

MAT-266: Elementos de álgebra matricial

Felipe Osorio

`fosorios.mat.utfsm.cl`

Departamento de Matemática, UTFSM



Definiciones básicas y propiedades

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

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

de números reales.

Observación:

Siempre consideraremos \mathbf{x} como un **vector columna**.

De este modo, podemos escribir

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



Definiciones básicas y propiedades

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

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

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



Definiciones básicas y propiedades

La suma de dos matrices del mismo orden es definida como

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

el producto de una matriz por un escalar λ es

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

Propiedades:

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

- (a) $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$,
- (b) $(\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C})$,
- (c) $(\lambda + \mu)\mathbf{A} = \lambda\mathbf{A} + \mu\mathbf{A}$,
- (d) $\lambda(\mathbf{A} + \mathbf{B}) = \lambda\mathbf{A} + \lambda\mathbf{B}$.



Definiciones básicas y propiedades

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

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

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

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

Propiedades:

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

- (a) $(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$,
- (b) $\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$,
- (c) $(\mathbf{A} + \mathbf{B})\mathbf{C} = \mathbf{AC} + \mathbf{BC}$.



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

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

Propiedades:

Para A y B matrices de órdenes apropiados. Tenemos

- (a) $(A^\top)^\top = A$,
- (b) $(A + B)^\top = A^\top + B^\top$,
- (c) $(AB)^\top = B^\top A^\top$.



Definiciones básicas y propiedades

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

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

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

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

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

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



Propiedades:

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

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

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

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

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

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

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



Resultado (Desigualdad de Cauchy-Schwarz):

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

Demostración:

Si $\mathbf{x} = \lambda \mathbf{y}$, el resultado es inmediato. Sino, note que

$$0 < \|\mathbf{x} - \lambda \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \lambda^2 \|\mathbf{y}\|^2 - 2\lambda \langle \mathbf{x}, \mathbf{y} \rangle, \quad \forall \lambda \in \mathbb{R},$$

de este modo el discriminante del polinomio cuadrático debe satisfacer $4\langle \mathbf{x}, \mathbf{y} \rangle^2 - 4\|\mathbf{x}\|^2 \|\mathbf{y}\|^2 < 0$.



Definiciones básicas y propiedades

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

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

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

Ejemplo:

Considere el **vector centrado**:

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

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



Definiciones básicas y propiedades

Note que

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

De este modo, podemos escribir

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

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

$$\mathbf{v} = \mathbf{y} - \bar{y} \mathbf{1}_n = \mathbf{C} \mathbf{y}.$$

De este modo,

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

es el **coeficiente de correlación** entre \mathbf{x} e \mathbf{y} .



Definiciones básicas y propiedades

En efecto,

$$\begin{aligned}\sum_{i=1}^n (x_i - \bar{x})^2 &= \sum_{i=1}^n x_i^2 - n\bar{x}^2 = \mathbf{x}^\top \mathbf{x} - n \left(\frac{1}{n} \mathbf{1}^\top \mathbf{x} \right)^2 \\ &= \mathbf{x}^\top \mathbf{x} - n \left(\frac{1}{n} \mathbf{x}^\top \mathbf{1} \right) \left(\frac{1}{n} \mathbf{1}^\top \mathbf{x} \right) \\ &= \mathbf{x}^\top \left(\mathbf{I} - \frac{1}{n} \mathbf{1} \mathbf{1}^\top \right) \mathbf{x} = \mathbf{x}^\top \mathbf{C} \mathbf{x}.\end{aligned}$$

Análogamente,

$$\begin{aligned}\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) &= \sum_{i=1}^n x_i y_i - n\bar{x}\bar{y} = \mathbf{x}^\top \mathbf{y} - n \left(\frac{1}{n} \mathbf{1}^\top \mathbf{x} \right) \left(\frac{1}{n} \mathbf{1}^\top \mathbf{y} \right) \\ &= \mathbf{x}^\top \mathbf{y} - n \left(\frac{1}{n} \mathbf{x}^\top \mathbf{1} \right) \left(\frac{1}{n} \mathbf{1}^\top \mathbf{y} \right) \\ &= \mathbf{x}^\top \left(\mathbf{I} - \frac{1}{n} \mathbf{1} \mathbf{1}^\top \right) \mathbf{y} = \mathbf{x}^\top \mathbf{C} \mathbf{y}.\end{aligned}$$



Definiciones básicas y propiedades

- ▶ Una matriz se dice **cuadrada** si tiene el mismo número de filas que de columnas.
- ▶ Una matriz cuadrada A es **triangular inferior** (superior) si $a_{ij} = 0$ para $i < j$ (si $a_{ij} = 0$ para $i > j$).
- ▶ Una matriz cuadrada $A = (a_{ij})$ se dice **simétrica** si $A^T = A$ y **sesgo-simétrica** si $A^T = -A$.
- ▶ Una matriz cuadrada se dice **ortogonal** si

$$AA^T = A^T A = I$$

- ▶ Una matriz $A \in \mathbb{R}^{n \times n}$, se dice **idempotente** si $A^2 = A$.
- ▶ Decimos que A es **matriz de proyección** si es simétrica e idempotente, esto es, $A^T = A$ y $A^2 = A$.
- ▶ Cualquier matriz B satisfaciendo $B^2 = A$ se dice **raíz cuadrada** de A y se denota como $A^{1/2}$.



Ejemplo:

Sea

$$C = I - \frac{1}{n}J_n, \quad J_n = \mathbf{1}\mathbf{1}^\top,$$

la **matriz de centrado**. Tenemos que $C^\top = C$, y

$$C^2 = \left(I - \frac{1}{n}J_n\right)\left(I - \frac{1}{n}J_n\right) = I - \frac{1}{n}J_n - \frac{1}{n}J_n + \frac{1}{n^2}J_n^2$$

pero $J_n^2 = nJ_n$, luego $C^2 = C$ es matriz idempotente y simétrica.



Resultado:

Suponga \mathbf{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 \mathbf{A} es simétrica e idempotente, tenemos

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

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 la simetría de \mathbf{A} , establece (c).



Definiciones básicas y propiedades

- ▶ Sea $\mathbf{a} \in \mathbb{R}^n$, $\mathbf{A} \in \mathbb{R}^{n \times n}$ y $\mathbf{B} \in \mathbb{R}^{n \times m}$. La expresión $\mathbf{a}^\top \mathbf{x}$ se dice una **forma lineal** y $\mathbf{x}^\top \mathbf{A} \mathbf{x}$ una **forma cuadrática**, mientras que $\mathbf{x}^\top \mathbf{B} \mathbf{y}$ es una **forma bilineal**.
- ▶ Se asumirá que la matriz asociada a la forma cuadrática $\mathbf{x}^\top \mathbf{A} \mathbf{x}$ es **simétrica**.
Note que

$$\mathbf{x}^\top \mathbf{B} \mathbf{x} = \frac{1}{2} \mathbf{x}^\top (\mathbf{A}^\top + \mathbf{A}) \mathbf{x},$$

en cuyo caso tenemos que \mathbf{B} es matriz simétrica.

- ▶ Decimos que una matriz simétrica \mathbf{A} es **definida positiva** (negativa) si $\mathbf{x}^\top \mathbf{A} \mathbf{x} > 0$ ($\mathbf{x}^\top \mathbf{A} \mathbf{x} < 0$) para todo $\mathbf{x} \neq \mathbf{0}$.
- ▶ Cuando $\mathbf{x}^\top \mathbf{A} \mathbf{x} \geq 0$ ($\mathbf{x}^\top \mathbf{A} \mathbf{x} \leq 0$) $\forall \mathbf{x}$ decimos que \mathbf{A} es **semidefinida positiva** (negativa).
- ▶ \mathbf{A} es (semi)definida negativa sólo si $-\mathbf{A}$ es (semi)definida positiva.
- ▶ Las matrices $\mathbf{B}^\top \mathbf{B}$ y $\mathbf{B} \mathbf{B}^\top$ son semidefinidas positivas



Definiciones básicas y propiedades

- ▶ Un conjunto de vectores x_1, \dots, x_n se dice **linealmente independiente** si $\sum_i \alpha_i x_i = 0$ implica que **todos** los $\alpha_i = 0$.
- ▶ Sea $A \in \mathbb{R}^{m \times n}$, el **rango** columna (fila) de A es el número de columnas (filas) linealmente independientes. Denotamos el rango de A como

$$\text{rg}(A).$$

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

Propiedades:

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

- (a) $\text{rg}(A) = \text{rg}(A^\top) = \text{rg}(A^\top A) = \text{rg}(AA^\top)$,
- (b) $\text{rg}(AB) \leq \min\{\text{rg}(A), \text{rg}(B)\}$,
- (c) $\text{rg}(BAC) = \text{rg}(A)$ si B y C son matrices de rango completo,
- (d) si $A \in \mathbb{R}^{m \times n}$ y $Ax = 0$ para algún $x \neq 0$, entonces $\text{rg}(A) \leq n - 1$.



Ejemplo:

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

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

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



Definiciones básicas y propiedades

Definición:

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

$$AB = BA = I_n.$$

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

Propiedades:

Siempre que todas las matrices inversas involucradas existan, tenemos que

- (a) $(A^{-1})^T = (A^T)^{-1}$.
- (b) $(AB)^{-1} = B^{-1}A^{-1}$.
- (c) $(\lambda A)^{-1} = \frac{1}{\lambda}A^{-1}$.
- (d) $P^{-1} = P^T$, si P es matriz ortogonal.
- (e) Si $A > 0$, entonces $A^{-1} > 0$.

¹Si $\text{rg}(A) < n$, decimos que A es **singular**



Propiedades:

Siempre que todas las matrices inversas involucradas existan, tenemos que

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

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

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

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

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

es conocida como la [fórmula de Sherman-Morrison](#).

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



Definiciones básicas y propiedades

Ejemplo:

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

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

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

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



Ejemplo:

Considere la matriz

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

note que

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

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



Definiciones básicas y propiedades

Definición:

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

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

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 \mathbf{A} como:

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

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

(a) $|A| = |A^T|$.

(b) $|AB| = |A| |B|$.

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

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

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

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



Definiciones básicas y propiedades

Ejemplo:

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

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

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

Ejemplo:

Tenemos que

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

de este modo,

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



Definición:

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

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

Propiedades:

Siempre que las operaciones matriciales están definidas

- (a) $\text{tr}(\mathbf{A} + \mathbf{B}) = \text{tr}(\mathbf{A}) + \text{tr}(\mathbf{B})$,
- (b) $\text{tr}(\lambda \mathbf{A}) = \lambda \text{tr}(\mathbf{A})$ si λ es un escalar,
- (c) $\text{tr}(\mathbf{A}^\top) = \text{tr}(\mathbf{A})$,
- (d) $\text{tr}(\mathbf{AB}) = \text{tr}(\mathbf{BA})$ (propiedad cíclica de la traza),²
- (e) $\text{tr}(\mathbf{A}) = 0$ si $\mathbf{A} = \mathbf{0}$.

²Aunque \mathbf{AB} y \mathbf{BA} son cuadradas, **no necesitan** ser del mismo orden.



Ejemplo:

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

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

Ejemplo:

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

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

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



Ejemplo:

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

$$q = \text{tr}(\mathbf{x}^\top \mathbf{A} \mathbf{x}) = \text{tr}(\mathbf{A} \mathbf{x} \mathbf{x}^\top)$$

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

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

de este modo, podemos definir una **norma matricial** como:

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

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



Definiciones básicas y propiedades

Si A y B son matrices reales del mismo orden, una matriz compleja Z puede ser definida como

$$Z = A + iB,$$

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

$$Z^H = A^\top - iB^\top.$$

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

³Equivalente complejo de una matriz **simétrica**

⁴Equivalente complejo de una matriz **ortogonal**



Definiciones básicas y propiedades

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

$$|\lambda I - A| = 0,$$

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

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

$$Av = \lambda v.$$

el vector v se denomina **vector propio** asociado al valor propio λ .

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



Resultado:

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

Resultado:

Si A es matriz cuadrada $n \times n$ y G es matriz no singular $n \times n$, entonces A y $G^{-1}AG$ 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 A matriz $n \times n$, entonces

- (a) A^T y $I - A$ son idempotentes sólo si A es idempotente,
- (b) si A es idempotente, entonces $\text{rg}(A) = \text{tr}(A) = r$. Si $\text{rg}(A) = n$, entonces $A = I$.

Ejemplo:

Sabemos que la matriz de centrado C es matriz de proyección, luego

$$\text{rg}(C) = \text{tr}(C) = \text{tr}\left(I - \frac{1}{n} \mathbf{1}\mathbf{1}^T\right) = n - 1.$$



Resultado:

Si $A \in \mathbb{C}^{n \times n}$ es matriz Hermitiana y v_1, v_2 son vectores propios asociados a λ_1 y λ_2 , respectivamente, donde $\lambda_1 \neq \lambda_2$. Entonces $v_1 \perp v_2$.

Resultado (Descomposición de Schur):

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

$$U^H A U = M.$$

Resultado (Descomposición espectral):

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

$$U^H A U = \Lambda,$$

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



Resultado:

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

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

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

Resultado (Descomposición de Schur):

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



Resultado:

Sea A matriz definida positiva y B semidefinida positiva. Entonces

$$|A + B| \geq |A|,$$

con la igualdad sólo si $B = 0$.

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

Resultado:

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

Resultado:

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



Definiciones básicas y propiedades

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

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

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

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

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

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

$$AC = \begin{pmatrix} A_{11}C_{11} + A_{12}C_{21} & A_{11}C_{12} + A_{12}C_{22} \\ A_{21}C_{11} + A_{22}C_{21} & A_{21}C_{12} + A_{22}C_{22} \end{pmatrix}.$$



La transpuesta de \mathbf{A} está dada por

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

Por otro lado,

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

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

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

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



Definiciones básicas y propiedades

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

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

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

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



Definiciones básicas y propiedades

Definición:

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

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

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

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

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

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

Resultado:

Para cada A , existe una única 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) $\text{rg}(A) = \text{rg}(A^+) = \text{rg}(AA^+) = \text{rg}(A^+A)$,
- (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.

