

Matrices ortogonales y unitarias

Proposición - Definición

- ① Sea $U \in \mathbb{R}^{n \times n}$. Son equivalentes
- a) las columnas de U forman una base ortonormal de $\mathbb{R}^n$
 - b) U es invertible y $U^{-1} = U^t (= U^*)$
 - c) las filas de U forman una base ortonormal de $\mathbb{R}^n$

En esas condiciones, se dice que $U \in \mathbb{R}^{n \times n}$ es una matriz **ortogonal**

- ② Sea $U \in \mathbb{C}^{n \times n}$. Son equivalentes
- a) las columnas de U forman una base ortonormal de $\mathbb{C}^n$
 - b) U es invertible y $U^{-1} = \overline{U}^t (= U^*)$
 - c) las filas de U forman una base ortonormal de $\mathbb{C}^n$

En esas condiciones, se dice que $U \in \mathbb{C}^{n \times n}$ es una matriz **unitaria**

Demostación Problemas ①, ② se hace en forma análoga

(a $\Leftrightarrow$ b)

$$U^t U = \begin{bmatrix} u_1^t \\ \vdots \\ u_n^t \end{bmatrix} \begin{bmatrix} | & & | \\ u_1 & & u_n \\ | & & | \end{bmatrix} = \begin{bmatrix} u_1^t u_1 & \dots & u_1^t u_n \\ \vdots & & \vdots \\ u_n^t u_1 & \dots & u_n^t u_n \end{bmatrix} = (\langle u_i, u_j \rangle)_{1 \leq i, j \leq n}$$

Luego $U^t U = Id_n \Leftrightarrow \langle u_i, u_j \rangle = \delta_{ij} \Leftrightarrow \{u_1, \dots, u_n\}$ base o.n. de $\mathbb{R}^n$

(b $\Leftrightarrow$ c)

$$U U^t = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} \begin{bmatrix} u_1^t & \dots & u_n^t \end{bmatrix} = (\langle u_i, u_j \rangle)_{1 \leq i, j \leq n}$$

Luego $U U^t = Id_n \Leftrightarrow$ las filas de U son base o.n. de $\mathbb{R}^n$ □

Por lo tanto:

- Sea $A \in \mathbb{R}^{n \times n}$ simétrica. Entonces $\exists U \in \mathbb{R}^{n \times n}$ ortogonal t.q.
 $D = U^t A U$ es una matriz diagonal real
- Sea $A \in \mathbb{C}^{n \times n}$ hermitiana. Entonces $\exists U \in \mathbb{C}^{n \times n}$ unitaria t.q.
 $D = \overline{U}^t A U$ es una matriz diagonal real.

Observación:

- $O(n) := \{U \in \mathbb{R}^{n \times n}, U \text{ ortogonal}\}$ es un grupo con $\cdot$.
- $U(n) := \{U \in \mathbb{C}^{n \times n}, U \text{ unitaria}\}$ es un grupo con $\cdot$.
- Si $U \in U(n)$, entonces $|\det(U)| = 1$ ¿por qué?
- Si $U \in O(n)$, entonces $\det(U) = \pm 1$.

Descomposición en Valores Singulares (SVD)

(2)

(James Joseph Sylvester, 1889, para matrices reales cuadradas. Y en su versión general actual: Carl Eckard - Gale Young, 1936
Cálculo eficiente: Gene Golub y otros ~1960)

Teorema (SVD)

Sea $A \in \mathbb{C}^{m \times n}$ una matriz rectangular. Entonces existen matrices unitarias $U \in \mathbb{C}^{m \times m}$ y $V \in \mathbb{C}^{n \times n}$ y una matriz diagonal $\Sigma = \begin{bmatrix} \sigma_1 & & \\ & \sigma_r & \\ & & 0 \end{bmatrix} \in \mathbb{R}^{m \times n}$ con $\sigma_1, \dots, \sigma_r$ reales, $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0$, tales que $A = U \Sigma V^*$
 $\sigma_1, \dots, \sigma_r$ se llaman los valores singulares de A (son únicos)

$$\boxed{A} = \boxed{U} \boxed{\begin{matrix} \sigma_1 & \\ & \sigma_r \end{matrix}} \boxed{V^*}$$

Claramente $r = \text{rg}(A)$
pues $A \equiv \Sigma$

$$\boxed{A} = \boxed{U} \boxed{\begin{matrix} \sigma_1 & \\ & \sigma_r \end{matrix}} \boxed{V^*}$$

Preparación

① Sea $A \in \mathbb{C}^{m \times n}$. Entonces $A^* A \in \mathbb{C}^{n \times n}$ es autoadjunta.
pues $(A^* A)^* = A^* (A^*)^* = A^* A$ ($A^* = \overline{A}^t \in \mathbb{C}^{n \times m}$)

② Los autovalores de $A^* A$ (que son todos reales) son todos ≥ 0 :

$$A^* A x = \lambda_0 x \Rightarrow \overline{x}^t A^* A x = \overline{x}^t \lambda_0 x \Rightarrow$$

$$(\overline{A x})^t A x = \lambda_0 \|x\|^2 \Rightarrow \|\overline{A x}\|^2 = \lambda_0 \|x\|^2 \Rightarrow \lambda_0 \geq 0$$

③ Sea $A \in \mathbb{C}^{m \times n}$. Entonces $N(A^* A) = N(A)$ y $\text{rg}(A^* A) = \text{rg}(A)$

pues $N(A) \subseteq N(A^* A)$ ✓

$$\text{y por otro lado si } A^* A x = 0 \Rightarrow (A^* A x)^t \cdot \overline{x} = 0$$

$$\Rightarrow (A x)^t (A^*)^t \cdot \overline{x} = 0 \Rightarrow (A x)^t \overline{A x} = 0 \Rightarrow \|A x\|^2 = 0$$

$$\Rightarrow A x = 0. \text{ Luego } x \in N(A); N(A^* A) \subseteq N(A)$$

Demostración La vamos a hacer en el caso real para evitar todos los (3)

conjugados. Pero la demostración en el caso complejo es exactamente la misma (cuidando de considerar número producto escalar complejo).
Intuición de la demostración.

Si es cierto que $A = U \Sigma V^t$, entonces

$$A^t A = (U \Sigma V^t)^t (U \Sigma V^t) = V \Sigma^t \cancel{U^t U} \Sigma V^t = V \underbrace{\Sigma^t \Sigma}_{\text{diagonal}} V^t$$

pero $\Sigma^t \Sigma = \begin{bmatrix} \sigma_1^2 & & \\ & \sigma_r^2 & \\ & & 0 \end{bmatrix}$ diagonal $\Rightarrow \Sigma^t \Sigma$ es la matriz diagonal D

correspondiente a $A^t A$; o sea $\Sigma^t \Sigma = \begin{bmatrix} \lambda_1 & & \\ & \lambda_r & \\ & & 0 \end{bmatrix} \leftarrow \text{autovalores de } A^t A$
Hagamos entonces el camino al revés: $\Rightarrow \sigma_j^2 = \lambda_j, 1 \leq j \leq r$

Existencia:

Sea $A \in \mathbb{R}^{m \times n}$. Consideremos $A^t A \in \mathbb{R}^{n \times n}$ que es simétrica, luego

diagonalizable en una base o.n., con $\text{rg}(A^t A) = \text{rg}(A) =: r$, y con autovalores reales ≥ 0 . Al tener rango r , tiene exactamente r autovalores $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_r > 0$ (los ordeno de mayor a menor) y está el autovalor 0 con multiplicidad $n-r$.

Sea $(v_1, \dots, v_n)$ una base ortonormal de autovectores asociados a $\lambda_1, \dots, \lambda_r, 0, \dots, 0$.

$$\Rightarrow A^t A = V D V^t \text{ con } V = [v_1 | \dots | v_n] \text{ y } D = \begin{bmatrix} \lambda_1 & & \\ & \lambda_r & \\ & & 0 \end{bmatrix}$$

Defino $\sigma_1 := \sqrt{\lambda_1}, \sigma_2 := \sqrt{\lambda_2}, \dots, \sigma_r := \sqrt{\lambda_r}$.

$$\text{y } \Sigma = \begin{bmatrix} \sigma_1 & & \\ & \sigma_r & \\ & & 0 \end{bmatrix} \begin{matrix} \uparrow \\ m \\ \downarrow \end{matrix} \begin{matrix} \leftarrow n \rightarrow \end{matrix} \text{ completo con ceros}$$

Ya tenemos V . Falta definir $U \in \mathbb{R}^{m \times m}$:

Si vale que $A = U \Sigma V^t$, entonces para $1 \leq j \leq r$,

$$A v_j = U \Sigma V^t v_j = U \Sigma \begin{bmatrix} v_1^t \\ \vdots \\ v_n^t \end{bmatrix} v_j = U \Sigma e_j = U \sigma_j e_j,$$

o sea debe valer $\underbrace{(U e_j)}_{\substack{\downarrow \\ j\text{-ésima columna de } U}} = \frac{A v_j}{\sigma_j}$

o sea haciendo el camino al revés,

(4)

Defino $u_j = \frac{A v_j}{\sigma_j}$ para $1 \leq j \leq r$ (bien definido pues $\sigma_j \neq 0, 1 \leq j \leq r$)

Para que estas sean columnas válidas de U , tenemos que probar que son ortogonales entre sí y tienen norma 1:

$$\begin{aligned} \langle u_i, u_j \rangle &= u_i^t \cdot u_j = \left(\frac{A v_i}{\sigma_i} \right)^t \cdot \left(\frac{A v_j}{\sigma_j} \right) = \frac{v_i^t A^t A v_j}{\sigma_i \sigma_j} = \frac{v_i^t \lambda_j v_j}{\sigma_i \sigma_j} \\ &= \frac{\lambda_j v_i^t v_j}{\sigma_i \sigma_j} = \begin{cases} \frac{\lambda_j}{\sigma_j^2} = 1 & \text{si } i=j \\ 0 & \text{si } i \neq j \end{cases} \end{aligned}$$

luego estos vectores son efectivamente ortogonales entre sí y de norma 1, y los completa a una base ortonormal de $\mathbb{R}^m$, con $u_{r+1}, \dots, u_m$ (como quisiere)

Sea entonces $U = \begin{bmatrix} | & | & u_r & | & u_{r+1} & | & \dots & | & u_m \\ | & | & & | & & | & & | & \end{bmatrix}$

Probamos que efectivamente $A = U \Sigma V^t$:

Para ello alcanza con probar que vale la igualdad cuando se aplican estas matrices a una base de $\mathbb{R}^m$: por ejemplo la base $\{v_1, \dots, v_m\}$.

Se tiene

• Para $1 \leq j \leq r$: $A v_j = \sigma_j u_j$ por def de u_j (=)
y $U \Sigma V^t v_j = U \Sigma e_j = U \sigma_j e_j = \sigma_j U e_j = \sigma_j u_j$

• Para $r+1 \leq j \leq m$: $A v_j = 0$ pues $N(A) = N(A^t A)$
(y v_j es autovector asociado al autovalor 0 de $A^t A$)

y $U \Sigma V^t v_j = U \Sigma e_j = U \cdot 0 = 0$. Coinciden también.

• La unicidad de los valores singulares sale de que son necesariamente las raíces cuadradas de los autovalores no nulos de $A^* A$ □

Observación De la demo sale que:

$A = U \Sigma V^*$ con $\text{rg}(A) = r$, entonces

• $N(A) = \langle v_{r+1}, \dots, v_m \rangle$ (las $m-r$ últimas columnas de V)

(pues $A v_{r+1} = \dots = A v_m = 0$ y son $m-r = \dim(N(A))$)

• $\text{Im}(A) = \langle u_1, \dots, u_r \rangle$ (las r primeras columnas de U)

(pues $\langle A v_1, \dots, A v_m \rangle = \langle \underset{\neq 0}{\sigma_1 u_1}, \dots, \underset{\neq 0}{\sigma_r u_r}, \underset{0}{0}, \dots, \underset{0}{0} \rangle = \langle u_1, \dots, u_r \rangle$)

Ejemplo

Sea $A = \begin{pmatrix} 1 & -1 \\ 0 & -1 \\ -1 & 0 \end{pmatrix}$. Entonces $A^t A = \begin{pmatrix} 1 & 0 & -1 \\ 0 & -1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 0 & -1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$

$\chi_{A^t A}(\lambda) = \lambda^2 - 4\lambda + 3 = (\lambda - \underset{0}{3})(\lambda - \underset{0}{1})$ (A tiene rango 2)

$N(3\text{Id}_2 - A^t A) = N \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = \langle (1, -1) \rangle = \langle (\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}) \rangle$

$N(\text{Id}_2 - A^t A) = N \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} = \langle (1, 1) \rangle = \langle (\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}) \rangle$

$\Rightarrow V = \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix}, \sigma_1 = \sqrt{\lambda_1} = \sqrt{3}, \sigma_2 = \sqrt{\lambda_2} = \sqrt{1} = 1$

Calculamos $U \in \mathbb{R}^{3 \times 3}$:

$u_1 = \frac{A v_1}{\sigma_1} = \begin{pmatrix} \sqrt{2}/\sqrt{3} \\ \sqrt{2}/2\sqrt{3} \\ -\sqrt{2}/2\sqrt{3} \end{pmatrix} = \begin{pmatrix} \sqrt{6}/3 \\ \sqrt{6}/6 \\ -\sqrt{6}/6 \end{pmatrix} \leftarrow \text{efectivamente } \|u_1\| = 1$

$u_2 = \frac{A v_2}{\sigma_2} = \begin{pmatrix} 0 \\ -\sqrt{2}/2 \\ -\sqrt{2}/2 \end{pmatrix} \leftarrow \text{efectivamente } \|u_2\| = 1 \text{ y } u_2 \perp u_1$

Completamos una base o.n. de $\mathbb{R}^3$ (hay 2 elecciones aquí) por ejemplo

$u_3 = \begin{pmatrix} \sqrt{3}/3 \\ -\sqrt{3}/3 \\ \sqrt{3}/3 \end{pmatrix}$. Queda $U = \begin{pmatrix} \sqrt{6}/3 & 0 & \sqrt{3}/3 \\ \sqrt{6}/6 & -\sqrt{2}/2 & -\sqrt{3}/3 \\ -\sqrt{6}/6 & -\sqrt{2}/2 & \sqrt{3}/3 \end{pmatrix}$

y $A = U \Sigma V^t$.

Observación:

Para las bases $\{u_1, \dots, u_m\}$ de $\mathbb{C}^m$ y $\{v_1, \dots, v_m\}$ de $\mathbb{C}^n$ se cumple:

$\begin{cases} A v_i = \sigma_i v_i, & 1 \leq i \leq r \\ A v_i = 0, & r+1 \leq i \leq n \end{cases}$

y $\begin{cases} A^* u_i = \sigma_i v_i, & 1 \leq i \leq r \\ A^* u_i = 0, & r+1 \leq i \leq m \end{cases}$

Los vectores v_i se llaman vectores singulares a derecha (de A)

y los vectores u_i se llaman vectores singulares a izquierda (de A)

De hecho, se puede probar que $\sigma \geq 0$ (aquí admitimos el 0) es valor singular de A si y sólo si existen vectores unitarios $u \in \mathbb{C}^m$ y $v \in \mathbb{C}^n$

t.q. $A v = \sigma u$ y $A^* u = \sigma v$.

(u y v son los vectores singulares a izq. y a der. para σ , respectivamente)