AULA 6

Sumário. Valores singulares: definição e caracterização. Decomposição SVD.

Teorema 6.1 (Decomposição dos valores singulares (SVD)). *Seja* $\mathbf{A} \in \mathbb{C}^{m\times n}$ *uma matriz arbitrária e seja* $r = r(\mathbf{A})$. *Então, existem matrizes unitárias* $\mathbf{U} \in \mathbb{C}^{m\times m}$ *e* $\mathbf{V} \in \mathbb{C}^{n\times n}$ *tais que*

$$\mathbf{UAV} = \begin{bmatrix} \mathbf{D} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \qquad \mathbf{D} = \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \sigma_r \end{bmatrix} \in \mathbb{R}^{r\times r},$$

onde $\sigma_1, \dots, \sigma_r \in \mathbb{R}$ *são tais que* $\sigma_1 \geqslant \sigma_2 \geqslant \cdots \geqslant \sigma_r > 0$.

Demonstração. O resultado é óbvio quando $\mathbf{A} = \mathbf{0}$, de modo que podemos admitir que $\mathbf{A} \neq \mathbf{0}$.

Em primeiro lugar, suponhamos que $m = 1$ (isto é, $\mathbf{A} \in \mathbb{C}^{1\times n}$ é uma matriz linha com comprimento n). Seja $\mathbf{v} = \mathbf{A}^* \in \mathbb{C}^{n\times 1}$ e seja $\sigma_1 = \|\mathbf{v}\| \in \mathbb{R}^+$ (porque $\mathbf{v} \neq \mathbf{0}$); assim,

$$\sigma_1^2 = \|\mathbf{v}\|^2 = \langle \mathbf{v}|\mathbf{v}\rangle = \mathbf{v}^*\mathbf{v} = (\mathbf{A}^*)^*\mathbf{A}^* = \mathbf{AA}^*.$$

Seja $\mathbf{V} \in \mathbb{C}^{n\times n}$ uma matriz unitária cuja primeira coluna é $\mathbf{v}_1 = \sigma_1^{-1}\mathbf{v} = \frac{1}{\|\mathbf{v}\|}\mathbf{v}$ e sejam $\mathbf{v}_2, \dots, \mathbf{v}_n \in \mathbb{C}^{n\times 1}$ as restantes colunas de $\mathbf{V}$. Então,

$$\mathbf{AV} = \begin{bmatrix} \mathbf{Av}_1 & \mathbf{Av}_2 & \cdots & \mathbf{Av}_n \end{bmatrix} = \begin{bmatrix} \sigma_1^{-1}\mathbf{v}^*\mathbf{v} & \mathbf{v}^*\mathbf{v}_2 & \cdots & \mathbf{v}^*\mathbf{v}_n \end{bmatrix}.$$

Como $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ é uma base ortonormada de $\mathbb{C}^{n\times 1}$, temos

$$\mathbf{v}^*\mathbf{v}_i = \langle \mathbf{v}|\mathbf{v}_i\rangle = 0, \qquad 2 \leqslant i \leqslant n,$$

pelo que

$$\mathbf{AV} = \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \end{bmatrix} = \begin{bmatrix} \mathbf{D} & \mathbf{0} \end{bmatrix}$$

onde $\mathbf{D} = \begin{bmatrix} \sigma_1 \end{bmatrix} \in \mathbb{R}^{1\times 1}$.

Sendo assim, podemos supor que $m \geqslant 2$. Procedemos por indução sobre n (isto é, sobre o número de colunas de $\mathbf{A}$). Quando $n = 1$, o argumento anterior, aplicado à matriz $\mathbf{A}^* \in \mathbb{C}^{1\times m}$, garante que existe uma matriz unitária $\mathbf{V} \in \mathbb{C}^{m\times m}$ tal que

$$\mathbf{A}^*\mathbf{V} = \begin{bmatrix} \sigma_1 \ \mathbf{0} \end{bmatrix}, \qquad \sigma_1 \in \mathbb{R}^+.$$

Por conseguinte, pondo $\mathbf{U} = \mathbf{V}^*$, obtemos uma matriz unitária $\mathbf{U} \in \mathbb{C}^{m\times m}$ tal que

$$\mathbf{UA} = \mathbf{V}^*(\mathbf{A}^*)^* = (\mathbf{A}^*\mathbf{V})^* = \begin{bmatrix}\sigma_1 \ \mathbf{0}\end{bmatrix}^* = \begin{bmatrix}\sigma_1 \\ \mathbf{0}\end{bmatrix}$$

e, portanto, o resultado é verdadeiro com $\mathbf{D} = \begin{bmatrix}\sigma_1\end{bmatrix}$.

Agora, suponhamos que $n \geqslant 2$ e que o resultado está provado para qualquer $\mathbf{A}_0 \in \mathbb{C}^{m\times(n-1)}$. Seja $\mathbf{A} \in \mathbb{C}^{m\times n}$ e sejam $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n$ os valores próprios de $\mathbf{A}^*\mathbf{A} \in \mathbb{C}^{n\times n}$ (notemos que $\mathbf{A}^*\mathbf{A}$ é uma matriz hermítica). Seja $\mathbf{v}_1 \in \mathbb{C}^{n\times 1}$ um vector próprio de $\mathbf{A}^*\mathbf{A}$ associado a λ_1 com norma $\|\mathbf{v}_1\| = 1$. Como

$$\lambda_1 = \max_{\substack{\mathbf{w}\in\mathbb{C}^{n\times 1}\\ \|\mathbf{w}\|=1}} \mathbf{w}^*\mathbf{A}^*\mathbf{A}\mathbf{w} = \max_{\substack{\mathbf{w}\in\mathbb{C}^{n\times 1}\\ \|\mathbf{w}\|=1}} (\mathbf{A}\mathbf{w})^*(\mathbf{A}\mathbf{w}) = \max_{\substack{\mathbf{w}\in\mathbb{C}^{n\times 1}\\ \|\mathbf{w}\|=1}} \|\mathbf{A}\mathbf{w}\|^2 \in \mathbb{R}^+,$$

podemos considerar $\sigma_1 = \sqrt{\lambda_1} \in \mathbb{R}^+$. Seja $\mathbf{V}_1 \in \mathbb{C}^{n\times n}$ uma matriz unitária cuja primeira coluna é $\mathbf{v}_1$, de modo que $\mathbf{V}_1 = \begin{bmatrix}\mathbf{v}_1 \ \mathbf{v}_2 \ \cdots \ \mathbf{v}_n\end{bmatrix}$ onde $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ é uma base ortonormada de $\mathbb{C}^{n\times 1}$. Então,

$$\mathbf{AV}_1 = \begin{bmatrix}\mathbf{A}\mathbf{v}_1 & \mathbf{A}\mathbf{v}_2 & \cdots & \mathbf{A}\mathbf{v}_n\end{bmatrix}.$$

Por outro lado, consideremos o vector $\mathbf{u}_1 = \sigma_1^{-1}\mathbf{A}\mathbf{v}_1 \in \mathbb{C}^{m\times 1}$, de modo que $\mathbf{A}\mathbf{v}_1 = \sigma_1\mathbf{u}_1$. Temos

$$\|\mathbf{u}_1\|^2 = \mathbf{u}_1^*\mathbf{u}_1 = (\sigma_1^{-1}\mathbf{A}\mathbf{v}_1)^*(\sigma_1^{-1}\mathbf{A}\mathbf{v}_1) = \sigma_1^{-2}\mathbf{v}_1^*\mathbf{A}^*\mathbf{A}\mathbf{v}_1 = \mathbf{v}_1^*\mathbf{v}_1 = \|\mathbf{v}_1\|^2 = 1$$

(porque $(\mathbf{A}^*\mathbf{A})\mathbf{v}_1 = \lambda_1\mathbf{v}_1$ e $\lambda_1 = \sigma_1^2$) e, portanto, existe uma matriz unitária $\mathbf{U}_1 \in \mathbb{C}^{m\times 1}$ cuja primeira coluna é $\mathbf{u}_1$. Temos

$$\mathbf{U}_1^*\mathbf{AV}_1 = \begin{bmatrix}\mathbf{U}_1^*\mathbf{A}\mathbf{v}_1 & \mathbf{U}_1^*\mathbf{A}\mathbf{v}_2 & \cdots & \mathbf{U}_1^*\mathbf{A}\mathbf{v}_n\end{bmatrix}.$$

Escolhendo $\mathbf{u}_2, \cdots, \mathbf{u}_m \in \mathbb{C}^{m\times 1}$ tais que $\mathbf{U}_1 = \begin{bmatrix}\mathbf{u}_1 \ \mathbf{u}_2 \ \cdots \ \mathbf{u}_n\end{bmatrix}$, verificamos que

$$\mathbf{U}_1^*\mathbf{A}\mathbf{v}_j = \begin{bmatrix}\mathbf{u}_1^* \\ \mathbf{u}_2^* \\ \vdots \\ \mathbf{u}_m^*\end{bmatrix}\mathbf{A}\mathbf{v}_j = \begin{bmatrix}\mathbf{u}_1^*\mathbf{A}\mathbf{v}_j \\ \mathbf{u}_2^*\mathbf{A}\mathbf{v}_j \\ \cdots \\ \mathbf{u}_m^*\mathbf{A}\mathbf{v}_j\end{bmatrix}, \qquad 1 \leqslant j \leqslant n.$$

Ora,

- $\mathbf{u}_1^*\mathbf{A}\mathbf{v}_1 = \sigma_1\mathbf{u}_1^*\mathbf{u}_1 = \sigma_1,$
- $\mathbf{u}_i^*\mathbf{A}\mathbf{v}_1 = \mathbf{u}_i^*(\sigma_1\mathbf{u}_1) = \sigma_1\mathbf{u}_i^*\mathbf{u}_1 = 0, \qquad 2 \leqslant i \leqslant m,$
- $\mathbf{u}_1^*\mathbf{A}\mathbf{v}_j = \sigma_1^{-1}\mathbf{v}_1^*\mathbf{A}^*\mathbf{A}\mathbf{v}_j = \sigma_1^{-1}(\mathbf{A}^*\mathbf{A}\mathbf{v}_1)^*\mathbf{v}_j = \sigma_1^{-1}\lambda_1\mathbf{v}_1^*\mathbf{v}_j = 0, \qquad 2 \leqslant j \leqslant n.$

Por conseguinte,

$$\mathbf{U}_1^*\mathbf{AV}_1 = \begin{bmatrix}\sigma_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_0\end{bmatrix}$$

para alguma matriz $\mathbf{A}_0 \in \mathbb{C}^{(m-1)\times(n-1)}$.

Por hipótese de indução, existem matrizes unitárias $\mathbf{U}_0 \in \mathbb{C}^{(m-1)\times(m-1)}$ e $\mathbf{V}_0 \in \mathbb{C}^{(n-1)\times(n-1)}$ tais que

$$\mathbf{U}_0\mathbf{A}_0\mathbf{V}_0 = \begin{bmatrix} \mathbf{D}_0 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$$

onde

$$\mathbf{D}_0 = \begin{bmatrix} \sigma_2 & 0 & \cdots & 0 \\ 0 & \sigma_3 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \sigma_r \end{bmatrix} \in \mathbb{R}^{(r-1)\times(r-1)}$$

em que $\sigma_1, \ldots, \sigma_r \in \mathbb{R}$ são tais que

$$\sigma_2 \geqslant \sigma_3 \geqslant \cdots \geqslant \sigma_r > 0;$$

notemos que $r(\mathbf{A}_0) = r(\mathbf{A}) - 1 = r - 1$ (porque $\sigma_1 \neq 0$). Pondo

$$\mathbf{U} = \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{U}_0 \end{bmatrix} \mathbf{U}_1^* \qquad \text{e} \qquad \mathbf{V} = \mathbf{V}_1 \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{V}_0 \end{bmatrix},$$

obtemos matrizes unitárias $\mathbf{U} \in \mathbb{C}^{m\times m}$ e $\mathbf{V} \in \mathbb{C}^{n\times n}$ tais que

$$\mathbf{UAV} = \begin{bmatrix} \sigma_1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_0 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}.$$

Para concluir a demonstração, falta provar que $\sigma_1 \geqslant \sigma_2 \geqslant \cdots \geqslant \sigma_r > 0$. Ora, temos

$$\begin{aligned} \mathbf{V}^*(\mathbf{A}^*\mathbf{A})\mathbf{V} &= \mathbf{V}^*\mathbf{A}^*(\mathbf{U}^*\mathbf{U})\mathbf{AV} = (\mathbf{UAV})^*(\mathbf{UAV}) \\ &= \begin{bmatrix} \sigma_1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_0 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \sigma_1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_0 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} = \begin{bmatrix} \sigma_1^2 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_0^2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} \end{aligned}$$

e, portanto, $\sigma_1^2, \ldots, \sigma_r^2$ são os valores próprios não-nulos da matriz $\mathbf{A}^*\mathbf{A}$ (porque $\mathbf{V}$ é invertível e $\mathbf{V}^{-1} = \mathbf{V}^*$). O resultado segue-se porque $\lambda_1 = \sigma_1^2$ e $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n \geqslant 0$. □

TEOREMA 6.2. *Seja* $\mathbf{A} \in \mathbb{C}^{m\times n}$ *uma matriz arbitrária e suponhamos que existem matrizes unitárias* $\mathbf{U} \in \mathbb{C}^{m\times m}$ *e* $\mathbf{V} \in \mathbb{C}^{n\times n}$ *tais que*

$$\mathbf{UAV} = \begin{bmatrix} \mathbf{D} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \qquad \mathbf{D} = \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \sigma_r \end{bmatrix} \in \mathbb{R}^{r\times r},$$

onde $\sigma_1, \ldots, \sigma_r \in \mathbb{R}$ *são tais que* $\sigma_1 \geqslant \sigma_2 \geqslant \cdots \geqslant \sigma_r > 0$. *Então,* $\sigma_1^2, \ldots, \sigma_r^2$ *são os valores próprios não-nulos da matriz* $\mathbf{A}^*\mathbf{A} \in \mathbb{C}^{n\times n}$*; em particular,* $\sigma_1, \ldots, \sigma_r$ *são univocamente determinados.*

Demonstração. Basta repetir o último parágrafo da demonstração anterior. $\square$

$\triangleright$ Definimos os valores singulares de uma matriz $\mathbf{A} \in \mathbb{C}^{m\times n}$ como sendo as raízes quadradas positivas dos valores próprios não-nulos da matriz hermítica $\mathbf{A}^*\mathbf{A} \in \mathbb{C}^{n\times n}$.

Proposição 6.3. *Seja $\mathbf{A} \in \mathbb{C}^{m\times n}$ uma matriz arbitrária. Então, os valores próprios das matrizes $\mathbf{A}^*\mathbf{A} \in \mathbb{C}^{n\times n}$ e $\mathbf{A}\mathbf{A}^* \in \mathbb{C}^{m\times m}$ são números reais não-negativos e, além disso, $\mathbf{A}^*\mathbf{A}$ e $\mathbf{A}\mathbf{A}^*$ têm os mesmos valores próprios não-nulos (contando multiplicidades); em particular, as matrizes $\mathbf{A}$ e $\mathbf{A}^*$ têm os mesmos valores singulares.*

Demonstração. Como $\mathbf{A}^*\mathbf{A} \in \mathbb{C}^{n\times n}$ é uma matriz hermítica, os seus valores próprios são números reais. Além disso, se $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n$ forem os valores próprios de $\mathbf{A}^*\mathbf{A}$, então o teorema de Courant-Fischer (ou a Proposição 5.2) garante que

$$\lambda_n = \min_{\substack{\mathbf{v}\in\mathbb{C}^{n\times 1}\\ \|\mathbf{v}\|=1}} \mathbf{v}^*\mathbf{A}^*\mathbf{A}\mathbf{v}.$$

Como

$$\mathbf{v}^*\mathbf{A}^*\mathbf{A}\mathbf{v} = \|\mathbf{A}\mathbf{v}\|^2 \in \mathbb{R}_0^+, \qquad \mathbf{v} \in \mathbb{C}^{n\times 1}$$

concluímos $\lambda_n \geqslant 0$ e, portanto, $\lambda_1, \ldots, \lambda_n \in \mathbb{R}_0^+$. De modo análogo, concluímos que os valores próprios de $\mathbf{A}\mathbf{A}^* = (\mathbf{A}^*)^*\mathbf{A}^*$ são números reais não-negativos.

Sejam $\sigma_1 \geqslant \sigma_2 \geqslant \cdots \geqslant \sigma_r > 0$ os valores singulares de $\mathbf{A}$ e sejam $\mathbf{U} \in \mathbb{C}^{m\times m}$ e $\mathbf{V} \in \mathbb{C}^{n\times n}$ matrizes unitárias tais que

$$\mathbf{U}\mathbf{A}\mathbf{V} = \begin{bmatrix} \mathbf{D} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \qquad \mathbf{D} = \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \sigma_r \end{bmatrix} \in \mathbb{R}^{r\times r}.$$

Então,

$$\mathbf{U}(\mathbf{A}\mathbf{A}^*)\mathbf{U}^* = \mathbf{U}\mathbf{A}(\mathbf{V}\mathbf{V}^*)\mathbf{A}^*\mathbf{U}^* = (\mathbf{U}\mathbf{A}\mathbf{V})(\mathbf{U}\mathbf{A}\mathbf{V})^* = \begin{bmatrix} \mathbf{D} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}\begin{bmatrix} \mathbf{D} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} = \begin{bmatrix} \mathbf{D}^2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix},$$

logo $\sigma_1^2 \geqslant \sigma_2^2 \geqslant \cdots \geqslant \sigma_r^2 > 0$ são os valores próprios não-nulos da matriz $\mathbf{A}\mathbf{A}^*$ (porque $\mathbf{V}$ é invertível e $\mathbf{V}^{-1} = \mathbf{V}^*$). O resultado segue-se. $\square$

Teorema 6.4 (max-min/min-max). *Seja $\mathbf{A} \in \mathbb{C}^{m\times n}$ e sejam $\sigma_1 \geqslant \sigma_2 \geqslant \cdots \geqslant \sigma_r$ os valores singulares de $\mathbf{A}$. Então, para qualquer $1 \leqslant i \leqslant r$,*

$$\sigma_i = \max_{\substack{\mathcal{V}\leqslant\mathbb{C}^{n\times 1}\\ \dim\mathcal{V}=i}} \min_{\substack{\mathbf{v}\in\mathcal{V}\\ \|\mathbf{v}\|=1}} \|\mathbf{A}\mathbf{v}\| = \min_{\substack{\mathcal{V}\leqslant\mathbb{C}^{n\times 1}\\ \dim\mathcal{V}=n-i+1}} \max_{\substack{\mathbf{v}\in\mathcal{V}\\ \|\mathbf{v}\|=1}} \|\mathbf{A}\mathbf{v}\|, \qquad 1 \leqslant i \leqslant r.$$

Equivalentemente,

$$\sigma_i = \max_{\substack{\mathcal{V}\leqslant\mathbb{C}^{n\times 1}\\ \dim\mathcal{V}=i}} \min_{\substack{\mathbf{v}\in\mathcal{V}\\ \mathbf{v}\neq\mathbf{0}}} \frac{\|\mathbf{A}\mathbf{v}\|}{\|\mathbf{v}\|} = \min_{\substack{\mathcal{V}\leqslant\mathbb{C}^{n\times 1}\\ \dim\mathcal{V}=n-i+1}} \max_{\substack{\mathbf{v}\in\mathcal{V}\\ \mathbf{v}\neq\mathbf{0}}} \frac{\|\mathbf{A}\mathbf{v}\|}{\|\mathbf{v}\|}, \qquad 1\leqslant i\leqslant r.$$

Demonstração. É consequência do teorema de Courant-Fischer, tendo em conta que

$$\sigma_1^2 \geqslant \cdots \geqslant \sigma_r^2 \geqslant \sigma_{r+1}^2 \geqslant \cdots \geqslant \sigma_n^2, \qquad \sigma_{r+1} = \cdots = \sigma_n = 0,$$

são os valores próprios da matriz (hermítica) $\mathbf{A}^*\mathbf{A} \in \mathbb{C}^{n\times n}$. □