Vol. XVIII, Nº 2, Diciembre (2010)
Matemáticas: 87–96

**Matemáticas:
Enseñanza Universitaria**

Independencia condicionada en una distribución gaussiana

Miguel A. Marmolejo L.
Universidad del Valle

Recibido Feb. 19, 2010 Aceptado Nov. 02, 2010

Abstract

In this paper we discuss some results about the gaussian conditional independence relation. In particular, we consider two theorems recently reported; one in Sullivant (2009) and other one in Marrelec and Benali (2008).

Keywords: Gaussian distribution, conditional independence.

MSC(2000): 60E05

Resumen

En este artículo se discuten algunos resultados sobre la relación de independencia condicionada en una distribución gaussiana. En particular, se consideran dos teoremas reportados recientemente; uno en Sullivant (2009) y otro en Marrelec y Benali (2008).

Palabras y frases claves: Distribución gaussiana, independencia condicionada.

1 Introducción

Como se sabe, la distribución gaussiana es una familia de distribuciones continuas de probabilidad, cuyos miembros están determinados por su vector de medias μ y su matriz de covarianzas $\mathbf{\Sigma}$, y que entre sus caraterísticas está el que sus distribuciones marginales y sus distribuciones condicionadas también son gaussianas (ver la Sección 13 del Capítulo II de Shiryaev [3]). Por sus aplicaciones, esta familia de distribuciones es, sin lugar a dudas, la más importante de las distribuciones de probabilidad.

En este artículo se discuten algunos resultados sobre la relación de independencia condicionada en esta familia de distribuciones. En particular, se consideran dos teoremas que se describen a continuación. Sea $\mathbb{X}$ un vector aleatorio con distribución gaussiana de matriz de covarianzas $\mathbf{\Sigma}$, donde $\mathbb{X}^T = \begin{bmatrix} X_1 & X_2 & \dots & X_n \end{bmatrix}$. El teorema principal de Sullivant [4] establece que, si se satisfacen las relaciones de independencia condicionada

$$X_1 \perp\!\!\!\perp X_2 \mid X_3, \quad X_2 \perp\!\!\!\perp X_3 \mid X_4, \quad \dots, \quad X_{n-1} \perp\!\!\!\perp X_n \mid X_1, \quad X_n \perp\!\!\!\perp X_1 \mid X_2,$$

entonces se satisfacen las relaciones de independencia

$$X_1 \perp\!\!\!\perp X_2, \quad X_2 \perp\!\!\!\perp X_3, \quad X_3 \perp\!\!\!\perp X_4, \quad \dots, \quad X_{n-1} \perp\!\!\!\perp X_n, \quad X_n \perp\!\!\!\perp X_1.$$

Además, si sólo se satisface un conjunto de las primeras condiciones, puede que no se cumpla ninguna de las segundas. Su demostración se apoya en la llamada descomposición primaria binomial. Aquí, se da una demostración elemental de que los dos conjuntos de relaciones son equivalentes.

De otra parte, el teorema principal de Marrelec y Benali[2] dice lo que sigue. Para dos elementos $i, j \in \mathbb{N}_n := \{1, 2, ..., n\}$, si la covarianza condicional entre X_i y X_j, dado cualquier conjunto condicionante $K \subseteq \mathbb{N} - \{i, j\}$, es cero, entonces $\mathbf{\Sigma}$ debe ser diagonal por bloques y los elementos i y j deben pertenecer a bloques diferentes. En la demostración usan inducción y el cálculo de la inversa de una matriz usando la matriz adjunta. Aquí, se presenta una demostración compacta usando matrices particionadas.

El artículo está organizado como sigue. En la Sección 2 se introduce la terminología y se establecen los resultados básicos, necesarios para el desarrollo del resto del trabajo; el Teorema 3 es la herramienta fundamental. La Sección 3 se dedica a los vectores cíclicos y markovianos; aquí, en el Teorema 4, se da una demostración elemental de una versión del teorema principal de Sullivant [4] En la Sección 4 se presentan dos resultados que involucran vectores gaussianos particionados; el Teorema 7 generaliza el teorema principal de Marrelec y Benali [2].

2 Resultados básicos

Sean $n \in \mathbb{N}$, $\mathbb{N}_n := \{1, 2, ..., n\}$, $\mu := [\mu_i]_{n\times 1} \in \mathbb{R}^{n\times 1}$ y $\mathbf{\Sigma} := [\sigma_{ij}]_{n\times n} \in \mathbb{R}^{n\times n}$ una matriz simétrica positiva definida. En lo que sigue, $\mathbb{X}$ es un vector aleatorio n-dimensional gaussiano de vector de medias μ y matriz de covarianzas $\mathbf{\Sigma}$, lo que se simboliza con $\mathbb{X} \sim \mathcal{N}(\mu, \mathbf{\Sigma})$. Así, la matriz de correlaciones de $\mathbb{X}$ es $\mathbf{R} := [\rho_{ij}]_{n\times n} = D\mathbf{\Sigma}D$, donde $D = diag(\sigma_{11}^{-1/2}, \sigma_{22}^{-1/2}, ..., \sigma_{nn}^{-1/2})$.

Si $A = \{a_1, a_2, ..., a_r\}$ es un subconjunto no vacío de $\mathbb{N}_n$, $\mathbb{X}_A$ indica el vector aleatorio r-dimensional tal que $\mathbb{X}_A^T = \begin{bmatrix} X_{a_1} & X_{a_2} & ... & X_{a_r}\end{bmatrix}$ y, de manera correspondiente, μ_A y $\mathbf{\Sigma}_{A,A}$ indican el vector de medias y la matriz de covarianzas de $\mathbb{X}_A$. Si $B = \{b_1, b_2, ..., b_s\}$ es otro subconjunto no vacío de $\mathbb{N}_n$ tal que $A \cap B = \emptyset$, se escribe $\mu_{(A|B)}$ y $\mathbf{\Sigma}_{(A|B)}$ para denotar el vector de medias y la matriz de covarianzas del vector aleatorio r-dimensional $(\mathbb{X}_A \mid \mathbb{X}_B = x_B)$. Por último, $\mathbf{\Sigma}_{A,B}$ denota la matriz de covarianzas entre $\mathbb{X}_A$ y $\mathbb{X}_B$; en particular, si $A = \{a\}$ es un conjunto unitario, se escribe $\mathbf{\Sigma}_{a,B}$ en lugar de $\mathbf{\Sigma}_{A,B}$.

Como es bien conocido (ver el contenido del siguiente Teorema), la independencia entre $\mathbb{X}_A$ y $\mathbb{X}_B$ equivale a la condición $\mathbf{\Sigma}_{A,B} = 0$, la cual se simboliza con $\mathbb{X}_A \perp\!\!\!\perp \mathbb{X}_B$. Si $C = \{c_1, c_2, ..., c_t\}$ es otro subconjunto no vacío de $\mathbb{N}_n$ tal que $C \cap A = \emptyset$ y $C \cap B = \emptyset$, la independencia condicionada entre $\mathbb{X}_A$ y $\mathbb{X}_B$ dada la condición $\mathbb{X}_C = x_C$ se simboliza con $\mathbb{X}_A \perp\!\!\!\perp \mathbb{X}_B \mid \mathbb{X}_C$.

El contenido del siguiente Teorema es bien conocido (ver por ejemplo el Capítulo 3 de Graybill [1]).

Teorema 1. *Suponga que* $\mathbb{X} \sim \mathcal{N}(\mu, \mathbf{\Sigma})$. *Si* $M = [m_{ij}]_{q\times n} \in \mathbb{R}^{q\times n}$, $m = [m_i]_{q\times 1} \in \mathbb{R}^{q\times 1}$ *y* A *y* B *son subconjuntos disjuntos de* $\mathbb{N}_n$, *entonces*

1. $\mathbb{Y} := M\mathbb{X} + m \sim \mathcal{N}(M\mu + m, M\mathbf{\Sigma}M^T)$.
2. *Si* P *es una matriz tal que* $\mathbf{\Sigma} = PP^T$, *entonces* $\mathbb{Z} := P^{-1}(\mathbb{X} - \mu) \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$. *En particular, si* $\mathbf{\Sigma}^{\frac{1}{2}}$ *es la única matriz simétrica positiva definida tal que* $\mathbf{\Sigma} = \mathbf{\Sigma}^{\frac{1}{2}}\mathbf{\Sigma}^{\frac{1}{2}}$, *entonces* $\mathbb{Z} := \mathbf{\Sigma}^{\frac{1}{2}^{-1}}(\mathbb{X} - \mu) \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$.
3. $\mathbb{X}_A \sim \mathcal{N}(\mu_A, \mathbf{\Sigma}_{A,A})$, *donde* $\mu_A^T = \begin{bmatrix} \mu_{a_1} & \mu_{a_2} & \dots & \mu_{a_r} \end{bmatrix}$ *y* $\mathbf{\Sigma}_{A,A} = [\sigma_{a_i a_j}]_{r \times r}$.
4. $(\mathbb{X}_A \mid \mathbb{X}_B = x_B) \sim \mathcal{N}(\mu_{(A|B)}, \mathbf{\Sigma}_{(A|B)})$, *donde*
$$\mu_{(A|B)} = \mu_A + \mathbf{\Sigma}_{A,B}\mathbf{\Sigma}_{B,B}^{-1}(x_B - \mu_B); \quad \mathbf{\Sigma}_{(A|B)} = \mathbf{\Sigma}_{A,A} - \mathbf{\Sigma}_{A,B}\mathbf{\Sigma}_{B,B}^{-1}\mathbf{\Sigma}_{B,A}.$$
5. $\mathbb{X}_A \perp\!\!\!\perp \mathbb{X}_B$ *si y sólo si* $\mathbf{\Sigma}_{A,B} = 0$.
6. *Si* $\mathbb{Y}_{\mathbb{A}} = (\mathbb{X}_A - \mu_A) - \mathbf{\Sigma}_{A,B}\mathbf{\Sigma}_{B,B}^{-1}(\mathbb{X}_B - \mu_B)$ *y* $\mathbb{Y}_B = (\mathbb{X}_B - \mu_B)$, *entonces* $\mathbb{Y}_A \perp\!\!\!\perp \mathbb{Y}_B$. *Más aún, el vector aleatorio* $\mathbb{Y}$ *tal que* $\mathbb{Y}^T = \begin{bmatrix} \mathbb{Y}_A^T & \mathbb{Y}_B^T \end{bmatrix}$ *es centrado de matriz de covarianzas*
$$\begin{bmatrix} \mathbf{\Sigma}_{(A|B)} & 0 \\ 0 & \mathbf{\Sigma}_{B,B} \end{bmatrix}.$$

De acuerdo con (4) del teorema anterior, el vector de medias y la matriz de covarianzas en una distribución condicionada involucra el cálculo de la inversa de una matriz; por ésto, el siguiente Lema es importante para lo que sigue.

Lema 1. *Suponga que* B *y* C *son subconjuntos disjuntos de* $\mathbb{N}_n$. *Si* $M = \mathbf{\Sigma}_{B\cup C, B\cup C}$ *es la matriz de covarianzas del vector aleatorio* $\mathbb{X}_{B\cup C}$; *es decir*
$$M = \begin{bmatrix} \mathbf{\Sigma}_{B,B} & \mathbf{\Sigma}_{B,C} \\ \mathbf{\Sigma}_{C,B} & \mathbf{\Sigma}_{C,C} \end{bmatrix},$$
entonces la matriz inversa de M *se puede expresar de las siguientes formas:*
$$M^{-1} = \begin{bmatrix} \mathbf{\Sigma}_{(B|C)}^{-1} & -\mathbf{\Sigma}_{B,B}^{-1}\mathbf{\Sigma}_{B,C}\mathbf{\Sigma}_{(C|B)}^{-1} \\ -\mathbf{\Sigma}_{C,C}^{-1}\mathbf{\Sigma}_{C,B}\mathbf{\Sigma}_{(B|C)}^{-1} & \mathbf{\Sigma}_{(C|B)}^{-1} \end{bmatrix}. \tag{1}$$
$$M^{-1} = \begin{bmatrix} 0 & 0 \\ 0 & \mathbf{\Sigma}_{C,C}^{-1} \end{bmatrix} + \begin{bmatrix} -I \\ \mathbf{\Sigma}_{C,C}^{-1}\mathbf{\Sigma}_{C,B} \end{bmatrix} \mathbf{\Sigma}_{(B|C)}^{-1} \begin{bmatrix} -I & \mathbf{\Sigma}_{B,C}\mathbf{\Sigma}_{C,C}^{-1} \end{bmatrix}. \tag{2}$$

Demostración. En cada caso, basta verificar la igualdad $MM^{-1} = I$. □

Teorema 2. *Suponga que* A, B *y* C *son subconjuntos disjuntos de* $\mathbb{N}_n$. *Entonces*

1. *La matriz de covarianzas entre* $\mathbb{X}_A$ *y* $\mathbb{X}_B$ *en la distribución del vector aleatorio condicionado* $(\mathbb{X}_{A\cup B} \mid \mathbb{X}_C = x_C)$ *es*
$$\mathbf{\Sigma}_{(A,B|C)} = \mathbf{\Sigma}_{A,B} - \mathbf{\Sigma}_{A,C}\mathbf{\Sigma}_{C,C}^{-1}\mathbf{\Sigma}_{C,B}.$$

2. *El vector aleatorio* $(\mathbb{X}_A \mid \begin{bmatrix}\mathbb{X}_B & \mathbb{X}_C\end{bmatrix}^T = \begin{bmatrix}x_B & x_C\end{bmatrix}^T)$ *es gaussiano de matriz de covarianzas*

$$\mathbf{\Sigma}_{(A|B\cup C)} = \mathbf{\Sigma}_{(A|C)} - \mathbf{\Sigma}_{(A,B|C)}\mathbf{\Sigma}^{-1}_{(B|C)}\mathbf{\Sigma}_{(B,A|C)}$$

y vector de medias

$$\mu_{(A|B\cup C)} = \mu_{(A|C)} + R(x_B, x_C),$$

donde

$$R(x_B, x_C) = \mathbf{\Sigma}_{(A,B|C)}\mathbf{\Sigma}^{-1}_{(B|C)}[\mu_{(B|C)} - x_B].$$

Además, $R(\mathbb{X}_B, \mathbb{X}_C)$ *es un vector aleatorio independiente de* $(\mathbb{X}_C - \mu_C)$ *de matriz de covarianzas*

$$\mathbf{\Sigma}_{(A,B|C)}\mathbf{\Sigma}^{-1}_{(B|C)}\mathbf{\Sigma}_{(B,A|C)}.$$

Demostración. (1) Por el numeral (4) del Teorema 1,

$$\begin{aligned}\mathbf{\Sigma}_{(A\cup B|C)} &= \mathbf{\Sigma}_{A\cup B, A\cup B} - \mathbf{\Sigma}_{A\cup B, C}\mathbf{\Sigma}^{-1}_{C,C}\mathbf{\Sigma}_{C,A\cup B}\\ &= \begin{bmatrix}\mathbf{\Sigma}_{A,A} & \mathbf{\Sigma}_{A,B}\\ \mathbf{\Sigma}_{B,A} & \mathbf{\Sigma}_{B,B}\end{bmatrix} - \begin{bmatrix}\mathbf{\Sigma}_{A,C}\\ \mathbf{\Sigma}_{B,C}\end{bmatrix}\mathbf{\Sigma}^{-1}_{C,C}\begin{bmatrix}\mathbf{\Sigma}_{C,A} & \mathbf{\Sigma}_{C,B}\end{bmatrix}.\end{aligned}$$

De esta igualdad se sigue que $\mathbf{\Sigma}_{(A,B|C)} = \mathbf{\Sigma}_{A,B} - \mathbf{\Sigma}_{A,C}\mathbf{\Sigma}^{-1}_{C,C}\mathbf{\Sigma}_{C,B}$.

(2) Otra vez, por el numeral (4) del Teorema 1 se tiene

$$\begin{aligned}\mathbf{\Sigma}_{(A|B\cup C)} &= \mathbf{\Sigma}_{A,A} - \mathbf{\Sigma}_{A,B\cup C}\mathbf{\Sigma}^{-1}_{B\cup C,B\cup C}\mathbf{\Sigma}_{B\cup C,A}\\ &= \mathbf{\Sigma}_{A,A} - \begin{bmatrix}\mathbf{\Sigma}_{A,B} & \mathbf{\Sigma}_{A,C}\end{bmatrix}\begin{bmatrix}\mathbf{\Sigma}_{B,B} & \mathbf{\Sigma}_{B,C}\\ \mathbf{\Sigma}_{C,B} & \mathbf{\Sigma}_{C,C}\end{bmatrix}^{-1}\begin{bmatrix}\mathbf{\Sigma}_{B,A}\\ \mathbf{\Sigma}_{C,A}\end{bmatrix}.\end{aligned}$$

Usando (2) del Lema 1 se llega a que

$$\begin{aligned}\mathbf{\Sigma}_{(A|B\cup C)} &= [\mathbf{\Sigma}_{A,A} - \mathbf{\Sigma}_{A,C}\mathbf{\Sigma}^{-1}_{C,C}\mathbf{\Sigma}_{C,B}]\\ &- [\mathbf{\Sigma}_{A,B} - \mathbf{\Sigma}_{A,C}\mathbf{\Sigma}^{-1}_{C,C}\mathbf{\Sigma}_{C,B}]\mathbf{\Sigma}^{-1}_{(B|C)}[\mathbf{\Sigma}_{B,A} - \mathbf{\Sigma}_{B,C}\mathbf{\Sigma}^{-1}_{C,C}\mathbf{\Sigma}_{C,A}]\\ &= \mathbf{\Sigma}_{(A|C)} - \mathbf{\Sigma}_{(A,B|C)}\mathbf{\Sigma}^{-1}_{(B|C)}\mathbf{\Sigma}_{(B,A|C)}.\end{aligned}$$

De otra parte,

$$\mu_{(A|B\cup C)} = \mu_A + \mathbf{\Sigma}_{A,B\cup C}\mathbf{\Sigma}^{-1}_{B\cup C,B\cup C}\begin{bmatrix}x_B - \mu_B\\ x_C - \mu_C\end{bmatrix}.$$

Puesto que

$$\boldsymbol{\Sigma}_{A,B\cup C}\boldsymbol{\Sigma}^{-1}_{B\cup C,B\cup C}\begin{bmatrix} x_B-\mu_B \\ x_C-\mu_C \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Sigma}_{A,B} & \boldsymbol{\Sigma}_{A,C} \end{bmatrix} \begin{bmatrix} \boldsymbol{\Sigma}_{B,B} & \boldsymbol{\Sigma}_{B,C} \\ \boldsymbol{\Sigma}_{C,B} & \boldsymbol{\Sigma}_{C,C} \end{bmatrix}^{-1} \begin{bmatrix} x_B-\mu_B \\ x_C-\mu_C \end{bmatrix},$$

usando la expresión (2) del Lema 1, se encuentra que

$$\begin{aligned} \mu_{(A|B\cup C)} &= [\mu_A + \boldsymbol{\Sigma}_{A,C}\boldsymbol{\Sigma}^{-1}_{C,C}(x_C-\mu_C)] \\ &+ [\boldsymbol{\Sigma}_{A,B} - \boldsymbol{\Sigma}_{A,C}\boldsymbol{\Sigma}^{-1}_{C,C}\boldsymbol{\Sigma}_{C,B}]\boldsymbol{\Sigma}^{-1}_{(B|C)}[-(x_B-\mu_B) + \boldsymbol{\Sigma}_{B,C}\boldsymbol{\Sigma}^{-1}_{C,C}(x_C-\mu_C)] \\ &= \mu_{(A|C)} + \boldsymbol{\Sigma}_{(A,B|C)}\boldsymbol{\Sigma}^{-1}_{(B|C)}[\mu_{(B|C)} - x_B] \\ &= \mu_{(A|C)} + R(x_B, x_C). \end{aligned}$$

La independecia entre $R(\mathbb{X}_B, \mathbb{X}_C)$ y $(\mathbb{X}_C - \mu_C)$ se sigue de (6) del Teorema 1. Por último, usando (1) del mismo Teorema se determina que la matriz de covarianzas de $R(\mathbb{X}_B, \mathbb{X}_C)$ corresponde a la expresión indicada. □

Teorema 3. *Suponga que A, B y C son subconjuntos disjuntos de $\mathbb{N}_n$. Las siguientes afirmaciones son equivalentes*

1. $\mathbb{X}_A \perp\!\!\!\perp \mathbb{X}_B \mid \mathbb{X}_C$.
2. $\boldsymbol{\Sigma}_{A,B} = \boldsymbol{\Sigma}_{A,C}\boldsymbol{\Sigma}^{-1}_{C,C}\boldsymbol{\Sigma}_{C,B}$.
3. *Para todo $a \in A$ y todo $b \in B$,* $\sigma_{ab} = \boldsymbol{\Sigma}_{a,C}\boldsymbol{\Sigma}^{-1}_{C,C}\boldsymbol{\Sigma}_{C,b}$.
4. *Para todo $a \in A$ y todo $b \in B$,* $X_a \perp\!\!\!\perp X_b \mid \mathbb{X}_C$.
5. $\boldsymbol{\Sigma}_{(A|B\cup C)} = \boldsymbol{\Sigma}_{(A|C)}$.
6. $\mu_{(A|B\cup C)} \equiv \mu_{(A|C)}$.

Demostración. De acuerdo con (1) del Teorema anterior, la matriz de covarianzas entre $\mathbb{X}_A$ y $\mathbb{X}_B$ en la distribución del vector aleatorio $(\mathbb{X}_{A\cup B} \mid \mathbb{X}_C = x_C)$ es

$$\boldsymbol{\Sigma}_{(A,B|C)} = \boldsymbol{\Sigma}_{A,B} - \boldsymbol{\Sigma}_{A,C}\boldsymbol{\Sigma}^{-1}_{C,C}\boldsymbol{\Sigma}_{C,B},$$

y la covarianza entre X_a y X_b en la distribución del vector aleatorio condicionado $(\mathbb{X}_{\{a,b\}} \mid \mathbb{X}_C = x_C)$ es

$$\sigma_{(ab|C)} = \sigma_{ab} - \boldsymbol{\Sigma}_{a,C}\boldsymbol{\Sigma}^{-1}_{C,C}\boldsymbol{\Sigma}_{C,b}.$$

Así, en virtud de (5) del Teorema 1 se tienen las equivalencias entre (1) y (2) y entre (3) y (4). La equivalencia entre (2) y (3) es obvia. Por último, las equivalencias entre (2) , (5) y (6) se siguen del Teorema anterior, teniendo en cuenta que la matriz $\boldsymbol{\Sigma}^{-1}_{(B|C)}$ es positiva definida. □

Corolario 1. *Sean j, k, m elementos distintos de $\mathbb{N}_n$. Las siguientes afirmaciones son equivalentes*

1. $X_j \perp\!\!\!\perp X_k \mid X_m$.
2. $\sigma_{jk} = (\sigma_{jm}\sigma_{mk})/\sigma_{mm}$.
3. $\rho_{jk} = \rho_{jm}\rho_{mk}$.
4. $\mathbf{\Sigma}_{(j|\{k,m\})} = \mathbf{\Sigma}_{(j|m)}$.
5. $\mu_{(j|\{k,m\})} \equiv \mu_{(j|m)}$. □

3 Vectores cíclicos, vectores markovianos

En esta Sección se dan condiciones necesarias y suficientes para que un vector aleatorio gaussiano sea cíclico o sea markoviano. El Teorema 4 es una versión del teorema principal de Sullivant [4], del cual se presenta aquí una demostración elemental

Definición 1. *(Vector cíclico). Suponga que $\mathbb{X} \sim \mathcal{N}(\mu, \mathbf{\Sigma})$, donde, como se estableció al comienzo, $\mathbb{X}^T = \begin{bmatrix} X_1 & X_2 & \ldots & X_n \end{bmatrix}$ y $\mathbf{\Sigma} = [\sigma_{ij}]_{n\times n}$ es una matriz simétrica positiva definida. Se dice que $\mathbb{X}$ es cíclico, si se cumplen las relaciones*

$$X_1 \perp\!\!\!\perp X_2, \quad X_2 \perp\!\!\!\perp X_3, \quad X_3 \perp\!\!\!\perp X_4, \quad \ldots, \quad X_{n-1} \perp\!\!\!\perp X_n, \quad X_n \perp\!\!\!\perp X_1.$$

Teorema 4. *Suponga que $\mathbb{X} \sim \mathcal{N}(\mu, \mathbf{\Sigma})$. Las siguientes afirmaciones son equivalentes.*

1. *$\mathbb{X}$ es cíclico.*
2. *Se cumplen las relaciones*

$$X_1 \perp\!\!\!\perp X_2 \mid X_3, \quad X_2 \perp\!\!\!\perp X_3 \mid X_4, \quad \ldots, \quad X_{n-1} \perp\!\!\!\perp X_n \mid X_1, \quad X_n \perp\!\!\!\perp X_1 \mid X_2.$$

Demostración. (2) ⇒ (1) Por la hipótesis y (2) del Corolario 1, se verifican las siguientes igualdades

$$\sigma_{12} = \frac{\sigma_{13}\sigma_{32}}{\sigma_{33}}, \quad \sigma_{23} = \frac{\sigma_{24}\sigma_{43}}{\sigma_{44}}, \quad \ldots, \sigma_{(n-1)n} = \frac{\sigma_{(n-1)1}\sigma_{1n}}{\sigma_{11}}, \quad \sigma_{n1} = \frac{\sigma_{n2}\sigma_{21}}{\sigma_{22}}.$$

Estas igualdades conducen a las que siguen:

$$\sigma_{12} = \alpha\sigma_{12}, \quad \sigma_{23} = \alpha\sigma_{23}, \quad \ldots, \quad \sigma_{(n-1)n} = \alpha\sigma_{(n-1)n}, \quad \sigma_{n1} = \alpha\sigma_{n1},$$

donde

$$\alpha = \frac{\sigma_{13}\sigma_{24}\sigma_{35}\cdots\sigma_{(n-1)1}\sigma_{n2}}{\sigma_{11}\sigma_{22}\sigma_{33}\cdots\sigma_{nn}}.$$

Puesto que $-1 < \alpha < 1$ ($\mathbf{\Sigma}$ es positiva definida), entonces

$$\sigma_{12} = 0, \quad \sigma_{23} = 0, \quad ..., \quad \sigma_{(n-1)n} = 0, \quad \sigma_{n1} = 0,$$

lo que, en virtud de (5) del Teorema 1, equivale a la afirmación (1).

La demostración de la otra implicación es inmediata. □

Observación 1. *Suponga que* $\mathbb{X} \sim \mathcal{N}(\mu, \mathbf{\Sigma})$ *es cíclico. Si* D *es una matriz diagonal invertible y* $b = [b_i]_{q\times 1} \in \mathbb{R}^{q\times 1}$ *, entonces* $\mathbb{Y} := D\mathbb{X} + b$ *también es cíclico.*

Observación 2. *Suponga que* $\mathbb{X} \sim \mathcal{N}(\mu, \mathbf{\Sigma})$ *es cíclico y que* $n = 4$*. Si* $B = [b_{ij}]_{4\times 4} \in \mathbb{R}^{4\times 4}$ *es una matriz invertible tal que* $b_{12} = b_{23} = b_{34} = b_{41} = b_{14} = b_{43} = b_{32} = b_{21} = 0$ *y* $b = [b_i]_{q\times 1} \in \mathbb{R}^{4\times 1}$*, entonces* $\mathbb{Y} := B\mathbb{X} + b$ *también es cíclico. En particular,* $\mathbb{Z} := \mathbf{\Sigma}\mathbb{X} + b$ *y* $\mathbb{W} := \mathbf{\Sigma}^{-1}\mathbb{X} + b$ *son cíclicos.*

Definición 2. *(Vector markoviano) Suponga que* $\mathbb{X} \sim \mathcal{N}(\mu, \mathbf{\Sigma})$*. Se dice que* $\mathbb{X}$ *es markoviano, si para todo* $k \in \{2, 3, ..., n-1\}$*, todo* $i \in K^- := \{1, 2, .., k-1\}$ *y todo* $j \in K^+ := \{k+1, k+2, ..., n\}$ *se verifica* $X_i \perp\!\!\!\perp X_j \mid X_k$*.*

Teorema 5. *Suponga que* $\mathbb{X} \sim \mathcal{N}(\mu, \mathbf{\Sigma})$*. Las siguientes afirmaciones son equivalentes.*

1. $\mathbb{X}$ *es markoviano.*
2. *Para todo* $k \in \{2, 3, ..., n-1\}$ *se cumple la relación* $\mathbb{X}_{K^-} \perp\!\!\!\perp \mathbb{X}_{K^+} \mid X_k$*.*
3. *Para todo* $k \in \{2, 3, ..., n-1\}$*, todo* $i \in K^-$ *y todo* $j \in K^+$ *se verifica* $\sigma_{ij} = (\sigma_{ik}\sigma_{kj})/\sigma_{kk}$*.*
4. *Para todo* $k \in \{2, 3, ..., n-1\}$*, todo* $i \in K^-$ *y todo* $j \in K^+$ *se verifica* $\rho_{ij} = \rho_{ik}\rho_{kj}$*.*

Demostración. Es consecuencia inmediata del Teorema 3 y del Corolario 1. □

Observación 3. *Por* (3) *del Teorema anterior, un vector markoviano centrado queda determinado por los* $(2n-1)$ *números* $\sigma_{11}, \sigma_{22}, ..., \sigma_{nn}, \sigma_{12}, \sigma_{23}, ..., \sigma_{(n-1)n}$*.*

Observación 4. *Suponga que* $\mathbb{X} \sim \mathcal{N}(\mu, \mathbf{\Sigma})$ *es markoviano. Si* D *es una matriz diagonal invertible y* $b = [b_i]_{q\times 1} \in \mathbb{R}^{q\times 1}$ *, entonces* $\mathbb{Y} := D\mathbb{X} + b$ *también es markoviano.*

Ejemplo 1. *Sean* $t_1, t_2, ..., t_n$ *números reales tales que* $0 < t_1 < t_2 < ... < t_n$*. Si* $\mathbb{X} \sim \mathcal{N}(\mu, \mathbf{\Sigma})$*, siendo* $\mu = \mathbf{0}$ *y* $\mathbf{\Sigma} = [t_i \wedge t_j]_{n\times n}$*, entonces* $\mathbb{X}$ *es markoviano.*

4 Vectores particionados

En esta Sección se presentan dos resultados que involucran independencia condicionada entre vectores aleatorios gaussianos. El teorema 7 es una generalización del teorema principal de Marrelec y Benali[2] a vectores particionados.

Teorema 6. *Suponga que* $\mathbb{X} \sim \mathcal{N}(\mu, \mathbf{\Sigma})$, *donde*

$$\mathbb{X}^T = \begin{bmatrix} X_1 & X_2 & \dots & X_n \end{bmatrix} = \begin{bmatrix} \mathbb{X}_A^T & \mathbb{X}_B^T & \mathbb{X}_C^T \end{bmatrix},$$

siendo A, B *y* C *son conjuntos disjuntos no vacíos tales que* $A \cup B \cup C = \mathbb{N}_n$. *Los vectores* $\mathbb{X}_A$, $\mathbb{X}_B$ *y* $\mathbb{X}_C$ *son independientes si y sólo si se cumplen las relaciones* $\mathbb{X}_A \perp\!\!\!\perp \mathbb{X}_B \mid \mathbb{X}_C$; $\mathbb{X}_B \perp\!\!\!\perp \mathbb{X}_C \mid \mathbb{X}_A$ *y* $\mathbb{X}_C \perp\!\!\!\perp \mathbb{X}_A \mid \mathbb{X}_B$.

Demostración. ($\Leftarrow$) Por el Teorema 3 y la hipótesis, se cumplen las igualdades

$$\mathbf{\Sigma}_{A,B} = \mathbf{\Sigma}_{A,C}\mathbf{\Sigma}_{C,C}^{-1}\mathbf{\Sigma}_{C,B}; \quad \mathbf{\Sigma}_{B,C} = \mathbf{\Sigma}_{B,A}\mathbf{\Sigma}_{A,A}^{-1}\mathbf{\Sigma}_{A,C}; \quad \mathbf{\Sigma}_{C,A} = \mathbf{\Sigma}_{C,B}\mathbf{\Sigma}_{B,B}^{-1}\mathbf{\Sigma}_{B,A}.$$

Usando las dos primeras se obtiene $\mathbf{\Sigma}_{A,B} = \mathbf{\Sigma}_{A,C}\mathbf{\Sigma}_{C,C}^{-1}\mathbf{\Sigma}_{C,A}\mathbf{\Sigma}_{A,A}^{-1}\mathbf{\Sigma}_{A,B}$, o lo que es lo mismo,

$$(\mathbf{\Sigma}_{A,A} - \mathbf{\Sigma}_{A,C}\mathbf{\Sigma}_{C,C}^{-1}\mathbf{\Sigma}_{C,A})\mathbf{\Sigma}_{A,A}^{-1}\mathbf{\Sigma}_{A,B} = 0.$$

Como la matriz $\mathbf{\Sigma}_{(A|C)} = \mathbf{\Sigma}_{A,A} - \mathbf{\Sigma}_{A,C}\mathbf{\Sigma}_{C,C}^{-1}\mathbf{\Sigma}_{C,A}$ es invertible, entonces $\mathbf{\Sigma}_{A,B} = 0$. Esto significa que $\mathbb{X}_A \perp\!\!\!\perp \mathbb{X}_B$. Un argumento similar conduce a $\mathbb{X}_B \perp\!\!\!\perp \mathbb{X}_C$ y $\mathbb{X}_C \perp\!\!\!\perp \mathbb{X}_A$.

La otra implicación es inmediata. □

Teorema 7. *Suponga que* $\mathbb{X} \sim \mathcal{N}(\mu, \mathbf{\Sigma})$, *donde*

$$\mathbb{X}^T = \begin{bmatrix} X_1 & X_2 & \dots & X_n \end{bmatrix} = \begin{bmatrix} \mathbb{X}_A^T & \mathbb{X}_B^T & \mathbb{X}_C^T \end{bmatrix},$$

siendo A, B *y* C *son conjuntos disjuntos no vacíos tales que* $A \cup B \cup C = \mathbb{N}_n$. *Suponga además que*

1. $\mathbb{X}_A \perp\!\!\!\perp \mathbb{X}_B$.
2. $\mathbb{X}_A \perp\!\!\!\perp \mathbb{X}_B \mid \mathbb{X}_M$, *para todo conjunto no vacío* $M \subseteq C$.

Entonces existen conjuntos disjuntos A^* *y* B^* *tales que* $A \subsetneqq A^*$ *o* $B \subsetneqq B^*$ *;* $A^* \cup B^* = \mathbb{N}_n$ *y*

$$\mathbb{X}_{A^*} \perp\!\!\!\perp \mathbb{X}_{B^*}.$$

Demostración. Por el Teorema 3 y la hipótesis, para cada conjunto no vacío $M \subseteq C$ se verifica $0 = \mathbf{\Sigma}_{A,B} = \mathbf{\Sigma}_{A,M}\mathbf{\Sigma}_{M,M}^{-1}\mathbf{\Sigma}_{M,B}$. En particular, para $M = \{c\}$; $c \in C$, se cumple

$$0 = \mathbf{\Sigma}_{A,B} = \mathbf{\Sigma}_{A,c}[\frac{1}{\sigma_{cc}}]\mathbf{\Sigma}_{c,B}.$$

Por tanto, para cada $c \in C$ se tiene $\mathbf{\Sigma}_{A,c}\mathbf{\Sigma}_{c,B} = 0$. Ahora bien; si $\mathbf{\Sigma}_{A,C} = 0$ o $\mathbf{\Sigma}_{C,B} = 0$, entonces se obtine el resultado. Basta tomar $A^* = A$ y $B^* = B \cup C$ en el primer caso o $A^* = A \cup C$ y $B^* = B$ en el segundo caso. Suponga que $\mathbf{\Sigma}_{A,C} \neq 0$ y que $\mathbf{\Sigma}_{C,B} \neq 0$. Considere la siguiente partición del conjunto C:

$$A_1 = \{c \in C : \mathbf{\Sigma}_{A,c} \neq 0, \mathbf{\Sigma}_{c,B} = 0\}; \quad B_1 = \{c \in C : \mathbf{\Sigma}_{A,c} = 0, \mathbf{\Sigma}_{c,B} \neq 0\};$$

$$C_1 = \{c \in C : \mathbf{\Sigma}_{A,c} = 0 = \mathbf{\Sigma}_{c,B}\}.$$

Así; $\mathbf{\Sigma}_{A_1,B} = 0$, $\mathbf{\Sigma}_{A,B_1} = 0$, $\mathbf{\Sigma}_{C_1,B} = 0 = \mathbf{\Sigma}_{C_1,B}$ y, como se muestra enseguida, se verifica la relación

$$\mathbb{X}_{A_1} \perp\!\!\!\perp \mathbb{X}_{B_1}.$$

Sean $c_1 \in A_1$ y $c_2 \in A_2$. Entonces existen $a \in A$ y $b \in B$ tales que $\sigma_{ac_1} \neq 0$ ($\mathbf{\Sigma}_{c_1,B} = 0$) y $\sigma_{c_2b} \neq 0$ ($\mathbf{\Sigma}_{c_2,B} = 0$). Por la hipótesis,

$$X_a \perp\!\!\!\perp X_b; \quad X_a \perp\!\!\!\perp X_b \mid \{X_{c_1}, X_{c_2}\}.$$

Por el Teorema 3,

$$0 = \sigma_{ab} = \begin{bmatrix} \sigma_{ac_1} & 0 \end{bmatrix} \begin{bmatrix} \sigma_{c_1c_1} & \sigma_{c_1c_2} \\ \sigma_{c_2c_1} & \sigma_{c_2c_2} \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ \sigma_{c_2b} \end{bmatrix}.$$

De esto se sigue que $\sigma_{c_1c_2} = 0$; es decir, $X_{c_1} \perp\!\!\!\perp X_{c_2}$. Ahora, si $C_1 = \emptyset$, entonces se obtiene el resultado tomando $A^* = A \cup A_1$ y $B^* = B \cup B_1$.

Suponga que $C_1 \neq \emptyset$. Se cumple entonces que para todo conjunto no vacío $M \subseteq C_1$:

$$\mathbb{X}_{A_1} \perp\!\!\!\perp \mathbb{X}_{B_1} \mid \mathbb{X}_M.$$

En efecto, sean como antes, $c_1 \in A_1$, $c_2 \in A_2$, $a \in A$ y $b \in B$ tales que $\sigma_{ac_1} \neq 0$ ($\mathbf{\Sigma}_{c_1,B} = 0$) y $\sigma_{c_2b} \neq 0$ ($\mathbf{\Sigma}_{c_2,B} = 0$). Sea, además, M un subconjunto no vacío de C_1. Por la hipótesis,

$$X_a \perp\!\!\!\perp X_b; \quad X_a \perp\!\!\!\perp X_b \mid \{X_{c_1}, X_{c_2}, \mathbb{X}_M\}.$$

Por el Teorema 3,

$$0 = \sigma_{ab} = \begin{bmatrix} \sigma_{ac_1} & 0 & 0 \end{bmatrix} \begin{bmatrix} \sigma_{c_1c_1} & 0 & \mathbf{\Sigma}_{c_1,M} \\ 0 & \sigma_{c_2c_2} & \mathbf{\Sigma}_{c_2,M} \\ \mathbf{\Sigma}_{M,c_1} & \mathbf{\Sigma}_{M,c_2} & \mathbf{\Sigma}_{M,M} \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ \sigma_{c_2b} \\ 0 \end{bmatrix}.$$

Una aplicación del Lema 1 permite concluir que

$$0 = \sigma_{c_1c_2} = \mathbf{\Sigma}_{c_1,M}\mathbf{\Sigma}_{M,M}^{-1}\mathbf{\Sigma}_{M,c_2},$$

es decir, $\mathbb{X}_{A_1} \perp\!\!\!\perp \mathbb{X}_{B_1} \mid \mathbb{X}_M$. En este caso, para $A' = A \cup A_1$, $B' = B \cup B_1$ y $C' = C_1$, se concluye que:

1. $\mathbb{X}_{A'} \perp\!\!\!\perp \mathbb{X}_{B'}$.
2. $\mathbb{X}_{A'} \perp\!\!\!\perp \mathbb{X}_{B'} \mid \mathbb{X}_M$, para todo conjunto no vacío $M \subseteq C'$,

esto es, se llega a obtener la hipótesis del teorema, salvo que ahora el cardinal de C' es estrictamente menor que el de C. Por tanto, después de un número finito de pasos se debe llegar a que existen conjuntos disjuntos A^* y B^* tales que $A \subsetneqq A^*$ o $B \subsetneqq B^*$; $A^* \cup B^* = \mathbb{N}_n$ y $\mathbb{X}_{A^*} \perp\!\!\!\perp \mathbb{X}_{B^*}$. □

Si en el Teorema anterior se empieza con $A = \{1\}$, $B = \{2\}$ y $C = \mathbb{N}_n - \{1, 2\}$, entonces se obtiene el siguiente Corolario, que es una versión del teorema principal de Marrelec y Benali[2].

Corolario 2. *Suponga que* $\mathbb{X} \sim \mathcal{N}(\mu, \mathbf{\Sigma})$; $\mathbb{X}^T = \begin{bmatrix} X_1 & X_2 & \ldots & X_n \end{bmatrix}$ *y que*

1. $X_1 \perp\!\!\!\perp X_2$.
2. $X_1 \perp\!\!\!\perp X_2 \mid \mathbb{X}_M$, *para todo conjunto no vacío* $M \subseteq \mathbb{N}_n - \{1, 2\}$.

Entonces existen conjuntos disjuntos A^* *y* B^* *tales que* $1 \in A^*$, $2 \in B^*$; $A^* \cup B^* = \mathbb{N}_n$ *y* $\mathbb{X}_{A^*} \perp\!\!\!\perp \mathbb{X}_{B^*}$.

5 Conclusiones

En este trabajo se han presentado algunas resultados sobre la relación de independencia condicionada en una distribución gaussiana. De manera especial, se ha dado una demostración elemental de una versión del teorema principal de Sullivant [4] y se ha generalizado el teorema principal de Marrelec y Benali[2] al caso de vectores particionados.

Referencias

[1] Graybill, F. A.: Introduction to the linear models, Duxbury Press, (1976)

[2] Marrelec, G., Benali, H.: Conditional independence between variables given any conditioning subset implies block diagonal covariance matrix for multivariate gaussian distributios, Statistics and Probability Letters 78, pg. 1922-1928, (2008)

[3] Shiryaev, A. N. : Probability. Second Edition, Springer, (1996)

[4] Sullivant, S.: Gaussian conditional independence relations have no finite complete characterization, Journal of Pure and Applied Algebra, Vol. 213, Issue 8, pg. 1502-1506, (2009)

Dirección del autor
Miguel A. Marmolejo L. — Departamento de Matemáticas, Universidad del Valle, Cali-Colombia
e-mail: `mimarmol@univalle.edu.co`