Descomposición en Valores Singulares

Cualquier matriz $A \in \mathbb{K}^{n \times m}$ puede escribirse como el producto de tres matriz U, Σ y V. A esta descomposición o factorización la llamamos desc. en valores singulares (SVD, por sus siglas en inglés). Veamos quiénes son estas matrices.

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

$A \in \mathbb{K}^{n \times m}$

$U \in \mathbb{K}^{n \times n}$ unitaria

$\Sigma \in \mathbb{R}^{n \times m}$ diagonal

$V \in \mathbb{K}^{m \times m}$ unitaria

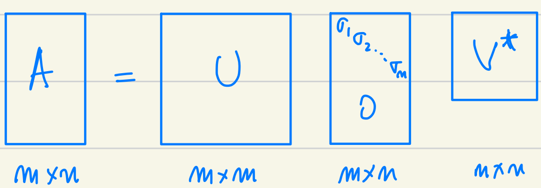
Los elementos de la diagonal de Σ los llamaremos σ_i y cumplen $\sigma_1 \geq \sigma_2 \geq \dots \geq 0$ y son los llamados valores singulares.

Observaciones

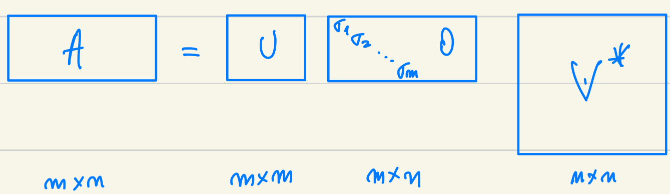
- Los valores singulares siempre son reales no negativos. (y están ordenados)
- Esta descomposición existe para cualquier matriz de cualquier tamaño.
- Al ser U y V unitarios: las columnas de U forman una b.o.n de $\mathbb{K}^m$ y las columnas de V forman una b.o.n. de $\mathbb{K}^n$.

¿Qué "forma" tiene tiene la SVD? Sea $A \in \mathbb{K}^{m\times n}$

$m \geq n$



$m \leq n$



Tratemos de entender qué significan V, Σ y U, y qué propiedades deberían cumplir sus componentes

Para ello, **asumamos que existe la SVD : $A = U\Sigma V^*$**

Llamemos u_i y v_i a las columnas de U y V respectivamente

$$\Rightarrow \quad A = \left(u_1 \,\middle|\, u_2 \,\middle|\, \cdots \,\middle|\, u_m \right) \begin{pmatrix} \sigma_1 & & \\ & \sigma_2 & \\ & & \ddots \end{pmatrix} \begin{pmatrix} v_1^* \\ \hline v_2^* \\ \hline \vdots \\ \hline v_n^* \end{pmatrix}$$

Obs. I

$$A^*A = (U\Sigma V^*)^*(U\Sigma V^*) = V\Sigma^* \underbrace{U^*U}_{= I} \Sigma V^* =$$
$$= V(\Sigma^*\Sigma)V^*$$

Como V es unitaria : $V^{-1} = V^*$

$\Sigma^*\Sigma$ es cuadrada y diagonal, además $\Sigma^* = \Sigma^t$

$\Sigma^*\Sigma = \Sigma^t\Sigma = D \in \mathbb{R}^{m\times m}$ (pues $\Sigma \in \mathbb{R}^{m\times n}$)

$$\rightarrow \underset{m\times m}{A^*A} = VDV^{-1} \rightarrow \text{ es una diagonalización}$$

Una diagonalización revela los autovalores y autovectores de la matriz
Efectivamente:

$$(A^*A)\,v_i = (V\Sigma^t\Sigma V^*)\,v_i = V\Sigma^t\Sigma e_i = V\Sigma^t \sigma_i e_i =$$
$$= V\Sigma^t e_i \sigma_i = V e_i \sigma_i^2 = \sigma_i^2 v_i$$
$$\Rightarrow (A^*A)\,v_i = \sigma_i^2 v_i$$

$\Rightarrow$ v_i son autovectores de A^*A con autovalor σ_i^2

Obs. 2 Podemos hacer un desarrollo similar para AA^*:

$$AA^* = (U\Sigma V^*)(U\Sigma V^*)^* = U\Sigma \underbrace{V^*V}_{= I}\Sigma^* U^* =$$
$$= U\Sigma\Sigma^t U^* = U\tilde{D}U^*$$

$AA^* \in \mathbb{K}^{n\times n}$ y $\tilde{D} \in \mathbb{R}^{n\times n}$ es diagonal y cuadrada

Notar que D y $\tilde{D}$ tienen diferente dimensión si $n \neq m$
En ambos casos $D_{ii} = \tilde{D}_{ii} = \sigma_i^2$, $1 \le i \le \min\{n, m\}$

Con un desarrollo similar:

$$(AA^*)u_i = (U\Sigma\Sigma^t U^*)u_i = U\Sigma\Sigma^t e_i = U\Sigma\sigma_i e_i =$$
$$= U e_i \sigma_i^2 = \sigma_i^2 u_i$$

$$\Rightarrow (AA^*)u_i = \sigma_i^2 u_i$$

$\Rightarrow$ u_i son autovectores de AA^* con autovalor σ_i^2

Obs 3

$$\overset{m}{A} = U \Sigma \overset{n}{V^*} \iff AV = U\Sigma$$

$$\Rightarrow A\left(v_1 | v_2 | \dots | v_n\right) = \left(u_1 | u_2 | \dots | u_m\right)\begin{pmatrix} \sigma_1 & & \\ & \sigma_2 & \\ & & \ddots \end{pmatrix}$$

A la izquierda de la igualdad tenemos una matriz con columnas Av_i

Una columna de la matriz de la derecha de la igualdad la podemos obtener mediante

$$U \cdot col_i(\Sigma) = U\begin{pmatrix} 0 \\ \vdots \\ 0 \\ \sigma_i \\ 0 \\ \vdots \\ 0 \end{pmatrix} = U e_i \sigma_i = u_i \sigma_i$$

↳ σ_i en posición i-ésima

Luego, se cumple que $A v_i = \sigma_i u_i$

Notemos que $\{v_1, \ldots, v_n\}$ es una base de $\mathbb{K}^n$

$\Rightarrow$ tenemos definida la transformación lineal determinada por A sobre una base.

Obs (4) similar a la observación anterior consideremos A^*

$$A^* = (U \Sigma V^*)^* = V \Sigma^* U^* = V \Sigma^t U^*$$

$$\Rightarrow A^* = V \Sigma^t U^* \Leftrightarrow A^* U = V \Sigma^t$$

con un desarrollo igual que antes obtenemos

$$A^* u_i = \sigma_i v_i$$

Las 4 observaciones anteriores son propiedades que logramos deducir a partir de asumir la existencia de la SVD

Escribamos el teorema que establece la existencia de la descomposición y mostremos que podemos ir calculando cada uno de los componentes.

teo Dada una matriz $A \in \mathbb{K}^{n \times m}$ existen matrices $U \in \mathbb{K}^{n \times n}$ y $V \in \mathbb{K}^{m \times m}$ unitarias y $\Sigma \in \mathbb{R}^{n \times m}$ diagonal (con $\Sigma_{ii} = \sigma_i$ y $\sigma_i \geq \sigma_j \geq 0$ para $j \geq i$) tales que

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

Llamamos valor singular a los σ_i.

Observación: los valores singulares que se encuentran en la diagonal de Σ siempre están ordenados en orden decreciente y son mayores o iguales a 0.

Justifiquemos cómo podemos ir construyendo la descomposición.

A partir de la obs. 1 vimos que v_i y σ_i^2 deberían corresponderse a los autovectores y autovalores de A^*A. Los n vectores v_i deberían formar una base de K^n.
¿Podemos asegurar entonces que existe dicha base de autovectores?
La respuesta es sí. Pues A^*A es una matriz Hermitiana y vimos que (x Schur) estas matrices tienen base de autovectores.

Por otro lado, A*A es semidefinida positiva

Esto quiere decir que $\forall x \in \mathbb{K}^n$:

$$x^* (A^*A) x \geq 0$$

Efectivamente: $x^* A^* A x = (Ax)^* (Ax) =$
$= \|Ax\|_2^2 \geq 0$

En particular, si $x = v_i$ (es decir, es autovector de A*A) entonces

$$v_i^* A^* A v_i = v_i^* \lambda_i v_i = \lambda_i \underbrace{\|v_i\|_2^2}_{\text{asumimos autovectores de norma 1}} = \lambda_i \geq 0$$

$\Rightarrow$ Por Schur vimos que las matrices Hermitianas tienen autovalores reales y acabamos de ver que además, para A*A, son ≥ 0

$\Rightarrow$ Podemos tomar $\sigma_i = \sqrt{\lambda_i}$ siendo λ_i autovalor de A*A asociado a v_i.

Luego, ya tenemos la matriz V y Σ, formado V con los autovectores de A^*A y definiendo $\sigma_i = \sqrt{\lambda_i}$ (de esta forma se cumple lo que vimos en la obs ① que los σ_i^2 eran autovalores de A^*A)

Nos resta construir la matriz $U \in K^{m \times m}$

Si bien observamos que los u_i deben ser autovectores de AA^* en la obs ②, también debe cumplirse la obs ③:

$$\underline{A v_i = \sigma_i u_i}$$

Entonces definamos $u_i = \dfrac{A v_i}{\sigma_i}$ (para $\sigma_i \neq 0$)

Con esta definición se cumple:

- $u_i \perp u_j$ $(i \neq j)$

$$i \neq j: \quad u_i^* u_j = \left(\frac{A v_i}{\sigma_i}\right)^* \left(\frac{A v_j}{\sigma_j}\right) = \frac{v_i^* A^* A v_j}{\sigma_i \sigma_j} = \lambda_j \underbrace{\frac{v_i^* v_j}{\sigma_i \sigma_j}}_{v_i \perp v_j} = 0$$

- $\|u_i\|_2 = 1$

$$\|u_i\|_2^2 = u_i^* u_i = \frac{v_i^* A^* A v_i}{\sigma_i^2} = \frac{\overset{\sigma_i^2}{\overset{\|}{\lambda_i}}}{\sigma_i^2} \overbrace{v_i^* v_i}^{\overset{\|v_i\|_2^2 = 1}{\|}} = 1$$

- u_i es autovector de AA^* (se cumple obs 2)

sabemos que $A^*A v_i = \lambda_i v_i \Rightarrow$

(multiplico x A) $\Rightarrow A A^* \underbrace{(A v_i)}_{\sigma_i u_i} = \lambda_i \underbrace{(A v_i)}_{\sigma_i u_i}$

$\Rightarrow AA^* \sigma_i u_i = \lambda_i \sigma_i u_i$ $\qquad \sigma_i \neq 0$

$\Rightarrow AA^* u_i = \lambda_i u_i$

- Se cumple obs 4 : $A^* u_i = \sigma_i v_i$

Por como definimos u_i: $u_i = \dfrac{A v_i}{\sigma_i} \Rightarrow A v_i = \sigma_i u_i$

mult. por A^* $\Longrightarrow \underbrace{A^* A v_i}_{\lambda_i v_i} = \sigma_i A^* u_i \Rightarrow \underbrace{\frac{\lambda_i}{\sigma_i}}_{\sigma_i} v_i = A^* u_i$

$\sigma_i v_i$

Hasta aquí tenemos:

- $v_1 \dots v_m$ son autovectores de A^*A
- $\sigma_i = \sqrt{\lambda_i}$ con λ_i autovalor de A^*A asociado con autovector v_i
- $u_i = \dfrac{A v_i}{\sigma_i}$ para los $\sigma_i \neq 0$

Estos candidatos cumplen con todas las observaciones realizadas al comienzo

Solo nos resta definir los u_i restantes cuando $\sigma_i = 0$

¿Cómo obtenemos los restantes?

→ extendemos a una b.o.n para obtener una base de $\mathbb{K}^m$ y construir la matriz U.

Observemos que los u_i que restan definir están asociados al autovalor $\lambda_i = 0$

Llamemos r a la cantidad de λ_i's no nulos

$\Rightarrow \quad A^*A\, v_i = \lambda_i v_i$ con $\lambda_1 \dots \lambda_r > 0$

y $\lambda_{r+1} \dots \lambda_n = 0$

$\Rightarrow \quad u_i = \dfrac{A v_i}{\sigma_i}$ para $i = 1, \dots, r$

y $\{u_1, \dots, u_r, u_{r+1}, \dots, u_m\}$ es b.o.n

Observar que $u_1, \dots, u_r$ están en $\operatorname{Im}(A)$

Ejemplo $A = \begin{pmatrix} 1 & 1 \\ 0 & 0 \\ 1 & 1 \end{pmatrix}$ Busco $A = U \Sigma V^t$

Buscamos $U \in \mathbb{R}^{3\times 3}$, $\Sigma \in \mathbb{R}^{3\times 2}$, $V \in \mathbb{R}^{2\times 2}$

PASO 1: calculo V y Σ. Calculo autoval. y autovect. de A^tA

$$A^tA = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix}$$

$$0 = \det \begin{pmatrix} 2-\lambda & 2 \\ 2 & 2-\lambda \end{pmatrix} = (2-\lambda)^2 - 4 = \lambda^2 - 4\lambda + 4 - 4$$
$$= \lambda(\lambda - 4)$$

$\lambda_1 = 4$

$$\begin{pmatrix} -2 & 2 \\ 2 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \qquad \begin{matrix} -2x_1 + 2x_2 = 0 \\ x_1 = x_2 \end{matrix}$$

$$\tilde{v}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$\lambda_2 = 0$

$$\begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \qquad \begin{matrix} 2x_1 + x_2 = 0 \\ x_1 = -x_2 \end{matrix}$$

$$\tilde{v}_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

Normalizo autovectores

$$v_1 = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix} \qquad v_2 = \begin{pmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$$

$$\sigma_1 = \sqrt{\lambda_1} = 2 \qquad \sigma_2 = 0$$

$$\Rightarrow V = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \frac{1}{\sqrt{2}} \qquad \Sigma = \begin{pmatrix} 2 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$$

<u>PASO 2</u> calculo $U \in \mathbb{R}^{3\times 3}$

Definimos $u_1 = \dfrac{A v_1}{\sigma_1} = \dfrac{\begin{pmatrix} 1 & 1 \\ 0 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}}{2} =$

$$= \begin{pmatrix} 2/\sqrt{2} \\ 0 \\ 2/\sqrt{2} \end{pmatrix} \frac{1}{2} = \begin{pmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{pmatrix}$$

$$U = \left(u_1 \,\middle|\, u_2 \,\middle|\, u_3 \right)$$

Buscamos u_2 y u_3 tal que $\{u_1, u_2, u_3\}$ es b.o.n

$u_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} 1/\sqrt{2}$ puedo elegir $u_2 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} 1/\sqrt{2}$

$u_3 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} 1/\sqrt{2}$

<u>otra opción</u> Buscar $Nu(A^t)$

(ya que $\underbrace{Nu(A^t) = Im(A)^{\perp}}_{\langle u_2, u_3 \rangle} \oplus \underbrace{Im(A)}_{\langle u_1 \rangle} = \mathbb{R}^3$)

$$Nu(A^t) = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \,/\, \begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0 \right\}$$

$$\Rightarrow x + z = 0 \Rightarrow x = -z \Rightarrow \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -z \\ y \\ z \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} -z \\ y \\ z \end{pmatrix} = z \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} + y \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

$$\Rightarrow Nu(A^t) = \left\langle \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\rangle$$

$$\Rightarrow \begin{pmatrix} 1 & 1 \\ 0 & 0 \\ 1 & 1 \end{pmatrix} = \underbrace{\begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \\ 1/\sqrt{2} & -1/\sqrt{2} & 0 \end{pmatrix}}_{U} \underbrace{\begin{pmatrix} 2 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}}_{\Sigma} \underbrace{\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \frac{1}{\sqrt{2}}}_{V^t}$$

observar que $U\Sigma = \left(\sigma_1 u_1 \middle| \; 0 \right) = \begin{pmatrix} 2/\sqrt{2} & 0 \\ 0 & 0 \\ 2/\sqrt{2} & 0 \end{pmatrix}$

u_2 y u_3 se anulan en $U\Sigma$

Cabe mencionar que así como calculamos la SVD comenzando por definir V y Σ y luego definir U, podríamos haber ido por el otro "camino": comenzamos calculando autovalores y autovectores de AA^* para encontrar primero los u_i y σ_i y luego definir los v_i de la siguiente manera a partir de la observación 4: $A^* u_i = \sigma_i v_i$

$$\Rightarrow \text{definimos } v_i = \frac{A^* u_i}{\sigma_i}, \quad \sigma_i \neq 0$$

Igual que antes, para los $r_i = 0$ restan
les completamos con v_i para formar
una B.O.N.

Veamos en el ejemplo anterior

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 0 \\ 1 & 1 \end{pmatrix}$$

comenzamos buscando autoval / autovec. de AA^t

$$AA^t = \begin{pmatrix} 1 & 1 \\ 0 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 0 & 2 \\ 0 & 0 & 0 \\ 2 & 0 & 2 \end{pmatrix}$$

$$\Rightarrow \det \begin{pmatrix} 2-\lambda & 0 & 2 \\ 0 & -\lambda & 0 \\ 2 & 0 & 2-\lambda \end{pmatrix} = (2-\lambda) \begin{vmatrix} -\lambda & 0 \\ 0 & 2-\lambda \end{vmatrix} + 2 \begin{vmatrix} 0 & 2 \\ -\lambda & 0 \end{vmatrix}$$

$$= (2-\lambda)(-\lambda)(2-\lambda) + 4\lambda = \lambda\left(4-(2-\lambda)^2\right) =$$

$$= -\lambda\left((2-\lambda)^2 - 4\right) = -\lambda\left(\lambda(\lambda-4)\right) = 0$$

$$\text{si } \lambda = \begin{cases} 0 \\ 4 \end{cases}$$

para $\lambda = 4$: $\begin{pmatrix} -2 & 0 & 2 \\ 0 & -4 & 0 \\ 2 & 0 & -2 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$

$$\begin{cases} -2x + 2z = 0 \\ -4y = 0 \\ 2x - 2z = 0 \end{cases} \rightarrow \begin{array}{l} x = z \\ y = 0 \end{array} \rightarrow \begin{pmatrix} z \\ 0 \\ z \end{pmatrix} = z\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

$$\Rightarrow u_1 = \begin{pmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{pmatrix}, \quad \sigma_1 = \sqrt{4} = 2$$

para $\lambda = 0$ $\begin{pmatrix} 2 & 0 & 2 \\ 0 & 0 & 0 \\ 2 & 0 & 2 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$

$$\left\{ 2x + 2z = 0 \right. \rightarrow x = -z \rightarrow \begin{pmatrix} -z \\ y \\ z \end{pmatrix} = z\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} + y\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

$$\rightarrow u_2 = \begin{pmatrix} -1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{pmatrix}, \quad u_3 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \sigma_2 = 0$$

$\Rightarrow$ Tenemos $U = \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} & 0 \\ 0 & 0 & 1 \\ 1/\sqrt{2} & 1/\sqrt{2} & 0 \end{pmatrix}$

$$\Sigma = \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Definimos: $v_1 = \dfrac{A^t u_1}{\sigma_1} = \dfrac{\begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{pmatrix}}{2} =$

$$= \frac{\begin{pmatrix} 2/\sqrt{2} \\ 2/\sqrt{2} \end{pmatrix}}{2} = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$$

y para obtener v_2 extendemos a una b.o.n

$$v_2 = \begin{pmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix} \Rightarrow V = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \frac{1}{\sqrt{2}}$$

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

Pasos para calcular SVD de $A \in \mathbb{K}^{m\times n}$

$U \in \mathbb{K}^{m\times m}$, $\Sigma \in \mathbb{R}^{m\times n}$, $V \in \mathbb{K}^{n\times n}$

A matriz de rango r (r autoval. no nulos de A^*A)

1) (λ_i, x_i) pares de (autoval, autovec) de A^*A
2) $\sigma_i = \sqrt{\lambda_i}$, $v_i = x_i / \|x_i\|_2$
3) $u_i = \dfrac{A v_i}{\sigma_i}$ para $i = 1, \dots, r$
4) Extender a base ortonormal: $u_{r+1}, \dots, u_m$

1) (λ_i, x_i) pares de (autoval, autovec) de AA^*
2) $\sigma_i = \sqrt{\lambda_i}$, $u_i = x_i / \|x_i\|_2$
3) $v_i = \dfrac{A^* u_i}{\sigma_i}$ para $i = 1 \dots r$
4) Extender a base ortonormal: $v_{r+1}, \dots, v_n$

Teo Sea $A \in \mathbb{K}^{m \times n}$, entonces $\|A\|_2 = \sigma_1$

dem $\|A\|_2 = \max_{\|x\|_2=1} \|Ax\|_2 = \max_{\|x\|_2=1} \|U\Sigma V^* x\|_2 = \triangle$

recordemos que para las matrices unitarias U se cumple $\forall x: \|Ux\|_2 = \|x\|_2$ (no cambian la norma 2 de los vectores).

$$\Rightarrow \|U(\Sigma V^* x)\|_2 = \|\Sigma V^* x\|_2$$

$$\triangle = \max_{\|x\|_2=1} \|\Sigma V^* x\|_2 \overset{\circledcirc}{=} \max_{\|y\|_2=1} \|\Sigma y\|_2 =$$

($\circledcirc$: $V^* x = y \Leftrightarrow Vy = x$ y $\|x\|_2 = \|y\|_2$)

$$= \max_{\|y\|_2=1} \left\| \begin{pmatrix} \sigma_1 y \\ \sigma_2 y \\ \vdots \end{pmatrix} \right\|_2 = \max_{\|y\|_2=1} \sqrt{\sum_{i=1}^{n} \sigma_i^2 |y_i|^2} \leq$$

$$\leq \max_{\|y\|_2=1} \sqrt{\sigma_1^2 \sum_{i=1}^{n} |y_i|^2} = \sigma_1 \max_{\|y\|_2=1} \underbrace{\sqrt{\sum_{i=1}^{n} |y_i|^2}}_{\|y\|_2} = \sigma_1$$

Veamos que también se cumple la igualdad.

Sea v_1 el autovector de A^tA con su autovalor $\lambda_1 = \sigma_1^2$

$$\Rightarrow \|Av_1\|_2 = \sqrt{(Av_1)^*(Av_1)} = \sqrt{v_1^* A^* A\, v_1} =$$

$$= \sqrt{v_1^* v_1 \lambda_1} = \underbrace{\sqrt{\lambda_1}}_{\sigma_1} \underbrace{\sqrt{v_1^* v_1}}_{\|v_1\|_2 = 1} = \sigma_1$$

$\therefore \|A\|_2 = 1$

Otra propiedad para cualquier $A \in \mathbb{K}^{n \times n}$ (invertible)

$$cond_2(A) = \sigma_1 / \sigma_r$$

Recordemos que $cond_2(A) = \|A\|_2 \|A^{-1}\|_2$

Acabamos de ver que $\|A\|_2 = \sigma_1$

Por otro lado se ve claramente que los valores singulares de A^{-1} son $1/\sigma_1, 1/\sigma_2, \dots, 1/\sigma_n$, con $\sigma_1, \dots, \sigma_n$ valor singular de A. El mayor de los valores singulares es la norma 2 $\Rightarrow$ para A^{-1} el mayor de todos es $1/\sigma_n$

$\Rightarrow \|A^{-1}\|_2 = 1/\sigma_n$

$\Rightarrow \text{cond}_2(A) = \sigma_1/\sigma_n$

Recordemos que $\|A\|_F = \sqrt{\sum_{i,j} |a_{ij}|^2}$

Esta norma de Frobenius también se puede expresar en función de los valores singulares de A:

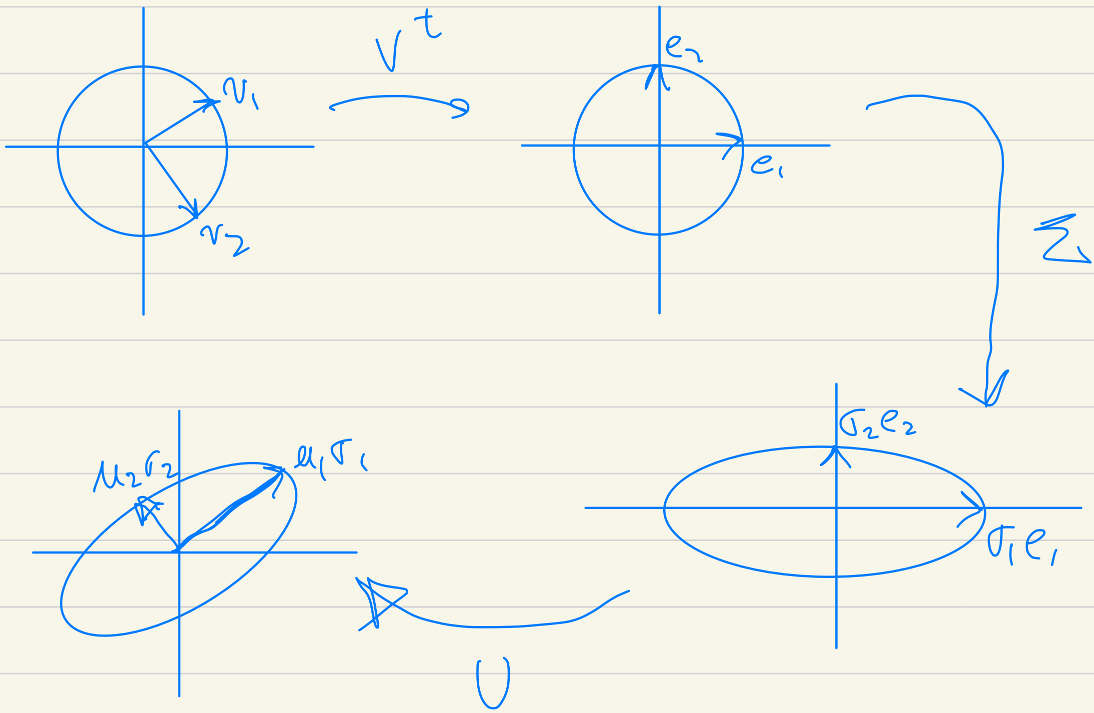
$$\|A\|_F = \sqrt{\sigma_1^2 + \sigma_2^2 + \dots + \sigma_r^2}$$

La norma de Frobenius es invariante ante matrices unitarias. Para cualquier $Q \in \mathbb{K}^{n \times n}$ unitaria y $B \in \mathbb{K}^{n \times n}$ se cumple: $\|B\|_F = \|QB\|_F = \|BQ\|_F$ (probarlo!!)

$$\Rightarrow \|A\|_F = \|U\Sigma V^*\|_F = \|\Sigma V^*\|_F = \|\Sigma\|_F$$

$$\|\Sigma\|_F = \sqrt{\sigma_1^2 + \dots + \sigma_r^2}\,.$$

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



$$A = \left(u_1 \mid u_2\right) \begin{pmatrix}\sigma_1 & 0 \\ 0 & \sigma_2\end{pmatrix} \begin{pmatrix} v_1^t \\ \hline v_2^t \end{pmatrix}$$

$$\Rightarrow Ax = U\Sigma V^t x = \left(u_1\sigma_1 \mid u_2\sigma_2\right)\begin{pmatrix} v_1^t x \\ v_2^t x\end{pmatrix} = u_1 \underbrace{\sigma_1 (v_1^t x)}_{\alpha_1} + u_2 \underbrace{\sigma_2 (v_2^t x)}_{\alpha_2}$$

teo Sea $A \in \mathbb{K}^{n \times n}$ inversible y $A = U \Sigma V^*$ su SVD.

Definimos $$\hat{\Sigma} = \begin{pmatrix} \sigma_1 & & & & \\ & \sigma_2 & & & \\ & & \ddots & & \\ & & & \sigma_{n-1} & \\ & & & & 0 \end{pmatrix}$$

Entonces la matriz $\hat{A} = U \hat{\Sigma} V^*$ es la matriz singular más cercana a A en norma 2.

Es decir, $\|A - \hat{A}\|_2 = \min\limits_{B \text{ singular}} \|A - B\|_2$

Además $\|A - \hat{A}\|_2 = \sigma_n$

dem Tratemos de encontrar una matriz $B \in \mathbb{K}^{n \times n}$ tal que $A + B$ sea singular y $\|B\|_2$ sea lo más chica posible.

Veamos que $A + B = A(I + A^{-1}B)$

Consideremos el siguiente lema

Sea $C \in \mathbb{K}^{n\times n}$. Si $\|C\| < 1 \Rightarrow I + C$ es inversible.

dem/ Supongamos que $I+C$ no es inversible

Entonces $\exists x \in \mathbb{R}^n$, $x \neq 0$ tal que $(I+C)x = 0$

$\Rightarrow x + Cx = 0 \Rightarrow Cx = -x$

$\Rightarrow \|Cx\| = \|-x\| = \|x\|$

$\Rightarrow \dfrac{\|Cx\|}{\|x\|} = 1 \Rightarrow \|C\| \geq 1$ abs!

(Consideramos $\|\cdot\| : \mathbb{K}^{n\times n} \to \mathbb{R}_{\geq 0}$ cualquier norma inducida)

Usando el lema anterior,

si $\|A^{-1}B\| < 1 \Rightarrow I + A^{-1}B$ es inversible

y entonces también lo será $A+B$

Por el contrario, si $A+B$ es singular y A es inversible $\Rightarrow I+A^{-1}B$ debe ser singular.

Pero $I+A^{-1}B$ singular $\Rightarrow \|A^{-1}B\| \geq 1$

$\Rightarrow 1 \leq \|A^{-1}B\| \leq \|A^{-1}\| \, \|B\|$

↳ prop. submultiplic. de norma inducida

$\Rightarrow \|B\| \geq \dfrac{1}{\|A^{-1}\|}$

Luego llegamos a que para que $A+B$ sea singular debe cumplirse $\|B\| \geq \dfrac{1}{\|A^{-1}\|}$

Sabemos que $\|A^{-1}\|_2 = \dfrac{1}{\sigma_n} \Rightarrow \dfrac{1}{\|A^{-1}\|} = \sigma_n$

siendo σ_n el n-ésimo valor singular de A.

$\rightarrow$ Debe ser $\underline{\|B\| \geq \sigma_n}$

Construyamos una B que cumpla la igualdad.

Sea $B = U E V^*$ con U y V las matrices de la SVD de A.

$$E = \begin{pmatrix} 0 & & & \\ & \ddots & & \\ & & 0 & \\ & & & -\sigma_n \end{pmatrix}$$

$$\|B\|_2 = \|E\|_2 = \sigma_n$$

$$\Rightarrow A+B = U(\Sigma + E)V^* = U \hat{\Sigma} V^*$$

Además se cumple $\|A-(A+B)\|_2 = \sigma_n$

Ejemplo: ¿Cuál es la matriz singular más cercana a $\begin{pmatrix} -2 & 0 & 0 \\ 0 & -5 & 0 \\ 0 & 0 & 3 \end{pmatrix}$?

Podemos calcular la SVD y aplicar el teorema.

Consideremos una matriz de permutación que "ordene" la diagonal (en módulo)

Sea $P = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \Rightarrow P^{-1} = P^t = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$

$$\Rightarrow P^t \begin{pmatrix} -2 & 0 & 0 \\ 0 & -5 & 0 \\ 0 & 0 & 3 \end{pmatrix} P = P^t \begin{pmatrix} 0 & 0 & -2 \\ -5 & 0 & 0 \\ 0 & 3 & 0 \end{pmatrix} = \begin{pmatrix} -5 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -2 \end{pmatrix}$$

$$\Rightarrow P^t \begin{pmatrix} -2 & 0 & 0 \\ 0 & -5 & 0 \\ 0 & 0 & 3 \end{pmatrix} P = \begin{pmatrix} -5 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -2 \end{pmatrix} \Leftrightarrow \boxed{\begin{pmatrix} -2 & 0 & 0 \\ 0 & -5 & 0 \\ 0 & 0 & 3 \end{pmatrix} = P \begin{pmatrix} -5 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -2 \end{pmatrix} P^t}$$

Por otro lado: $\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

$$\Rightarrow \begin{pmatrix} -2 & 0 & 0 \\ 0 & -5 & 0 \\ 0 & 0 & 3 \end{pmatrix} = \underbrace{P \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}} \underbrace{\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} -5 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -2 \end{pmatrix}} P^t =$$

$$= \begin{pmatrix} 0 & 0 & -1 \\ -1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} = U \Sigma V^t$$

Luego, ya tenemos una SVD para $\begin{pmatrix} -2 & 0 & 0 \\ 0 & -5 & 0 \\ 0 & 0 & 3 \end{pmatrix}$

Obs: podríamos haber calculado la SVD por el camino tradicional: calculando autovalores y autovectores de A^tA para calcular V y Σ y luego definiendo los valores para U.

$$\Rightarrow \begin{pmatrix} -2 & 0 & 0 \\ 0 & -5 & 0 \\ 0 & 0 & 3 \end{pmatrix} = \begin{pmatrix} 0 & 0 & -1 \\ -1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} = U\Sigma V^t$$

$\Rightarrow$ aplicando el teorema:

$$\hat{A} = \begin{pmatrix} 0 & 0 & -1 \\ -1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \underbrace{\begin{pmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 0 \end{pmatrix}}_{\hat{\Sigma}} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -5 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

es la matriz singular más cercana a A

y $\|A - \hat{A}\|_2 = 2 = \sigma_3$

Este resultado se puede generalizar para obtener la matriz más cercana a otra de un rango específico.

Teo Sea $A \in \mathbb{K}^{m\times n}$ de rango $r>0$, con $A = U\Sigma V^*$ su SVD. La matriz $A_k \in \mathbb{K}^{m\times n}$ de rango $k<r$ más cercana a A (en norma 2) es $A_k = U\Sigma_k V^*$, con

$$\Sigma_k = \begin{pmatrix} \sigma_1 & & & & & \\ & \sigma_2 & & & & \\ & & \ddots & & & \\ & & & \sigma_k & & \\ & & & & 0 & \\ & & & & & \ddots \\ & & & & & & 0 \end{pmatrix}$$

Es decir,

$$\sigma_{k+1} = \|A - A_k\|_2 = \min\{\|A-B\|_2 \mid \text{rango}(B) \leq k\}$$

→ Σ_k es la matriz que resulta de poner en cero los elementos de la diagonal de Σ de la posición $(k+1, k+1)$ en adelante. El resto de los valores singulares desde 1 a k quedan iguales.