Apéndice A

Elementos de Álgebra Matricial

En este Apéndice se introduce la notación, definiciones y resultados básicos de álgebra lineal y matricial, esenciales para el estudio de modelos estadísticos multivariados y de regresión lineal. El material presentado a continuación puede ser hallado en textos como Graybill (1983), Ravishanker y Dey (2002) y Magnus y Neudecker (2007).

A.1. Vectores y matrices

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. Note que $\boldsymbol{x}$ está *orientado* como un vector "columna", y por tanto la transpuesta de $\boldsymbol{x}$ es un vector fila,

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

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 $\boldsymbol{A}$.

A.2. Definiciones básicas y propiedades

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

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

Resultado A.1 (Propiedades de la suma matricial). *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) $\lambda(\boldsymbol{A} + \boldsymbol{B}) = \lambda\boldsymbol{A} + \lambda\boldsymbol{B}$,
(e) $\lambda\mu\boldsymbol{A} = (\lambda\mu)\boldsymbol{A}$.

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

$$\boldsymbol{AB} = \boldsymbol{C}, \quad \text{donde}, \quad c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj},$$

para $i = 1, \dots, m$ y $j = 1, \dots, p$.

Resultado A.2 (Propiedades del producto de matrices). *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) $(\boldsymbol{A} + \boldsymbol{B})\boldsymbol{C} = \boldsymbol{AC} + \boldsymbol{BC}$.

Note que la existencia de $\boldsymbol{AB}$ *no* implica la existencia de $\boldsymbol{BA}$ y cuando ambos productos existen, en general no son iguales.

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

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

Resultado A.3 (Propiedades de la transpuesta). *Tenemos*

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

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

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

Resultado A.4 (Propiedades del producto interno). *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* $\boldsymbol{a} = \mathbf{0}$,
(e) $\|\boldsymbol{a} \pm \boldsymbol{b}\|^2 = \|\boldsymbol{a}\|^2 + \|\boldsymbol{b}\|^2 \pm 2\langle \boldsymbol{a}, \boldsymbol{b} \rangle$,
(f) $\|\boldsymbol{a} + \boldsymbol{b}\| \leq \|\boldsymbol{a}\| + \|\boldsymbol{b}\|$.

Proposición A.5 (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* $\lambda \in \mathbb{R}$.

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

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

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

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 $\boldsymbol{x}^\top \boldsymbol{y} = 0$.

El *producto externo* entre dos vectores $\boldsymbol{x} \in \mathbb{R}^m$ y $\boldsymbol{y} \in \mathbb{R}^n$ es la matriz $m \times n$

$$\boldsymbol{x} \wedge \boldsymbol{y} = \boldsymbol{x}\boldsymbol{y}^\top = (x_i y_j).$$

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 $\boldsymbol{A}^\top = -\boldsymbol{A}$. Para cualquier matriz cuadrada $\boldsymbol{A} = (a_{ij})$ se define $\operatorname{diag}(\boldsymbol{A})$ como

$$\operatorname{diag}(\boldsymbol{A}) = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix} = \operatorname{diag}(a_{11}, a_{22}, \dots, a_{nn}).$$

Si $\boldsymbol{A} = \operatorname{diag}(\boldsymbol{A})$, decimos que $\boldsymbol{A}$ es *matriz diagonal*. Un tipo particular de matriz diagonal es la identidad

$$\boldsymbol{I} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} = (\delta_{ij}),$$

donde $\delta_{ij} = 1$ si $i = j$ y $\delta_{ij} = 0$ si $i \neq j$ (δ_{ij} se denomina *delta de Kronecker*). Tenemos que para $\boldsymbol{A} \in \mathbb{R}^{m \times n}$

$$\boldsymbol{I}_m \boldsymbol{A} = \boldsymbol{A}\boldsymbol{I}_n = \boldsymbol{A}.$$

Una matriz cuadrada se dice *ortogonal* si

$$\boldsymbol{A}\boldsymbol{A}^\top = \boldsymbol{A}^\top \boldsymbol{A} = \boldsymbol{I}$$

y sus columnas son ortonormales. Note que, si

$$\boldsymbol{A} = (\boldsymbol{a}_1, \dots, \boldsymbol{a}_n) \quad \text{con } \boldsymbol{a}_j \in \mathbb{R}^n,$$

entonces $\boldsymbol{A}$ tiene columnas *ortonormales* si

$$\boldsymbol{a}_i^\top \boldsymbol{a}_j = \begin{cases} 1, & \text{si } i = j, \\ 0, & \text{si } i \neq j, \end{cases} \qquad i, j = 1, \dots, n.$$

Una matriz rectangular $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ puede tener la propiedad $\boldsymbol{A}\boldsymbol{A}^\top = \boldsymbol{I}_m$ ó $\boldsymbol{A}^\top \boldsymbol{A} = \boldsymbol{I}_n$ pero no ambas, en cuyo caso tal matriz se denomina semi-ortogonal.

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 $\boldsymbol{A}^{1/2}$ tal matriz *no* necesita ser única.

A.2.1. Formas lineales y cuadráticas. 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* en $\boldsymbol{x}$ y $\boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x}$ una *forma cuadrática*, mientras que $\boldsymbol{x}^\top \boldsymbol{B}\boldsymbol{y}$ es una forma bilineal.

Sin pérdida de generalidad se asumirá que la matriz asociada a la forma cuadrática $\boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x}$ es simétrica. Note que siempre es posible

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

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).

Note que las matrices $\boldsymbol{B}^\top \boldsymbol{B}$ y $\boldsymbol{B}\boldsymbol{B}^\top$ son semidefinidas positivas y que $\boldsymbol{A}$ es (semi)definida negativa sólo si $-\boldsymbol{A}$ es (semi)definida positiva.

RESULTADO A.6. *Sea $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, $\boldsymbol{B} \in \mathbb{R}^{n\times p}$ y $\boldsymbol{C} \in \mathbb{R}^{n\times p}$ y $\boldsymbol{x}$ vector n-dimensional. Entonces*

(a) $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0} \Leftrightarrow \boldsymbol{A}^\top \boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$,
(b) $\boldsymbol{A}\boldsymbol{B} = \boldsymbol{0} \Leftrightarrow \boldsymbol{A}^\top \boldsymbol{A}\boldsymbol{B} = \boldsymbol{0}$,
(c) $\boldsymbol{A}^\top \boldsymbol{A}\boldsymbol{B} = \boldsymbol{A}^\top \boldsymbol{A}\boldsymbol{C} \Leftrightarrow \boldsymbol{A}\boldsymbol{B} = \boldsymbol{A}\boldsymbol{C}$.

DEMOSTRACIÓN. (a) Claramente $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0} \Rightarrow \boldsymbol{A}^\top \boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$. Por otro lado, si $\boldsymbol{A}^\top \boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$, entonces $\boldsymbol{x}^\top \boldsymbol{A}^\top \boldsymbol{A}\boldsymbol{x} = (\boldsymbol{A}\boldsymbol{x})^\top \boldsymbol{A}\boldsymbol{x} = 0$ y de ahí que $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$. (b) sigue desde (a). Finalmente, (c) sigue desde (b) mediante substituir $\boldsymbol{B} - \boldsymbol{C}$ por $\boldsymbol{B}$ en (c). □

RESULTADO A.7. *Sean $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ y $\boldsymbol{B}$, $\boldsymbol{C}$ matrices $n\times n$ con $\boldsymbol{B}$ simétrica. Entonces*

(a) $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$, $\forall \boldsymbol{x} \in \mathbb{R}^n$ *sólo si* $\boldsymbol{A} = \boldsymbol{0}$,
(b) $\boldsymbol{x}^\top \boldsymbol{B}\boldsymbol{x} = 0$, $\forall \boldsymbol{x} \in \mathbb{R}^n$ *sólo si* $\boldsymbol{B} = \boldsymbol{0}$,
(c) $\boldsymbol{x}^\top \boldsymbol{C}\boldsymbol{x} = 0$, $\forall \boldsymbol{x} \in \mathbb{R}^n$ *sólo si* $\boldsymbol{C}^\top = -\boldsymbol{C}$.

A.2.2. Rango de una matriz. Un conjunto de vectores $\boldsymbol{x}_1, \dots, \boldsymbol{x}_n$ se dice *linealmente independiente* si $\sum_i \alpha_i \boldsymbol{x}_i = \boldsymbol{0}$ implica que *todos* los $\alpha_i = 0$. Si $\boldsymbol{x}_1, \dots, \boldsymbol{x}_n$ *no* son linealmente independientes, ellos se dicen *linealmente dependientes*.

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

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

note que

$$\operatorname{rg}(\boldsymbol{A}) \leq \operatorname{mín}(m, n).$$

Si $\operatorname{rg}(\boldsymbol{A}) = n$ decimos que $\boldsymbol{A}$ tiene rango columna completo. Si $\operatorname{rg}(\boldsymbol{A}) = 0$, entonces $\boldsymbol{A}$ es la matriz nula. Por otro lado, si $\boldsymbol{A} = \boldsymbol{0}$, entonces $\operatorname{rg}(\boldsymbol{A}) = 0$.

RESULTADO A.8 (Propiedades del rango). *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) $\operatorname{rg}(\boldsymbol{AB}) \leq \min\{\operatorname{rg}(\boldsymbol{A}), \operatorname{rg}(\boldsymbol{B})\}$,
(c) $\operatorname{rg}(\boldsymbol{BAC}) = \operatorname{rg}(\boldsymbol{A})$ *si* $\boldsymbol{B}$ *y* $\boldsymbol{C}$ *son matrices de rango completo,*
(d) $\operatorname{rg}(\boldsymbol{A}+\boldsymbol{B}) \leq \operatorname{rg}(\boldsymbol{A}) + \operatorname{rg}(\boldsymbol{B})$,
(e) *si* $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ *y* $\boldsymbol{Ax} = \boldsymbol{0}$ *para algún* $\boldsymbol{x} \neq \boldsymbol{0}$, *entonces* $\operatorname{rg}(\boldsymbol{A}) \leq n-1$.

El *espacio columna* de $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, denotado por $\mathcal{M}(\boldsymbol{A})$, es el conjunto de vectores

$$\mathcal{M}(\boldsymbol{A}) = \{\boldsymbol{y} : \boldsymbol{y} = \boldsymbol{Ax} \text{ para algún } \boldsymbol{x} \in \mathbb{R}^n\}.$$

De este modo, $\mathcal{M}(\boldsymbol{A})$ es el espacio vectorial generado por las columnas de $\boldsymbol{A}$. La dimensión de este espacio es $\operatorname{rg}(\boldsymbol{A})$. Se tiene que

$$\mathcal{M}(\boldsymbol{A}) = \mathcal{M}(\boldsymbol{AA}^\top)$$

para cualquier matriz $\boldsymbol{A}$.

El *espacio nulo*, $\mathcal{N}(\boldsymbol{A})$, de una matriz $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ consiste de todos los vectores n-dimensionales $\boldsymbol{x}$, tal que $\boldsymbol{Ax} = \boldsymbol{0}$, esto es,

$$\mathcal{N}(\boldsymbol{A}) = \{\boldsymbol{x} \in \mathbb{R}^n \text{ tal que } \boldsymbol{Ax} = \boldsymbol{0}\}.$$

Note que, el espacio nulo es el conjunto de todas las soluciones de el sistema lineal homogéneo $\boldsymbol{Ax} = \boldsymbol{0}$. $\mathcal{N}(\boldsymbol{A})$ es un subespacio de $\mathbb{R}^n$ y su dimensión se denomina *nulidad* de $\boldsymbol{A}$. Además $\mathcal{N}(\boldsymbol{A}) = \{\mathcal{M}(\boldsymbol{A})\}^\perp$. Finalmente, considere la siguiente proposición

Resultado A.9. *Para cualquier matriz* $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, *entonces* $n = \dim(\mathcal{N}(\boldsymbol{A})) + \operatorname{rg}(\boldsymbol{A})$.

A.2.3. Matriz inversa. Sea $\boldsymbol{A}$ una matriz cuadrada de orden $n\times n$. Decimos que $\boldsymbol{A}$ es *no singular* si $\operatorname{rg}(\boldsymbol{A}) = n$, y que $\boldsymbol{A}$ es *singular* si $\operatorname{rg}(\boldsymbol{A}) < n$. De este modo, si $\boldsymbol{A}$ es 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 $\boldsymbol{A}$.

Resultado A.10 (Propiedades de la inversa). *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* $\boldsymbol{A}^{-1} > 0$,
(f) $(\boldsymbol{A}+\boldsymbol{BCD})^{-1} = \boldsymbol{A}^{-1} - \boldsymbol{A}^{-1}\boldsymbol{B}(\boldsymbol{C}^{-1} + \boldsymbol{DA}^{-1}\boldsymbol{B})^{-1}\boldsymbol{DA}^{-1}$, *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 (Teorema de Sherman-Morrison-Woodbury),*
(g) *si* $1 \pm \boldsymbol{v}^\top\boldsymbol{A}^{-1}\boldsymbol{u} \neq 0$, *entonces*

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

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

A.2.4. Determinante de una matriz. 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 $(j_1,\dots,j_n)$ (una transposición consiste en intercambiar dos números).

Una *submatriz* de $\boldsymbol{A}$ es un arreglo rectangular obtenido mediante eliminar filas y columnas de $\boldsymbol{A}$. Un *menor* es el determinante de una submatriz cuadrada de $\boldsymbol{A}$. El menor asociado al elemento a_{ij} es el determinante de la submatriz de $\boldsymbol{A}$ obtenida por eliminar su i-ésima fila y j-ésima columna. El *cofactor* de a_{ij}, digamos c_{ij} es $(-1)^{i+j}$ veces el menor de a_{ij}. La matriz $\boldsymbol{C} = (c_{ij})$ se denomina matriz cofactor de $\boldsymbol{A}$. La transpuesta de $\boldsymbol{C}$ es llamada *adjunta* de $\boldsymbol{A}$ y se denota $\boldsymbol{A}^{\#}$. Tenemos que

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

Resultado A.11 (Propiedades del determinante). *Sea $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ y λ un escalar. Entonces*

(a) $|\boldsymbol{A}| = |\boldsymbol{A}^\top|$,
(b) $|\boldsymbol{A}\boldsymbol{B}| = |\boldsymbol{A}|\,|\boldsymbol{B}|$,
(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) *el resultado en (e) también es válido para $\boldsymbol{A} = \operatorname{diag}(\boldsymbol{A})$, note también que* $|\boldsymbol{I}_n| = 1$,
(g) *si $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ y $\boldsymbol{B} \in \mathbb{R}^{n\times m}$, entonces* $|\boldsymbol{I}_m + \boldsymbol{A}\boldsymbol{B}| = |\boldsymbol{I}_n + \boldsymbol{B}\boldsymbol{A}|$.

A.2.5. La traza de una matriz. 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:

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

Resultado A.12 (Propiedades de la traza). *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) $\operatorname{tr}(\boldsymbol{A}\boldsymbol{B}) = \operatorname{tr}(\boldsymbol{B}\boldsymbol{A})$ *(propiedad cíclica de la traza),*
(e) $\operatorname{tr}(\boldsymbol{A}) = 0$ *si* $\boldsymbol{A} = \boldsymbol{0}$.

Note en (d) que aunque ambas $\boldsymbol{A}\boldsymbol{B}$ y $\boldsymbol{B}\boldsymbol{A}$ son cuadradas, no necesitan ser del mismo orden.
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 (Euclidiana) 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 $\boldsymbol{A} = \boldsymbol{0}$.

A.2.6. Valores y vectores propios. 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* si $\boldsymbol{Z}^H = \boldsymbol{Z}$ (equivalente complejo de una matriz simétrica) y *unitaria* si $\boldsymbol{Z}^H\boldsymbol{Z} = \boldsymbol{I}$ (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 A.13. *Si $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ es matriz Hermitiana, entonces todos sus valores propios son reales*

RESULTADO A.14. *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)*

DEMOSTRACIÓN. Note que

$$|\lambda\boldsymbol{I} - \boldsymbol{G}^{-1}\boldsymbol{A}\boldsymbol{G}| = |\lambda\boldsymbol{G}^{-1}\boldsymbol{G} - \boldsymbol{G}^{-1}\boldsymbol{A}\boldsymbol{G}| = |\boldsymbol{G}^{-1}||\lambda\boldsymbol{I} - \boldsymbol{A}||\boldsymbol{G}| = |\lambda\boldsymbol{I} - \boldsymbol{A}|$$

□

RESULTADO A.15. *Una matriz singular tiene al menos un valor propio cero*

DEMOSTRACIÓN. Si $\boldsymbol{A}$ es matriz singular, entonces $\boldsymbol{A}\boldsymbol{v} = \boldsymbol{0}$ para algún $\boldsymbol{v} \neq \boldsymbol{0}$, luego desde $\boldsymbol{A}\boldsymbol{v} = \lambda\boldsymbol{v}$, tenemos que $\lambda = 0$. □

RESULTADO A.16. *Una matriz simétrica es definida positiva (semidefinida positiva) sólo si todos sus valores propios son positivos (no-negativos).*

DEMOSTRACIÓN. Si $\boldsymbol{A}$ es definida positiva y $\boldsymbol{A}\boldsymbol{v} = \lambda\boldsymbol{v}$, entonces $\boldsymbol{v}^\top\boldsymbol{A}\boldsymbol{v} = \lambda\boldsymbol{v}^\top\boldsymbol{v}$. Ahora, como $\boldsymbol{v}^\top\boldsymbol{A}\boldsymbol{v} > 0$ y $\boldsymbol{v}^\top\boldsymbol{v} > 0$ implica $\lambda > 0$. La conversa no será probada aquí. □

Resultado A.17. *Una matriz idempotente sólo tiene valores propios 0 ó 1. Todos los valores propios de una matriz unitaria tienen modulo 1*

Demostración. Sea $\boldsymbol{A}$ matriz idempotente, esto es, $\boldsymbol{A}^2 = \boldsymbol{A}$. De este modo, si $\boldsymbol{A}\boldsymbol{v} = \lambda\boldsymbol{v}$, entonces

$$\lambda\boldsymbol{v} = \boldsymbol{A}\boldsymbol{v} = \boldsymbol{A}^2\boldsymbol{v} = \lambda\boldsymbol{A}\boldsymbol{v} = \lambda^2\boldsymbol{v}$$

y de ahí que $\lambda = \lambda^2$, esto implica que $\lambda = 0$ ó $\lambda = 1$.
Por otro lado, si $\boldsymbol{A}$ es unitaria, entonces $\boldsymbol{a}^H\boldsymbol{A} = \boldsymbol{I}$. De este modo, si $\boldsymbol{A}\boldsymbol{v} = \lambda\boldsymbol{v}$, entonces

$$\boldsymbol{v}^H\boldsymbol{A}^H = \overline{\lambda}\boldsymbol{v}^H,$$

luego

$$\boldsymbol{v}^H\boldsymbol{v} = \boldsymbol{v}^H\boldsymbol{A}^H\boldsymbol{A}\boldsymbol{v} = \overline{\lambda}\lambda\boldsymbol{v}^H\boldsymbol{v}.$$

Como $\boldsymbol{v}^H\boldsymbol{v} \neq 0$, obtenemos que $\overline{\lambda}\lambda = 1$ y de ahí que $|\lambda| = 1$. □

Resultado A.18 (Propiedades de la matrices idempotentes). *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 $\boldsymbol{A} = \boldsymbol{I}$.*

Resultado A.19. *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 $\boldsymbol{v}_1 \perp \boldsymbol{v}_2$.*

El resultado anterior muestra que si todos los valores propios de una matriz Hermitiana $\boldsymbol{A}$ son distintos, entonces existe una base ortonormal de vectores propios tal que $\boldsymbol{A}$ es diagonalizable.

Proposición A.20 (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*

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

Proposición A.21 (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 $\boldsymbol{A}$.

Para aplicaciones en Estadística siempre haremos uso de la Proposición A.21 considerando $\boldsymbol{A}$ matriz simétrica, en cuyo caso todos sus valores propios serán reales y $\boldsymbol{U}$ será una matriz ortogonal. Para $\boldsymbol{Q} \in \mathbb{R}^{n\times n}$ matriz ortogonal, denotamos el grupo de matrices ortogonales como

$$\mathcal{O}_n = \{\boldsymbol{Q} \in \mathbb{R}^{n\times n} : \boldsymbol{Q}^\top\boldsymbol{Q} = \boldsymbol{I}\}$$

Note que si $\boldsymbol{A}$ es matriz simétrica y definida positiva, entonces

$$\boldsymbol{A} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^\top = (\boldsymbol{U}\boldsymbol{\Lambda}^{1/2})(\boldsymbol{U}\boldsymbol{\Lambda}^{1/2})^\top = (\boldsymbol{U}\boldsymbol{\Lambda}^{1/2}\boldsymbol{U}^\top)^2$$

donde $\boldsymbol{\Lambda} = \operatorname{diag}(\boldsymbol{\lambda})$ y $\boldsymbol{\Lambda}^{1/2} = \operatorname{diag}(\boldsymbol{\lambda}^{1/2})$. Por tanto,

$$\boldsymbol{A} = \boldsymbol{M}\boldsymbol{M}^\top, \quad \text{con} \quad \boldsymbol{M} = \boldsymbol{U}\boldsymbol{\Lambda}^{1/2},$$

o bien,

$$\boldsymbol{A} = \boldsymbol{B}^2, \quad \text{con} \quad \boldsymbol{B} = \boldsymbol{U}\boldsymbol{\Lambda}^{1/2}\boldsymbol{U}^\top,$$

esto es, $\boldsymbol{B}$ es una matriz raíz cuadrada de $\boldsymbol{A}$.

Resultado A.22. *Sea $\boldsymbol{A}$ matriz simétrica $n \times n$, con valores propios $\lambda_1, \dots, \lambda_n$. Entonces*

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

Demostración. Usando que $\boldsymbol{A} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^\top$. Tenemos

$$\operatorname{tr}(\boldsymbol{A}) = \operatorname{tr}(\boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^\top) = \operatorname{tr}(\boldsymbol{\Lambda}\boldsymbol{U}^\top\boldsymbol{U}) = \operatorname{tr}(\boldsymbol{\Lambda}) = \sum_{i=1}^n \lambda_i$$

y

$$|\boldsymbol{A}| = |\boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^\top| = |\boldsymbol{U}||\boldsymbol{\Lambda}||\boldsymbol{U}^\top| = |\boldsymbol{\Lambda}| = \prod_{i=1}^n \lambda_i$$

□

Resultado A.23. *Si $\boldsymbol{A}$ es una matriz simétrica con r valores propios distintos de cero, entonces* $\operatorname{rg}(\boldsymbol{A}) = r$.

Demostración. Tenemos que $\boldsymbol{U}^\top\boldsymbol{A}\boldsymbol{U} = \boldsymbol{\Lambda}$ y de ahí que

$$\operatorname{rg}(\boldsymbol{A}) = \operatorname{rg}(\boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^\top) = \operatorname{rg}(\boldsymbol{\Lambda}) = r$$

□

A.2.7. Matrices (semi)definidas positivas.

Proposición A.24. *Sea $\boldsymbol{A}$ matriz definida positiva y $\boldsymbol{B}$ semidefinida positiva. Entonces*

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

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

Demostración. Tenemos $\boldsymbol{U}^\top\boldsymbol{A}\boldsymbol{U} = \boldsymbol{\Lambda}$, con $\boldsymbol{\Lambda} = \operatorname{diag}(\boldsymbol{\lambda})$ y $\boldsymbol{U}^\top\boldsymbol{U} = \boldsymbol{U}\boldsymbol{U}^\top = \boldsymbol{I}$. Luego,

$$\boldsymbol{A} + \boldsymbol{B} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^\top + \boldsymbol{B} = \boldsymbol{U}\boldsymbol{\Lambda}^{1/2}(\boldsymbol{I} + \boldsymbol{\Lambda}^{-1/2}\boldsymbol{U}^\top\boldsymbol{B}\boldsymbol{U}\boldsymbol{\Lambda}^{-1/2})\boldsymbol{\Lambda}^{1/2}\boldsymbol{U}^\top,$$

de este modo

$$\begin{aligned}
|\boldsymbol{A} + \boldsymbol{B}| &= |\boldsymbol{U}\boldsymbol{\Lambda}^{1/2}||\boldsymbol{I} + \boldsymbol{\Lambda}^{-1/2}\boldsymbol{U}^\top\boldsymbol{B}\boldsymbol{U}\boldsymbol{\Lambda}^{-1/2}||\boldsymbol{\Lambda}^{1/2}\boldsymbol{U}^\top| \\
&= |\boldsymbol{U}\boldsymbol{\Lambda}^{1/2}\boldsymbol{\Lambda}^{1/2}\boldsymbol{U}^\top||\boldsymbol{I} + \boldsymbol{\Lambda}^{-1/2}\boldsymbol{U}^\top\boldsymbol{B}\boldsymbol{U}\boldsymbol{\Lambda}^{-1/2}| \\
&= |\boldsymbol{A}||\boldsymbol{I} + \boldsymbol{\Lambda}^{-1/2}\boldsymbol{U}^\top\boldsymbol{B}\boldsymbol{U}\boldsymbol{\Lambda}^{-1/2}|.
\end{aligned}$$

Si $\boldsymbol{B} = \boldsymbol{0}$, tenemos $|\boldsymbol{A} + \boldsymbol{B}| = |\boldsymbol{A}|$. Por otro lado, si $\boldsymbol{B} \neq \boldsymbol{0}$. Entonces la matriz $\boldsymbol{I} + \boldsymbol{\Lambda}^{-1/2}\boldsymbol{U}^\top\boldsymbol{B}\boldsymbol{U}\boldsymbol{\Lambda}^{-1/2}$ tendrá al menos un valor propio no nulo y por tanto, $|\boldsymbol{I} + \boldsymbol{\Lambda}^{-1/2}\boldsymbol{U}^\top\boldsymbol{B}\boldsymbol{U}\boldsymbol{\Lambda}^{-1/2}| > 1$, esto es $|\boldsymbol{A} + \boldsymbol{B}| > |\boldsymbol{A}|$. □

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 A.25. *Sean $\boldsymbol{A}$, $\boldsymbol{B}$ matrices definidas positivas $n \times n$. Entonces $\boldsymbol{A} > \boldsymbol{B}$ sólo si $\boldsymbol{B}^{-1} > \boldsymbol{A}^{-1}$.*

PROPOSICIÓN A.26. *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* $\boldsymbol{A} = \boldsymbol{B}$.

DEMOSTRACIÓN. Sea $\boldsymbol{C} = \boldsymbol{A}-\boldsymbol{B}$. Como $\boldsymbol{B}$ es definida positiva y $\boldsymbol{C}$ es semidefinida positiva, tenemos por la Proposición A.24 que $|\boldsymbol{B}+\boldsymbol{C}| \geq |\boldsymbol{B}|$, con la igualdad sólo si $\boldsymbol{C} = \boldsymbol{0}$. □

A.2.8. Descomposiciones matriciales.

PROPOSICIÓN A.27 (Descomposición LDL). *Si* $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ *es matriz simétrica y no singular, entonces existe* $\boldsymbol{L}$ *matriz triangular inferior y* $\boldsymbol{D} = \operatorname{diag}(d_1, \dots, d_n)$, *tal que*

$$\boldsymbol{A} = \boldsymbol{L}\boldsymbol{D}\boldsymbol{L}^\top.$$

PROPOSICIÓN A.28 (Descomposición Cholesky). *Si* $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ *es simétrica y definida positiva, entonces existe una única matriz triangular inferior* $\boldsymbol{G} \in \mathbb{R}^{n\times n}$ *(factor Cholesky) con elementos diagonales positivos, tal que*

$$\boldsymbol{A} = \boldsymbol{G}\boldsymbol{G}^\top.$$

PROPOSICIÓN A.29 (Descomposición ortogonal-triangular). *Sea* $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, *entonces existe* $\boldsymbol{Q} \in \mathcal{O}_m$ *y* $\boldsymbol{R} \in \mathbb{R}^{m\times n}$, *tal que*

$$\boldsymbol{A} = \boldsymbol{Q}\boldsymbol{R},$$

donde

$$\boldsymbol{R} = \begin{pmatrix} \boldsymbol{R}_1 \\ \boldsymbol{0} \end{pmatrix}$$

con $\boldsymbol{R}_1 \in \mathbb{R}^{n\times n}$ *matriz triangular superior, aquí suponemos que* $m \geq n$. *Si* $\operatorname{rg}(\boldsymbol{A}) = r$, *entonces las primeras* n *columnas de* $\boldsymbol{Q}$ *forman una base ortonormal para* $\mathcal{M}(\boldsymbol{A})$.

Note que, si $\boldsymbol{A} = \boldsymbol{Q}\boldsymbol{R}$ entonces

$$\boldsymbol{A}^\top\boldsymbol{A} = \boldsymbol{R}^\top\boldsymbol{Q}^\top\boldsymbol{Q}\boldsymbol{R} = \boldsymbol{R}^\top\boldsymbol{R} = \boldsymbol{R}_1^\top\boldsymbol{R}_1,$$

y $\boldsymbol{R}_1$ corresponde al factor Cholesky de $\boldsymbol{A}^\top\boldsymbol{A}$.

PROPOSICIÓN A.30 (Descomposición valor singular). *Sea* $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ *con* $\operatorname{rg}(\boldsymbol{A}) = r$, *entonces existen matrices* $\boldsymbol{U} \in \mathcal{O}_m$, $\boldsymbol{V} \in \mathcal{O}_n$, *tal que*

$$\boldsymbol{A} = \boldsymbol{U}\begin{pmatrix} \boldsymbol{D}_r & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix}\boldsymbol{V}^\top,$$

donde $\boldsymbol{D}_r = \operatorname{diag}(\delta_1, \dots, \delta_r)$ *con* $\delta_i > 0$ *para* $i = 1, \dots, r$, *llamados* ***valores singulares*** *de* $\boldsymbol{A}$.

A.2.9. Matrices particionadas.

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

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} \end{pmatrix}, \tag{A.1}$$

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

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

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

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

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

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

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 como en (A.1) 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

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

Considere el determinante

$$\begin{vmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{0} & \boldsymbol{A}_{22} \end{vmatrix} = |\boldsymbol{A}_{11}||\boldsymbol{A}_{22}| = \begin{vmatrix} \boldsymbol{A}_{11} & \boldsymbol{0} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} \end{vmatrix},$$

si $\boldsymbol{A}_{11}$ y $\boldsymbol{A}_{22}$ son matrices cuadradas.

Ahora, para una matriz particionada como en (A.1) con $m_1 = n_1$ y $m_2 = n_2$, tenemos

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

si $\boldsymbol{A}_{11}$ y $\boldsymbol{A}_{22}$ son matrices no singulares.

A.3. Inversa generalizada y sistemas de ecuaciones lineales

En esta sección se generaliza el concepto de invertibilidad para matrices singulares así como para matrices rectangulares. En particular, introducimos la inversa Moore-Penrose (MP), generalización que permite resolver de forma explícita un sistema de ecuaciones lineales.

A.3.1. Inversa Moore-Penrose. Sea $\boldsymbol{A} \in \mathbb{R}^{m \times n}$, la inversa Moore-Penrose, $\boldsymbol{G} \in \mathbb{R}^{n \times m}$ debe satisfacer las siguientes condiciones

$$\boldsymbol{AGA} = \boldsymbol{A}, \tag{A.2}$$

$$\boldsymbol{GAG} = \boldsymbol{G}, \tag{A.3}$$

$$(\boldsymbol{AG})^{\top} = \boldsymbol{AG}, \tag{A.4}$$

$$(\boldsymbol{GA})^{\top} = \boldsymbol{GA}. \tag{A.5}$$

La inversa MP de $\boldsymbol{A}$ se denota comunmente como $\boldsymbol{A}^{+}$. Si $\boldsymbol{G}$ satisface sólo la condición en (A.2) entonces decimos que $\boldsymbol{G}$ es *una* inversa generalizada y la denotamos por $\boldsymbol{A}^{-}$.

Proposición A.31 (Unicidad de la inversa MP). *Para cada* $\boldsymbol{A}$*, existe una única* $\boldsymbol{A}^{+}$*.*

Resultado A.32 (Propiedades de la inversa MP).

(a) $\boldsymbol{A}^{+} = \boldsymbol{A}^{-1}$ *para* $\boldsymbol{A}$ *matriz no singular,*
(b) $(\boldsymbol{A}^{+})^{+} = \boldsymbol{A}$*,*
(c) $(\boldsymbol{A}^{\top})^{+} = (\boldsymbol{A}^{+})^{\top}$*,*
(d) $\boldsymbol{A}^{+} = \boldsymbol{A}$ *si* $\boldsymbol{A}$ *es simétrica e idempotente,*
(e) $\boldsymbol{A}\boldsymbol{A}^{+}$ *y* $\boldsymbol{A}^{+}\boldsymbol{A}$ *son idempotentes,*
(f) $\operatorname{rg}(\boldsymbol{A}) = \operatorname{rg}(\boldsymbol{A}^{+}) = \operatorname{rg}(\boldsymbol{A}\boldsymbol{A}^{+}) = \operatorname{rg}(\boldsymbol{A}^{+}\boldsymbol{A})$*,*
(g) $\boldsymbol{A}^{\top}\boldsymbol{A}\boldsymbol{A}^{+} = \boldsymbol{A} = \boldsymbol{A}^{+}\boldsymbol{A}\boldsymbol{A}^{\top}$*,*
(h) $\boldsymbol{A}^{\top}\boldsymbol{A}^{+^{\top}}\boldsymbol{A}^{+} = \boldsymbol{A}^{+} = \boldsymbol{A}^{+}\boldsymbol{A}^{+^{\top}}\boldsymbol{A}^{\top}$*,*
(i) $\boldsymbol{A}^{+} = (\boldsymbol{A}^{\top}\boldsymbol{A})^{+}\boldsymbol{A}^{\top} = \boldsymbol{A}^{\top}(\boldsymbol{A}\boldsymbol{A}^{\top})^{+}$*,*
(j) $\boldsymbol{A}^{+} = (\boldsymbol{A}^{\top}\boldsymbol{A})^{-1}\boldsymbol{A}^{\top}$*, si* $\boldsymbol{A}$ *tiene rango columna completo,*
(k) $\boldsymbol{A}^{+} = \boldsymbol{A}^{\top}(\boldsymbol{A}\boldsymbol{A}^{\top})^{-1}$*, si* $\boldsymbol{A}$ *tiene rango fila completo.*

A.3.2. Solución de sistemas de ecuaciones lineales. La solución general de un sistema de ecuaciones homegéneo $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$ es

$$\boldsymbol{x} = (\boldsymbol{I} - \boldsymbol{A}^{+}\boldsymbol{A})\boldsymbol{q},$$

con $\boldsymbol{q}$ un vector arbitrário. La solución de $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$ es única sólo si $\boldsymbol{A}$ tiene rango columna completo, esto es, $\boldsymbol{A}^{\top}\boldsymbol{A}$ es no singular. El sistema homogéneo $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$ siempre tiene al menos una solución, digamos $\boldsymbol{x} = \boldsymbol{0}$.

El sistema no homogéneo

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b},$$

tendrá al menos una solución si es *consistente*.

Proposición A.33. *Sea* $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ *y* $\boldsymbol{b}$ *vector* $m \times 1$*. Entonces son equivalentes:*

(a) *la ecuación* $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ *tiene una solución para* $\boldsymbol{x}$*,*
(b) $\boldsymbol{b} \in \mathcal{M}(\boldsymbol{A})$*,*
(c) $\operatorname{rg}(\boldsymbol{A} : \boldsymbol{b}) = \operatorname{rg}(\boldsymbol{A})$*,*
(d) $\boldsymbol{A}\boldsymbol{A}^{+}\boldsymbol{b} = \boldsymbol{b}$*.*

Proposición A.34. *Una condición necesaria y suficiente para que la ecuación* $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ *tenga una solución es que*

$$\boldsymbol{A}\boldsymbol{A}^{+}\boldsymbol{b} = \boldsymbol{b},$$

en cuyo caso la solución general está dada por

$$\boldsymbol{x} = \boldsymbol{A}^{+}\boldsymbol{b} + (\boldsymbol{I} - \boldsymbol{A}^{+}\boldsymbol{A})\boldsymbol{q},$$

donde $\boldsymbol{q}$ *es un vector arbitrário.*

Si el sistema $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ es consistente, entonces tendrá *solución única* sólo si $\boldsymbol{A}$ es de rango completo, en cuyo caso la solución está dada por $\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{b}$.

Proposición A.35. *Una condición necesaria y suficiente para que la ecuación matricial* $\boldsymbol{A}\boldsymbol{X}\boldsymbol{B} = \boldsymbol{C}$ *tenga una solución es que*

$$\boldsymbol{A}\boldsymbol{A}^{+}\boldsymbol{C}\boldsymbol{B}^{+}\boldsymbol{B} = \boldsymbol{C},$$

en cuyo caso la solución general es

$$\boldsymbol{X} = \boldsymbol{A}^{+}\boldsymbol{C}\boldsymbol{B}^{+} + \boldsymbol{Q} - \boldsymbol{A}^{+}\boldsymbol{A}\boldsymbol{Q}\boldsymbol{B}\boldsymbol{B}^{+},$$

donde $\boldsymbol{Q}$ *es una matriz arbitrária de órdenes apropiados.*