Introducción al Análisis Multivariado

SEMANA-2

Raúl Alberto Pérez

Universidad Nacional de Colombia, Escuela de Estadística, 2021-I

Vectores y Matrices Particionados

Particionamiento del Vector de Medias-Poblacionales

Sea $\underline{\mathbf{x}}$-un vector-aleatorio p-dimensional, entonces se tienen los siguientes particionamientos para el vector de variables aleatoria $\underline{\mathbf{x}}$, el vector de medias poblacionales $\underline{\mu}$ y la matriz de var-cov poblacionales $\mathbf{\Sigma}$, dadas por:

$$\underline{\mathbf{x}} = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \end{bmatrix} = \begin{bmatrix} X_1 \\ \vdots \\ X_q \\ \cdots \\ X_{q+1} \\ \vdots \\ X_p \end{bmatrix} = \begin{bmatrix} \underline{\mathbf{x}}^{(1)} \\ \cdots \\ \underline{\mathbf{x}}^{(2)} \end{bmatrix} \quad \text{y} \quad \underline{\mu} = E[\underline{\mathbf{x}}] = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_q \\ \cdots \\ \mu_{q+1} \\ \vdots \\ \mu_p \end{bmatrix} = \begin{bmatrix} \underline{\mu}^{(1)} \\ \cdots \\ \underline{\mu}^{(2)} \end{bmatrix}$$

Particionamiento de la Matriz de Var-Cov Poblacional Σ

$$\Sigma = \begin{bmatrix} \sigma_{11} & \cdots & \sigma_{1q} & | & \sigma_{1,q+1} & \cdots & \sigma_{1p} \\ \vdots & & \vdots & | & \vdots & & \vdots \\ \sigma_{q1} & \cdots & \sigma_{qq} & | & \sigma_{q,q+1} & \cdots & \sigma_{qp} \\ -- & -- & -- & -- & -- & -- & -- \\ \sigma_{q+1,1} & \cdots & \sigma_{q+1,q} & | & \sigma_{q+1,q+1} & \cdots & \sigma_{q+1,p} \\ \vdots & & \vdots & | & \vdots & & \vdots \\ \sigma_{p,1} & \cdots & \sigma_{p,q} & | & \sigma_{p,q+1} & \cdots & \sigma_{p,p} \end{bmatrix}$$

$$= \begin{array}{cc} & \begin{array}{ccc} q & & p-q \end{array} \\ \begin{array}{c} q \\ \\ p-q \end{array} & \begin{bmatrix} \Sigma_{11} & | & \Sigma_{12} \\ -- & -- & -- \\ \Sigma_{21} & | & \Sigma_{22} \end{bmatrix} \end{array}$$

La anterior matriz particionada se obtiene como sigue:

$$\left[\left(\underline{\mathbf{x}}^{(1)} - \underline{\mu}^{(1)}\right)\left(\underline{\mathbf{x}}^{(2)} - \underline{\mu}^{(2)}\right)^t\right] = \begin{bmatrix} X_1 - \mu_1 \\ X_2 - \mu_2 \\ \vdots \\ X_q - \mu_q \end{bmatrix} \begin{bmatrix} X_{q+1} - \mu_{q+1} & \cdots & X_p - \mu_p \end{bmatrix}$$

$$= \begin{bmatrix} \left(X_1 - \mu_1\right)\left(X_{q+1} - \mu_{q+1}\right) & \cdots & \left(X_1 - \mu_1\right)\left(X_p - \mu_p\right) \\ \left(X_2 - \mu_2\right)\left(X_{q+1} - \mu_{q+1}\right) & \cdots & \left(X_2 - \mu_2\right)\left(X_p - \mu_p\right) \\ \vdots & & \vdots \\ \left(X_q - \mu_q\right)\left(X_{q+1} - \mu_{q+1}\right) & \cdots & \left(X_q - \mu_q\right)\left(X_p - \mu_p\right) \end{bmatrix}$$

Ahora, tomando el valor esperado a ambos lados se tiene:

$$E\left[\left(\underline{\mathbf{x}}^{(1)}-\underline{\mu}^{(1)}\right)\left(\underline{\mathbf{x}}^{(2)}-\underline{\mu}^{(2)}\right)^t\right]$$

$$=\begin{bmatrix} E\left(X_1-\mu_1\right)\left(X_{q+1}-\mu_{q+1}\right) & \cdots & E\left(X_1-\mu_1\right)\left(X_p-\mu_p\right) \\ E\left(X_2-\mu_2\right)\left(X_{q+1}-\mu_{q+1}\right) & \cdots & E\left(X_2-\mu_2\right)\left(X_p-\mu_p\right) \\ \vdots & & \vdots \\ E\left(X_q-\mu_q\right)\left(X_{q+1}-\mu_{q+1}\right) & \cdots & E\left(X_q-\mu_q\right)\left(X_p-\mu_p\right) \end{bmatrix}$$

$$=\begin{bmatrix} \sigma_{1,q+1} & \sigma_{1,q+2} & \cdots & \sigma_{1p} \\ \vdots & \vdots & \vdots & \\ \sigma_{q,q+1} & \sigma_{q,q+2} & \cdots & \sigma_{qp} \end{bmatrix} = \mathbf{\Sigma}_{12}$$

Similarmente para $\mathbf{\Sigma}_{11}, \mathbf{\Sigma}_{22}$ y $\mathbf{\Sigma}_{21} = \mathbf{\Sigma}^t_{12}$

La matriz particionada completa se obtiene multiplicando los siguientes vectores particionados:

$$\left[\left(\underline{\mathbf{x}}-\underline{\mu}\right)\left(\underline{\mathbf{x}}-\underline{\mu}\right)^t\right] = \left[\begin{pmatrix}\underline{\mathbf{x}}^{(1)}\\ \cdots \\ \underline{\mathbf{x}}^{(2)}\end{pmatrix} - \begin{pmatrix}\underline{\mu}^{(1)}\\ \cdots \\ \underline{\mu}^{(2)}\end{pmatrix}\right]\left[\begin{pmatrix}\underline{\mathbf{x}}^{(1)}\\ \cdots \\ \underline{\mathbf{x}}^{(2)}\end{pmatrix} - \begin{pmatrix}\underline{\mu}^{(1)}\\ \cdots \\ \underline{\mu}^{(2)}\end{pmatrix}\right]^t$$

$$= \begin{bmatrix}\left(\underline{\mathbf{x}}^{(1)}-\underline{\mu}^{(1)}\right)\\ \\ \left(\underline{\mathbf{x}}^{(2)}-\underline{\mu}^{(2)}\right)\end{bmatrix}\left[\left(\underline{\mathbf{x}}^{(1)}-\underline{\mu}^{(1)}\right)^t \;\vdots\; \left(\underline{\mathbf{x}}^{(2)}-\underline{\mu}^{(2)}\right)^t\right]$$

$$= \begin{bmatrix}\left(\underline{\mathbf{x}}^{(1)}-\underline{\mu}^{(1)}\right)\left(\underline{\mathbf{x}}^{(1)}-\underline{\mu}^{(1)}\right)^t & \left(\underline{\mathbf{x}}^{(1)}-\underline{\mu}^{(1)}\right)\left(\underline{\mathbf{x}}^{(2)}-\underline{\mu}^{(2)}\right)^t\\ \\ \left(\underline{\mathbf{x}}^{(2)}-\underline{\mu}^{(2)}\right)\left(\underline{\mathbf{x}}^{(1)}-\underline{\mu}^{(1)}\right)^t & \left(\underline{\mathbf{x}}^{(2)}-\underline{\mu}^{(2)}\right)\left(\underline{\mathbf{x}}^{(2)}-\underline{\mu}^{(2)}\right)^t\end{bmatrix},$$ luego

$$\mathbf{\Sigma} = E\left[\left(\underline{\mathbf{x}}-\underline{\mu}\right)\left(\underline{\mathbf{x}}-\underline{\mu}\right)^t\right] = \begin{bmatrix}\mathbf{\Sigma}_{11} & \mathbf{\Sigma}_{12}\\ \mathbf{\Sigma}_{21} & \mathbf{\Sigma}_{22}\end{bmatrix}$$

Propiedades Sobre la Media y Varianza de Combinaciones Lineales

Sean X_1 y X_2 variables aleatorias y $a, b, c \in \mathbb{R}$, entonces:

1. $\left[cX_1\right] = cE[X_1]$

2. $E\left[aX_1 + bX_2\right] = E[X_1] + bE[X_2] = a\mu_1 + b\mu_2$.

   Notar que si, $aX_1 + bX_2 = \left[a \;\; b\right]\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = \underline{\mathbf{c}}^t\underline{\mathbf{x}}$, de donde:

   $$E\left[aX_1 + bX_2\right] = E\left[\underline{\mathbf{c}}^t\underline{\mathbf{x}}\right] = a\mu_1 + b\mu_2 = \left[a \;\; b\right]\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} = \underline{\mathbf{c}}^t\underline{\mu}$$

   con $\underline{\mathbf{c}} = (a \;\; b)^t$ y $\underline{\mu} = (\mu \;\; \mu_2)^t$.

3. $Var\left[cX_1\right] = c^2 Var[X_1]$

4. $Cov\left[aX_1\ ,\ bX_2\right] = abCov\left[X_1\ ,\ X_2\right] = ab\sigma_{12}$

5.

$$Var\left[aX_1 + bX_2\right] = a^2 Var[X_1] + b^2 Var[X_2] + 2abCov[X_1\ ,\ X_2]$$

$$= a^2\sigma_{11} + b^2\sigma_{22} + 2ab\sigma_{12}$$

Si, $\mathbf{\Sigma} = \begin{bmatrix} \mathbf{\Sigma}_{11} & \mathbf{\Sigma}_{12} \\ \mathbf{\Sigma}_{21} & \mathbf{\Sigma}_{22} \end{bmatrix}$, entonces:

$$\underline{\mathbf{c}}^t \mathbf{\Sigma} \underline{\mathbf{c}} = \begin{bmatrix} a & b \end{bmatrix} \begin{bmatrix} \mathbf{\Sigma}_{11} & \mathbf{\Sigma}_{12} \\ \mathbf{\Sigma}_{21} & \mathbf{\Sigma}_{22} \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = a^2\sigma_{11} + b^2\sigma_{22} + 2ab\sigma_{12}$$

En resumen:

$$Var\left[aX_1 + bX_2\right] = Var\left(\underline{\mathbf{c}}^t \underline{\mathbf{x}}\right) = \underline{\mathbf{c}}^t \mathbf{\Sigma} \underline{\mathbf{c}}$$

En general, dado vector aleatorio $\underline{\mathbf{x}} = \left(X_1, X_2, \cdots, X_p\right)^t$, con matriz de Var-Cov dada por $\mathbf{\Sigma}_{p\times p}$, si $\underline{\mathbf{c}} = \left(c_1, c_2, \cdots, c_p\right)^t$ es un vector de constantes, entonces:

$$E\left[\underline{\mathbf{c}}^t\underline{\mathbf{x}}\right] = E\left[c_1X_1 + c_2X_2 + \cdots + c_pX_p\right]$$

$$= \underline{\mathbf{c}}^t\underline{\mu},$$

y

$$Var\left[\underline{\mathbf{c}}^t\underline{\mathbf{x}}\right] = Var\left[c_1X_1 + c_2X_2 + \cdots + c_pX_p\right]$$

$$= \underline{\mathbf{c}}^t\mathbf{\Sigma}\underline{\mathbf{c}}.$$

6. Sea $\mathbf{C} = \Big[(c_{ij})\Big]_{q\times p}$ una matriz de constantes y $\underline{\mathbf{x}} = \Big(X_1, X_2, \cdots, X_p\Big)^t$ un vector aleatorio con vector de medias $\underline{\mu}$ y matriz de Var-Cov $\mathbf{\Sigma}_{p\times p}$. Para el vector aleatorio $\underline{\mathbf{z}}_{q\times 1}$ definido como: $\underline{\mathbf{z}} = \mathbf{C}\underline{\mathbf{x}}$, se tiene que:

$$E\Big[\underline{\mathbf{z}}\Big] = E\Big[\mathbf{C}\underline{\mathbf{x}}\Big] = \mathbf{C}\underline{\mu}$$

$$Var\Big[\underline{\mathbf{z}}\Big] = Var\Big[\mathbf{C}\underline{\mathbf{x}}\Big] = \mathbf{C}\mathbf{\Sigma}\mathbf{C}^t$$

7. Sea $\mathbf{A} = \Big[(a_{ij})\Big]_{q_1\times p}$ y $\mathbf{B} = \Big[(b_{ij})\Big]_{q_2\times p}$ matrices de constantes y $\underline{\mathbf{x}} = \Big(X_1, X_2, \cdots, X_p\Big)^t$ un vector aleatorio con vector de medias $\underline{\mu}$ y matriz de Var-Cov $\mathbf{\Sigma}_{p\times p}$. Para los vectores aleatorios: $\underline{\mathbf{z}}_1 = \mathbf{A}\underline{\mathbf{x}}$ y $\underline{\mathbf{z}}_2 = \mathbf{B}\underline{\mathbf{x}}$, se tiene que:

$$Cov\Big[\underline{\mathbf{z}}_1\ ,\ \underline{\mathbf{z}}_2\Big] = Cov\Big[\mathbf{A}\underline{\mathbf{x}}\ ,\ \mathbf{B}\underline{\mathbf{x}}\Big] = \underbrace{\mathbf{A}\mathbf{\Sigma}\mathbf{B}^t}_{q_1\times q_2},\ \text{ similarmente}$$

$$Cov\Big[\underline{\mathbf{z}}_1\ ,\ \underline{\mathbf{z}}_2\Big] = Cov\Big[\mathbf{A}\underline{\mathbf{x}}\ ,\ \mathbf{B}\underline{\mathbf{y}}\Big] = \underbrace{\mathbf{A}\mathbf{\Sigma}_{XY}\mathbf{B}^t}_{q_1\times q_2},\ \text{con }\ \underline{\mathbf{x}}, \underline{\mathbf{y}} - vect - aleats.$$

Ejemplo: Sea $\underline{\mathbf{x}} = \begin{pmatrix} X_1 & X_2 \end{pmatrix}^t$-un vector aleatorio con media $\underline{\mu} = \begin{pmatrix} \mu_1 & \mu_2 \end{pmatrix}$ y matriz de Var-Cov dada por: $\mathbf{\Sigma} = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix}$.

Sea el vector aleatorio $\underline{\mathbf{z}}_{2\times 1} = \begin{pmatrix} Z_1 & Z_2 \end{pmatrix}^t$ cuyas componentes están dadas por: $Z_1 = X_1 - X_2$ y $Z_2 = X_1 + X_2$, calcular la media y la matriz de Var-Cov de $\underline{\mathbf{z}}$.

Solución: El vector $\underline{\mathbf{z}}$ se puede escribir como sigue:

$$\underline{\mathbf{z}} = \begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = \mathbf{C}\underline{\mathbf{x}}$$

luego, usando el resultado anterior se tiene que:

$$E[\underline{\mathbf{z}}] = E[\mathbf{C}\underline{\mathbf{x}}] = \mathbf{C}\underline{\mu} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} = \begin{bmatrix} \mu_1 - \mu_2 \\ \mu_1 + \mu_2 \end{bmatrix}$$

$$\begin{aligned} Var[\underline{\mathbf{z}}] &= Var[\mathbf{C}\underline{\mathbf{x}}] = \mathbf{C\Sigma C}^t \\ &= \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \\ &= \begin{bmatrix} \sigma_{11} - \sigma_{12} & \sigma_{12} - \sigma_{22} \\ \sigma_{11} + \sigma_{12} & \sigma_{12} + \sigma_{22} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \\ &= \begin{bmatrix} \sigma_{11} - \sigma_{12} - \sigma_{12} + \sigma_{22} & \sigma_{11} - \sigma_{12} + \sigma_{12} - \sigma_{22} \\ \sigma_{11} + \sigma_{12} - \sigma_{12} - \sigma_{22} & \sigma_{11} + \sigma_{12} + \sigma_{12} + \sigma_{22} \end{bmatrix} \end{aligned}$$

$$= \begin{bmatrix} \sigma_{11} - 2\sigma_{12} + \sigma_{22} & \sigma_{11} - \sigma_{22} \\ \sigma_{11} - \sigma_{22} & \sigma_{11} + 2\sigma_{12} + \sigma_{22} \end{bmatrix} = \begin{bmatrix} Var[Z_1] & Cov[Z_1\ ,\ Z_2] \\ Cov[Z_2\ ,\ Z_1] & Var[Z_2] \end{bmatrix}$$

Particionamiento del Vector de Medias Muestrales

Sea $\underline{\mathbf{x}}$-un vector-aleatorio p-dimensional, entonces se tienen los siguientes particionamientos para el vector de variables aleatoria $\underline{\mathbf{x}}$, el vector de medias muestrales $\overline{\mathbf{x}}$ y la matriz de var-cov muestrales $\mathbf{S}$, dadas por:

$$\underline{\mathbf{x}} = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \end{bmatrix} = \begin{bmatrix} X_1 \\ \vdots \\ X_q \\ \cdots \\ X_{q+1} \\ \vdots \\ X_p \end{bmatrix} = \begin{bmatrix} \underline{\mathbf{x}}^{(1)} \\ \cdots \\ \underline{\mathbf{x}}^{(2)} \end{bmatrix} \quad \text{y} \quad \underline{\overline{\mathbf{x}}} = \begin{bmatrix} \overline{X}_1 \\ \vdots \\ \overline{X}_q \\ \cdots \\ \overline{X}_{q+1} \\ \vdots \\ \overline{X}_p \end{bmatrix} = \begin{bmatrix} \underline{\overline{\mathbf{x}}}^{(1)} \\ \cdots \\ \underline{\overline{\mathbf{x}}}^{(2)} \end{bmatrix}$$

Particionamiento de la Matriz de Var-Cov Muestrales

$$
\mathbf{S} = \left[\begin{array}{ccccccc}
s_{11} & \cdots & s_{1q} & | & s_{1,q+1} & \cdots & s_{1p} \\
\vdots & \ddots & \vdots & | & \vdots & \ddots & \vdots \\
s_{q1} & \cdots & s_{qq} & | & s_{q,q+1} & \cdots & s_{qp} \\
-- & -- & -- & -- & -- & -- & -- \\
s_{q+1,1} & \cdots & s_{q+1,q} & | & s_{q+1,q+1} & \cdots & s_{q+1,p} \\
\vdots & \ddots & \vdots & | & \vdots & \ddots & \vdots \\
s_{p,1} & \cdots & s_{p,q} & | & s_{p,q+1} & \cdots & s_{p,p}
\end{array}\right]
$$

$$
= \begin{array}{cc}
 & \begin{array}{ccc} q & & p-q \end{array} \\
\begin{array}{c} q \\ \\ p-q \end{array} & \left[\begin{array}{ccc}
S_{11} & | & S_{12} \\
-- & -- & -- \\
S_{21} & | & S_{22}
\end{array}\right]
\end{array}
$$

Algunas formas matriciales Eficientes

Para la matriz de datos

$$\mathbf{X}_{n\times p} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix}, \quad \mathbf{X}^T = \begin{bmatrix} x_{11} & x_{21} & \cdots & x_{n1} \\ x_{12} & x_{22} & \cdots & x_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ x_{1p} & x_{2p} & \cdots & x_{np} \end{bmatrix}$$

y

$$\mathbf{1} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}_{n\times 1}$$

Se tienen las siguientes expresiones para ciertas estadísticas de resúmenes:

1. $\underline{\bar{\mathbf{x}}} = \begin{bmatrix} \bar{X}_1 \\ \bar{X}_2 \\ \vdots \\ \bar{X}_p \end{bmatrix} = \frac{1}{n}\mathbf{X}^T\mathbf{1} = \frac{1}{n}\begin{bmatrix} \sum_{i=1}^n x_{i1} \\ \sum_{i=1}^n x_{i2} \\ \vdots \\ \sum_{i=1}^n x_{ip} \end{bmatrix}_{p\times 1}$

2. $\mathbf{S} = \frac{1}{n-1}\mathbf{X}^T\left[\mathbf{I}_n - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n^T\right]\mathbf{X} = \frac{1}{n-1}\mathbf{X}^T\mathbf{H}\mathbf{X}$, con $\mathbf{H} = \mathbf{I}_n - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n^T$

3. $\mathbf{R} = D^{-\frac{1}{2}}\mathbf{S}D^{-\frac{1}{2}}$, en donde $D^{-\frac{1}{2}}$, es una matriz diagonal cuyos elementos son los inversos de las desviaciones estándar muestrales, es decir que:

$$D^{-\frac{1}{2}} = \begin{bmatrix} \frac{1}{\sqrt{s_{11}}} & 0 & \cdots & 0 \\ 0 & \frac{1}{\sqrt{s_{22}}} & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \frac{1}{\sqrt{s_{pp}}} \end{bmatrix}$$

También se cumple que: $\mathbf{S} = D^{\frac{1}{2}}\mathbf{R}D^{\frac{1}{2}}$

4. Se define la Varianza-Generalizada de $\mathbf{X}$ como el determinante de $\mathbf{S}$,

$$VG = |\mathbf{S}|,$$

la cual representa una medida de variabilidad del vector de variables aleatorias $\underline{\mathbf{x}}$

5. Se define la Varianza Total de $\mathbf{X}$ como la traza de $\mathbf{S}$, ie

$$VT = Tr(\mathbf{S})$$

Muestra Aleatoria de Distribuciones p-Variadas

Sea un vector p-variado de variables aleatorias $\underline{\mathbf{x}} = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \end{bmatrix}$, con función de distribución multivariada representada por $f(\underline{\mathbf{x}})$, con media $\underline{\mu}$ y var-cov $\mathbf{\Sigma}$, ie, $E[\underline{\mathbf{x}}] = \underline{\mu}$ y Var$(\underline{\mathbf{x}}) = \mathbf{\Sigma}$.

Una muestra aleatoria de tamaño n de esta distribución es un conjunto de n-vectores aleatorios p-variados, $\underline{\mathbf{x}}_1, \underline{\mathbf{x}}_2, \cdots, \underline{\mathbf{x}}_n$, independientes e identicamente distribuídos con distribución $f(\underline{\mathbf{x}})$, es decir que:

$$h\left(\underline{\mathbf{x}}_1, \underline{\mathbf{x}}_2, \cdots, \underline{\mathbf{x}}_n\right) = h_1(\underline{\mathbf{x}}_1) h_2(\underline{\mathbf{x}}_2) \ldots h_n(\underline{\mathbf{x}}_n) = \prod_{i=1}^{n} f(\underline{\mathbf{x}}_i)$$

con $E[\underline{\mathbf{x}}_i] = \underline{\mu}$, Var$\left(\underline{\mathbf{x}}_i\right) = \mathbf{\Sigma}$ y $Cov\left(\underline{\mathbf{x}}_j \,,\, \underline{\mathbf{x}}_k\right) = 0 \;\forall\; j \neq k$

Teorema: Propiedades de $\overline{\underline{\mathbf{x}}}$

Dada una m.a de tamaña n, $\underline{\mathbf{x}}_1, \underline{\mathbf{x}}_2, \cdots, \underline{\mathbf{x}}_n$ de una distribución p-variada, con vector de media $\underline{\mu}$ y matriz de Var-Cov $\mathbf{\Sigma}$, ie. $E[\underline{\mathbf{x}}_i] = \underline{\mu}$ y $Var[\underline{\mathbf{x}}_i] = \mathbf{\Sigma}$, se cumple lo siguiente:

1. $E[\overline{\underline{\mathbf{x}}}] = \underline{\mu}$

2. Var$[\overline{\underline{\mathbf{x}}}] = \frac{1}{n}\mathbf{\Sigma}$

3. $E[\mathrm{S}] = \mathbf{\Sigma}$ y $E[\mathrm{S}_n] = \left(\frac{n-1}{n}\right)\mathbf{\Sigma}$

Dm: Para la parte (i) observe que:

$$\underline{\overline{\mathbf{x}}}_{p\times 1} = \begin{bmatrix} \bar{X}_1 \\ \bar{X}_2 \\ \vdots \\ \bar{X}_p \end{bmatrix} = \frac{1}{n}\begin{bmatrix} \sum_{i=1}^{n} x_{i1} \\ \sum_{i=1}^{n} x_{i2} \\ \vdots \\ \sum_{i=1}^{n} x_{ip} \end{bmatrix} = \frac{1}{n}\begin{bmatrix} x_{11}+x_{21}+\cdots+x_{n1} \\ x_{12}+x_{22}+\cdots+x_{n2} \\ \vdots \\ x_{1p}+x_{2p}+\cdots+x_{np} \end{bmatrix}_{p\times 1}$$

$$= 1/n\Big[\underline{\mathbf{x}}_1 + \underline{\mathbf{x}}_2 + \cdots + \underline{\mathbf{x}}_n\Big]$$

Luego de lo anterior se tiene que:

$$E[\underline{\overline{\mathbf{x}}}] = \frac{1}{n}E\Big[\underline{\mathbf{x}}_1 + \underline{\mathbf{x}}_2 + \cdots + \underline{\mathbf{x}}_n\Big]$$

$$= \frac{1}{n}\Big(E[\underline{\mathbf{x}}_1] + E[\underline{\mathbf{x}}_2] + \cdots + E[\underline{\mathbf{x}}_n]\Big)$$

$$E[\underline{\overline{\mathbf{x}}}] = \frac{1}{n}\Big(\underline{\mu} + \underline{\mu} + \cdots + \underline{\mu}\Big) = \frac{1}{n}n\underline{\mu} = \underline{\mu}$$

Para la parte (ii), oberve que:

$$Var(\underline{\overline{\mathbf{x}}}) = E\left[\left(\underline{\overline{\mathbf{x}}} - E[\underline{\overline{\mathbf{x}}}]\right)\left(\underline{\overline{\mathbf{x}}} - E[\underline{\overline{\mathbf{x}}}]\right)^T\right] = E\left[\left(\underline{\overline{\mathbf{x}}} - \underline{\mu}\right)\left(\underline{\overline{\mathbf{x}}} - \underline{\mu}\right)^T\right]$$

$$= E\left[\left(\frac{1}{n}\sum_{j=1}^{n}\left(\underline{\mathbf{x}}_j - \underline{\mu}\right)\right)\left(\frac{1}{n}\sum_{k=1}^{n}\left(\underline{\mathbf{x}}_k - \underline{\mu}\right)\right)^T\right]$$

$$= \frac{1}{n^2}\sum_{j=1}^{n}\sum_{k=1}^{n} E\left(\underline{\mathbf{x}}_j - \underline{\mu}\right)\left(\underline{\mathbf{x}}_k - \underline{\mu}\right)^T$$

$$= \frac{1}{n^2}\left[\sum_{j=1}^{n} E\left(\underline{\mathbf{x}}_j - \underline{\mu}\right)\left(\underline{\mathbf{x}}_j - \underline{\mu}\right)^T + \sum_{j,k,j\neq k} E\left(\underline{\mathbf{x}}_j - \underline{\mu}\right)\left(\underline{\mathbf{x}}_k - \underline{\mu}\right)^T\right]$$

$$= \frac{1}{n^2}\left[\sum_{j=1}^{n} \mathrm{Var}(\underline{\mathbf{x}}_j) + \sum_{j,k,j\neq k} \mathrm{Cov}(\underline{\mathbf{x}}_j, \underline{\mathbf{x}}_k)\right]$$

$$= \frac{1}{n^2}[\mathbf{\Sigma} + \mathbf{\Sigma} + \cdots + \mathbf{\Sigma} + 0 + 0 + \cdots + 0] = \frac{1}{n^2}n\mathbf{\Sigma} = \frac{1}{n}\mathbf{\Sigma}$$

Para la parte (iii), observe que:

$$\sum_{j=1}^{n}\left(\underline{\mathbf{x}}_j-\overline{\underline{\mathbf{x}}}\right)\left(\underline{\mathbf{x}}_j-\overline{\underline{\mathbf{x}}}\right)^T=\sum_{j=1}^{n}\left(\underline{\mathbf{x}}_j-\overline{\underline{\mathbf{x}}}\right)(\underline{\mathbf{x}}_j^T-\overline{\underline{\mathbf{x}}}^T)$$

$$=\sum_{j=1}^{n}\left(\underline{\mathbf{x}}_j-\overline{\underline{\mathbf{x}}}\right)\underline{\mathbf{x}}_j^T-\sum_{j=1}^{n}\left(\underline{\mathbf{x}}_j-\overline{\underline{\mathbf{x}}}\right)\overline{\underline{\mathbf{x}}}^T$$

$$=\sum_{j=1}^{n}\underline{\mathbf{x}}_j\underline{\mathbf{x}}_j^T-n\overline{\underline{\mathbf{x}}}\;\overline{\underline{\mathbf{x}}}^T,$$

lo anterior debido a que se cumple lo siguiente:

$$\sum_{j=1}^{n}\left(\underline{\mathbf{x}}_j-\overline{\underline{\mathbf{x}}}\right)=0 \quad \text{y} \quad \sum_{j=1}^{n}\underline{\mathbf{x}}_j^T=n\overline{\underline{\mathbf{x}}}^T$$

Tomando valor esperado en lo anterior se tiene que:

$$E\left[\sum_{j=1}^{n}\left(\underline{\mathbf{x}}_j-\overline{\underline{\mathbf{x}}}\right)\left(\underline{\mathbf{x}}_j-\overline{\underline{\mathbf{x}}}\right)^T\right]=E\left[\sum_{j=1}^{n}\underline{\mathbf{x}}_j\underline{\mathbf{x}}_j^T-n\overline{\underline{\mathbf{x}}}\,\overline{\underline{\mathbf{x}}}^T\right]$$

$$=\sum_{j=1}^{n}E\left[\underline{\mathbf{x}}_j\underline{\mathbf{x}}_j^T\right]-nE\left[\overline{\underline{\mathbf{x}}}\,\overline{\underline{\mathbf{x}}}^T\right],$$

pero por propiedades del valor esperado y de la matriz de Var-Cov de un vector aleatorio $\underline{\mathbf{x}}$ se tiene que:

$$Var(\ \underline{\mathbf{x}}\ )=\mathbf{\Sigma}=E\left[\underline{\mathbf{x}\mathbf{x}}^T\right]-E\left[\underline{\mathbf{x}}\right]E\left[\underline{\mathbf{x}}\right]^T,$$

es decir que,

$$E\left[\underline{\mathbf{x}\mathbf{x}}^T\right]=\mathbf{\Sigma}+E\left[\underline{\mathbf{x}}\right]E\left[\underline{\mathbf{x}}\right]^T,$$

de donde:

$$\sum_{j=1}^{n} E\left[\underline{\mathbf{x}}_j\underline{\mathbf{x}}_j^T\right] - nE\left[\overline{\underline{\mathbf{x}}}\,\overline{\underline{\mathbf{x}}}^T\right]$$

$$= \sum_{j=1}^{n} \left[Var(\ \underline{\mathbf{x}}_j\ ) + E\left[\underline{\mathbf{x}}_j\right] E\left[\underline{\mathbf{x}}_j\right]^T\right] - n\left[Var(\ \overline{\underline{\mathbf{x}}}\ ) + E\left[\overline{\underline{\mathbf{x}}}\right] E\left[\overline{\underline{\mathbf{x}}}\right]^T\right]$$

$$= \sum_{j=1}^{n} \left(\mathbf{\Sigma} + \underline{\mu}\ \underline{\mu}^{\ T}\right) - n\left(\frac{1}{n}\mathbf{\Sigma} + \underline{\mu}\ \underline{\mu}^{\ T}\right)$$

$$= n\mathbf{\Sigma} + n\underline{\mu}\ \underline{\mu}^{\ T} - \mathbf{\Sigma} - n\underline{\mu}\ \underline{\mu}^{\ T} = (n-1)\mathbf{\Sigma},$$

es decir que,

$$E\left[\sum_{j=1}^{n}\left(\underline{\mathbf{x}}_j - \overline{\underline{\mathbf{x}}}\right)\left(\underline{\mathbf{x}}_j - \overline{\underline{\mathbf{x}}}\right)^T\right] = (n-1)\mathbf{\Sigma}$$

Pero

$$\mathbf{S} = \frac{1}{n-1}\sum_{j=1}^{n}\left(\underline{\mathbf{x}}_j - \overline{\underline{\mathbf{x}}}\right)\left(\underline{\mathbf{x}}_j - \overline{\underline{\mathbf{x}}}\right)^T,$$

y por lo tanto se tiene que:

$$E[\mathbf{S}] = \frac{1}{n-1}E\left[\sum_{j=1}^{n}\left(\underline{\mathbf{x}}_j - \overline{\underline{\mathbf{x}}}\right)\left(\underline{\mathbf{x}}_j - \overline{\underline{\mathbf{x}}}\right)^T\right]$$

$$= \frac{1}{n-1}(n-1)\mathbf{\Sigma}$$

$$E[\mathbf{S}] = \mathbf{\Sigma}$$

Resumen: EL teorema anterior dice que el vector de medias muestrales $\overline{\underline{\mathbf{x}}}$ es un estimador insesgado del vector de medias poblacionales $\underline{\mu}$ y que la matriz de Var-Cov muestrales $\mathbf{S}$ también es un estimador insesgado de la matriz de Var-Cov poblacionales $\mathbf{\Sigma}$, pero $\mathbf{S}_n$ es sesgado.