Preliminarios

Graciela Boente

Definiciones

- $\mathbf{I}_p$ Matriz identidad en $\mathbb{R}^{p\times p}$
- $\mathbf{1}_{pq} \in \mathbb{R}^{p\times q}$ la matriz con sus elementos iguales a 1
- $\mathbf{1}_p \in \mathbb{R}^p$ el vector con sus elementos iguales a 1
- $\mathbf{A}$ es definida no negativa $\mathbf{A} \geq 0 \Leftrightarrow \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x} \geq \mathbf{0}\ \forall \mathbf{x}$
- $\mathbf{A}$ es definida positiva $\mathbf{A} > \mathbf{0} \Leftrightarrow \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x} > 0\ \forall \mathbf{x} \neq 0$
- $\mathbf{A} \geq \mathbf{B}$ sii $\mathbf{A} - \mathbf{B} \geq \mathbf{0}$
- Dada $\mathbf{A} \in \mathbb{R}^{p\times p}$, $\mathrm{tr}(\mathbf{A}) = \sum_{i=1}^{p} a_{ii}$

Propiedades

1. $\mathrm{tr}(\mathbf{A}+\mathbf{B}) = \mathrm{tr}(\mathbf{A}) + \mathrm{tr}(\mathbf{B})$
2. $\mathrm{tr}(\mathbf{AB}) = \mathrm{tr}(\mathbf{BA})$

3. Sea $\mathbf{A} \in \mathbb{R}^{p\times p}$ simétrica. Sean $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p$ los autovalores de $\mathbf{A}$.
   - Existe $\mathbf{T} = (\mathbf{t}_1, \ldots, \mathbf{t}_p)$ ortogonal tal que

     $$\mathbf{T}^{\mathrm{T}}\mathbf{A}\mathbf{T} = \mathrm{diag}(\lambda_1, \ldots, \lambda_p)$$

     Por lo tanto,

     $$\mathbf{A} = \sum_{i=1}^{p} \lambda_i \mathbf{t}_i\, \mathbf{t}_i^{\mathrm{T}}$$

   - $\mathrm{tr}(\mathbf{A}) = \sum_{i=1}^{p} \lambda_i$, $|\mathbf{A}| = \det(\mathbf{A}) = \prod_{i=1}^{p} \lambda_i$
   - $\det(\mathbf{I}_p \pm \mathbf{A}) = \prod_{i=1}^{p} (1 \pm \lambda_i)$

Propiedades

4. Sean $\mathbf{A} \in \mathbb{R}^{p \times p}$, $\mathbf{D} \in \mathbb{R}^{q \times q}$, simétricas entonces

$$\begin{vmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{vmatrix} = \begin{cases} |\mathbf{D}| \, \left|\mathbf{A} - \mathbf{B}\,\mathbf{D}^{-1}\mathbf{C}\right| & \text{si } \mathbf{D}^{-1} \text{ existe} \\ |\mathbf{A}| \, \left|\mathbf{D} - \mathbf{C}\,\mathbf{A}^{-1}\mathbf{B}\right| & \text{si } \mathbf{A}^{-1} \text{ existe} \end{cases}$$

5. **Descomposición de Cholesky**
   Sea $\mathbf{A} \in \mathbb{R}^{p \times p}$ simétrica. Existe $\mathbf{C}$ triangular inferior tal que

$$\mathbf{A} = \mathbf{C}\mathbf{C}^{\mathrm{T}}$$

Propiedades

4. Sean $\mathbf{A} \in \mathbb{R}^{p\times p}$, $\mathbf{D} \in \mathbb{R}^{q\times q}$, simétricas entonces

$$\begin{vmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{vmatrix} = \begin{cases} |\mathbf{D}| \; \left|\mathbf{A} - \mathbf{B}\,\mathbf{D}^{-1}\mathbf{C}\right| & \text{si } \mathbf{D}^{-1} \text{ existe} \\ |\mathbf{A}| \; \left|\mathbf{D} - \mathbf{C}\,\mathbf{A}^{-1}\mathbf{B}\right| & \text{si } \mathbf{A}^{-1} \text{ existe} \end{cases}$$

5. **Descomposición de Cholesky**
   Sea $\mathbf{A} \in \mathbb{R}^{p\times p}$ simétrica. Existe $\mathbf{C}$ triangular inferior tal que

$$\mathbf{A} = \mathbf{C}\mathbf{C}^{\mathrm{T}}$$

- $\mathbf{P}$ se dice idempotente si $\mathbf{P}^2 = \mathbf{P}$.
- Una matriz de proyección es una matriz simétrica idempotente.

Propiedades

Sea $\mathbf{P} \in \mathbb{R}^{p \times p}$.

7. Si $\mathbf{P}$ es simétrica, entonces $\mathbf{P}$ es idempotente de rango $r \Leftrightarrow$ tiene r autovalores 1 y $p - r$ autovalores 0

8. Si $\mathbf{P}$ es una matriz de proyección de rango r, entonces

$$\mathbf{P} = \sum_{i=1}^{r} \mathbf{t}_i \mathbf{t}_i^{\mathrm{T}}$$

   con $\mathbf{t}_1, \ldots, \mathbf{t}_r$ ortonormales.

9. Si $\mathbf{P}$ es una matriz de proyección $\mathrm{tr}(\mathbf{P}) = \mathrm{rango}(\mathbf{P})$.

10. Si $\mathbf{P}$ es idempotente $\Longrightarrow \mathbf{I}_p - \mathbf{P}$ lo es.

Más Propiedades

Sean $\mathbf{A}, \mathbf{V} \in \mathbb{R}^{p \times p}$, $\mathbf{A}$ simétrica, $\mathbf{x}_j \in \mathbb{R}^p$, $1 \leq j \leq n$, y

$$\mathbf{X} = \begin{pmatrix} \mathbf{x}_1^{\mathrm{T}} \\ \vdots \\ \mathbf{x}_n^{\mathrm{T}} \end{pmatrix} = \left(\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(p)}\right) .$$

11. $\displaystyle\sum_{j=1}^{n} \mathbf{x}_j^{\mathrm{T}} \mathbf{A} \mathbf{x}_j = \operatorname{tr}\left(\mathbf{A} \sum_{j=1}^{n} \mathbf{x}_j \mathbf{x}_j^{\mathrm{T}}\right) = \operatorname{tr}\left(\mathbf{A}\, \mathbf{X}^{\mathrm{T}} \mathbf{X}\right)$
12. $\displaystyle\frac{\partial}{\partial \mathbf{x}}(\mathbf{x}^{\mathrm{T}} \mathbf{A} \mathbf{x}) = 2\mathbf{A}\mathbf{x} \qquad \frac{\partial}{\partial \mathbf{A}}(\mathbf{x}^{\mathrm{T}} \mathbf{A} \mathbf{x}) = \mathbf{x}\mathbf{x}^{\mathrm{T}}$
13. $\displaystyle\frac{\partial}{\partial \mathbf{A}} \operatorname{tr}(\mathbf{A}\mathbf{V}) = \mathbf{V}^{\mathrm{T}}$
14. $\displaystyle\frac{\partial}{\partial \mathbf{A}} \det(\mathbf{A}) = \det(\mathbf{A})\, \mathbf{A}^{-1}$

Definiciones

- Sea $\mathbf{X} = (x_{ij})$ una matriz de variables aleatorias, se define $\mathbb{E}\mathbf{X} = (\mathbb{E}x_{ij})$

- Sean $\mathbf{x} \in \mathbb{R}^p$, $\mathbf{y} \in \mathbb{R}^q$ dos vectores aleatorios, se define la matriz de covarianzas de $\mathbf{x}$ e $\mathbf{y}$ como

$$\begin{aligned}\boldsymbol{\Sigma}_{\mathbf{x},\mathbf{y}} &= \mathrm{Cov}\,(\mathbf{x},\mathbf{y}) = (\mathrm{Cov}(x_i, y_j))_{1\leq i\leq p, 1\leq j\leq q} \\ &= \mathbb{E}(\mathbf{x}\,\mathbf{y}^{\mathrm{T}}) - \mathbb{E}(\mathbf{x})\,\mathbb{E}(\mathbf{y})^{\mathrm{T}}\,.\end{aligned}$$

- Dado $\mathbf{x} \in \mathbb{R}^p$,

$$\boldsymbol{\Sigma}_{\mathbf{x},\mathbf{x}} = \mathrm{Cov}\,(\mathbf{x},\mathbf{x}) = \mathrm{Cov}(\mathbf{x}) = \mathrm{Var}(\mathbf{x})$$

es una matriz simétrica definida no–negativa.

Propiedades

Sean $\mathbf{X} = (x_{ij})$ una matriz de variables aleatorias, $\mathbf{x} \in \mathbb{R}^p$ e $\mathbf{y} \in \mathbb{R}^q$ dos vectores aleatorios

- $\mathbb{E}(\mathbf{AXB} + \mathbf{C}) = \mathbf{A}\,\mathbb{E}(\mathbf{X})\,\mathbf{B} + \mathbf{C}$.
- Si $\mathbf{x}$ e $\mathbf{y}$ son independientes, entonces $\mathrm{Cov}\,(\mathbf{x}, \mathbf{y}) = \mathbf{0}$.
- $\mathrm{Cov}\,(\mathbf{A}\,\mathbf{x}, \mathbf{B}\,\mathbf{y}) = \mathbf{A}\,\mathrm{Cov}\,(\mathbf{x}, \mathbf{y})\ \mathbf{B}^{\mathrm{T}}$.
- $\mathrm{Cov}(\mathbf{A}\,\mathbf{x}) = \mathbf{A}\,\mathrm{Cov}(\mathbf{x})\,\mathbf{A}^{\mathrm{T}}$.

Definición

Sean $\mathbf{x}_i \in \mathbb{R}^p$, $1 \leq j \leq n$, independientes, $\mathbb{E}(\mathbf{x}_i) = \boldsymbol{\mu}_i$, $\mathrm{Cov}(\mathbf{x}_i) = \boldsymbol{\Sigma}_i$

$$\mathbf{X} = \begin{pmatrix} \mathbf{x}_1^{\mathrm{T}} \\ \vdots \\ \mathbf{x}_n^{\mathrm{T}} \end{pmatrix} = \left(\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(p)}\right) .$$

Se define la matriz de covarianza de $\mathbf{X}$ como la matriz de covarianza de

$$\mathbf{y} = \mathrm{vec}(\mathbf{X}) = \begin{pmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_n \end{pmatrix}$$

Propiedad

Se define el producto de Kronecker de las matrices $\mathbf{A} \in \mathbb{R}^{p\times p}$ y $\mathbf{B} \in \mathbb{R}^{n\times n}$ como la matriz $\mathbf{A}\otimes\mathbf{B} \in \mathbb{R}^{pn\times pn}$,

$$\mathbf{A}\otimes\mathbf{B} = \begin{pmatrix} a_{11}\mathbf{B} & \dots & a_{1p}\mathbf{B} \\ \vdots & & \vdots \\ a_{p1}\mathbf{B} & \dots & a_{pp}\mathbf{B} \end{pmatrix}$$

Propiedad

Sean $\mathbf{x}_i \in \mathbb{R}^p$, $1 \le j \le n$, independientes, $\mathbb{E}(\mathbf{x}_i) = \boldsymbol{\mu}$, $\mathrm{Cov}(\mathbf{x}_i) = \boldsymbol{\Sigma}$, entonces

$$\mathbb{E}(\mathbf{X}) = \mathbf{1}_n\boldsymbol{\mu}^{\mathrm{T}} \qquad \mathrm{Cov}(\mathbf{X}) = \begin{pmatrix} \boldsymbol{\Sigma} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ & & \ddots & & \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \boldsymbol{\Sigma} \end{pmatrix} = \mathbf{I}_n \otimes \boldsymbol{\Sigma}$$

Propiedad

Más aún, si $\mathbf{P} \in \mathbb{R}^{d\times p}$, $\mathbf{Q} \in \mathbb{R}^{M\times n}$,

$$\mathrm{Cov}(\mathbf{QXP}^{\mathrm{T}}) = \mathbf{QI}_n\mathbf{Q}^{\mathrm{T}} \otimes \mathbf{P\Sigma P}^{\mathrm{T}}$$

Luego, si $M = n$ y $\mathbf{Q} \in \mathbb{R}^{n\times n}$ es ortogonal,

$$\mathrm{Cov}(\mathbf{QXP}^{\mathrm{T}}) = \mathbf{I}_n \otimes \mathbf{P\Sigma P}^{\mathrm{T}}$$

Sean $\mathbf{x}_i \in \mathbb{R}^p$, $1 \leq j \leq n$, independientes, $\mathbb{E}(\mathbf{x}_i) = \boldsymbol{\mu}_i$, $\mathrm{Cov}(\mathbf{x}_i) = \boldsymbol{\Sigma}_i$

Lema

Consideremos la forma cuadrática

$$\mathbf{X}^{\mathrm{T}}\mathbf{A}\mathbf{X} = \sum_{1\leq i,j\leq n} a_{ij}\mathbf{x}_i\,\mathbf{x}_j^{\mathrm{T}}\,.$$

- $\mathbb{E}\left(\mathbf{X}^{\mathrm{T}}\mathbf{A}\mathbf{X}\right) = \sum_{i=1}^{n} a_{ii}\boldsymbol{\Sigma}_i + \mathbb{E}\left(\mathbf{X}\right)^{\mathrm{T}}\,\mathbf{A}\,\mathbb{E}\left(\mathbf{X}\right)$
- Si $\boldsymbol{\Sigma}_i = \boldsymbol{\Sigma}$ para todo i

$$\mathbb{E}\left(\mathbf{X}^{\mathrm{T}}\mathbf{A}\mathbf{X}\right) = \mathrm{tr}(\mathbf{A})\boldsymbol{\Sigma} + \mathbb{E}\left(\mathbf{X}\right)^{\mathrm{T}}\,\mathbf{A}\,\mathbb{E}\left(\mathbf{X}\right)$$

Sean $\mathbf{x}_i \in \mathbb{R}^p$, $1 \leq j \leq n$, i.i.d $\mathbb{E}(\mathbf{x}_i) = \boldsymbol{\mu}$, $\mathrm{Cov}(\mathbf{x}_i) = \boldsymbol{\Sigma}$.

Por analogía con el caso univariado definamos

$$
\begin{aligned}
\overline{\mathbf{x}} &= \frac{1}{n}\sum_{i=1}^{n} \mathbf{x}_i = \frac{1}{n}\mathbf{X}^{\mathrm{T}}\mathbf{1}_n \\
\mathbf{S} &= \frac{1}{n-1}\sum_{i=1}^{n}(\mathbf{x}_i - \overline{\mathbf{x}})(\mathbf{x}_i - \overline{\mathbf{x}})^{\mathrm{T}} = \frac{1}{n-1}\mathbf{Q} \\
&= \frac{1}{n-1}\left\{\sum_{i=1}^{n}(\mathbf{x}_i - \boldsymbol{\mu})(\mathbf{x}_i - \boldsymbol{\mu})^{\mathrm{T}} - n\left(\overline{\mathbf{x}} - \boldsymbol{\mu}\right)\left(\overline{\mathbf{x}} - \boldsymbol{\mu}\right)^{\mathrm{T}}\right\}
\end{aligned}
$$

Sean $\mathbf{x}_i \in \mathbb{R}^p$, $1 \leq j \leq n$, i.i.d $\mathbb{E}(\mathbf{x}_i) = \boldsymbol{\mu}$, $\mathrm{Cov}(\mathbf{x}_i) = \boldsymbol{\Sigma}$.

Por analogía con el caso univariado definamos

$$
\begin{aligned}
\overline{\mathbf{x}} &= \frac{1}{n}\sum_{i=1}^{n}\mathbf{x}_i = \frac{1}{n}\mathbf{X}^{\mathrm{T}}\mathbf{1}_n \\
\mathbf{S} &= \frac{1}{n-1}\sum_{i=1}^{n}(\mathbf{x}_i - \overline{\mathbf{x}})(\mathbf{x}_i - \overline{\mathbf{x}})^{\mathrm{T}} = \frac{1}{n-1}\mathbf{Q} \\
&= \frac{1}{n-1}\left\{\sum_{i=1}^{n}(\mathbf{x}_i - \boldsymbol{\mu})(\mathbf{x}_i - \boldsymbol{\mu})^{\mathrm{T}} - n\left(\overline{\mathbf{x}} - \boldsymbol{\mu}\right)\left(\overline{\mathbf{x}} - \boldsymbol{\mu}\right)^{\mathrm{T}}\right\} \\
&= \frac{1}{n-1}\widetilde{\mathbf{X}}^{\mathrm{T}}\widetilde{\mathbf{X}}
\end{aligned}
$$

donde

$$
\widetilde{\mathbf{X}}^{\mathrm{T}} = (\mathbf{x}_1 - \overline{\mathbf{x}}, \ldots, \mathbf{x}_n - \overline{\mathbf{x}}) = \mathbf{X}^{\mathrm{T}}(\mathbf{I}_n - \mathbf{P}_1) = \mathbf{X}^{\mathrm{T}}(\mathbf{I}_n - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n^{\mathrm{T}})
$$

Luego,

$$\mathbf{Q} \quad = \quad \mathbf{X}^{\mathrm{T}}(\mathbf{I}_n - \mathbf{P}_1)\mathbf{X}$$

de donde como $\mathbf{P}_1$ es idempotente de rango 1 y $(\mathbf{I}_n - \mathbf{P}_1)\mathbf{1}_n = \mathbf{0}$ obtenemos del Lema que

$$\mathbb{E}(\mathbf{Q}) = (n-1)\boldsymbol{\Sigma}$$

Además, $\mathbb{E}(\overline{\mathbf{x}}) = \boldsymbol{\mu}$ por lo que $\overline{\mathbf{x}}$ y $\mathbf{S}$ son estimadores insesgados de $\boldsymbol{\mu}$ y de $\boldsymbol{\Sigma}$, respectivamente.

Por otra parte,

$$\mathrm{Cov}(\overline{\mathbf{x}}) = \frac{1}{n}\boldsymbol{\Sigma} \to 0$$

es decir, $\overline{\mathbf{x}}$ es un estimador débilmente consistente de $\boldsymbol{\mu}$.

Matriz de correlación

Sea $\mathbf{x}$ tal que $\mathrm{Cov}(\mathbf{x}) = \boldsymbol{\Sigma} = (\sigma_{ij})_{1\leq i,j\leq p}$.

La matriz de correlación del vector $\mathbf{x}$ está dada por

$$\mathrm{Corr}(\mathbf{x}) = \boldsymbol{\Delta}^{-\frac{1}{2}}\boldsymbol{\Sigma}\boldsymbol{\Delta}^{-\frac{1}{2}}$$

donde

$$\boldsymbol{\Delta} = \mathrm{diag}\left(\sigma_{11}, \ldots, \sigma_{pp}\right) .$$

Un estimador de $\mathrm{Corr}(\mathbf{x})$ es

$$\mathbf{R} = \mathbf{D}^{-\frac{1}{2}}\mathbf{S}\mathbf{D}^{-\frac{1}{2}}$$

donde $\mathbf{D} = \mathrm{diag}\left(s_{11}, \ldots, s_{pp}\right)$

$$\mathbf{S} = \frac{1}{n-1}\sum_{i=1}^{n}(\mathbf{x}_i - \overline{\mathbf{x}})(\mathbf{x}_i - \overline{\mathbf{x}})^{\mathrm{T}} = (s_{ij})_{1\leq i,j\leq p}$$

Luego,

$$r_{jk} = \frac{s_{jk}}{\sqrt{s_{jj}\, s_{kk}}} = \frac{\sum_{i=1}^{n}(x_{ij} - \overline{x}_j)(x_{ik} - \overline{x}_k)}{\left(\sum_{i=1}^{n}(x_{ij} - \overline{x}_j)^2 \, \sum_{i=1}^{n}(x_{ik} - \overline{x}_k)^2\right)^{\frac{1}{2}}}$$