IECD-325: Elementos de álgebra matricial

Felipe Osorio

felipe.osorio@uv.cl

Definiciones básicas y propiedades

Sea $\mathbb{R}^n$ el espacio Euclidiano n-dimensional, de este modo $\boldsymbol{x} \in \mathbb{R}^n$ representa la n-upla

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix},$$

de números reales.

Observación:

Siempre consideraremos $\boldsymbol{x}$ como un vector columna.

De este modo, podemos escribir

$$\boldsymbol{x} = (x_1, \dots, x_n)^\top.$$

Una matriz $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ es un arreglo de números reales

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix},$$

y escribimos $\boldsymbol{A} = (a_{ij})$. Los números reales a_{ij} son llamados elementos de $\boldsymbol{A}$.

La suma de dos matrices del mismo orden es definida como

$$\boldsymbol{A} + \boldsymbol{B} = (a_{ij}) + (b_{ij}) = (a_{ij} + b_{ij}),$$

el producto de una matriz por un escalar λ es

$$\lambda \boldsymbol{A} = \boldsymbol{A} \lambda = (\lambda a_{ij})$$

Propiedades:

Sean $\boldsymbol{A}, \boldsymbol{B}$ y $\boldsymbol{C}$ matrices del mismo orden y λ, μ escalares. Entonces:

(a) $\boldsymbol{A} + \boldsymbol{B} = \boldsymbol{B} + \boldsymbol{A}$,

(b) $(\boldsymbol{A} + \boldsymbol{B}) + \boldsymbol{C} = \boldsymbol{A} + (\boldsymbol{B} + \boldsymbol{C})$,

(c) $(\lambda + \mu)\boldsymbol{A} = \lambda \boldsymbol{A} + \mu \boldsymbol{A}$,

(d) $\lambda(\boldsymbol{A} + \boldsymbol{B}) = \lambda \boldsymbol{A} + \lambda \boldsymbol{B}$.

Una matriz cuyos elementos son todos cero se denomina matriz nula y se denota por $\mathbf{0}$. Tenemos que

$$\boldsymbol{A} + (-1)\boldsymbol{A} = \mathbf{0}.$$

Si $\boldsymbol{A}$ y $\boldsymbol{B}$ son matrices $m \times n$ y $n \times p$, respectivamente, se define el producto de $\boldsymbol{A}$ y $\boldsymbol{B}$ como

$$\boldsymbol{AB} = \boldsymbol{C}, \qquad c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}, \quad i = 1, \ldots, m; j = 1, \ldots, p.$$

Propiedades:

Sean $\boldsymbol{A}, \boldsymbol{B}$ y $\boldsymbol{C}$ matrices de órdenes apropiados. Entonces:

(a) $(\boldsymbol{AB})\boldsymbol{C} = \boldsymbol{A}(\boldsymbol{BC})$,

(b) $\boldsymbol{A}(\boldsymbol{B} + \boldsymbol{C}) = \boldsymbol{AB} + \boldsymbol{AC}$,

(c) $(\boldsymbol{A} + \boldsymbol{B})\boldsymbol{C} = \boldsymbol{AC} + \boldsymbol{BC}$.

La transpuesta de una matriz $\boldsymbol{A} = (a_{ij}) \in \mathbb{R}^{m \times n}$ es la matriz $n \times m$, $\boldsymbol{A}^\top$ cuyo elemento ij está dado por a_{ji}, esto es

$$\boldsymbol{A}^\top = (a_{ji}).$$

Propiedades:

Para $\boldsymbol{A}$ y $\boldsymbol{B}$ matrices de órdenes apropiados. Tenemos

(a) $(\boldsymbol{A}^\top)^\top = \boldsymbol{A}$,

(b) $(\boldsymbol{A} + \boldsymbol{B})^\top = \boldsymbol{A}^\top + \boldsymbol{B}^\top$,

(c) $(\boldsymbol{AB})^\top = \boldsymbol{B}^\top \boldsymbol{A}^\top$.

Definiciones básicas y propiedades

Definimos el producto interno entre dos vectores $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ como

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \boldsymbol{x}^\top \boldsymbol{y} = \sum_{i=1}^{n} x_i y_i.$$

asociado al producto interno tenemos la norma Euclidiana (o largo) de un vector $\boldsymbol{x}$ definida como

$$\|\boldsymbol{x}\| = \langle \boldsymbol{x}, \boldsymbol{x} \rangle^{1/2} = \Big(\sum_{i=1}^{n} x_i^2 \Big)^{1/2},$$

finalmente, la distancia Euclidiana entre dos vectores $\boldsymbol{a}$ y $\boldsymbol{b}$ se define como

$$d(\boldsymbol{a}, \boldsymbol{b}) = \|\boldsymbol{a} - \boldsymbol{b}\|.$$

Propiedades:

Sean $\boldsymbol{a}$, $\boldsymbol{b}$ y $\boldsymbol{c}$ vectores n-dimensionales y λ un escalar, entonces

(a) $\langle \boldsymbol{a}, \boldsymbol{b} \rangle = \langle \boldsymbol{b}, \boldsymbol{a} \rangle$,

(b) $\langle \boldsymbol{a}, \boldsymbol{b} + \boldsymbol{c} \rangle = \langle \boldsymbol{a}, \boldsymbol{b} \rangle + \langle \boldsymbol{a}, \boldsymbol{c} \rangle$,

(c) $\lambda \langle \boldsymbol{a}, \boldsymbol{b} \rangle = \langle \lambda \boldsymbol{a}, \boldsymbol{b} \rangle = \langle \boldsymbol{a}, \lambda \boldsymbol{b} \rangle$,

(d) $\langle \boldsymbol{a}, \boldsymbol{a} \rangle \geq 0$ con la igualdad sólo si $\boldsymbol{a} = \boldsymbol{0}$,

(e) $\|\boldsymbol{a} \pm \boldsymbol{b}\|^2 = \|\boldsymbol{a}\|^2 + \|\boldsymbol{b}\|^2 \pm 2\langle \boldsymbol{a}, \boldsymbol{b} \rangle$,

(f) $\|\boldsymbol{a} + \boldsymbol{b}\| \leq \|\boldsymbol{a}\| + \|\boldsymbol{b}\|$.

Resultado (Desigualdad de Cauchy-Schwarz):

$|\langle \boldsymbol{x}, \boldsymbol{y} \rangle| \leq \|\boldsymbol{x}\| \, \|\boldsymbol{y}\|$, $\forall \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ con la igualdad sólo si $\boldsymbol{x} = \lambda \boldsymbol{y}$, para algún $\lambda \in \mathbb{R}$.

El ángulo θ entre dos vectores no nulos $\boldsymbol{x}, \boldsymbol{y}$ se define en términos de su producto interno como

$$\cos\theta = \frac{\langle \boldsymbol{x}, \boldsymbol{y}\rangle}{\|\boldsymbol{x}\|\,\|\boldsymbol{y}\|} = \frac{\boldsymbol{x}^\top \boldsymbol{y}}{\sqrt{\boldsymbol{x}^\top \boldsymbol{x}}\sqrt{\boldsymbol{y}^\top \boldsymbol{y}}},$$

dos vectores se dicen ortogonales sólo si $\boldsymbol{x}^\top \boldsymbol{y} = 0$.

Ejemplo:

Considere el vector centrado:

$$\boldsymbol{u} = \begin{pmatrix} x_1 - \overline{x} \\ x_2 - \overline{x} \\ \vdots \\ x_n - \overline{x} \end{pmatrix} = \boldsymbol{x} - \overline{x}\,\mathbf{1}_n,$$

con $\mathbf{1}_n = (1, \dots, 1)^\top$ vector de 1's de dimensión $n \times 1$.

Note que

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{n}\mathbf{1}^\top \boldsymbol{x}.$$

De este modo, podemos escribir

$$\boldsymbol{u} = \boldsymbol{x} - \overline{x}\,\mathbf{1}_n = \boldsymbol{x} - \frac{1}{n}\mathbf{1}\mathbf{1}^\top \boldsymbol{x} = \boldsymbol{C}\boldsymbol{x},$$

con $\boldsymbol{C} = \boldsymbol{I} - \frac{1}{n}\mathbf{1}\mathbf{1}^\top$ la matriz de centrado, y análogamente considere

$$\boldsymbol{v} = \boldsymbol{y} - \overline{y}\,\mathbf{1}_n = \boldsymbol{C}\boldsymbol{y}.$$

De este modo,

$$\frac{\boldsymbol{u}^\top \boldsymbol{v}}{\sqrt{\boldsymbol{u}^\top \boldsymbol{u}}\sqrt{\boldsymbol{v}^\top \boldsymbol{v}}} = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\{\sum_{i=1}^{n}(x_i - \overline{x})^2 \sum_{i=1}^{n}(y_i - \overline{y})^2\}^{1/2}},$$

es el coeficiente de correlación entre $\boldsymbol{x}$ e $\boldsymbol{y}$.

En efecto,

$$\begin{aligned}\sum_{i=1}^{n}(x_i-\overline{x})^2 = \sum_{i=1}^{n}x_i^2 - n\overline{x}^2 &= \boldsymbol{x}^\top\boldsymbol{x} - n\Big(\frac{1}{n}\boldsymbol{1}^\top\boldsymbol{x}\Big)^2 \\ &= \boldsymbol{x}^\top\boldsymbol{x} - n\Big(\frac{1}{n}\boldsymbol{x}^\top\boldsymbol{1}\Big)\Big(\frac{1}{n}\boldsymbol{1}^\top\boldsymbol{x}\Big) \\ &= \boldsymbol{x}^\top\Big(\boldsymbol{I} - \frac{1}{n}\boldsymbol{1}\boldsymbol{1}^\top\Big)\boldsymbol{x} = \boldsymbol{x}^\top\boldsymbol{C}\boldsymbol{x}.\end{aligned}$$

Análogamente,

$$\begin{aligned}\sum_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y}) = \sum_{i=1}^{n}x_iy_i - n\overline{xy} &= \boldsymbol{x}^\top\boldsymbol{y} - n\Big(\frac{1}{n}\boldsymbol{1}^\top\boldsymbol{x}\Big)\Big(\frac{1}{n}\boldsymbol{1}^\top\boldsymbol{y}\Big) \\ &= \boldsymbol{x}^\top\boldsymbol{y} - n\Big(\frac{1}{n}\boldsymbol{x}^\top\boldsymbol{1}\Big)\Big(\frac{1}{n}\boldsymbol{1}^\top\boldsymbol{y}\Big) \\ &= \boldsymbol{x}^\top\Big(\boldsymbol{I} - \frac{1}{n}\boldsymbol{1}\boldsymbol{1}^\top\Big)\boldsymbol{y} = \boldsymbol{x}^\top\boldsymbol{C}\boldsymbol{y}.\end{aligned}$$

- Una matriz se dice cuadrada si tiene el mismo número de filas que de columnas.
- Una matriz cuadrada $\boldsymbol{A}$ es triangular inferior (superior) si $a_{ij} = 0$ para $i < j$ (si $a_{ij} = 0$ para $i > j$).
- Una matriz cuadrada $\boldsymbol{A} = (a_{ij})$ se dice simétrica si $\boldsymbol{A}^\top = \boldsymbol{A}$ y sesgo-simétrica si $\boldsymbol{A}^\top = -\boldsymbol{A}$.
- Una matriz cuadrada se dice ortogonal si
$$\boldsymbol{A}\boldsymbol{A}^\top = \boldsymbol{A}^\top\boldsymbol{A} = \boldsymbol{I}$$
- Una matriz $\boldsymbol{A} \in \mathbb{R}^{n\times n}$, se dice idempotente si $\boldsymbol{A}^2 = \boldsymbol{A}$.
- Decimos que $\boldsymbol{A}$ es matriz de proyección si es simétrica e idempotente, esto es, $\boldsymbol{A}^\top = \boldsymbol{A}$ y $\boldsymbol{A}^2 = \boldsymbol{A}$.
- Cualquier matriz $\boldsymbol{B}$ satisfaciendo $\boldsymbol{B}^2 = \boldsymbol{A}$ se dice raíz cuadrada de $\boldsymbol{A}$ y se denota como $\boldsymbol{A}^{1/2}$.

Ejemplo:

Sea

$$\boldsymbol{C} = \boldsymbol{I} - \frac{1}{n}\boldsymbol{J}_n, \qquad \boldsymbol{J}_n = \boldsymbol{1}\boldsymbol{1}^\top,$$

la matriz de centrado. Tenemos que $\boldsymbol{C}^\top = \boldsymbol{C}$, y

$$\boldsymbol{C}^2 = \Big(\boldsymbol{I} - \frac{1}{n}\boldsymbol{J}_n\Big)\Big(\boldsymbol{I} - \frac{1}{n}\boldsymbol{J}_n\Big) = \boldsymbol{I} - \frac{1}{n}\boldsymbol{J}_n - \frac{1}{n}\boldsymbol{J}_n + \frac{1}{n^2}\boldsymbol{J}_n^2$$

pero $\boldsymbol{J}_n^2 = n\boldsymbol{J}_n$, luego $\boldsymbol{C}^2 = \boldsymbol{C}$ es matriz idempotente y simétrica.

Resultado:

Suponga $\boldsymbol{A}$ matriz $m \times m$, simétrica e idempotente. Entonces,

(a) $a_{ii} \geq 0$, $i = 1, \dots, n$.

(b) $a_{ii} \leq 1$, $i = 1, \dots, n$.

(c) $a_{ij} = a_{ji} = 0$, para todo $j \neq i$, si $a_{ii} = 0$ o $a_{ii} = 1$.

Demostración:

Como $\boldsymbol{A}$ es simétrica e idempotente, tenemos

$$\boldsymbol{A} = \boldsymbol{A}^2 = \boldsymbol{A}^\top \boldsymbol{A},$$

de ahí que

$$a_{ii} = \sum_{j=1}^{n} a_{ji}^2,$$

que claramente es no negativo.

Además, podemos escribir

$$a_{ii} = a_{ii}^2 + \sum_{j \neq i} a_{ji}^2.$$

Por tanto, $a_{ii} \geq a_{ii}^2$ y de este modo (b) es satisfecha. Si $a_{ii} = 0$ o bien $a_{ii} = 1$, entonces $a_{ii} = a_{ii}^2$ y debemos tener

$$\sum_{j \neq i} a_{ji}^2 = 0,$$

lo que junto con las simetría de $\boldsymbol{A}$, establece (c).

Definiciones básicas y propiedades

- Sea $\boldsymbol{a} \in \mathbb{R}^n$, $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ y $\boldsymbol{B} \in \mathbb{R}^{n \times m}$. La expresión $\boldsymbol{a}^\top \boldsymbol{x}$ se dice una forma lineal y $\boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x}$ una forma cuadrática, mientras que $\boldsymbol{x}^\top \boldsymbol{B}\boldsymbol{y}$ es una forma bilineal.

- Se asumirá que la matriz asociada a la forma cuadrática $\boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x}$ es simétrica. Note que
$$\boldsymbol{x}^\top \boldsymbol{B}\boldsymbol{x} = \tfrac{1}{2}\boldsymbol{x}^\top (\boldsymbol{A}^\top + \boldsymbol{A})\boldsymbol{x},$$
en cuyo caso tenemos que $\boldsymbol{B}$ es matriz simétrica.

- Decimos que una matriz simétrica $\boldsymbol{A}$ es definida positiva (negativa) si $\boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x} > 0$ ($\boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x} < 0$) para todo $\boldsymbol{x} \neq \boldsymbol{0}$.

- Cuando $\boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x} \geq 0$ ($\boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x} \leq 0$) $\forall \boldsymbol{x}$ decimos que $\boldsymbol{A}$ es semidefinida positiva (negativa).

- $\boldsymbol{A}$ es (semi)definida negativa sólo si $-\boldsymbol{A}$ es (semi)definida positiva.

- Las matrices $\boldsymbol{B}^\top \boldsymbol{B}$ y $\boldsymbol{B}\boldsymbol{B}^\top$ son semidefinidas positivas

- Un conjunto de vectores $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ se dice linealmente independiente si $\sum_i \alpha_i \boldsymbol{x}_i = \mathbf{0}$ implica que todos los $\alpha_i = 0$.

- Sea $\boldsymbol{A} \in \mathbb{R}^{m \times n}$, el rango columna (fila) de $\boldsymbol{A}$ es el número de columnas (filas) linealmente independientes. Denotamos el rango de $\boldsymbol{A}$ como

$$\operatorname{rg}(\boldsymbol{A}).$$

  Si $\operatorname{rg}(\boldsymbol{A}) = n$ decimos que $\boldsymbol{A}$ tiene rango columna completo.

Propiedades:

Sea $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ y $\boldsymbol{B}$, $\boldsymbol{C}$ matrices de órdenes apropiados, entonces

(a) $\operatorname{rg}(\boldsymbol{A}) = \operatorname{rg}(\boldsymbol{A}^\top) = \operatorname{rg}(\boldsymbol{A}^\top \boldsymbol{A}) = \operatorname{rg}(\boldsymbol{A}\boldsymbol{A}^\top)$,

(b) $\operatorname{rg}(\boldsymbol{A}\boldsymbol{B}) \leq \min\{\operatorname{rg}(\boldsymbol{A}), \operatorname{rg}(\boldsymbol{B})\}$,

(c) $\operatorname{rg}(\boldsymbol{B}\boldsymbol{A}\boldsymbol{C}) = \operatorname{rg}(\boldsymbol{A})$ si $\boldsymbol{B}$ y $\boldsymbol{C}$ son matrices de rango completo,

(d) si $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ y $\boldsymbol{A}\boldsymbol{x} = \mathbf{0}$ para algún $\boldsymbol{x} \neq \mathbf{0}$, entonces $\operatorname{rg}(\boldsymbol{A}) \leq n - 1$.

Ejemplo:

Sea $\boldsymbol{C} = \boldsymbol{I} - \frac{1}{n}\mathbf{1}\mathbf{1}^\top$ matriz de centrado $n \times n$. Entonces,

$$\boldsymbol{C}\mathbf{1} = \mathbf{1} - \frac{1}{n}\mathbf{1}\mathbf{1}^\top\mathbf{1} = \mathbf{1} - \mathbf{1} = \mathbf{0},$$

luego, tenemos que $\operatorname{rg}(\boldsymbol{C}) \leq n - 1$.

Definición:

Sea $\boldsymbol{A}$ una matriz cuadrada de orden $n \times n$. Decimos que $\boldsymbol{A}$ es no singular si $\operatorname{rg}(\boldsymbol{A}) = n$[1]. De este modo, para $\boldsymbol{A}$ no singular, entonces existe una matriz no singular $\boldsymbol{B}$ tal que

$$\boldsymbol{AB} = \boldsymbol{BA} = \boldsymbol{I}_n.$$

La matriz $\boldsymbol{B}$, denotada $\boldsymbol{A}^{-1}$ es única y se denomina inversa de $\boldsymbol{A}$.

Propiedades:

Siempre que todas las matrices inversas involucradas existan, tenemos que

(a) $(\boldsymbol{A}^{-1})^\top = (\boldsymbol{A}^\top)^{-1}$.

(b) $(\boldsymbol{AB})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$.

(c) $(\lambda \boldsymbol{A})^{-1} = \frac{1}{\lambda}\boldsymbol{A}^{-1}$.

(d) $\boldsymbol{P}^{-1} = \boldsymbol{P}^\top$, si $\boldsymbol{P}$ es matriz ortogonal.

(e) Si $\boldsymbol{A} > 0$, entonces $\boldsymbol{A}^{-1} > 0$.

[1] Si $\operatorname{rg}(\boldsymbol{A}) < n$, decimos que $\boldsymbol{A}$ es singular

Propiedades:

Siempre que todas las matrices inversas involucradas existan, tenemos que

(f) (Teorema de Sherman-Morrison-Woodbury)

$$(\boldsymbol{A}+\boldsymbol{BCD})^{-1}=\boldsymbol{A}^{-1}-\boldsymbol{A}^{-1}\boldsymbol{B}(\boldsymbol{C}^{-1}+\boldsymbol{DA}^{-1}\boldsymbol{B})^{-1}\boldsymbol{DA}^{-1},$$

donde $\boldsymbol{A},\boldsymbol{B},\boldsymbol{C}$ y $\boldsymbol{D}$ son matrices $m\times m$, $m\times n$, $n\times n$ y $n\times m$, respectivamente.

(g) Si $1\pm\boldsymbol{v}^\top\boldsymbol{A}^{-1}\boldsymbol{u}\neq 0$, entonces

$$(\boldsymbol{A}\pm\boldsymbol{u}\boldsymbol{v}^\top)^{-1}=\boldsymbol{A}^{-1}\mp\frac{\boldsymbol{A}^{-1}\boldsymbol{u}\boldsymbol{v}^\top\boldsymbol{A}^{-1}}{1\pm\boldsymbol{v}^\top\boldsymbol{A}^{-1}\boldsymbol{u}},$$

es conocida como la fórmula de Sherman-Morrison.

(h) $(\boldsymbol{I}+\lambda\boldsymbol{A})^{-1}=\boldsymbol{I}+\sum_{i=1}^{\infty}(-1)^i\lambda^i\boldsymbol{A}^i$.

Ejemplo:

Considere la matriz de correlación intra-clase $\boldsymbol{R}(\boldsymbol{\tau}) \in \mathbb{R}^{n\times n}$, la que también se denomina matriz de equicorrelación, definida por

$$\boldsymbol{R} = \phi \begin{pmatrix} 1 & \rho & \cdots & \rho \\ \rho & 1 & \cdots & \rho \\ \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \cdots & 1 \end{pmatrix} = \phi[(1-\rho)\boldsymbol{I} + \rho\boldsymbol{1}\boldsymbol{1}^\top], \qquad \boldsymbol{\tau} = (\phi, \rho)^\top,$$

donde $\rho \in (-1, 1)$ y $\phi > 0$. De este modo, $\boldsymbol{R}^{-1} = \phi^{-1}[(1-\rho)\boldsymbol{I} + \rho\boldsymbol{1}\boldsymbol{1}^\top]^{-1}$ y usando la Propiedad (f) con $\boldsymbol{A} = (1-\rho)\boldsymbol{I}$, $\boldsymbol{u} = \rho\boldsymbol{1}$ y $\boldsymbol{v} = \boldsymbol{1}$, tenemos que

$$\begin{aligned} \boldsymbol{R}^{-1} &= \frac{1}{\phi}\Big[\frac{1}{1-\rho}\boldsymbol{I} - \frac{\rho}{(1-\rho)^2}\frac{1}{1+n\rho(1-\rho)^{-1}}\boldsymbol{1}\boldsymbol{1}^\top\Big] \\ &= \frac{1}{\phi(1-\rho)}\Big[\boldsymbol{I} - \frac{\rho}{1+(n-1)\rho}\boldsymbol{1}\boldsymbol{1}^\top\Big] \end{aligned}$$

Ejemplo:

Considere la matriz

$$\boldsymbol{A} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix},$$

note que

$$\boldsymbol{A}\boldsymbol{A}^\top = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

y análogamente $\boldsymbol{A}^\top\boldsymbol{A} = \boldsymbol{I}$. Es decir, $\boldsymbol{A}$ es matriz ortogonal, y por tanto $\boldsymbol{A}^{-1} = \boldsymbol{A}^\top$.

Definición:

El determinante de una matriz corresponde a la función $\det : \mathbb{R}^{n\times n} \to \mathbb{R}$, denotada comúnmente como $|\boldsymbol{A}| = \det(\boldsymbol{A})$ y definida como

$$|\boldsymbol{A}| = \sum (-1)^{\sigma(j_1,\dots,j_n)} \prod_{i=1}^{n} a_{ij_i}$$

donde la sumatoria es tomada sobre todas las permutaciones $(j_1, \dots, j_n)$ del conjunto de enteros $(1, \dots, n)$, y $\sigma(j_1, \dots, j_n)$ es el número de transposiciones necesarias para cambiar $(1, \dots, n)$ en $(j_1, \dots, j_n)$

Observación:

Podemos escribir el determinante de $\boldsymbol{A}$ como:

$$|\boldsymbol{A}| = \sum_{j=1}^{n} a_{ij}c_{ij} = \sum_{j=1}^{n} a_{jk}c_{jk}, \quad \text{para } i, k = 1, \dots, n,$$

donde c_{ij} es el cofactor de a_{ij}, es decir, c_{ij} es $(-1)^{i+j}$ veces el menor de a_{ij}.

Propiedades:

Sea $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ y λ un escalar. Entonces

(a) $|\boldsymbol{A}| = |\boldsymbol{A}^\top|$.

(b) $|\boldsymbol{AB}| = |\boldsymbol{A}|\,|\boldsymbol{B}|$.

(c) $|\lambda \boldsymbol{A}| = \lambda^n |\boldsymbol{A}|$.

(d) $|\boldsymbol{A}^{-1}| = |\boldsymbol{A}|^{-1}$, si $\boldsymbol{A}$ es no singular.

(e) Si $\boldsymbol{A}$ es matriz triangular, entonces $|\boldsymbol{A}| = \prod_{i=1}^n a_{ii}$.

(f) Si $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ y $\boldsymbol{B} \in \mathbb{R}^{n \times m}$, entonces $|\boldsymbol{I}_m + \boldsymbol{AB}| = |\boldsymbol{I}_n + \boldsymbol{BA}|$.

Ejemplo:

Considere $\boldsymbol{A}$ matriz ortogonal, esto es, $\boldsymbol{A}^\top \boldsymbol{A} = \boldsymbol{A}\boldsymbol{A}^\top = \boldsymbol{I}$. Entonces

$$|\boldsymbol{A}^\top \boldsymbol{A}| = |\boldsymbol{A}\boldsymbol{A}^\top| = 1,$$

luego, $|\boldsymbol{A}|^2 = 1$ y por tanto, $|\boldsymbol{A}| = \pm 1$.

Ejemplo:

Tenemos que

$$\boldsymbol{R} = \phi[(1-\rho)\boldsymbol{I}_n + \rho \mathbf{1}\mathbf{1}^\top] = \phi(1-\rho)[\boldsymbol{I}_n + \rho(1-\rho)^{-1}\mathbf{1}\mathbf{1}^\top],$$

de este modo,

$$\begin{aligned}|\boldsymbol{R}| &= \phi^n(1-\rho)^n[1 + \rho(1-\rho)^{-1}\mathbf{1}^\top\mathbf{1}] = \phi^n(1-\rho)^{n-1}(1-\rho+n\rho) \\ &= \phi^n(1-\rho)^{n-1}[1+\rho(n-1)].\end{aligned}$$

Definición:

La traza de una matriz cuadrada $\boldsymbol{A} \in \mathbb{R}^{n \times n}$, denotada por $\operatorname{tr}(\boldsymbol{A})$, es la suma de sus elementos diagonales:

$$\operatorname{tr}(\boldsymbol{A}) = \sum_{i=1}^{n} a_{ii}.$$

Propiedades:

Siempre que las operaciones matriciales están definidas

(a) $\operatorname{tr}(\boldsymbol{A} + \boldsymbol{B}) = \operatorname{tr}(\boldsymbol{A}) + \operatorname{tr}(\boldsymbol{B})$,

(b) $\operatorname{tr}(\lambda \boldsymbol{A}) = \lambda \operatorname{tr}(\boldsymbol{A})$ si λ es un escalar,

(c) $\operatorname{tr}(\boldsymbol{A}^\top) = \operatorname{tr}(\boldsymbol{A})$,

(d) $\operatorname{tr}(\boldsymbol{AB}) = \operatorname{tr}(\boldsymbol{BA})$ (propiedad cíclica de la traza),[2]

(e) $\operatorname{tr}(\boldsymbol{A}) = 0$ si $\boldsymbol{A} = \boldsymbol{0}$.

[2] Aunque $\boldsymbol{AB}$ y $\boldsymbol{BA}$ son cuadradas, no necesitan ser del mismo orden.

Ejemplo:

Considere $\boldsymbol{C} = \boldsymbol{I} - \frac{1}{n}\boldsymbol{1}\boldsymbol{1}^\top$, entonces

$$\operatorname{tr}(\boldsymbol{C}) = \operatorname{tr}(\boldsymbol{I}) - \frac{1}{n}\operatorname{tr}(\boldsymbol{1}\boldsymbol{1}^\top) = n - \frac{1}{n}\boldsymbol{1}^\top\boldsymbol{1} = n - 1.$$

Ejemplo:

Sea $\boldsymbol{X} \in \mathbb{R}^{n\times p}$ con $\operatorname{rg}(\boldsymbol{X}) = p$ y considere $\boldsymbol{H} = \boldsymbol{X}(\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top$, luego

$$\operatorname{tr}\boldsymbol{H} = \operatorname{tr}\boldsymbol{X}(\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top = \operatorname{tr}(\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top\boldsymbol{X} = \operatorname{tr}\boldsymbol{I}_p = p,$$

note además que $\operatorname{tr}(\boldsymbol{I} - \boldsymbol{H}) = n - p$.

Ejemplo:

Considere $q = \boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x}$, tenemos que

$$q = \operatorname{tr}(\boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x}) = \operatorname{tr}(\boldsymbol{A}\boldsymbol{x}\boldsymbol{x}^\top)$$

Además, es directo que la normal vectorial (Euclidiana), satisface

$$\|\boldsymbol{x}\| = (\boldsymbol{x}^\top \boldsymbol{x})^{1/2} = (\operatorname{tr} \boldsymbol{x}\boldsymbol{x}^\top)^{1/2},$$

de este modo, podemos definir una normal matricial como:

$$\|\boldsymbol{A}\| = (\operatorname{tr} \boldsymbol{A}^\top \boldsymbol{A})^{1/2}.$$

En efecto, se tiene que $\operatorname{tr}(\boldsymbol{A}^\top \boldsymbol{A}) \geq 0$ con la igualdad sólo si $\boldsymbol{A} = \boldsymbol{0}$.

Si $\boldsymbol{A}$ y $\boldsymbol{B}$ son matrices reales del mismo orden, una matriz compleja $\boldsymbol{Z}$ puede ser definida como

$$\boldsymbol{Z} = \boldsymbol{A} + i\boldsymbol{B},$$

donde i denota la unidad imaginaria que satisface $i^2 = -1$. El conjugado complejo de $\boldsymbol{Z}$, denotado por $\boldsymbol{Z}^H$, se define como

$$\boldsymbol{Z}^H = \boldsymbol{A}^\top - i\boldsymbol{B}^\top.$$

Una matriz $\boldsymbol{Z} \in \mathbb{C}^{n\times n}$ se dice Hermitiana[3] si $\boldsymbol{Z}^H = \boldsymbol{Z}$ y unitaria[4] si $\boldsymbol{Z}^H\boldsymbol{Z} = \boldsymbol{I}$.

[3] Equivalente complejo de una matriz simétrica

[4] Equivalente complejo de una matriz ortogonal

Sea $\boldsymbol{A}$ una matriz cuadrada $n \times n$. Los valores propios de $\boldsymbol{A}$ son definidos como las raíces de la ecuación característica

$$|\lambda \boldsymbol{I} - \boldsymbol{A}| = 0,$$

la ecuación anterior tiene n raíces, en general complejas y posiblemente con algunas repeticiones (multiplicidad).

Sea λ un valor propio de $\boldsymbol{A}$, entonces existe un vector $\boldsymbol{v} \neq \boldsymbol{0} \in \mathbb{C}^n$ tal que $(\lambda \boldsymbol{I} - \boldsymbol{A})\boldsymbol{v} = \boldsymbol{0}$, esto es,

$$\boldsymbol{A}\boldsymbol{v} = \lambda \boldsymbol{v}.$$

el vector $\boldsymbol{v}$ se denomina vector propio asociado al valor propio λ.

Note que, si $\boldsymbol{v}$ es un vector propio, también lo es $\alpha \boldsymbol{v}$, $\forall \alpha \in \mathbb{C}$, y en particular $\boldsymbol{v}/\|\boldsymbol{v}\|$ es un vector propio normalizado.

Resultado:

Si $\boldsymbol{A} \in \mathbb{C}^{n \times n}$ es matriz Hermitiana, entonces todos sus valores propios son reales.

Resultado:

Si $\boldsymbol{A}$ es matriz cuadrada $n \times n$ y $\boldsymbol{G}$ es matriz no singular $n \times n$, entonces $\boldsymbol{A}$ y $\boldsymbol{G}^{-1}\boldsymbol{A}\boldsymbol{G}$ tienen el mismo conjunto de valores propios (con las mismas multiplicidades)

Resultado:

Una matriz singular tiene al menos un valor propio cero.

Resultado:

Una matriz simétrica es definida positiva (semidefinida positiva) sólo si todos sus valores propios son positivos (no-negativos).

Resultado:

Una matriz idempotente sólo tiene valores propios 0 ó 1. Todos los valores propios de una matriz unitaria tienen modulo 1

Propiedades:

Sea $\boldsymbol{A}$ matriz $n \times n$, entonces

(a) $\boldsymbol{A}^\top$ y $\boldsymbol{I} - \boldsymbol{A}$ son idempotentes sólo si $\boldsymbol{A}$ es idempotente,

(b) si $\boldsymbol{A}$ es idempotente, entonces $\operatorname{rg}(\boldsymbol{A}) = \operatorname{tr}(\boldsymbol{A}) = r$. Si $\operatorname{rg}(\boldsymbol{A}) = n$, entonces $\boldsymbol{A} = \boldsymbol{I}$.

Ejemplo:

Sabemos que la matriz de centrado $\boldsymbol{C}$ es matriz de proyección, luego

$$\operatorname{rg}(\boldsymbol{C}) = \operatorname{tr}(\boldsymbol{C}) = \operatorname{tr}(\boldsymbol{I} - \tfrac{1}{n}\boldsymbol{1}\boldsymbol{1}^\top) = n - 1.$$

Resultado:

Si $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ es matriz Hermitiana y $\boldsymbol{v}_1$, $\boldsymbol{v}_2$ son vectores propios asociados a λ_1 y λ_2, respectivamente, donde $\lambda_1 \neq \lambda_2$. Entonces $\boldsymbol{v}_1 \perp \boldsymbol{v}_2$.

Resultado (Descomposición de Schur):

Sea $\boldsymbol{A} \in \mathbb{C}^{n\times n}$. Entonces existe una matriz unitaria $\boldsymbol{U} \in \mathbb{C}^{n\times n}$ y una matriz triangular $\boldsymbol{M}$ cuyos elementos diagonales son los valores propios de $\boldsymbol{A}$, tal que

$$\boldsymbol{U}^H\boldsymbol{A}\boldsymbol{U} = \boldsymbol{M}.$$

Resultado (Descomposición espectral):

Sea $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ matriz Hermitiana. Entonces existe una matriz unitaria $\boldsymbol{U} \in \mathbb{C}^{n\times n}$ tal que

$$\boldsymbol{U}^H\boldsymbol{A}\boldsymbol{U} = \boldsymbol{\Lambda},$$

donde $\boldsymbol{\Lambda} = \operatorname{diag}(\boldsymbol{\lambda})$ es matriz diagonal cuyos elementos diagonales son los valores propios de $\boldsymbol{A}$.

Resultado:

Sea $\boldsymbol{A}$ matriz simétrica $n \times n$, con valores propios $\lambda_1, \ldots, \lambda_n$. Entonces

(a) $\operatorname{tr}(\boldsymbol{A}) = \sum_{i=1}^{n} \lambda_i$,

(b) $|\boldsymbol{A}| = \prod_{i=1}^{n} \lambda_i$.

Resultado (Descomposición de Schur):

Si $\boldsymbol{A}$ es una matriz simétrica con r valores propios distintos de cero, entonces $\operatorname{rg}(\boldsymbol{A}) = r$.

Resultado:

Sea $\boldsymbol{A}$ matriz definida positiva y $\boldsymbol{B}$ semidefinida positiva. Entonces

$$|\boldsymbol{A}+\boldsymbol{B}| \geq |\boldsymbol{A}|,$$

con la igualdad sólo si $\boldsymbol{B}=\mathbf{0}$.

Para dos matrices simétricas $\boldsymbol{A}$ y $\boldsymbol{B}$, escribimos $\boldsymbol{A} \geq \boldsymbol{B}$ si $\boldsymbol{A}-\boldsymbol{B}$ es semidefinida positiva. Análogamente, escribimos $\boldsymbol{A} > \boldsymbol{B}$ si $\boldsymbol{A}-\boldsymbol{B}$ es definida positiva.

Resultado:

Sean $\boldsymbol{A}$, $\boldsymbol{B}$ matrices definidas positivas $n \times n$. Entonces $\boldsymbol{A} > \boldsymbol{B}$ sólo si $\boldsymbol{B}^{-1} > \boldsymbol{A}^{-1}$.

Resultado:

Sean $\boldsymbol{A}$ y $\boldsymbol{B}$ matrices definidas positivas y $\boldsymbol{A}-\boldsymbol{B} \geq 0$. Entonces $|\boldsymbol{A}| \geq |\boldsymbol{B}|$ con la igualdad sólo si $\boldsymbol{A}=\boldsymbol{B}$.

Sea $\boldsymbol{A}$ una matriz $m \times n$. Considere particionar $\boldsymbol{A}$ como sigue

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} \end{pmatrix},$$

donde $\boldsymbol{A}_{11} \in \mathbb{R}^{m_1 \times n_1}$, $\boldsymbol{A}_{12} \in \mathbb{R}^{m_1 \times n_2}$, $\boldsymbol{A}_{21} \in \mathbb{R}^{m_2 \times n_1}$, $\boldsymbol{A}_{22} \in \mathbb{R}^{m_2 \times n_2}$, y $m_1 + m_2 = m$, $n_1 + n_2 = n$.

Sea $\boldsymbol{B} \in \mathbb{R}^{m \times n}$ particionada de manera análoga a $\boldsymbol{A}$, entonces

$$\boldsymbol{A} + \boldsymbol{B} = \begin{pmatrix} \boldsymbol{A}_{11} + \boldsymbol{B}_{11} & \boldsymbol{A}_{12} + \boldsymbol{B}_{12} \\ \boldsymbol{A}_{21} + \boldsymbol{B}_{21} & \boldsymbol{A}_{22} + \boldsymbol{B}_{22} \end{pmatrix}.$$

Suponga $\boldsymbol{C} \in \mathbb{R}^{n \times p}$ particionada en submatrices $\boldsymbol{C}_{ij}$, para $i, j = 1, 2$ con dimensiones adecuadas, entonces

$$\boldsymbol{A}\boldsymbol{C} = \begin{pmatrix} \boldsymbol{A}_{11}\boldsymbol{C}_{11} + \boldsymbol{A}_{12}\boldsymbol{C}_{21} & \boldsymbol{A}_{11}\boldsymbol{C}_{12} + \boldsymbol{A}_{12}\boldsymbol{C}_{22} \\ \boldsymbol{A}_{21}\boldsymbol{C}_{11} + \boldsymbol{A}_{22}\boldsymbol{C}_{21} & \boldsymbol{A}_{21}\boldsymbol{C}_{12} + \boldsymbol{A}_{22}\boldsymbol{C}_{22} \end{pmatrix}.$$

La transpuesta de $\boldsymbol{A}$ está dada por

$$\boldsymbol{A}^\top = \begin{pmatrix} \boldsymbol{A}_{11}^\top & \boldsymbol{A}_{21}^\top \\ \boldsymbol{A}_{12}^\top & \boldsymbol{A}_{22}^\top \end{pmatrix}.$$

Por otro lado,

$$|\boldsymbol{A}| = |\boldsymbol{A}_{11}||\boldsymbol{A}_{22} - \boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12}| = |\boldsymbol{A}_{22}||\boldsymbol{A}_{11} - \boldsymbol{A}_{12}\boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{21}|,$$

siempre que $\boldsymbol{A}_{11}$ y $\boldsymbol{A}_{22}$ sean matrices no singulares.

Si $\boldsymbol{A}_{12}$ y $\boldsymbol{A}_{21}$ son matrices nulas y si ambas $\boldsymbol{A}_{11}$ y $\boldsymbol{A}_{22}$ son matrices no singulares, entonces la inversa de $\boldsymbol{A}$ es

$$\boldsymbol{A}^{-1} = \begin{pmatrix} \boldsymbol{A}_{11}^{-1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{A}_{22}^{-1} \end{pmatrix}.$$

En general, si $\boldsymbol{A}$ es matriz no singular particionada y $\boldsymbol{D} = \boldsymbol{A}_{22} - \boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12}$ también es no singular, entonces

$$\boldsymbol{A}^{-1} = \begin{pmatrix} \boldsymbol{A}_{11}^{-1} + \boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12}\boldsymbol{D}^{-1}\boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1} & -\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12}\boldsymbol{D}^{-1} \\ -\boldsymbol{D}^{-1}\boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1} & \boldsymbol{D}^{-1} \end{pmatrix}.$$

Por otro lado, si $\boldsymbol{A}$ es no singular y $\boldsymbol{E} = \boldsymbol{A}_{11} - \boldsymbol{A}_{12}\boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{21}$ es no singular, entonces

$$\boldsymbol{A}^{-1} = \begin{pmatrix} \boldsymbol{E}^{-1} & -\boldsymbol{E}^{-1}\boldsymbol{A}_{12}\boldsymbol{A}_{22}^{-1} \\ -\boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{21}\boldsymbol{E}^{-1} & \boldsymbol{A}_{22}^{-1} + \boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{21}\boldsymbol{E}^{-1}\boldsymbol{A}_{12}\boldsymbol{A}_{22}^{-1} \end{pmatrix}.$$

Definición:

Sea $\boldsymbol{A} \in \mathbb{R}^{m\times n}$. La inversa Moore-Penrose (MP), $\boldsymbol{G} \in \mathbb{R}^{n\times m}$ debe satisfacer las siguientes condiciones:

$$\boldsymbol{AGA} = \boldsymbol{A}, \tag{1}$$

$$\boldsymbol{GAG} = \boldsymbol{G}, \tag{2}$$

$$(\boldsymbol{AG})^\top = \boldsymbol{AG}, \tag{3}$$

$$(\boldsymbol{GA})^\top = \boldsymbol{GA}. \tag{4}$$

La inversa MP de $\boldsymbol{A}$ se denota comunmente como $\boldsymbol{A}^+$. Si $\boldsymbol{G}$ satisface sólo la condición en (1) entonces decimos que $\boldsymbol{G}$ es una inversa generalizada y la denotamos por $\boldsymbol{A}^-$.

Resultado:

Para cada $\boldsymbol{A}$, existe una única $\boldsymbol{A}^+$.

Propiedades:

(a) $A^+ = A^{-1}$ para A matriz no singular,

(b) $(A^+)^+ = A$,

(c) $(A^\top)^+ = (A^+)^\top$,

(d) $A^+ = A$ si A es simétrica e idempotente,

(e) AA^+ y A^+A son idempotentes,

(f) $\mathrm{rg}(A) = \mathrm{rg}(A^+) = \mathrm{rg}(AA^+) = \mathrm{rg}(A^+A)$,

(g) $A^\top AA^+ = A = A^+AA^\top$,

(h) $A^\top {A^+}^\top A^+ = A^+ = A^+ {A^+}^\top A^\top$,

(i) $A^+ = (A^\top A)^+ A^\top = A^\top (AA^\top)^+$,

(j) $A^+ = (A^\top A)^{-1} A^\top$, si A tiene rango columna completo,

(k) $A^+ = A^\top (AA^\top)^{-1}$, si A tiene rango fila completo.