DEPARTAMENTO DE COMPUTACION
Facultad de Ciencias Exactas y Naturales - UBA

Métodos Numéricos modo virtual
(pandemia COVID-19)
Material Complementario

Descomposición en valores singulares- versión 1.0

Es material complementario de las diapos de la clase de descomposición en valores singulares usadas durante el dictado virtual (pandemia-COVID-19). En este documento presentamos la demostración de la existencia de esta descomposición y algunas propiedades de los valores singulares.

Proposición: *Sea $A \in \mathbb{R}^{m\times n}$ y $r = rango(A)$ Existen $U \in \mathbb{R}^{m\times m}$ y $V \in \mathbb{R}^{n\times n}$ matrices ortogonales y $\Sigma \in \mathbb{R}^{m\times n}$ tal que*

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

con $\Sigma = \begin{bmatrix} \sigma^1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \sigma^2 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma^r & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{bmatrix}$

con $\sigma^1 \geq \sigma^2 \geq, \dots, \geq \sigma^r > 0$ *(llamados valores singulares de A).*

Demostración:
Notemos $u_1, \dots, u_m$ a las columnas de la matriz U y $v_1, \dots, v_n$ a las columnas de V. Si la descomposición existe, de la expresión $A = U\Sigma V^t$ ($A^t = V\Sigma^t U^t$), realizando el producto por columnas obtenemos que deben satisfacerse las siguientes relaciones entre los vectores columnas de las matrices U y V:

$A = U\Sigma V^t \Leftrightarrow AV = U\Sigma \Rightarrow$
$Av_i = \sigma^i u_i$ para $i = 1, \dots, r$
$Av_i = 0$ para $i = r+1, \dots, n$

$A^t = V\Sigma^t U^t \Leftrightarrow A^t U = V\Sigma \Rightarrow$
$A^t u_i = \sigma^i v_i$ para $i = 1, \dots, r$
$A^t u_i = 0$ para $i = r+1, \dots, m$
Si multiplicamos la primera relación por A^t y usamos la tercera, obtenemos que debe cumplirse que

$$A^t A v_i = \sigma^i A^t u_i = (\sigma^i)^2 v_i \text{ para } i = 1, \dots, r$$

. Si también multiplicamos la segunda por A^t, debe cumplirse que

$$A^t A v_i = 0 \text{ para } i = r+1, \dots, n$$

Deducimos entonces que los vectores v_i deberían ser autovectores de la matriz $A^t A$.
La matriz $A^t A$ es simétrica semidefinida positiva y el $rango(A^t A) = r$. Por la condición de simetría sabemos que existe una base ortonormal de autovectores. Además, por ser semidefinida positiva sus autovalores son ≥ 0 y como el rango es r, existen r autovalores no nulos y el 0 es autovalores con multiplicidad $n - r$. Sean entonces $\lambda^1, \dots, \lambda^r$ los autovalores > 0, $v_1, \dots, v_r$ los autovectores asociados y $v_{r+1}, \dots, v_n$ los autovectores asociados al autovalor 0. Sabemos que $v_1, \dots, v_n$ es base ortonormal. Estos vectores son los *candidatos* a conformar las columnas de V y definimos $\sigma^i = \sqrt{\lambda^i} > 0$.

De la relación $Av_i = \sigma^i u_i$ para $i = 1, \dots, r$ definimos $u_i = \frac{1}{\sigma^i} A v_i$. Para que esta definición sea correcta, debemos ver que $u_1, \dots, u_r$ son ortonormales:

$$u_i{}^t u_j = \frac{1}{\sigma^i}(Av_i)^t \frac{1}{\sigma^j} A v_j = \frac{1}{\sigma^i}\frac{1}{\sigma^j} v_i{}^t A^t A v_j \underset{\substack{\downarrow \\ v_j \text{ autovector de } A^t A}}{=} \frac{1}{\sigma^i}\frac{1}{\sigma^j} v_i{}^t \lambda^j v_j = \frac{1}{\sigma^i}\frac{1}{\sigma^j} \lambda^j v_i{}^t v_j \underset{\substack{\downarrow \\ \text{los } v_i \text{ son ortonormales}}}{=} 0$$

$$u_i{}^t u_i = \frac{1}{\sigma^i}(Av_i)^t \frac{1}{\sigma^i} A v_i = \frac{1}{\sigma^i}\frac{1}{\sigma^i} v_i{}^t A^t A v_i \underset{\substack{\downarrow \\ v_i \text{ autovector de } A^t A}}{=} \frac{1}{\sigma^i}\frac{1}{\sigma^i} v_i{}^t \lambda^i v_i = \frac{1}{\sigma^i}\frac{1}{\sigma^i} \lambda^i v_i{}^t v_i \underset{\substack{\downarrow \\ \text{los } v_i \text{ son ortonormales}}}{=} \frac{1}{\sigma^i}\frac{1}{\sigma^i} \lambda^i \overset{\substack{\text{por def de } \sigma^i \\ \uparrow}}{=} 1$$

Como $dim(Im(A)) = r$ y los u_i ($u_i = \frac{1}{\sigma^i} A v_i$ para $i = 1, \dots, r$) pertenecen a $Im(A)$, entonces conforman una base ortonormal de $Im(A)$. Nos falta aún definir el resto de los u_i $i = r+1, \dots, m$.
Sabemos de un resultado de álgebra lineal que $Im(A) \oplus Nu(A^t) = \mathbb{R}^m$ y $Nu(A^t) = Im(A)^{\perp}$. Entonces toda base ortonormal de $Im(A)$ se puede extender a una base ortonormal de todo el espacio con vectores ortonormales que pertenecen a $Nu(A^t)$. Sean $u_{r+1}, \dots, u_m$ dicha extensión.
En definitiva, tenemos hasta el momento dos bases ortonormales: $u_1, \dots, u_m$ y $v_1, \dots, v_n$ *candidatas* a conformar las columnas de U y V y valores $\sigma^1, \dots, \sigma^r$ para definir a Σ. Verifiquemos que son una descomposición de la matriz A.

- $Av_i = \sigma^i u_i$ para $i = 1, \dots, r$. Se cumple por la definición de los u_i.
- $Av_i = 0$ para $i = r+1, \dots, n$. Como v_i es autovector de $A^t A$ del autovalor 0, entonces $A^t A v_i = 0 \Rightarrow v_i^t A^t A v_i = 0 \Rightarrow ||A^t v_i||_2{}^2 = 0 \Rightarrow A^t v_i = 0$.
- $A^t u_i = \sigma^i v_i$ para $i = 1, \dots, r$. Se cumple por la definición de los u_i.
- $A^t u_i = 0$ para $i = r+1, \dots, m$. Se cumple porque los $u_i \in Nu(A^t)$.

Observar que un análisis similar se puede hacer para los vectores u_i y concluir que son los autovectores de AA^t.

En conclusión $A = U\Sigma V^t$ donde las columnas de U son una base ortonormal de autovectores de AA^t, las columnas de V son una base ortonormal de autovectores de $A^t A$ y σ^i son las raíces cuadradas de los autovalores de $A^t A$ (AA^t). ■

Veamos ahora algunas propiedades de los valores singulares.

- $||\mathbf{A}||_\mathbf{2} = \sigma^\mathbf{1}$

$$||A||_2 = \max_{||x||_2=1} ||Ax||_2 = \max_{||x||_2=1} ||U\Sigma V^t x||_2 \underset{\substack{\downarrow \\ U \text{ ortogonal}}}{=} \max_{||x||_2=1} ||\Sigma V^t x||_2 \underset{\substack{\downarrow \\ V \text{ ortogonal}}}{=} \max_{||x||_2=1} ||\Sigma x||_2 = \max_{||x||_2=1} \sqrt{(\sigma^1 x_1)^2 + \ldots + (\sigma^r x_r)^2}$$

$$||A||_2 = \max_{||x||_2=1} \sqrt{(\sigma^1 x_1)^2 + \ldots + (\sigma^r x_r)^2} \leq \max_{||x||_2=1} \sigma^1 \sqrt{(x_1)^2 + \ldots + (x_r)^2} \leq \sigma^1$$

Además

$$||A||_2 = \max_{||x||_2=1} ||Ax||_2 \geq ||Av_1||_2 = \sqrt{v_1^t A^t A v_1} \underset{\substack{\downarrow \\ v_1 \text{ autov } A^t A}}{=} \sqrt{v_1^t (\sigma^1)^2 v_1} \underset{\substack{\downarrow \\ ||v_1||_2 = 1}}{=} \sigma^1$$

Concluimos que $||A||_2 = \sigma^1$.

- $||\mathbf{A}||_{\mathbf{F}} = \sqrt{(\sigma^{\mathbf{1}})^{\mathbf{2}} + \ldots, (\sigma^{\mathbf{r}})^{\mathbf{2}}}$

  $$||A||_F = ||U\Sigma V^t||_F \underset{\substack{\downarrow \\ U \text{ ortog}}}{=} ||\Sigma V^t||_F \underset{\substack{\downarrow \\ V \text{ ortog}}}{=} ||\Sigma||_F = \sqrt{(\sigma^1)^2 + \ldots, (\sigma^r)^2}$$

- $\mathbf{A} \in \mathbb{R}^{\mathbf{n}\times\mathbf{n}}$ inversible entonces $\kappa(\mathbf{A})_{\mathbf{2}} = \frac{\sigma^{\mathbf{1}}}{\sigma^{\mathbf{n}}}$

  Como A es inversible, su rango es n y tiene n valores singulares no nulos. Además, los valores singulares de A^{-1} son $\frac{1}{\sigma^n}, \ldots, \frac{1}{\sigma^1}$. Por definición $\kappa(A)_2 = ||A||_2||A^{-1}||_2$ que por la propiedad anterior $\kappa(A) = \frac{\sigma^1}{\sigma^n}$