol{x}_2 - \boldsymbol{\mu}_2), \qquad \boldsymbol{\Sigma}_{11\cdot 2} = \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-}\boldsymbol{\Sigma}_{21},$$

con $\boldsymbol{\Sigma}_{22}^{-}$ una matriz que satisface[2]

$$\boldsymbol{\Sigma}_{22}\boldsymbol{\Sigma}_{22}^{-}\boldsymbol{\Sigma}_{22} = \boldsymbol{\Sigma}_{22}.$$

[2] $\boldsymbol{\Sigma}_{22}^{-}$ es una inversa generalizada de $\boldsymbol{\Sigma}_{22}$.

Definición 2 (Distribución normal matricial)

Se dice que una matriz aleatoria $\boldsymbol{Y} \in \mathbb{R}^{n \times p}$ tiene una distribución normal matricial si su función característica es de la forma:

$$\varphi_Y(\boldsymbol{H}) = \exp\Big(i \operatorname{tr} \boldsymbol{H}^\top \boldsymbol{M} - \frac{1}{2} \operatorname{tr} \boldsymbol{H}^\top \boldsymbol{\Omega} \boldsymbol{H} \boldsymbol{\Sigma}\Big), \qquad \boldsymbol{H} \in \mathbb{R}^{n \times p}, \tag{1}$$

donde $\boldsymbol{M} \in \mathbb{R}^{n \times p}$, $\boldsymbol{\Omega} \geq 0$ y $\boldsymbol{\Sigma} \geq 0$ son matrices semidefinidas positivas de órdenes $n \times n$ y $p \times p$, respectivamente.

Una manera de introducir la distribución normal matricial es considerar $\boldsymbol{z}_1,\dots,\boldsymbol{z}_n$ vectores aleatorios independientes, tales que $\boldsymbol{z}_i \sim \mathsf{N}_p(\boldsymbol{0},\boldsymbol{I})$ para $i = 1,\dots,n$. De este modo

$$\begin{aligned}\varphi_Z(\boldsymbol{H}) &= \mathsf{E}\{\exp(i\operatorname{tr}\boldsymbol{H}^\top\boldsymbol{Z})\} = \prod_{i=1}^n \exp\big(-\tfrac{1}{2}\boldsymbol{h}_i^\top\boldsymbol{h}_i\big)\\ &= \exp\Big(-\frac{1}{2}\sum_{i=1}^n \boldsymbol{h}_i^\top\boldsymbol{h}_i\Big) = \exp\big(-\tfrac{1}{2}\operatorname{tr}\boldsymbol{H}^\top\boldsymbol{H}\big), \qquad (2)\end{aligned}$$

donde

$$\boldsymbol{Z} = \begin{pmatrix}\boldsymbol{z}_1^\top\\ \vdots\\ \boldsymbol{z}_n^\top\end{pmatrix}, \qquad \boldsymbol{H} = \begin{pmatrix}\boldsymbol{h}_1^\top\\ \vdots\\ \boldsymbol{h}_n^\top\end{pmatrix},$$

son ámbas matrices $n \times p$.

Considere la transformación

$$\boldsymbol{Y} = \boldsymbol{M} + \boldsymbol{\Omega}^{1/2}\boldsymbol{Z}\boldsymbol{\Sigma}^{1/2}.$$

De este modo, la función característica de $\boldsymbol{Y}$ adopta la forma:

$$\begin{aligned}
\varphi_Y(\boldsymbol{H}) &= \mathsf{E}\{\exp(i \operatorname{tr} \boldsymbol{H}^\top \boldsymbol{Y})\} \\
&= \mathsf{E}\left\{ \exp\left(i \operatorname{tr} \boldsymbol{H}^\top (\boldsymbol{M} + \boldsymbol{\Omega}^{1/2}\boldsymbol{Z}\boldsymbol{\Sigma}^{1/2})\right)\right\} \\
&= \exp(i \operatorname{tr} \boldsymbol{H}^\top \boldsymbol{M}) \, \mathsf{E}\{\exp(i \operatorname{tr} \boldsymbol{H}^\top \boldsymbol{\Omega}^{1/2}\boldsymbol{Z}\boldsymbol{\Sigma}^{1/2})\},
\end{aligned}$$

como

$$\operatorname{tr} \boldsymbol{H}^\top \boldsymbol{\Omega}^{1/2}\boldsymbol{Z}\boldsymbol{\Sigma}^{1/2} = \operatorname{tr} \boldsymbol{\Sigma}^{1/2}\boldsymbol{H}^\top \boldsymbol{\Omega}^{1/2}\boldsymbol{Z},$$

y haciendo $\boldsymbol{T}^\top = \boldsymbol{\Sigma}^{1/2}\boldsymbol{H}^\top\boldsymbol{\Omega}^{1/2}$ tenemos que

$$\operatorname{tr} \boldsymbol{T}^\top \boldsymbol{T} = \operatorname{tr} \boldsymbol{\Sigma}^{1/2}\boldsymbol{H}^\top \boldsymbol{\Omega}^{1/2}\boldsymbol{\Omega}^{1/2}\boldsymbol{H}\boldsymbol{\Sigma}^{1/2} = \operatorname{tr} \boldsymbol{H}^\top \boldsymbol{\Omega}\boldsymbol{H}\boldsymbol{\Sigma},$$

usando (2), sigue la función característica definida en (1).

Observación:

Cuando una matriz aleatoria tiene función característica dada por (1), anotamos

$$\boldsymbol{Y} \sim \mathsf{N}_{n,p}(\boldsymbol{M}, \boldsymbol{\Omega}, \boldsymbol{\Sigma}).$$

Resultado 7 (Función de densidad)

Suponga que $\boldsymbol{Y} \sim \mathsf{N}_{n,p}(\boldsymbol{M}, \boldsymbol{\Omega}, \boldsymbol{\Sigma})$ donde $\boldsymbol{\Omega}$ y $\boldsymbol{\Sigma}$, donde $\boldsymbol{\Omega}$ y $\boldsymbol{\Sigma}$ son definidas positivas. Entonces la función de densidad de $\boldsymbol{Y}$ asume la forma:

$$f(\boldsymbol{Y}) = (2\pi)^{-np/2}|\boldsymbol{\Omega}|^{-p/2}|\boldsymbol{\Sigma}|^{-n/2} \exp\big\{-\tfrac{1}{2}\operatorname{tr}\boldsymbol{\Omega}^{-1}(\boldsymbol{Y}-\boldsymbol{M})\boldsymbol{\Sigma}^{-1}(\boldsymbol{Y}-\boldsymbol{M})^\top\big\}.$$

Demostración:

En efecto, sea $\boldsymbol{y} = \operatorname{vec}(\boldsymbol{Y}^\top)$. Tenemos que $\boldsymbol{y} \sim \mathsf{N}_{np}(\boldsymbol{\mu}, \boldsymbol{\Omega} \otimes \boldsymbol{\Sigma})$ con $\boldsymbol{\mu} = \operatorname{vec}(\boldsymbol{M}^\top)$. La distribución conjunta de los elementos de $\boldsymbol{y}$ es dada por

$$f(\boldsymbol{y}) = |2\pi\boldsymbol{\Omega} \otimes \boldsymbol{\Sigma}|^{-1/2} \exp\big\{ -\tfrac{1}{2}(\boldsymbol{y} - \boldsymbol{\mu})^\top (\boldsymbol{\Omega} \otimes \boldsymbol{\Sigma})^{-1} (\boldsymbol{y} - \boldsymbol{\mu})\big\}.$$

El resultado sigue luego de notar que

$$|2\pi\boldsymbol{\Omega} \otimes \boldsymbol{\Sigma}| = (2\pi)^{np} |\boldsymbol{\Omega} \otimes \boldsymbol{\Sigma}| = (2\pi)^{np} |\boldsymbol{\Omega}|^p |\boldsymbol{\Sigma}|^n,$$

y,[3]

$$\begin{aligned}(\boldsymbol{y} - \boldsymbol{\mu})^\top (\boldsymbol{\Omega} \otimes \boldsymbol{\Sigma})^{-1} (\boldsymbol{y} - \boldsymbol{\mu}) &= (\operatorname{vec}(\boldsymbol{Y} - \boldsymbol{M})^\top)^\top (\boldsymbol{\Omega}^{-1} \otimes \boldsymbol{\Sigma}^{-1}) \operatorname{vec}(\boldsymbol{Y} - \boldsymbol{M})^\top \\ &= \operatorname{tr} \boldsymbol{\Sigma}^{-1} (\boldsymbol{Y} - \boldsymbol{M})^\top \boldsymbol{\Omega}^{-1} (\boldsymbol{Y} - \boldsymbol{M}).\end{aligned}$$

[3] $\operatorname{tr} \boldsymbol{ABCD} = (\operatorname{vec} \boldsymbol{D}^\top)^\top (\boldsymbol{C}^\top \otimes \boldsymbol{A}) \operatorname{vec} \boldsymbol{B} = (\operatorname{vec} \boldsymbol{D})^\top (\boldsymbol{A} \otimes \boldsymbol{C}^\top) \operatorname{vec} \boldsymbol{B}^\top$.

Observación:

Es decir, sea $\boldsymbol{Y}$ una matriz aleatoria $n \times p$ y $\boldsymbol{y} = \operatorname{vec}(\boldsymbol{Y}^\top)$. Entonces

$$\boldsymbol{Y} \sim \mathsf{N}_{n,p}(\boldsymbol{M}, \boldsymbol{\Omega}, \boldsymbol{\Sigma}),$$

si y solo si $\boldsymbol{y} \sim \mathsf{N}_{np}(\operatorname{vec}(\boldsymbol{M}^\top), \boldsymbol{\Omega} \otimes \boldsymbol{\Sigma})$.[4]

Resultado 8 (Momentos)

La esperanza y covarianza de una matriz aleatoria con distribución $\mathsf{N}_{n,p}(\boldsymbol{M}, \boldsymbol{\Omega}, \boldsymbol{\Sigma})$ son dadas por:

$$\mathsf{E}(\boldsymbol{Y}) = \boldsymbol{M},$$

$$\mathsf{Cov}(\boldsymbol{Y}) := \mathsf{Cov}(\operatorname{vec}(\boldsymbol{Y}^\top)) = \boldsymbol{\Omega} \otimes \boldsymbol{\Sigma}$$

[4] En ocasiones, abusaremos de la notación, escribiendo $\boldsymbol{Y} \sim \mathsf{N}_{n,p}(\boldsymbol{M}, \boldsymbol{\Omega} \otimes \boldsymbol{\Sigma})$.

Resultado 9 (Transformaciones lineales)

Sea $\boldsymbol{X} \sim \mathsf{N}_{n,p}(\boldsymbol{M}, \boldsymbol{\Omega}, \boldsymbol{\Sigma})$ y $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, $\boldsymbol{B} \in \mathbb{R}^{p\times q}$ y $\boldsymbol{C} \in \mathbb{R}^{m\times q}$. Entonces

$$\boldsymbol{A}\boldsymbol{X}\boldsymbol{B} + \boldsymbol{C} \sim \mathsf{N}_{m,q}(\boldsymbol{A}\boldsymbol{M}\boldsymbol{B} + \boldsymbol{C}, \boldsymbol{A}\boldsymbol{\Omega}\boldsymbol{A}^\top, \boldsymbol{B}^\top\boldsymbol{\Sigma}\boldsymbol{B})$$

Demostración:

Sea $\boldsymbol{X} = \boldsymbol{A}\boldsymbol{Y}\boldsymbol{B} + \boldsymbol{C}$. Así

$$\begin{aligned}
\varphi_X(\boldsymbol{H}) &= \mathsf{E}\{\exp(i \operatorname{tr} \boldsymbol{H}^\top \boldsymbol{X})\} \\
&= \mathsf{E}\{\exp(i \operatorname{tr} \boldsymbol{H}^\top (\boldsymbol{A}\boldsymbol{Y}\boldsymbol{B} + \boldsymbol{C}))\} \\
&= \mathsf{E}(i \operatorname{tr} \boldsymbol{H}^\top \boldsymbol{C})\, \mathsf{E}\{\exp(i \operatorname{tr} \boldsymbol{B}\boldsymbol{H}^\top \boldsymbol{A}\boldsymbol{Y})\}.
\end{aligned}$$

Es decir,

$$\begin{aligned}
\varphi_X(\boldsymbol{H}) = &\exp(i \operatorname{tr} \boldsymbol{H}^\top (\boldsymbol{A}\boldsymbol{M}\boldsymbol{B} + \boldsymbol{C})) \\
&\times \exp\big(-\tfrac{1}{2} \operatorname{tr} \boldsymbol{H}^\top \boldsymbol{A}\boldsymbol{\Omega}\boldsymbol{A}^\top \boldsymbol{H}\boldsymbol{B}^\top \boldsymbol{\Sigma}\boldsymbol{B}\big).
\end{aligned}$$

Usando el resultado anterior es fácil notar que, para $\boldsymbol{Y} \sim \mathsf{N}_{n,p}(\boldsymbol{M}, \boldsymbol{\Omega}, \boldsymbol{\Sigma})$ tenemos

$$(\boldsymbol{Y} - \boldsymbol{M})\boldsymbol{\Sigma}^{-1/2} \sim \mathsf{N}_{n,p}(\boldsymbol{0}, \boldsymbol{\Omega}, \boldsymbol{I}_p),$$

$$\boldsymbol{\Omega}^{-1/2}(\boldsymbol{Y} - \boldsymbol{M}) \sim \mathsf{N}_{n,p}(\boldsymbol{0}, \boldsymbol{I}_n, \boldsymbol{\Sigma}),$$

$$\boldsymbol{\Omega}^{-1/2}(\boldsymbol{Y} - \boldsymbol{M})\boldsymbol{\Sigma}^{-1/2} \sim \mathsf{N}_{n,p}(\boldsymbol{0}, \boldsymbol{I}_n, \boldsymbol{I}_p).$$

Adicionalmente

$$\operatorname{vec}(\boldsymbol{\Omega}^{-1/2}(\boldsymbol{Y} - \boldsymbol{M})\boldsymbol{\Sigma}^{-1/2}) \sim \mathsf{N}_{np}(\boldsymbol{0}, \boldsymbol{I}_{np}).$$

es decir,

$$\begin{aligned}\operatorname{vec}(\boldsymbol{\Omega}^{-1/2}(\boldsymbol{Y} - \boldsymbol{M})\boldsymbol{\Sigma}^{-1/2}) &= (\boldsymbol{\Sigma}^{-1/2} \otimes \boldsymbol{\Omega}^{-1/2}) \operatorname{vec}(\boldsymbol{Y} - \boldsymbol{M}) \\ &\sim \mathsf{N}_{np}(\boldsymbol{0}, \boldsymbol{I}_{np}).\end{aligned}$$

Esto lleva al siguiente resultado.

Resultado 10

Sea $\boldsymbol{X} \sim \mathsf{N}_{n,p}(\boldsymbol{M}, \boldsymbol{\Omega}, \boldsymbol{\Sigma})$. Entonces

$$\begin{aligned}
(\operatorname{vec} \boldsymbol{\Omega}^{-1/2}(\boldsymbol{Y} - \boldsymbol{M})\boldsymbol{\Sigma}^{-1/2})^\top \operatorname{vec} \boldsymbol{\Omega}^{-1/2}(\boldsymbol{Y} - \boldsymbol{M})\boldsymbol{\Sigma}^{-1/2}& \\
&= (\operatorname{vec}(\boldsymbol{Y} - \boldsymbol{M}))^\top (\boldsymbol{\Sigma}^{-1/2} \otimes \boldsymbol{\Omega}^{-1/2})(\boldsymbol{\Sigma}^{-1/2} \otimes \boldsymbol{\Omega}^{-1/2}) \operatorname{vec}(\boldsymbol{Y} - \boldsymbol{M}) \\
&= (\operatorname{vec}(\boldsymbol{Y} - \boldsymbol{M}))^\top (\boldsymbol{\Sigma}^{-1/2}\boldsymbol{\Sigma}^{-1/2} \otimes \boldsymbol{\Omega}^{-1/2}\boldsymbol{\Omega}^{-1/2}) \operatorname{vec}(\boldsymbol{Y} - \boldsymbol{M}) \\
&= (\operatorname{vec}(\boldsymbol{X} - \boldsymbol{M}))^\top (\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{\Omega}^{-1}) \operatorname{vec}(\boldsymbol{X} - \boldsymbol{M}) \\
&= \operatorname{tr} \boldsymbol{\Omega}^{-1}(\boldsymbol{X} - \boldsymbol{M})\boldsymbol{\Sigma}^{-1}(\boldsymbol{X} - \boldsymbol{M})^\top \\
&\sim \chi^2_{np}.
\end{aligned}$$

Propiedades del promedio muestral

Sea $\boldsymbol{y}_1, \dots, \boldsymbol{y}_n$ vectores aleatorios independientes desde $\mathsf{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Asumiremos que $\boldsymbol{\Sigma} > 0$. Sea

$$\boldsymbol{Y} = \begin{pmatrix} \boldsymbol{y}_1^\top \\ \vdots \\ \boldsymbol{y}_n^\top \end{pmatrix}.$$

Entonces

$$\mathsf{E}(\boldsymbol{Y}) = \begin{pmatrix} \boldsymbol{\mu}^\top \\ \vdots \\ \boldsymbol{\mu}^\top \end{pmatrix} = \boldsymbol{1}\boldsymbol{\mu}^\top, \qquad \mathsf{Cov}(\operatorname{vec}(\boldsymbol{Y}^\top)) = \boldsymbol{I}_n \otimes \boldsymbol{\Sigma}.$$

Es decir, $\boldsymbol{Y} \sim \mathsf{N}_{n,p}(\boldsymbol{1}\boldsymbol{\mu}^\top, \boldsymbol{I}_n, \boldsymbol{\Sigma})$.[5]

[5] o bien $\boldsymbol{Y} \sim \mathsf{N}_{np}(\boldsymbol{1}\boldsymbol{\mu}^\top, \boldsymbol{I}_n \otimes \boldsymbol{\Sigma})$.

Propiedades del promedio muestral

Considere el vector de medias muestrales

$$\overline{\boldsymbol{y}} = \frac{1}{n}\sum_{i=1}^{n} \boldsymbol{y}_i = \frac{1}{n}\,\boldsymbol{Y}^\top \boldsymbol{1},$$

y la matriz de covarianza

$$\boldsymbol{S} = \frac{1}{n-1}\,\boldsymbol{Q},$$

donde

$$\boldsymbol{Q} = \sum_{i=1}^{n} (\boldsymbol{y}_i - \overline{\boldsymbol{y}})(\boldsymbol{y}_i - \overline{\boldsymbol{y}})^\top,$$

son estimadores insesgados de $\boldsymbol{\mu}$ y $\boldsymbol{\Sigma}$, respectivamente.

Resultado 11 (Independencia de $\overline{y}$ con S)

Considere $\boldsymbol{Y} \sim \mathsf{N}_{n,p}(\mathbf{1}\boldsymbol{\mu}^\top, \boldsymbol{I}_n \otimes \boldsymbol{\Sigma})$. Entonces $\overline{\boldsymbol{y}}$ y $\boldsymbol{Q}$ son independientes, y

$$\overline{\boldsymbol{y}} \sim \mathsf{N}_p\Big(\boldsymbol{\mu}, \frac{1}{n}\boldsymbol{\Sigma}\Big),$$

mientras que $\boldsymbol{Q}$ tiene la misma distribución que $\boldsymbol{Z}^\top \boldsymbol{Z}$ donde $\boldsymbol{Z} \sim \mathsf{N}_{n,p}(\mathbf{0}, \boldsymbol{I}_n \otimes \boldsymbol{\Sigma})$.[6]

[6] Es decir, las filas de $\boldsymbol{Z}$ son IID desde $\mathsf{N}_p(\mathbf{0}, \boldsymbol{\Sigma})$.

Demostración:

Como $\boldsymbol{y}_1, \dots, \boldsymbol{y}_n$ son independientes, tenemos que

$$\begin{aligned}\varphi_{\overline{y}}(\boldsymbol{h}) &= \mathsf{E}\{\exp(i\boldsymbol{h}^\top \overline{\boldsymbol{y}})\} = \mathsf{E}\left\{\exp\left(i\boldsymbol{h}^\top \frac{1}{n}\sum_{i=1}^n \boldsymbol{y}_i\right)\right\} \\ &= \mathsf{E}\left\{\prod_{i=1}^n \exp(i\boldsymbol{h}^\top \boldsymbol{y}_i/n)\right\}.\end{aligned}$$

Sea $\boldsymbol{t} = \boldsymbol{h}/n$, luego

$$\begin{aligned}\varphi_{\overline{y}}(\boldsymbol{h}) &= \prod_{i=1}^n \mathsf{E}\{\exp(i\boldsymbol{t}^\top \boldsymbol{y}_i)\} = \prod_{i=1}^n \exp\left(i\boldsymbol{t}^\top \boldsymbol{\mu} - \tfrac{1}{2}\boldsymbol{t}^\top \boldsymbol{\Sigma}\boldsymbol{t}\right) \\ &= \exp\left(i\,n\boldsymbol{t}^\top \boldsymbol{\mu} - \tfrac{n}{2}\boldsymbol{t}^\top \boldsymbol{\Sigma}\boldsymbol{t}\right) = \exp\left(i\boldsymbol{h}^\top \boldsymbol{\mu} - \tfrac{1}{2n}\boldsymbol{h}^\top \boldsymbol{\Sigma}\boldsymbol{h}\right).\end{aligned}$$

Es decir, $\overline{\boldsymbol{y}} \sim \mathsf{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}/n)$.

Tenemos $\boldsymbol{y}_1, \dots, \boldsymbol{y}_n$ muestra aleatoria desde $\mathsf{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Luego, la densidad conjunta asume la forma:

$$
\begin{aligned}
f(\boldsymbol{Y}) &= (2\pi)^{-np/2}|\boldsymbol{\Sigma}|^{-n/2} \exp\Big\{-\frac{1}{2}\sum_{i=1}^{n}(\boldsymbol{y}_i-\boldsymbol{\mu})^\top\boldsymbol{\Sigma}^{-1}(\boldsymbol{y}_i-\boldsymbol{\mu})\Big\} \\
&= (2\pi)^{-np/2}|\boldsymbol{\Sigma}|^{-n/2} \exp\Big\{-\frac{1}{2}\operatorname{tr}\boldsymbol{\Sigma}^{-1}\sum_{i=1}^{n}(\boldsymbol{y}_i-\boldsymbol{\mu})(\boldsymbol{y}_i-\boldsymbol{\mu})^\top\Big\}.
\end{aligned}
$$

Ahora,[7]

$$
\begin{aligned}
\sum_{i=1}^{n}(\boldsymbol{y}_i-\boldsymbol{\mu})(\boldsymbol{y}_i-\boldsymbol{\mu})^\top &= \sum_{i=1}^{n}(\boldsymbol{y}_i-\overline{\boldsymbol{y}}+\overline{\boldsymbol{y}}-\boldsymbol{\mu})(\boldsymbol{y}_i-\overline{\boldsymbol{y}}+\overline{\boldsymbol{y}}-\boldsymbol{\mu})^\top \\
&= \sum_{i=1}^{n}(\boldsymbol{y}_i-\overline{\boldsymbol{y}})(\boldsymbol{y}_i-\overline{\boldsymbol{y}})^\top + n(\overline{\boldsymbol{y}}-\boldsymbol{\mu})(\overline{\boldsymbol{y}}-\boldsymbol{\mu})^\top \\
&= \boldsymbol{Q} + n(\overline{\boldsymbol{y}}-\boldsymbol{\mu})(\overline{\boldsymbol{y}}-\boldsymbol{\mu})^\top.
\end{aligned}
$$

[7] pues $\sum_{i=1}^{n}(\boldsymbol{y}_i-\overline{\boldsymbol{y}})(\overline{\boldsymbol{y}}-\boldsymbol{\mu})^\top = \boldsymbol{0}$.

Independencia de $\overline{\boldsymbol{y}}$ con S

Sea $\boldsymbol{z}_i = \boldsymbol{y}_i - \overline{\boldsymbol{y}}$, y considere

$$\boldsymbol{Z} = \begin{pmatrix} \boldsymbol{z}_1^\top \\ \vdots \\ \boldsymbol{z}_n^\top \end{pmatrix},$$

luego,

$$\boldsymbol{Q} = \sum_{i=1}^n (\boldsymbol{y}_i - \overline{\boldsymbol{y}})(\boldsymbol{y}_i - \overline{\boldsymbol{y}})^\top = \sum_{i=1}^n \boldsymbol{z}_i \boldsymbol{z}_i^\top = \boldsymbol{Z}^\top \boldsymbol{Z}.$$

De este modo la densidad conjunta de $\boldsymbol{Y}$ puede ser escrita como:

$$\begin{aligned} f(\boldsymbol{Y}) &= (2\pi)^{-np/2} |\boldsymbol{\Sigma}|^{-n/2} \exp\big\{ -\tfrac{1}{2} \operatorname{tr} \boldsymbol{\Sigma}^{-1} (\boldsymbol{Z}^\top \boldsymbol{Z} + n(\overline{\boldsymbol{y}} - \boldsymbol{\mu})(\overline{\boldsymbol{y}} - \boldsymbol{\mu})^\top) \big\} \\ &= (2\pi)^{-mp/2} |\boldsymbol{\Sigma}|^{-m/2} \exp\big\{ -\tfrac{1}{2} \operatorname{tr} \boldsymbol{\Sigma}^{-1} \boldsymbol{Z}^\top \boldsymbol{Z} \big\} \\ &\quad \times (2\pi)^{-p/2} |\boldsymbol{\Sigma}|^{-1/2} \exp\big\{ -\tfrac{n}{2} (\overline{\boldsymbol{y}} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\overline{\boldsymbol{y}} - \boldsymbol{\mu}) \big\}, \end{aligned}$$

con $m = n - 1$. Sabemos que $\overline{\boldsymbol{y}} \sim \mathsf{N}_p(\boldsymbol{\mu}, \frac{1}{n}\boldsymbol{\Sigma})$, luego sigue que $\boldsymbol{Z} \sim \mathsf{N}_{n,p}(\boldsymbol{0}, \boldsymbol{I} \otimes \boldsymbol{\Sigma})$ que es independiente de $\overline{\boldsymbol{y}}$.