Descomposición «ortogonal» de $\mathbb{R}^m$

Sea $\mathbf{A}$ una matriz $m \times n$ de números reales. Consideramos los subespacios $\operatorname{nul} \mathbf{A}^{\mathrm{T}}$ y $\operatorname{col} \mathbf{A}$ de $\mathbb{R}^m$. Se verifica

$$\mathbb{R}^m = \operatorname{nul} \mathbf{A}^{\mathrm{T}} \oplus \operatorname{col} \mathbf{A} \,. \tag{1}$$

Obtenemos esta descomposición de $\mathbb{R}^m$ como consecuencia del siguiente

Lema. *Para todo* $\mathbf{x} \in \operatorname{nul} \mathbf{A}^{\mathrm{T}}$ *y todo* $\mathbf{y} \in \operatorname{col} \mathbf{A}$ *se verifica*

$$\mathbf{x}^{\mathrm{T}} \mathbf{y} = 0 \,. \tag{2}$$

En particular,

$$\operatorname{nul} \mathbf{A}^{\mathrm{T}} \cap \operatorname{col} \mathbf{A} = \{\mathbf{0}\} \,. \tag{3}$$

Demostración. Como es $\mathbf{y} = \mathbf{A}\mathbf{z}$ para algún $\mathbf{z} \in \mathbb{R}^n$, resulta

$$\mathbf{x}^{\mathrm{T}} \mathbf{y} = \mathbf{x}^{\mathrm{T}} \mathbf{A} \mathbf{z} = \mathbf{0}\, \mathbf{z} = 0 \,.$$

Para demostrar (3) : Si $\mathbf{y} \in \operatorname{nul} \mathbf{A}^{\mathrm{T}} \cap \operatorname{col} \mathbf{A}$ entonces

$$0 = \mathbf{y}^{\mathrm{T}} \mathbf{y} = \sum_{i=1}^{m} y_i^2 \,,$$

luego ha de ser $y_1 = y_2 = \cdots = y_m = 0$.

Demostración de (1). Sea $r = \operatorname{rango} \mathbf{A}$. Sabemos que $\dim \operatorname{nul} \mathbf{A}^{\mathrm{T}} = m - r$. Por consiguiente,

$$\dim \left(\operatorname{nul} \mathbf{A}^{\mathrm{T}} + \operatorname{col} \mathbf{A}\right) = (m-r) + r - 0 = m \,,$$

luego tenemos

$$\mathbb{R}^m = \operatorname{nul} \mathbf{A}^{\mathrm{T}} + \operatorname{col} \mathbf{A}$$

además de (3).

Definición. *Se define el* producto escalar euclideo *de dos vectores* $\mathbf{x}$ *e* $\mathbf{y}$ *de* $\mathbb{R}^m$ *mediante*

$$\langle \mathbf{x} \,, \mathbf{y} \rangle_2 = x_1\, y_1 + x_2\, y_2 + \cdots + x_m\, y_m \,.$$

Es decir, $\langle \mathbf{x}\,, \mathbf{y}\rangle_2$ es el único elemento de la matriz $\mathbf{x}^{\mathrm{T}}\mathbf{y}$, que es una matriz 1×1

$$\mathbf{x}^{\mathrm{T}}\mathbf{y} = \begin{bmatrix} x_1\,, x_2\,, \ldots\,, x_m \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix} = \begin{bmatrix} x_1\, y_1 + x_2\, y_2 + \cdots + x_m\, y_m \end{bmatrix} .$$

Identificamos las matrices $\mathbb{R}^{1\times 1}$ con los números reales $\mathbb{R}$. En consecuencia, utilizaremos $\mathbf{x}^{\mathrm{T}}\mathbf{y}$ para calcular el producto escalar euclideo de los vectores $\mathbf{x}$ e $\mathbf{y}$.

Observación. Tal y como hemos observado en la demostración de (3), se verifica

1. Para todo $\mathbf{x} \in \mathbb{R}^m$ es $\langle \mathbf{x}\,, \mathbf{x}\rangle_2 \geq 0$.
2. $\mathbf{x} \in \mathbb{R}^m$ satisface $\langle \mathbf{x}\,, \mathbf{x}\rangle_2 = 0$ si y solamente si $\mathbf{x} = \mathbf{0}$.

Definición. *Se llama* norma[1] *del vector* $\mathbf{x}$ *a*

$$\|\mathbf{x}\|_2 = \sqrt{\langle \mathbf{x}\,, \mathbf{x}\rangle_2} = \sqrt{x_1^2 + x_2^2 + \cdots + x_m^2}\,.$$

Ortogonalidad

Definición. *Decimos que los vectores* $\mathbf{x}$ *e* $\mathbf{y}$ *de* $\mathbb{R}^m$ *son* ortogonales *cuando satisfacen*

$$\langle \mathbf{x}\,, \mathbf{y}\rangle_2 = 0\,.$$

Decimos que los subespacios E *y* F *de* $\mathbb{R}^m$ *son* ortogonales *cuando para todo* $\mathbf{x} \in E$ *y todo* $\mathbf{y} \in F$ *se verifica que* $\mathbf{x}$ *e* $\mathbf{y}$ *son ortogonales.*

Observación. La propiedad (2) expresa que $\operatorname{nul} \mathbf{A}^{\mathrm{T}}$ y $\operatorname{col} \mathbf{A}$ son subespacios ortogonales en $\mathbb{R}^m$.

La descomposición (1) se llama «descomposición ortogonal» de $\mathbb{R}^m$ asociada a $\mathbf{A}$ y escribimos

$$\mathbb{R}^m = \operatorname{nul} \mathbf{A}^{\mathrm{T}} \oplus \operatorname{col} \mathbf{A} \quad \text{y} \quad \operatorname{nul} \mathbf{A}^{\mathrm{T}} \perp_2 \operatorname{col} \mathbf{A}\,. \tag{4}$$

Teorema. *Si* E *y* F *son subespacios ortogonales, entonces* $E \cap F = \{\mathbf{0}\}$.

Demostración. Si $\mathbf{x} \in E \cap F$ entonces

$$0 = \langle \mathbf{x}\,, \mathbf{x}\rangle_2 = \sum_{i=1}^{m} x_i^2$$

[1] Asociada al producto escalar euclideo.

luego ha de ser $x_1 = x_2 = \cdots = x_m = 0$.

Teorema. *Si $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k$ son vectores no nulos y ortogonales dos a dos, entonces son vectores linealmente independientes.*

Demostración. Supongamos que

$$\mathbf{0} = x_1\,\mathbf{u}_1 + x_2\,\mathbf{u}_2 + \cdots + x_k\,\mathbf{u}_k\,.$$

Para cada $j = 1, 2, \ldots, k$ hacemos el producto escalar por $\mathbf{u}_j$ y resulta

$$0 = \langle \mathbf{0}, \mathbf{u}_j \rangle_2 = \sum_{i=1}^{k} x_i \,\langle \mathbf{u}_i, \mathbf{u}_j \rangle_2 = x_j \,\langle \mathbf{u}_j, \mathbf{u}_j \rangle_2\,,$$

de donde se deduce que $x_j = 0$ ya que $\mathbf{u}_j \neq \mathbf{0}$.

Definición. *Decimos que los vectores $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m$ forman una* base ortonormal *de $\big(\mathbb{R}^m, \langle \cdot\,, \cdot \rangle_2\big)$ cuando satisfacen, para cada i, j ,*

$$\langle \mathbf{u}_i, \mathbf{u}_j \rangle_2 = \begin{cases} 0\,, & si \quad i \neq j\,, \\ 1\,, & si \quad i = j\,. \end{cases}$$

Es decir, los vectores son ortogonales dos a dos y además todos tienen norma 1 .

Observación. Si $\mathcal{B} = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m\}$ es una base ortonormal de $\big(\mathbb{R}^m, \langle \cdot\,, \cdot \rangle_2\big)$ entonces para todo vector $\mathbf{x} \in \mathbb{R}^m$ se verifica

$$\big[\mathbf{x}\big]_{\mathcal{B}} = \begin{bmatrix} \langle \mathbf{x}, \mathbf{u}_1 \rangle_2 \\ \langle \mathbf{x}, \mathbf{u}_2 \rangle_2 \\ \vdots \\ \langle \mathbf{x}, \mathbf{u}_m \rangle_2 \end{bmatrix} = \mathbf{P}^{\mathsf{T}}\mathbf{x}\,,$$

donde $\mathbf{P}$ es la matriz $\mathbf{P} = \big[\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m\big]$.

Teorema. *Sea $\mathbf{P}$ una matriz $m \times m$ de números reales. Son equivalentes:*

1. *$\mathbf{P}$ es invertible y $\mathbf{P}^{-1} = \mathbf{P}^{\mathsf{T}}$*
2. *Las columnas de $\mathbf{P}$ forman una base ortonormal de $\big(\mathbb{R}^m, \langle \cdot\,, \cdot \rangle_2\big)$.*

Demostración. Basta observar que

$$\big(\mathbf{P}^{\mathsf{T}}\mathbf{P}\big)_{i,j} = (\mathbf{P}_{:,i})^{\mathsf{T}}\mathbf{P}_{:,j} = \langle \mathbf{P}_{:,i}, \mathbf{P}_{:,j} \rangle_2\,.$$

Definición. *Decimos que la matriz $\mathbf{P} \in \mathbb{R}^{m\times m}$ es* ortogonal *cuando sus columnas forman una base ortonormal de $\mathbb{R}^m$.*

Factorización ortogonal de $\mathbf{A} \in \mathbb{R}^{m\times n}$

Sea $\mathbf{A}$ una matriz $\mathbb{R}^{m\times n}$ con $\operatorname{rango}\mathbf{A} = r$. Utilizando la descomposición ortogonal de $\left(\mathbb{R}^m, \langle\cdot\,,\cdot\rangle_2\right)$ asociada a $\mathbf{A}$ obtenemos

$$\mathbb{R}^m = \operatorname{nul}\mathbf{A}^{\mathrm{T}} \oplus \operatorname{col}\mathbf{A} \quad \text{y} \quad \operatorname{nul}\mathbf{A}^{\mathrm{T}} \perp_2 \operatorname{col}\mathbf{A} \tag{5}$$

y utilizando la descomposición ortogonal de $\left(\mathbb{R}^n, \langle\cdot\,,\cdot\rangle_2\right)$ asociada a $\mathbf{A}^{\mathrm{T}}$,

$$\mathbb{R}^n = \operatorname{nul}\mathbf{A} \oplus \operatorname{col}\mathbf{A}^{\mathrm{T}} \quad \text{y} \quad \operatorname{nul}\mathbf{A} \perp_2 \operatorname{col}\mathbf{A}^{\mathrm{T}}. \tag{6}$$

Sean $\{\mathbf{u}_1, \dots, \mathbf{u}_{m-r}\}$ una base ortonormal de $\operatorname{nul}\mathbf{A}^{\mathrm{T}}$ y $\{\mathbf{u}'_1, \dots, \mathbf{u}'_r\}$ una base ortonormal de $\operatorname{col}\mathbf{A}$. Entonces, $\mathcal{B}_m = \{\mathbf{u}_1, \dots, \mathbf{u}_{m-r}; \mathbf{u}'_1, \dots, \mathbf{u}'_r\}$ es una base ortonormal de $\mathbb{R}^m$. O igualmente, la matriz

$$\mathbf{U} = \left[\mathbf{u}_1, \dots, \mathbf{u}_{m-r}; \mathbf{u}'_1, \dots, \mathbf{u}'_r\right]$$

es una matriz $m \times m$ y ortogonal.

De igual forma, sean $\{\mathbf{v}_1, \dots, \mathbf{v}_{n-r}\}$ una base ortonormal de $\operatorname{nul}\mathbf{A}$ y $\{\mathbf{v}'_1, \dots, \mathbf{v}'_r\}$ una base ortonormal de $\operatorname{col}\mathbf{A}^{\mathrm{T}}$. Entonces, $\mathcal{B}_n = \{\mathbf{v}_1, \dots, \mathbf{v}_{n-r}; \mathbf{v}'_1, \dots, \mathbf{v}'_r\}$ es una base ortonormal de $\mathbb{R}^n$. Es decir, la matriz

$$\mathbf{V} = \left[\mathbf{v}_1, \dots, \mathbf{v}_{n-r}; \mathbf{v}'_1, \dots, \mathbf{v}'_r\right]$$

es una matriz $n \times n$ y ortogonal.

Ahora calculamos $\mathbf{U}^{\mathrm{T}}\mathbf{A}\mathbf{V}$. Las primeras $m-r$ filas de este producto son de la forma

$$\mathbf{u}_i^{\mathrm{T}}\mathbf{A}\mathbf{V} = \mathbf{0}\mathbf{V} = \mathbf{0},$$

ya que cada $\mathbf{u}_i \in \operatorname{nul}\mathbf{A}^{\mathrm{T}}$. Las primeras $n-r$ columnas son de la forma

$$\mathbf{U}^{\mathrm{T}}\mathbf{A}\mathbf{v}_j = \mathbf{U}^{\mathrm{T}}\mathbf{0} = \mathbf{0},$$

pues cada $\mathbf{v}_j \in \operatorname{nul}\mathbf{A}$. Por consiguiente,

$$\mathbf{U}^{\mathrm{T}}\mathbf{A}\mathbf{V} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{C} \end{bmatrix}$$

donde la matriz $\mathbf{C}$ es $r \times r$ y de $\operatorname{rango}\mathbf{C} = r$ pues $\mathbf{U}$ y $\mathbf{V}$ son invertibles.

En definitiva, hemos obtenido

Teorema. *Para cada $\mathbf{A}$ matriz $m \times n$ de números reales existen matrices $\mathbf{U} \in \mathbb{R}^{m\times m}$ y $\mathbf{V} \in \mathbb{R}^{n\times n}$ ortogonales y tales que*

$$\mathbf{A} = \mathbf{U}\begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{C} \end{bmatrix}\mathbf{V}^{\mathrm{T}}, \tag{7}$$

donde $\mathbf{C}$ es $r \times r$ e invertible, siendo $r = \operatorname{rango}\mathbf{A}$.

Observación. El recíproco de este teorema también es cierto. Supongamos que $\mathbf{A}$ admite una factorización como la obtenida en el teorema. De

$$\mathbf{A}\mathbf{V} = \mathbf{U}\begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{C} \end{bmatrix} = \left[\mathbf{0}, \mathbf{U}\begin{bmatrix} \mathbf{0} \\ \mathbf{C} \end{bmatrix}\right]$$

resulta que las primeras $n-r$ columnas de $\mathbf{V}$ son vectores de $\operatorname{nul}\mathbf{A}$. Como las columnas de $\mathbf{V}$ forman una base ortonormal de $\mathbb{R}^n$, que a su vez se descompone según (6), las últimas r columnas de $\mathbf{V}$ han de ser vectores de $\operatorname{col}\mathbf{A}^{\mathrm{T}}$.

De igual forma,

$$\mathbf{U}^{\mathrm{T}}\mathbf{A} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{C} \end{bmatrix} \mathbf{V}^{\mathrm{T}} = \begin{bmatrix} & \mathbf{0} \\ \begin{bmatrix} \mathbf{0} & \mathbf{C} \end{bmatrix} & \mathbf{V}^{\mathrm{T}} \end{bmatrix}$$

nos dice que las primeras $m-r$ columnas de $\mathbf{U}$ son vectores de $\operatorname{nul}\mathbf{A}^{\mathrm{T}}$. Como $\mathbf{U}$ es ortogonal, sus primeras $m-r$ columnas son vectores ortonormales y son tantos vectores como dimensión tiene $\operatorname{nul}\mathbf{A}^{\mathrm{T}}$. La descomposición (5) obliga entonces a que las últimas r columnas de $\mathbf{U}$ sean vectores de $\operatorname{col}\mathbf{A}$, cuya dimensión es r .

Factorización ortogonal de matrices cuadradas y semejanza de matrices

Consideramos ahora las matrices cuadradas $\mathbf{A} \in \mathbb{R}^{n\times n}$ que se pueden factorizar en la forma

$$\mathbf{A} = \mathbf{U} \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{C} \end{bmatrix} \mathbf{U}^{\mathrm{T}},$$

siendo $\mathbf{U}$ una matriz ortogonal. En particular, esta factorización implica que

$$\mathbf{A} \approx \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{C} \end{bmatrix}$$

ya que $\mathbf{U}^{\mathrm{T}} = \mathbf{U}^{-1}$.

He aquí una caracterización de este tipo de matrices

Teorema. *Sea* $\mathbf{A}$ *una matriz* $n \times n$ *de números reales, con* $r = \operatorname{rango}\mathbf{A}$. *Son equivalentes*:

1. *Existe una matriz ortogonal* $\mathbf{U}$ *tal que*

$$\mathbf{A}\mathbf{U} = \mathbf{U} \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{C} \end{bmatrix},$$

donde $\mathbf{C}$ *es* $r \times r$ *e invertible.*

2. $\operatorname{col}\mathbf{A} = \operatorname{col}\mathbf{A}^{\mathrm{T}}$.
3. $\operatorname{nul}\mathbf{A} = \operatorname{nul}\mathbf{A}^{\mathrm{T}}$.
4. $\operatorname{nul}\mathbf{A} \perp_2 \operatorname{col}\mathbf{A}$.

Demostración. La equivalencia de 2., 3. y 4. es consecuencia de las descomposiciones ortogonales de $\mathbb{R}^n$ asociadas a $\mathbf{A}$ y a $\mathbf{A}^{\mathrm{T}}$. Estas condiciones implican 1., pues permiten tomar $\mathbf{U} = \mathbf{V}$ en la factorización ortogonal (7) de $\mathbf{A}$.

Demostración de «1. implica 3.» De 1. se deduce

$$\mathbf{A} \sim_f \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{C} \end{bmatrix} \mathbf{U}^{\mathrm{T}},$$

que es lo mismo que

$$\operatorname{nul} \mathbf{A} = \operatorname{nul} \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{C} \end{bmatrix} \mathbf{U}^{\mathrm{T}} .$$

Escribiendo $\mathbf{U} = \left[\mathbf{U}_1 \, , \mathbf{U}_2 \right]$ tenemos

$$\operatorname{nul} \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{C} \end{bmatrix} \mathbf{U}^{\mathrm{T}} = \operatorname{nul} \mathbf{C} \, \mathbf{U}_2^{\mathrm{T}} = \operatorname{nul} \mathbf{U}_2^{\mathrm{T}}$$

pues $\mathbf{C}$ es invertible. De igual forma,

$$\mathbf{A}^{\mathrm{T}} \sim_f \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{C}^{\mathrm{T}} \end{bmatrix} \mathbf{U}^{\mathrm{T}}$$

y

$$\operatorname{nul} \mathbf{A}^{\mathrm{T}} = \operatorname{nul} \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{C}^{\mathrm{T}} \end{bmatrix} \mathbf{U}^{\mathrm{T}} = \operatorname{nul} \mathbf{C}^{\mathrm{T}} \, \mathbf{U}_2^{\mathrm{T}} = \operatorname{nul} \mathbf{U}_2^{\mathrm{T}} .$$

Corolario. *Si* $\mathbf{A}$ *es* $n \times n$ *y satisface* $\operatorname{nul} \mathbf{A} \perp_2 \operatorname{col} \mathbf{A}$, *entonces* $\text{índice}\, \mathbf{A} \leq 1$.

Definición. *Decimos que la matriz* $\mathbf{A} \in \mathbb{R}^{n \times n}$ *es* normal *cuando*

$$\mathbf{A} \, \mathbf{A}^{\mathrm{T}} = \mathbf{A}^{\mathrm{T}} \mathbf{A} .$$

Observamos que toda matriz normal cumple las condiciones del teorema anterior, pues sabemos que toda matriz satisface $\operatorname{col} \mathbf{A} = \operatorname{col} \mathbf{A}^{\mathrm{T}} \mathbf{A}$.

Sin embargo, la matriz

$$\begin{bmatrix} 1 & 2 & \\ 3 & 4 & \\ & & 0 \end{bmatrix}$$

satisface las condiciones del teorema anterior y no es normal.