Descomposición en valores singulares (DVS)

Aplicaciones

Introducción

Toda matriz $A \in \mathbb{R}^{mxn}$ puede descomponerse como un producto $A = U\Sigma V^T$ en donde $U \in \mathbb{R}^{mxm}$ es ortogonal, V$\in \mathbb{R}^{nxn}$ es ortogonal y $\Sigma \in \mathbb{R}^{mxn}$ es tal que $\Sigma_{ii} = \sigma_i$, real y no negativo (son los valores singulares de A) y $\Sigma_{ij} = 0, i \neq j$

Por ejemplo: $\boldsymbol{A} \in \mathbb{R}^{4x3}$, $\boldsymbol{U} \in \mathbb{R}^{4x4}$, $\boldsymbol{\Sigma} \in \mathbb{R}^{4x3}$ **y** V$\in \mathbb{R}^{3x3}$

$$\underbrace{\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \end{pmatrix}}_{A} = \underbrace{\begin{pmatrix} u_{11} & u_{12} & u_{13} & u_{14} \\ u_{21} & u_{22} & u_{23} & u_{24} \\ u_{31} & u_{32} & u_{33} & u_{34} \\ u_{41} & u_{42} & u_{43} & u_{44} \end{pmatrix}}_{U} \underbrace{\begin{pmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \\ 0 & 0 & 0 \end{pmatrix}}_{\Sigma} \underbrace{\begin{pmatrix} v_{11} & v_{21} & v_{31} \\ v_{12} & v_{22} & v_{32} \\ v_{13} & v_{23} & v_{33} \end{pmatrix}}_{V^T}$$

Proposición

Dada la matriz $A \in \mathbb{R}^{mxn}$ la matriz A^TA es simétrica porque

$$(A^TA)^T = A^T(A^T)^T = A^TA$$

Y semidefinida positiva ya que:

$$\forall X \in \mathbb{R}^n\colon X^T(A^TA)X = (AX)^TAX = \|AX\|^2 \geq 0_{\mathbb{R}}$$

Observación:

en el caso de que las columnas de A sean linealmente independientes se tiene que $X^T(A^TA)X = 0_{\mathbb{R}} \Leftrightarrow \|AX\|^2=0_{\mathbb{R}} \Leftrightarrow AX = 0_{\mathbb{R}^m} \Leftrightarrow$ $X = 0_{\mathbb{R}^n} \Leftrightarrow A^TA$ es definida positiva

Definición

Si A es una matriz de mxn los valores singulares de A son las raíces cuadradas de los autovalores de $A^T A$.
Si $\lambda_1, \lambda_2, \ldots, \lambda_n$, son los autovalores de $A^T A$ ordenados en forma decreciente $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n \geq 0$ entonces
$\sigma_i = \sqrt{\lambda_i}$ es el i-ésimo valor singular de A.
Es decir : $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n \geq 0$

Teorema

Sea $A \in \mathbb{R}^{mxn}$ y $\lambda_1, \lambda_2, \ldots, \lambda_n$, los autovalores de $A^T A$ que suponemos $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_r > \lambda_{r+1} = \cdots = \lambda_n = 0$. Como $A^T A$ es simétrica se puede diagonalizar ortogonalmente, es decir existe una base ortonormal de $\mathbb{R}^n$ $\{v_1, v_2, \ldots, v_n\}$ tal que $A^T A\, v_i = \lambda_i v_i$ entonces:

(1) $\{Av_1, Av_2, \dots, Av_n\}$ es un conjunto ortogonal y $\|Av_i\| = \sqrt{\lambda_i} = \sigma_i \;\; \forall i = 1,2,..,n$

Lo probamos:

$< Av_i, Av_j > = (Av_j)^T Av_i = {v_j}^T (A^T A) v_i = {v_j}^T \lambda_i v_i = \lambda_i v_j^T v_i$

Si $i \neq j$ $< Av_i, Av_j >= 0$ entonces $Av_i \perp Av_j$

Si $i = j$ $< Av_i, Av_i >= \lambda_i v_i^T v_i = \lambda_i$entonces $\|Av_i\|^2 = \lambda_i$ de donde $\|Av_i\| = \sqrt{\lambda_i} = \sigma_i \;\; \forall i = 1,2,..,n$

(2) $\left\{\frac{Av_1}{\sigma_1}, \frac{Av_2}{\sigma_2}, \dots, \frac{Av_r}{\sigma_r}\right\}$ es una base ortonormal del espacio $Col(A)$

Por la propiedad (1) $\|Av_i\| = \sqrt{\lambda_i} = \sigma_i$ y hay r valores singulares no nulos $Av_i \neq 0_{\mathbb{R}^m} \Leftrightarrow 1 \leq i \leq r,$ luego $\{Av_1, Av_2, \dots, Av_r\}$ es linealmente independiente y están en el espacio $Col(A)$. Luego $\forall y \in Col(A)$:

$y = Au$ con $u \in \mathbb{R}^n$, y u se puede expresar como combinación lineal de

los vectores de la base: $u = \sum_{i=1}^{n} \alpha_i v_i$,
y como a partir de $r + 1$ los A$v_i = 0_{\mathbb{R}^m}$ se tiene que
$y = Au = \sum_{i=1}^{r} \alpha_i\, Av_i$ entonces $y \in gen\{Av_1, Av_2, \ldots, Av_r\}$ lo que demuestra que $\{Av_1, Av_2, \ldots, Av_r\}$ es una base ortogonal del espacio $Col(A)$ de aquí el $rg(A) = \dim\big(Col(A)\big) = r$
Además como $\|Av_i\| = \sigma_i$ el conjunto $\left\{\frac{Av_1}{\sigma_1}, \frac{Av_2}{\sigma_2}, \ldots, \frac{Av_r}{\sigma_r}\right\}$ resulta una BON del espacio $Col(A)$

3) $\{v_{r+1}, v_{r+2}, \ldots, v_n\}$ es una bon del espacio $Nul(A)$

por hipótesis el conjunto es ortonormal y $rg(A) + \dim\big(Nul(A)\big) = n$
Entonces $\dim\big(Nul(A)\big) = n - r$ y como $\|Av_i\| = 0, \forall i \geq r + 1$
los v_i con $i \geq r + 1$ pertenecen al espacio $Nul(A)$.

Descomposición en valores singulares

La matriz $A \in \mathbb{R}^{mxn}$ se puede factorizar $A = U\Sigma V^T$ en donde $U \in \mathbb{R}^{mxm}$ y V$\in \mathbb{R}^{nxn}$ son ortogonales y $\Sigma \in \mathbb{R}^{mxn}$ con

$$\Sigma = \begin{bmatrix} \Sigma_r & 0_{rx(n-r)} \\ 0_{(m-r)xr} & 0_{(m-r)x(n-r)} \end{bmatrix} \text{ y } \Sigma_r = \begin{bmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_r \end{bmatrix} \text{ siendo}$$

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r$$

Para hallar la matriz ortogonal V hallamos una base ortonormal $\{v_1, v_2, \dots, v_n\}$ de $\mathbb{R}^n$ tal que $A^T A\, v_i = \lambda_i v_i$

$$V = [v_1 \quad v_2 \quad \dots \quad v_n]$$

Para hallar la matriz ortogonal U usamos el resultado demostrado en el teorema, ítem (2) $u_i = \frac{Av_i}{\sigma_i}, 1 \leq i \leq r$ y en el caso que $r < m$ se completa a una base ortonormal de $\mathbb{R}^m$ se tiene entonces $U = [u_1 \quad u_2 \quad \dots \quad u_m]$

Demostrar que $A = U\Sigma V^T$ como $V^T = V^{-1}$ equivale a demostrar $AV = U\Sigma$

Se tiene que $Av_i = \sigma_i u_i \ , 1 \leq i \leq r$ y $Av_i = 0_{\mathbb{R}^m}, \forall i \geq r+1$

$$AV = [Av_1 \quad Av_2 \quad \dots \quad Av_n] = [Av_1 \quad Av_2 \quad \dots \quad Av_r \ 0 \dots 0] =$$

$$[\sigma_1 u_1 \quad \sigma u_2 \quad \dots \quad \sigma_r u_r \ 0 \dots 0] = [u_1 \quad u_2 \quad \dots \quad u_m] \begin{bmatrix} \sigma_1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & \sigma_r & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \end{bmatrix} =$$

$$U\Sigma$$

$$\Sigma = \begin{bmatrix} \Sigma_r & 0_{rx(n-r)} \\ 0_{(m-r)xr} & 0_{(m-r)x(n-r)} \end{bmatrix} \text{ y } \Sigma_r = \begin{bmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_r \end{bmatrix} \text{ siendo}$$

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r$$

Ejemplo 1: Hallar una DVS de la matriz $A \in \mathbb{R}^{3x2}$: $A = \begin{bmatrix} 7 & 1 \\ 0 & 0 \\ 5 & 5 \end{bmatrix}$

$A = U\Sigma V^T$ en donde $U \in \mathbb{R}^{3x3}$ es ortogonal, V$\in \mathbb{R}^{2x2}$ es ortogonal y $\Sigma \in \mathbb{R}^{3x2}$

Calculamos los autovalores de $A^T A = \begin{bmatrix} 74 & 32 \\ 32 & 26 \end{bmatrix}$ y los ordenamos de mayor a menor: $\lambda_1 = 90, \lambda_2 = 10$, los valores singulares son $\sigma_1 = \sqrt{90} = 3\sqrt{10}$, $\sigma_2 = \sqrt{10}$

Calculamos los autovectores asociados a $\lambda_1 = 90, \lambda_2 = 10$ y formamos una base ortonormal de $\mathbb{R}^2, S_{\lambda_1}(A^T A) = gen\{[2 \;\; 1]^T\}\; S_{\lambda_2}(A^T A) = gen\{[-1 \;\; 2]^T\}$

$B = \left\{ v_1 = \begin{bmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{bmatrix}^T, v_2 = \begin{bmatrix} -\frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{bmatrix}^T \right\}$ los vectores v_1 y v_2 formarán las columnas de la matriz $V = \begin{bmatrix} \frac{2}{\sqrt{5}} & -\frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{bmatrix}$

Ahora calculamos las columnas de la matriz U:

$$u_1 = \frac{Av_1}{\sigma_1} = \frac{\begin{bmatrix} 7 & 1 \\ 0 & 0 \\ 5 & 5 \end{bmatrix} \begin{bmatrix} \frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \end{bmatrix}}{3\sqrt{10}} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix} \qquad u_2 = \frac{Av_2}{\sigma_2} = \frac{\begin{bmatrix} 7 & 1 \\ 0 & 0 \\ 5 & 5 \end{bmatrix} \begin{bmatrix} -\frac{1}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} \end{bmatrix}}{\sqrt{10}} = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix}$$

u_3 se obtiene completando a una base ortonormal de $\mathbb{R}^3$: $u_3 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$

Luego una DVS de la matriz A es: $A = U\Sigma V^T$

$$A = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \end{bmatrix} \begin{bmatrix} 3\sqrt{10} & 0 \\ 0 & \sqrt{10} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ -\frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{bmatrix}$$

Ejemplo 2: Hallar una DVS de la matriz C $\in \mathbb{R}^{2x2}$: $C = \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix}$

C$= U\Sigma V^T$ en donde $U \in \mathbb{R}^{2x2}$ es ortogonal, V$\in \mathbb{R}^{2x2}$ es ortogonal y $\Sigma \in \mathbb{R}^{2x2}$

Calculamos los autovalores de $C^T C = \begin{bmatrix} 2 & 4 \\ 4 & 8 \end{bmatrix}$ y los ordenamos de mayor a menor: $\lambda_1 = 10, \lambda_2 = 0$, los valores singulares son $\sigma_1 = \sqrt{10},\ \sigma_2 = 0$

Calculamos los autovectores asociados a $\lambda_1 = 10, \lambda_2 = 0$ y formamos una base ortonormal de $\mathbb{R}^2$, $S_{\lambda_1}(C^T C) = gen\{[1 \;\; 2]^T\}$ $S_{\lambda_2}(C^T C) = gen\{[-2 \;\; 1]^T\}$

$B = \left\{ v_1 = \left[\frac{1}{\sqrt{5}} \;\; \frac{2}{\sqrt{5}}\right]^T, v_2 = \left[-\frac{2}{\sqrt{5}} \;\; \frac{1}{\sqrt{5}}\right]^T \right\}$ los vectores v_1 y v_2 formarán las columnas de la matriz $V = \begin{bmatrix} \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{bmatrix}$

Ahora calculamos las columnas de la matriz U:

$u_1 = \frac{Cv_1}{\sigma_1} = \frac{\begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} \frac{1}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} \end{bmatrix}}{\sqrt{10}} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$, u_2 se obtiene completando a una base

ortonormal de $\mathbb{R}^2$: $u_2 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} \end{bmatrix}$

Luego una DVS de la matriz C es: $C = U\Sigma V^T$

$$C=\begin{bmatrix}\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}}\end{bmatrix}\begin{bmatrix}\sqrt{10} & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}\frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}}\\ -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}}\end{bmatrix}$$

Descomposición en valores singulares reducida

Sea $A \in \mathbb{R}^{mxn}$ una DVS de $A = U\Sigma V^T$ en donde $U \in \mathbb{R}^{mxm}$ y V$\in \mathbb{R}^{nxn}$ son ortogonales y $\Sigma \in \mathbb{R}^{mxn}$ con $V = [v_1 \quad v_2 \quad \dots \quad v_n]$ y $U = [u_1 \quad u_2 \quad \dots \quad u_m]$
Si rg(A)=r definimos las matrices V $= [V_r \quad V_{n-r}]$ y U $= [U_r \quad U_{m-r}]$. Luego:

A $= [U_r \quad U_{m-r}]\begin{bmatrix}\Sigma_r & 0_{rx(n-r)}\\ 0_{(m-r)xr} & 0_{(m-r)x(n-r)}\end{bmatrix}\begin{bmatrix}V_r^T\\ V_{n-r}^T\end{bmatrix} = U_r\Sigma_r V_r^T$, observar que D es regular porque contiene los valores singulares mayores que cero.

Observaciones

teniendo en cuenta las definiciones de las matrices

V $= [V_r \quad V_{n-r}]$ y U $= [U_r \quad U_{m-r}]$ se tienen los siguientes resultados:

1) $\{u_1, u_2, \dots, u_r\}$ es una base ortonormal para $\boldsymbol{Col}(\boldsymbol{A})$

2) $\{u_{r+1}, u_{r+2}, \dots, u_m\}$ es una base ortonormal para Nul$(\boldsymbol{A}^T)$

3) $\{v_1, v_2, \dots, v_r\}$ es una base ortonormal para Fil$(\boldsymbol{A})$

4)$\{v_{r+1}, v_{r+2}, \dots, n\}$ es una base ortonormal para Nul$(\boldsymbol{A})$

Para el ejemplo 1

$$A = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \end{bmatrix} \begin{bmatrix} 3\sqrt{10} & 0 \\ 0 & \sqrt{10} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ -\frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{bmatrix} DVS$$

$$A = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ 0 & 0 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 3\sqrt{10} & 0 \\ 0 & \sqrt{10} \end{bmatrix} \begin{bmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ -\frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{bmatrix} \text{DVS reducida}$$

Para el ejemplo 2

$$C=\begin{bmatrix}\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}}\end{bmatrix}\begin{bmatrix}\sqrt{10} & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}\frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}}\\ -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}}\end{bmatrix}\text{ DVS}$$

$$C=\begin{bmatrix}\frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}}\end{bmatrix}\begin{bmatrix}\sqrt{10}\end{bmatrix}\begin{bmatrix}\frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}}\end{bmatrix}\text{ DVS reducida}$$

Definición

Sea $A=U_r\Sigma_r V_r^T$ una DVS reducida de A. La matriz $A^+=V_r\Sigma_r^{-1}U_r^T$ es la matriz pseudoinversa de Moore-Penrose de A.

Teorema

Sea A $\in \mathbb{R}^{mxn}$ y sea A^+la matriz pseudoinversa de Moore-Penrose de A entonces $\forall b\ \in\ \mathbb{R}^m$: $\hat{x}=A^+$b es la única solución por cuadrados mínimos de longitud mínima de la ecuación $Ax=b$ y pertenece al espacio $Fil(A)$

Aplicaciones de la DVS

1. Solución de norma mínima por cuadrados mínimos

Sea $A = \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix}$ hallar la solución por cuadrados mínimos de norma mínima del sistema $Ax = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

Una DVS reducida de la matriz A es $A = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \sqrt{10} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{bmatrix}$, calculamos la pseudoinversa $A^+ = V_r \Sigma_r^{-1} U_r^T = \begin{bmatrix} \frac{1}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{10}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} \frac{1}{10} & \frac{1}{10} \\ \frac{1}{5} & \frac{1}{5} \end{bmatrix}$

Y la solución por cuadrados mínimos de norma mínima es: $\hat{x} = A^+$b

$$\hat{x} = \begin{bmatrix} \frac{1}{10} & \frac{1}{10} \\ \frac{1}{5} & \frac{1}{5} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{10} \\ \frac{1}{5} \end{bmatrix} \in Fil(A)$$

2. Normas matriciales

La norma inducida en $\mathbb{R}^{n\times n}$ por el producto interno: $\langle A, B\rangle = \operatorname{tr}\left(B^T A\right)$ como:

$$\|\cdot\|_F : \mathbb{R}^{n\times n} \to \mathbb{R}_0^+ : \|A\|_F = \sqrt{\langle A, A\rangle}$$

llamada *norma de Frobenius*.

Ejemplo. Sea $A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$. Entonces:

$$\|A\|_F = \sqrt{\operatorname{tr}\left(A^T A\right)} = \sqrt{\operatorname{tr}\begin{pmatrix} 5 & 4 \\ 4 & 5 \end{pmatrix}} = \sqrt{10}$$

Este concepto puede extenderse a matrices $A \in \mathbb{R}^{m\times n}$.

Definición. Una *norma matricial* es una función que asocia a cada matriz $A \in \mathbb{R}^{m\times n}$ un número real $\|A\|$ tal que, $\forall A, B \in \mathbb{R}^{m\times n}$ se cumple que:

1. $\|A\| \geq 0$ y $\|A\| = 0 \Leftrightarrow A = 0_{\mathbb{R}^{m\times n}}$
2. $\|cA\| = |c|\|A\|$
3. $\|A + B\| \leq \|A\| + \|B\|$
4. $\|AB\| \leq \|A\|\|B\|$

Se dice que una norma matricial es *compatible* con una norma vectorial $\|\cdot\|$ definida para vectores $x \in \mathbb{R}^n$, si $\forall A \in \mathbb{R}^{m\times n}$ y todos los vectores $x \in \mathbb{R}^n$ se verifica que $\|Ax\| \leq \|A\|\|x\|$.

Proposiciones.

1. La norma de Frobenius es compatible con la norma euclídea definida en $\mathbb{R}^n$ como $\|\cdot\| : \mathbb{R}^n \to \mathbb{R}_0^+ : \|x\| = \sqrt{\sum_{i=1}^n x_i^2}, \forall x \in \mathbb{R}^n$.
2. La norma de Frobenius es una norma matricial.
3. Si dos matrices $A, B \in \mathbb{R}^{n\times n}$ son ortogonalmente semejantes entonces $\|A\|_F = \|B\|_F$.

La descomposición en valores singulares proporciona una expresión simple para ciertas expresiones que involucran normas matriciales. En particular, la norma de Frobenius está completamente determinada por los valores singulares de una matriz, como lo demuestra el siguiente

Teorema. Sea $A \in \mathbb{R}^{m\times n}$ y sean $\sigma_1, \cdots, \sigma_r$ todos los valores singulares distintos de cero de A. Entonces:

$$\|A\|_F = \sqrt{\sigma_1^2 + \cdots + \sigma_r^2}$$

La demostración de este resultado se basa en que si $A \in \mathbb{R}^{m\times n}$ y Q es ortogonal de $m \times m$, entonces $\|QA\|_F = \|A\|_F$. En efecto:

$$\begin{aligned}
\|QA\|_F^2 &= \operatorname{tr}\left((QA)^T(QA)\right) \\
&= \operatorname{tr}\left(A^T(Q^TQ)A\right) \\
&= \operatorname{tr}\left(A^TA\right) \\
&= \|A\|_F^2
\end{aligned}$$

Sea $A = U\Sigma V^T$. Entonces:

$$\|A\|_F^2 = \|U\Sigma V^T\|_F^2 = \|\Sigma V^T\|_F^2 = \|(\Sigma V^T)^T\|_F^2 = \|V\Sigma^T\|_F^2 = \sigma_1^2 + \cdots + \sigma_r^2$$

que conduce al resultado.

Observar que la norma de Frobenius de una matriz de $m \times n$ es igual a la de su transpuesta, como se deduce de su definición.

Continuando con el ejemplo anterior, los valores singulares de $A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ son $\sigma_1 = 3, \sigma_2 = 1$.

Luego, una forma alternativa de calcular la norma de Frobenius de A es:

$$\|A\|_F^2 = \sqrt{\sigma_1^2 + \sigma_2^2} = \sqrt{3^2 + 1^2} = \sqrt{10}$$

3. La forma de producto externo de la DVS

Sea la matriz $A \in \mathbb{R}^{m\times n}$, con valores singulares $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$ y $\sigma_{r+1} = \cdots = \sigma_n = 0$, donde r es el rango de la matriz A. A partir de la descomposición en valores singulares reducida

$$A = U_r \Sigma_r V_r^T = \begin{bmatrix} u_1 & \cdots u_r \end{bmatrix} \begin{bmatrix} \sigma_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_r \end{bmatrix} \begin{bmatrix} v_1^T \\ \vdots \\ v_r^T \end{bmatrix} =$$

efectuando el producto se obtiene:

$$= \begin{bmatrix} \sigma_1 u_1 & \cdots \sigma_r u_r \end{bmatrix} \begin{bmatrix} v_1^T \\ \vdots \\ v_r^T \end{bmatrix} = \sigma_1 u_1 v_1^T + \cdots + \sigma_r u_r v_r^T$$

Esta expresión se conoce con el nombre de *forma de producto externo* de la descomposición en valores singulares. Cada uno de los sumandos es una matriz de rango 1.

Ejemplo. Sea la matriz $A = \begin{bmatrix} 2 & 0 \\ 0 & -3 \\ 0 & 0 \end{bmatrix} = \underbrace{\begin{bmatrix} 0 & 1 \\ -1 & 0 \\ 0 & 0 \end{bmatrix}}_{U_r} \underbrace{\begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}}_{\Sigma_r} \underbrace{\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}}_{V_r^T} =$

$$= \underbrace{3 \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \end{bmatrix}}_{A_1} + \underbrace{2 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \end{bmatrix}}_{A_2} = \underbrace{\begin{bmatrix} 0 & 0 \\ 0 & -3 \\ 0 & 0 \end{bmatrix}}_{\mathrm{rg(A_1)=1}} + \underbrace{\begin{bmatrix} 2 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}}_{\mathrm{rg(A_2)=1}}$$

4. Descomposición polar

Toda matriz $A \in \mathbb{R}^{n\times n}$ se puede escribir como el producto $A = RQ$, donde R es simétrica semidefinida positiva y Q es ortogonal.

En primer lugar, a partir de una descomposición en valores singulares de A:

$$A = U\Sigma V^T = U\Sigma\left(U^T U\right)V^T = \underbrace{\left(U\Sigma U^T\right)}_{R}\underbrace{\left(UV^T\right)}_{Q}$$

A continuación probaremos que ambas matrices tienen las propiedades requeridas.

Recordemos que una matriz $C \in \mathbb{R}^{n\times n}$ es semidefinida positiva si y sólo si $C^T = C$ y $x^T C x \geq 0 \quad \forall x \in \mathbb{R}^n$.

$$R^T = \left(U\Sigma U^T\right)^T = U\Sigma^T U^T$$

Como $R \in \mathbb{R}^{n\times n}$, la matriz $\Sigma \in \mathbb{R}^{n\times n}$ y es diagonal; por lo tanto $\Sigma^T = \Sigma$; luego:

$$R^T = U\Sigma U^T = R$$

Ahora, dado que la matriz Σ tiene elementos positivos o cero en su diagonal principal, resulta:

$$\begin{aligned} x^T R x &= x^T U\Sigma U^T x = x^T U\left(\sqrt{\Sigma}\right)^T\sqrt{\Sigma}U^T x = \left(\sqrt{\Sigma}U^T x\right)^T\left(\sqrt{\Sigma}U^T x\right) \\ &= \|\sqrt{\Sigma}U^T x\|^2 \geq 0 \quad \forall x \in \mathbb{R}^n \end{aligned}$$

Con respecto a la matriz Q, se trata del producto de dos matrices ortogonales U y V^T, que es también ortogonal.

Ejemplo. Una DVS de $A = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$ es:

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \sqrt{2} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$$

Determinamos $R = U\Sigma U^T = \begin{bmatrix} \sqrt{2} & 0 \\ 0 & 0 \end{bmatrix}$

y $Q = UV^T = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$

Finalmente:

$$A = \begin{bmatrix} \sqrt{2} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$$

es la descomposición polar para A.