Descomposicón en valores singulares de una transformación lineal entre espacios euclídeos

Eugenio Hernández, Maria-Angeles Zurro

26 de octubre de 2020

Índice

1. El teorema espectral

Fijaremos en esta sección un espacio vectorial euclídeo $\mathbf{E} = (E, \langle , \rangle)$ con E un espacio vectorial sobre el cuerpo $\mathbb{R}$ de dimensión finita $n > 0$. Consideremos una transformación lineal $A : E \to E$. Esta transformación admite una factorización según los siguientes resultados.

Teorema 1.1 (Teorema espectral). *1. Dada una transformación ortogonal O en $O(n;\mathbb{R})$, existen una transformación ortogonal $C \in O(n;\mathbb{R})$ y una transformación del tipo*

$$J = \begin{pmatrix} I_p & 0 & 0 & \cdots & 0 \\ 0 & -I_q & 0 & \cdots & 0 \\ 0 & 0 & G_{\alpha_1} & \cdots & 0 \\ 0 & 0 & \cdots & \ddots & 0 \\ 0 & 0 & 0 & \cdots & G_{\alpha_k} \end{pmatrix}, \text{ donde } G_{\alpha_i} \text{ es el giro de ángulo } \alpha_i \ , \quad (1)$$

tales que se da la factorización:

$$O = CJC^t \ . \tag{2}$$

2. *Dada una transformación autoadjunta S en E, entonces existen Una transformación ortogonal $C \in O(n;\mathbb{R})$, y una matriz diagonal $D \in \mathbb{M}_{n,n}(\mathbb{R})$ tal que*

$$S = CDC^t \ . \tag{3}$$

Demostración. Véase [1], página 397. □

Teorema 1.2. *Dada una transformación invertible $A \in GL(n;\mathbb{R})$ se tiene que:*

1. *Existen una transformación ortogonal s $O \in O(n;\mathbb{R})$, y una transformación autoadjunta S tal que se da la factorización:*

$$A = OS \ . \tag{4}$$

2. *Existen dos transformaciones ortogonales $O_1, O_2 \in O(n;\mathbb{R})$, y una transformación diagonal D tal que se da la factorización:*

$$A = O_1 D O_2 \ . \tag{5}$$

Demostración. La demostración de la fórmula (4) se encuentra en [1], página 400. Para demostrar (5), basta observar que, por (4), $A = OS$, y como, por (3), $S = CDC^t$, concluímos que:

$$A = OS = (OC)D(C^t), \tag{6}$$

luego, basta definir $O_1 = OC$ y $O_2 = C^t$ para obtener la fórmula requerida (5). □

2. Factorización ortogonal de una transformación lineal

La generalización de la fórmula (5) para matrices $m \times n$, no necesariamente cuadradas, se conoce con el nombre de **descomposición en valores singulares** de la matriz (o abreviadamente por sus siglas en inglés como *SVD decomposition*). Tiene muchas aplicaciones prácticas (véase por ejemplo `https://en.wikipedia.org/wiki/Singular_value_decomposition`), por esta razón en la sección 3 aprenderemos cómo calcularla para cualquier matriz real.

Teorema 2.1 (Descomposición en valores singulares)**.** *Dada una matriz no nula $A \in \mathbb{M}_{m.n}(\mathbb{R})$, existen matrices ortogonales $U \in O(m;\mathbb{R})$ y $V \in O(n;\mathbb{R})$ tales que*

$$A = U\Sigma V^t \ , \tag{7}$$

donde Σ es una matriz maximal diagonal, es decir,

$$\Sigma = (D \ \mathbf{0}_{n-m}) \quad si \ n \geq m, \ \grave{o} \quad \Sigma = \begin{pmatrix} D \\ \mathbf{0}_{m-n} \end{pmatrix} \quad en \ caso \ contrario \ , \tag{8}$$

para una cierta matriz diagonal D.

Definición 2.1. *Dada una transformación lineal $A : \mathbb{R}^n \to \mathbb{R}^m$, la factorización dada en el teorema 2.1 se llama la* **descomposición en valores singulares** *de A. Los elementos de la diagonal principal de la matriz diagonal D en (8) se llaman los* **valores singulares** *de A.*

A continuación procederemos a dar una demostración detallada del teorema 2.1.

Demostración. Comencemos fijando bases ortonormales, $\beta_1 = \{e_1, \ldots, e_n\}$, $\gamma_1 = \{w_1, \ldots, w_m\}$ en $\mathbb{R}^n$ y $\mathbb{R}^m$ respectivamente. Sea A la matriz de la transformación dada en estas bases, $A : \mathbb{R}^n_{\beta_1} \to \mathbb{R}^m_{\gamma_1}$.

A continuación procederemos a buscar bases apropiadas, β_2 de $\mathbb{R}^n$ y γ_2 de $\mathbb{R}^m$, tales que la transformación A tenga una matriz en estas bases de la forma (8). Pare esto seguiremos los siguientes pasos.

1. Definimos la matriz simétrica $S = A^tA \in \mathbb{M}_{n,n}(\mathbb{R})$. Por la fórmula (3) se tiene que $S = V\Delta V^t$ donde Δ es una matriz diagonal no nula , y las columnas de V, $\vec{v}_i$, forman una base ortonormal $\mathbb{R}^n$ formada por autovectores de S. Pongamos $\beta_2 = \{\vec{v}_1, \ldots, \vec{v}_n\}$. Además los autovalores no nulos de S son no negativos ya que:
$$0 \leq ||A(\vec{v}_i)||^2 = \langle \vec{v}_i, S(\vec{v}_i)\rangle = \langle \vec{v}_i, \lambda_i\vec{v}_i\rangle = \lambda_i||\vec{v}_i||^2 = \lambda_i \ . \tag{9}$$
En consecuencia $S(\vec{v}_i) = \lambda_i\vec{v}_i$, y podemos asumir la reordenación de β_2 dada por el criterio
$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r > \lambda_{r+1} = \cdots = \lambda_n = 0 \ . \tag{10}$$

2. Definimos los siguientes escalares
$$\sigma_i = \sqrt{\lambda_i}, \quad i = 1, \ldots, r, \tag{11}$$

y los siguientes vectores:

$$\vec{u}_i = \frac{1}{\sigma_i} A(\vec{v}_i), \quad i = 1, \ldots, r. \tag{12}$$

Observa que $\mathcal{C} = \{\vec{u}_1, \ldots, \vec{u}_r\}$ es un sistema libre ortonormal en $\mathbb{R}^m$ ya que

- $||\vec{u}_i||^2 = \langle \vec{u}_i, \vec{u}_i \rangle = \frac{1}{\sigma_i^2} ||A(\vec{v}_i)||^2 = \frac{\lambda_i}{\sigma_i^2} = 1.$
- $\langle \vec{u}_i, \vec{u}_j \rangle = \frac{1}{\sigma_i \sigma_j} \langle A(\vec{v}_i), A(\vec{v}_j) \rangle = \frac{1}{\sigma_i \sigma_j} \langle \vec{v}_i, S(\vec{v}_j) \rangle = \frac{\lambda_j}{\sigma_i \sigma_j} \langle \vec{v}_i, \vec{v}_j \rangle = 0.$

3. Tomamos vectores unitarios $\vec{u}_{r+1}, \ldots, \vec{u}_m$ tales que $\gamma_2 = \mathcal{C} \cup \{\vec{u}_{r+1}, \ldots, \vec{u}_m\}$ forme una base ortonormal de $\mathbb{R}^m$.

4. Definimos la matriz $m \times n$

$$\Sigma := U^t A V \ . \tag{13}$$

La matriz $\Sigma = (\Sigma_{ij})$ satisface la identidad (8) para la matriz diagonal D dada por

$$D = \begin{pmatrix} \sigma_1 & 0 & \ldots & 0 & 0 & \ldots & 0 \\ 0 & \sigma_2 & \ldots & 0 & 0 & \ldots & 0 \\ \vdots & \ldots & \ddots & \ldots & \vdots & \ldots & \vdots \\ 0 & \ldots & 0 & \sigma_r & 0 & \ldots & 0 \\ 0 & \ldots & \ldots & 0 & 0 & \ldots & 0 \\ \vdots & \ldots & \ldots & \vdots & \vdots & \ddots & \vdots \\ 0 & \ldots & \ldots & 0 & 0 & \ldots & 0 \end{pmatrix}, \tag{14}$$

y cuyo tamaño es $\nu \times \nu$ para $\nu = \text{mín}\{n, m\}$. En efecto, si $\Sigma = (\Sigma_{ij})$, entonces

$$\Sigma_{ij} = \langle \vec{u}_i, A(\vec{v}_j) \rangle = \begin{cases} \langle \vec{u}_i, \vec{0} \rangle = 0 & \text{si } j > r \\ \\ \langle \vec{u}_i, \sigma_j \vec{u}_j \rangle = \begin{cases} \sigma_i & i = j \\ 0 & i \neq j \end{cases} & \text{si } j \leq r \end{cases}, \tag{15}$$

lo que da la fórmula (8) para D la matriz diagonal definida en (14).

□

3. Ejemplos

4. Solución aproximada de sistemas lineales

https://www.youtube.com/watch?v=DG7YTlGnCEo
https://www.youtube.com/watch?v=9vJDjkx825k

Referencias

[1] Hernández, E., Vázquez, M. J., Zurro, M. A., 2012. Álgebra Lineal y Geometría. Pearson.
