7.4 Décomposition en valeurs singulières (SVD)

Dans le cas plus général des matrices rectangulaires $A \in M_{m\times n}(\mathbb{R})$, il n'existe pas de décomposition $A = PDP^{-1}$ avec P inversible et D diagonale. Mais on peut factoriser A en

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

où Σ est *diagonale par blocs* et U et V sont orthogonales.

$U \in M_{m\times m}(\mathbb{R})$
$V \in M_{n\times n}(\mathbb{R})$
$\Sigma \in M_{m\times n}(\mathbb{R})$

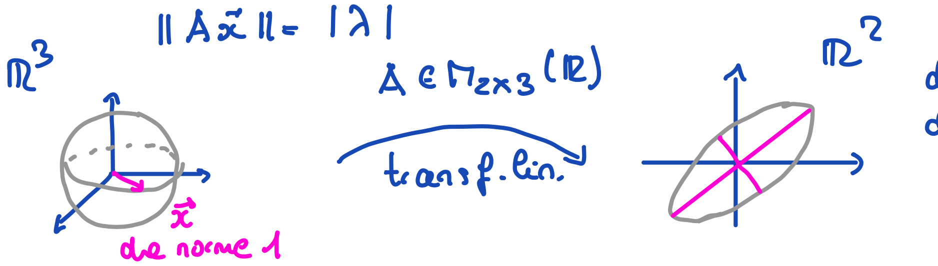
Valeurs singulières

Soit $A \in M_{n\times n}(\mathbb{R})$ une matrice symétrique. La valeur absolue des valeurs propres de A mesure la façon dont A étire ou comprime les vecteurs propres : Si $A\vec{x} = \lambda\vec{x}$ alors

$$\|A\vec{x}\| = \|\lambda\vec{x}\| = |\lambda|\cdot\|\vec{x}\|$$

et pour $\|\vec{x}\| = 1$, on aura

$$\|A\vec{x}\| = |\lambda|$$



détermination de valeur max. et min.

Soit $A \in M_{m\times n}(\mathbb{R})$ une matrice rectangulaire. On sait que $\overset{\in M_{n\times n}}{A^TA}$ est symétrique, donc diagonalisable en base orthogonale. Par conséquent, il existe une base $(\vec{v}_1, \dots, \vec{v}_n)$ orthonormale de $\mathbb{R}^n$ formée de vecteurs propres de A^TA.

Soient $\lambda_1, \dots, \lambda_n$ les valeurs propres correspondantes. Il se peut qu'on ait $\lambda_i = \lambda_j$ pour certaines valeurs de i et de j.

On a

$$A^TA\,\vec{v}_i = \lambda_i\vec{v}_i$$

$$\|A\vec{v}_i\|^2 = A\vec{v}_i \overset{\text{prod. scalaire}}{\cdot} A\vec{v}_i = (A\vec{v}_i)^{\overset{\text{produit de matrice}}{T}}(A\vec{v}_i)$$

$$= (\vec{v}_i^T A^T)(A\vec{v}_i) = \vec{v}_i^T A^TA\vec{v}_i$$

$$= \lambda_i \vec{v}_i^T\vec{v}_i = \lambda_i \underbrace{\vec{v}_i\cdot\vec{v}_i}_{=1} = \lambda_i$$

Donc $\lambda_i = \|A\vec{v}_i\|^2 \geq 0 \quad \forall\, 1 \leq i \leq n$

on renumérote les valeurs propres dans l'ordre décroissant

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \geq 0.$$

Définition 72 (valeurs singulières)**.**
Soit $A \in M_{m\times n}(\mathbb{R})$ une matrice. Les *valeurs singulières* de A sont les racines des valeurs propres de $A^T A$. On les note

$$\sigma_1, \dots, \sigma_n \quad \text{avec} \quad \sigma_i = \sqrt{\lambda_i} \quad \forall 1 \leq i \leq n$$

On a $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n \geq 0.$

Remarques

Comme $\lambda_i = \|A\vec{v}_i\|^2$ où $\vec{v}_i$ est un vecteur propre associé à λ_i, valeur propre de A^TA, on

$$\sigma_i = \|A\vec{v}_i\|$$

Donc les valeurs singulières de A sont les longueurs des vecteurs $A\vec{v}_1, \dots, A\vec{v}_n$.

Exemple $A = \begin{pmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{pmatrix}$ $\quad A^TA = \begin{pmatrix} 9 & -9 \\ -9 & 9 \end{pmatrix}$

$\Rightarrow \lambda \in \{0, 18\}$

$\lambda_1 = 18 \geq 0$ et les valeurs singulières de A

$\lambda_2 = 0$ sont $\sigma_1 = \sqrt{18} = 3\sqrt{2}$ et $\sigma_2 = 0$.

$\operatorname{Ker}(A^TA - 18 I_2) = \operatorname{span}\left\{\begin{pmatrix} 1 \\ -1 \end{pmatrix}\right\} \Rightarrow \vec{v}_1 = \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix}$

$\operatorname{Ker}(A^TA) = \operatorname{span}\left\{\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right\} \Rightarrow \vec{v}_2 = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$

$A\vec{v}_1 = \begin{pmatrix} -1 & 1 \\ -2 & 2 \\ 2 & -2 \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix} = \frac{2}{\sqrt{2}} \begin{pmatrix} 1 \\ -2 \\ 2 \end{pmatrix}$ $\qquad \|A\vec{v}_1\| = 3\sqrt{2} = \sigma_1$

$A\vec{v}_2 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ $\qquad \|A\vec{v}_2\| = 0 = \sigma_2$

Théorème 76. *Soit $A \in M_{m\times n}(\mathbb{R})$ une matrice et $(\vec{v}_1, \dots, \vec{v}_n)$ une base orthonormée de $\mathbb{R}^n$ formée de vecteurs propres de A^TA. Soient $\lambda_1 \geq \cdots \geq \lambda_n \geq 0$ les valeurs propres correspondantes. Supposons que A possède r valeurs singulières non nulles. Alors*

$(A\vec{v}_1, \dots, A\vec{v}_r)$ est une base orthogonale de $\mathrm{Im}(A)$ et $\mathrm{rang}(A) = r$.

Preuve: Soit $1 \leq i \neq j \leq r$

- $\vec{v}_i \cdot \vec{v}_j = 0$. Donc $A\vec{v}_i \cdot A\vec{v}_j = (A\vec{v}_i)^T A\vec{v}_j$

  $= \vec{v}_i^{\,T} A^T A \vec{v}_j = \lambda_j \vec{v}_i^{\,T} \vec{v}_j = \lambda_j \vec{v}_i \cdot \vec{v}_j = 0$

  Donc $(A\vec{v}_1, \dots, A\vec{v}_r)$ est une famille orthogonale.

- Pour $1 \leq i \leq r$, on a $\lambda_i \neq 0$ par hypothèse.

  Donc $\|A\vec{v}_i\| = \sigma_i \neq 0$ et donc $A\vec{v}_i \neq \vec{0}$.

$\Rightarrow$ $A\vec{v}_1, \dots, A\vec{v}_r$ sont lin. indépendants appartenant a $\mathrm{Im}(A)$

- $\vec{w} \in \mathrm{Im}(A)$. Alors $\vec{w} = A\vec{x}$ pour un $\vec{x} \in \mathbb{R}^n$.

  $\vec{x}$ peut s'écrire $\vec{x} = \sum_{i=1}^{n} \alpha_i \vec{v}_i$.

  $$\vec{w} = A\vec{x} = A\left(\sum_{i=1}^{n} \alpha_i \vec{v}_i\right) = \sum_{i=1}^{n} \alpha_i A\vec{v}_i = \sum_{i=1}^{r} \alpha_i A\vec{v}_i$$

  car les valeurs sing. λ_i avec $i > r$ sont supposées nulles.

  Donc $\vec{w} \in \mathrm{span}\{A\vec{v}_1, \dots, A\vec{v}_r\}$

Remarque Pratiquement, pour estimer le rang d'une matrice de grande taille, on compte les valeurs sing. non nulles. On considère comme nulles les valeurs propres très petites pour éviter les problèmes d'arrondi.

Décomposition en valeurs singulières (SVD)

Théorème 77. *Soit $A \in M_{m\times n}(\mathbb{R})$ une matrice de rang r. Il existe une matrice $\Sigma \in M_{m\times n}(\mathbb{R})$ diagonale par blocs, une matrice $U \in M_{m\times m}(\mathbb{R})$ orthogonale et $V \in M_{n\times n}(\mathbb{R})$ orthogonale telles que*

$$A = U\Sigma V^T.$$

De plus, Σ est de la forme

$$\Sigma = \left(\begin{array}{c|c} D & 0 \\ \hline \underbrace{0}_{r} & \underbrace{0}_{n-r} \end{array}\right)\begin{array}{l} \}\, r \\ \}\, m-r \end{array}$$

où $D \in M_{r\times r}(\mathbb{R})$ est diagonale avec les valeurs singulières $\sigma_1 \geq \cdots \sigma_r > 0$ non nulles sur la diagonale.

Remarque La décomposition SVD de A n'est pas unique, mais Σ l'est.

Preuve Soit $(\vec{v}_1, \dots, \vec{v}_n)$ une base orthonormée de $\mathbb{R}^n$ formée de vecteurs propres de A^TA. Le th. 76 affirme que

$(A\vec{v}_1, \dots, A\vec{v}_r)$ est une base orthogonale de $\mathrm{Im}(A)$ $(r = \mathrm{rang}(A))$.

Soit $$\vec{u}_i = \frac{1}{\|A\vec{v}_i\|} A\vec{v}_i = \frac{A\vec{v}_i}{\sigma_i} \qquad (1 \leq i \leq r)$$

(donc $A\vec{v}_i = \sigma_i \vec{u}_i$).

On complète ensuite en une base orthonormée de $\mathbb{R}^m$:

$$(\vec{u}_1, \dots, \vec{u}_r, \vec{u}_{r+1}, \dots, \vec{u}_m)$$

et on pose $U = (\vec{u}_1 \dots \vec{u}_m) \in M_{m \times m}(\mathbb{R})$.

U est une matrice orthogonale. On pose aussi:

$V = (\vec{v}_1, \dots, \vec{v}_n)$ (base orthonormée formée des vecteurs propres de A^TA).

V est aussi une matrice orthogonale, $V \in M_{n \times n}(\mathbb{R})$

$$AV = (A\vec{v}_1 \dots A\vec{v}_n) = (\sigma_1 \vec{u}_1 \dots \sigma_r \vec{u}_r \underbrace{0 \dots 0}_{n-r}).$$

Définissons $\Sigma = \begin{array}{c} \left(\begin{array}{c|c} D & 0 \\ \hline 0 & 0 \end{array}\right) \begin{array}{l} \}r \\ \}m-r \end{array} \\ \begin{array}{cc} \underbrace{}_{r} & \underbrace{}_{n-r} \end{array} \end{array}$ avec D diagonale avec

$\sigma_1, \dots, \sigma_r$ sur la diagonale

$$\underset{m \times m}{U} \underset{m \times n}{\Sigma} = (\underbrace{\sigma_1 \vec{u}_1 \dots \sigma_r \vec{u}_r}_{r} \underbrace{0 \dots 0}_{n-r}) = AV.$$ Comme U est orth.

on a donc $A \underbrace{VV^T}_{I_n} = U \Sigma V^T$

Exemple

$A = \begin{pmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{pmatrix}$ $\left.\begin{array}{l} \sigma_1 = 3\sqrt{2} \\ \sigma_2 = 0 \end{array}\right\}$

- Les vecteurs propres correspondants sont

$\vec{v}_1 = \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix}$ et $\vec{v}_2 = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$

$\| A\vec{v}_1 \| = 3\sqrt{2}$ $\| A\vec{v}_2 \| = 0$

- Comme U sera 3×3, on doit compléter $(A\vec{v}_1)$ en une base orthogonale de $\mathbb{R}^3$:

$$\vec{u}_1 = \frac{1}{\sigma_1} A\vec{v}_1 = \frac{1}{3\sqrt{2}} \begin{pmatrix} 2/\sqrt{2} \\ -4/\sqrt{2} \\ 4/\sqrt{2} \end{pmatrix} = \begin{pmatrix} 1/3 \\ -2/3 \\ 2/3 \end{pmatrix}.$$

On cherche $\vec{u}_2$ et $\vec{u}_3$ avec $\vec{u}_1 \cdot \vec{u}_2 = 0$ et $\vec{u}_2 \cdot \vec{u}_3 = 0$.

$\vec{u}_1 \cdot \vec{u}_3 = 0$

$\vec{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ avec $\vec{x} \cdot \vec{u}_1 = 0$ satisfait $\frac{1}{3} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \begin{pmatrix} 1 \\ -2 \\ 2 \end{pmatrix} = 0$

$\Leftrightarrow x - 2y + 2z = 0$ pour $\vec{u}_2$ et $\vec{u}_3$.

Prenons $\vec{u_2} = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} \cdot \frac{1}{\sqrt{5}} = \begin{pmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \\ 0 \end{pmatrix}$

"$\vec{u_3}$" $= \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix} \cdot \frac{1}{\sqrt{5}} = \begin{pmatrix} -2/\sqrt{5} \\ 0 \\ 1/\sqrt{5} \end{pmatrix}$

Comme $\vec{u_2} \cdot \vec{u_3} \neq 0$, on applique GS sur $(\vec{u_2}, \vec{u_3})$

d'où $\vec{u_3} = \begin{pmatrix} -2/\sqrt{5} \\ 0 \\ 1/\sqrt{5} \end{pmatrix} - \dfrac{\begin{pmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \\ 0 \end{pmatrix} \begin{pmatrix} -2/\sqrt{5} \\ 0 \\ 1/\sqrt{5} \end{pmatrix}}{\begin{pmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \\ 0 \end{pmatrix} \begin{pmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \\ 0 \end{pmatrix}} \begin{pmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \\ 0 \end{pmatrix}$

$$= \begin{pmatrix} -2/\sqrt{5} \\ 0 \\ 1/\sqrt{5} \end{pmatrix} - \frac{4/5}{1} \begin{pmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \\ 0 \end{pmatrix} = \frac{1}{\sqrt{5}} \begin{pmatrix} -2/5 \\ 4/5 \\ 1 \end{pmatrix}$$

normaliser

$$\Rightarrow \quad U = \begin{pmatrix} 1/3 & 2/\sqrt{5} & -2/\sqrt{45} \\ -2/3 & 1/\sqrt{5} & 4/\sqrt{45} \\ 2/3 & 0 & 5/\sqrt{45} \end{pmatrix} \qquad V = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}$$

$$\Sigma = \left(\begin{array}{cc|c} D & 0 & 0 \\ 0 & 0 & \\ \hline \multicolumn{2}{c|}{0} & 0 \end{array}\right) \qquad D \in M_{1\times1}(\mathbb{R}) \quad D = (3\sqrt{2})$$

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

On appelle $(\vec{u_1}, \dots, \vec{u_m})$ les vecteurs singuliers à gauche et $(\vec{v_1}, \dots, \vec{v_n})$ " " " à droite.

7.5. Caractérisation des matrices inversibles et pseudo-inverse

---

$(\vec{u_1}, \dots, \vec{u_r})$ est une base orthonormée de $Im(A)$.

Comme $Im(A)^{\perp} = Ker(A^T)$ (th. 59),

$(\vec{u}_{r+1}, \dots, \vec{u_m})$ est une base orthonormée de $Ker(A^T)$.

$\| A \vec{v_i} \| = \sigma_i$ pour $1 \leq i \leq \sigma_n$ et $\sigma_i = 0$ pour $i > r$.

$\Rightarrow$ $\vec{v}_{r+1}, \dots, \vec{v_n}$ engendrent un SEV de $Ker(A)$ de dim. égale à $n-r$.

Par le th. du rang, $\dim Ker(A) = n - rang(A) = n-r$.

$\Rightarrow$ $(\vec{v}_{r+1}, \dots, \vec{v}_n)$ est une base orthonormée de Ker (A).

$$(\underbrace{\vec{u}_1, \dots, \vec{u}_r}_{\text{Im}(A)}, \underbrace{\vec{u}_{r+1}, \dots, \vec{u}_m}_{\text{Ker}(A^T)})$$

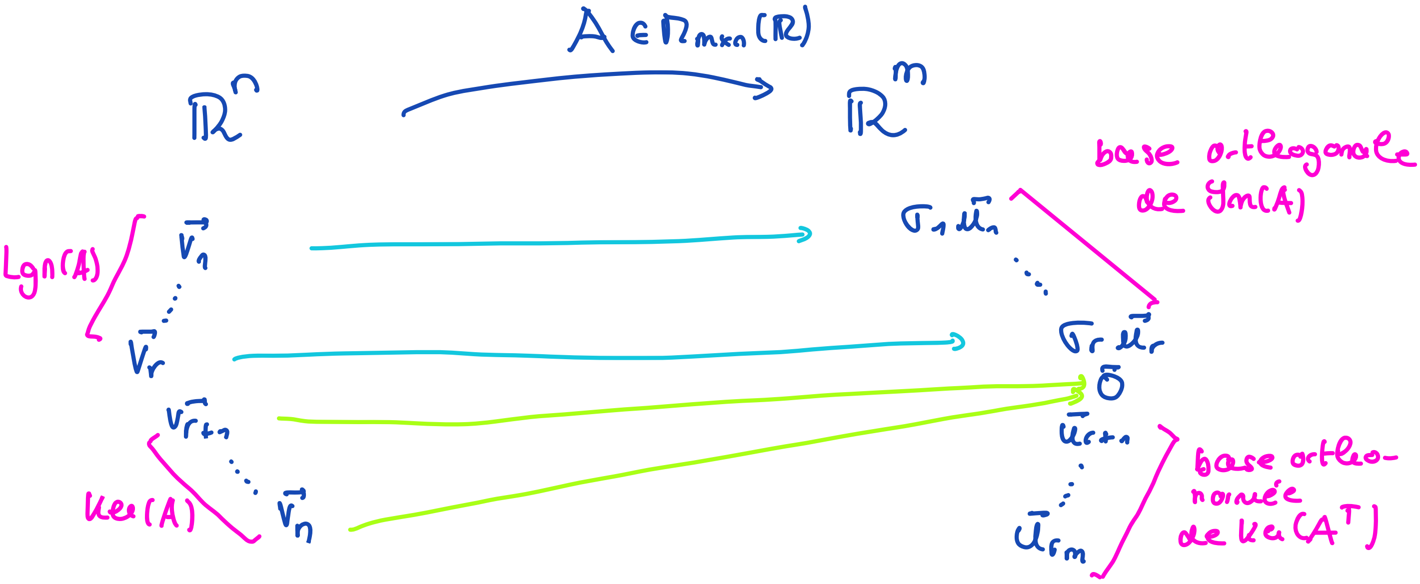
$$(\underbrace{\vec{v}_1, \dots, \vec{v}_r}_{\text{Lgn}(A)}, \underbrace{\vec{v}_{r+1}, \dots, \vec{v}_n}_{\text{Ker}(A)})$$

$\Rightarrow$ $(\vec{v}_1, \dots, \vec{v}_r)$ est une base orthonormée de Lgn(A).

Th. 59:

$$\text{Ker}(A) = \text{Im}(A^T)^{\perp} = \text{Lgn}(A)^{\perp}$$

$$\text{Ker}(A)^{\perp} = \text{Lgn}(A)$$



Th. 78 Soit $A \in M_{n\times n}(\mathbb{R})$. Les propriétés suivantes sont équivalentes :

(1) A est inversible

(2) $(\text{Im}(A))^{\perp} = \{0\}$

(3) $(\text{Ker}(A))^{\perp} = \mathbb{R}^n$

(4) $\text{Lgn}(A) = \mathbb{R}^n$

(5) A admet n valeurs singulières non nulles.

Ecrivons la décomposition SVD par blocs pour $A \in M_{m\times n}(\mathbb{R})$

$$A = U \Sigma V^T \quad \text{avec} \quad \Sigma = \left(\begin{array}{c|c} D & 0 \\ \hline 0 & 0 \end{array}\right)$$

$$= (U_r | U_{m-r}) \left(\begin{array}{c|c} D & 0 \\ \hline 0 & 0 \end{array}\right) \left(\begin{array}{c} V_r^T \\ \hline V_{n-r}^T \end{array}\right)$$

$= U_r D V_r^T$ (décomposition en valeurs singulières <u>réduite</u> de A).

Les coeff. diagonaux de D sont non nuls ⇒ D est inversible. On définit

$$A^+ = V_r D^{-1} U_r^T$$

qui est appelée <u>pseudo-inverse</u> de A.

$$AA^+ = (U_r D V_r^T)(V_r D^{-1} U_r^T) =$$

$$= U_r D \underbrace{V_r^T V_r}_{I_r} D^{-1} U_r^T = U_r \underbrace{D D^{-1}}_{I_r} U_r^T$$

$$= U_r U_r^T$$

<u>Lien avec les moindres carrés:</u>

$A\vec{x} = \vec{b}$

Posons $\hat{x} = A^+ \vec{b}$. On a alors

$$A\hat{x} = AA^+\vec{b} = U_r U_r^T \vec{b}$$

$U_r U_r^T$ est la projection orthogonale de $\vec{b}$ sur $\mathrm{Im}(A)$.
(th. 66 car $(\vec{u}_1, \dots, \vec{u}_r)$ est une base orthonormée de $\mathrm{Im}(A)$).

⇒ $\hat{x}$ est solution au sens des moindres carrés de $A\vec{x} = \vec{b}$.