1 Rappels d'algèbre linéaire

Matrice de permutation

$$P_\sigma := \left(\sum_{j=1}^{n} E_{\sigma(i),j}\right)_i$$

avec $E_{i,j}$ la matrice élémentaire (i, j).

Théorème de Schur $\forall A \in \mathcal{M}_n(\mathbb{C}), \exists U \in \mathcal{M}_n(\mathbb{C}), \underbrace{U^* = U^{-1}}_{\text{unitaire}} \wedge U^*AU \in \mathcal{T}_{n,s}(\mathbb{C})$

1.1 Méthodes itératives

Principe On cherche une suite $x^{(p)}$ de $\mathbb{R}^n$ convergeant vers la solution de $Ax = b$:

$$\forall x^{(0)} \quad x^{(p+1)} = H(x^{(p)})$$

Propriétés

- A n'est jamais modifiée
- Problème du suivi de la convergence et du choix du test d'arrêt
- Solution obtenue inexacte
- Matrice doit vérifier conditions de convergence
- Vitesse de convergence dépend de la matrice

2 Décomposition en valeurs singulières

Objectif Soit $A \in \mathcal{M}_n(\mathbb{R})$, $n \geq 1$.

On pose $A = U\Sigma V^\top$ avec $\begin{cases} U \in \mathcal{O}_n(\mathbb{R}) \\ V \in \mathcal{O}_n(\mathbb{R}) \\ \Sigma \in \mathcal{M}_n(\mathbb{R}) \end{cases}$

On a

$$\Sigma = \left(\begin{array}{ccc|ccc} \sigma_i & 0 & \cdots & 0 & \cdots & 0 \\ 0 & \sigma_i & \cdots & 0 & \cdots & 0 \\ 0 & 0 & \ddots & 0 & \cdots & 0 \\ \hline 0 & 0 & \cdots 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots 0 & 0 & \cdots & 0 \end{array}\right)$$

2.1 SVD d'une matrice

SVD Singular value decomposition

Soit $A \in \mathcal{M}_n(\mathbb{R})$ telle que $rgA \geq 1$

2.1.1 Propriétés

1. A est symétrique réelle semi-définie-positive i.e. $\forall x \in \mathbb{R}^n,\ x^\top(A^\top A)x \geq 0$ i.e. $\forall x \in \mathbb{R}^n,\ \|Ax\|^2 \geq 0$
2. $\mathrm{Sp}(A) \subset \mathbb{R}_+$
3. $A^\top A$ est orthoDZ: $\begin{cases} \ker(A^\top A) & = \ker A \\ \mathrm{im}(A^\top A) & = \mathrm{im}(A^\top) \end{cases}$

   Preuve ***(de 3.)***

   - $\ker A \subset \ker(A^\top A)$: $\mathbb{TRIVIAL}$
   - $\ker A \supset \ker(A^\top A)$ Soit $x \in \ker(A^\top A)$.
     On a
     $$\begin{aligned} & A^\top Ax = 0 \\ \implies & x^\top(A^\top Ax) = 0 \\ \implies & \|Ax\|^2 = 0 \\ \implies & x \in \ker A \end{aligned}$$
   - $\mathrm{im}\, A \subset \mathrm{im}(A^\top A)$
     $$\begin{aligned} \dim\ker(A^\top A) + \mathrm{rg}(A^\top A) = n & \\ \implies \mathrm{rg}(A^\top A) &= n - \dim\ker(A^\top A) \\ &= n - \dim\ker A \qquad \text{car } \ker A = \ker(A^\top A) \\ &= \mathrm{rg}\, A \\ &= \mathrm{rg}(A^\top) \\ \implies \mathrm{im}(A^\top A) &\subset \mathrm{im}\, A \end{aligned}$$
   - $\mathrm{im}\, A \supset \mathrm{im}(A^\top A)$: $\mathbb{TRIVIAL}$
4. $A^\top A \in GL_n(\mathbb{R}) \iff \mathrm{rg}\, A = n$

   Preuve ***(de 4.)***

   $$\begin{aligned} A^\top A \in GL_n(\mathbb{R}) &\iff \ker(A^\top A) = \{0\} \\ &\iff \ker A = \{0\} \\ &\iff \mathrm{rg}\, A = n \qquad \text{par théorème du rang} \end{aligned}$$

2.2 Construction de la SVD de A

Soit $A \in \mathcal{M}_{m,n}(\mathbb{R})$ tel que $\operatorname{rg} A \geq 1$. On note $r := \operatorname{rg} A$.

- $\dim \ker(A^\top A) = n - r$ par théorème du rang.
  Donc $0 \in \operatorname{Sp}(A^\top A)$ et $\operatorname{mult}_{A^\top A}(0) = n - r$
- On note $(\lambda_i)_{i \in [\![1,n]\!]} = \operatorname{Sp}(A^\top A) \subset \mathbb{R}_+^*$, **telles que** $i < j \implies \lambda_i \leq \lambda_j$
  On note $(v_i)_{i \in [\![1,r]\!]}$ une bond DZante de $A^\top A$ associée à $(\lambda_i)_{i \in [\![1,n]\!]}$
  Et on pose enfin

$$\mathcal{E} = \left(\underbrace{v_1, \dots, v_r}_{\text{bon de } (\ker A)^\top}, \underbrace{v_{r+1}, \dots, v_n}_{\text{bon de } \ker A} \right)$$

- On pose $V = (v_1, \dots, v_r) \in \mathcal{O}_n(\mathbb{R})$ Et

$$\forall i \in [\![1,n]\!] u_i = \frac{1}{\sqrt{\lambda_i}} A v_i \in \mathbb{R}^n$$

$$\begin{aligned}
\forall (i,j) \in [\![1,r]\!]^2 \quad \langle U_i, U_j \rangle &= \frac{1}{\sqrt{\lambda_i \lambda_j}} \langle A v_i, A v_j \rangle \\
&= \frac{1}{\sqrt{\lambda_i \lambda_j}} \left\langle \underbrace{A^\top A v_i}_{\lambda_i v_i}, A^\top A v_j \right\rangle \\
&= \sqrt{\frac{\lambda_i}{\lambda_j}} \underbrace{\langle v_i, v_j \rangle}_{\delta_{ij}} && \text{car } (v_i)_i \text{ orthonormée} \\
&= \delta_{ij} && \text{car si } i = j \text{, ça fait 1, sinon ça fait 0}
\end{aligned}$$

Donc $(U_i)_i$ est une famille orthonormée de $\mathbb{R}^m$
Donc $\operatorname{rg}(U_i)_{i \in [\![1,n]\!]} = r$
Or $(u_i)_i \subset \operatorname{im} A$
D'où $(u_i)_i \subset \operatorname{im} A$ et $(\operatorname{rg} u_i)_i = \operatorname{rg} A$
Donc $(U_i)_i = \operatorname{im} A$
et $(U_i)_i$ bon de $\operatorname{im} A$.
D'où $0 \in \operatorname{Sp}(A^\top A)$ avec $\operatorname{im} A$
On la complète en $(U_i)_{i \in [\![1,m]\!]}$ bon de $\mathbb{R}^m$

$$\mathcal{F} = \left(\underbrace{u_1, \dots, u_r}_{\text{bon de im } A}, \underbrace{u_{r+1}, \dots, u_m}_{\text{bon de } (\text{im } A)^\top} \right)$$

Posons $U = (U_1, \dots, U_m) \in \mathcal{O}_n(\mathbb{R})$

Remarque $\forall i \in [\![1, r]\!]$

$$Bu_i = \mu_i u_i$$

Avec B, μ_i à determiner.

$$\begin{aligned}
AA^\top U_i &= \frac{1}{\sqrt{\lambda_i}} \underbrace{}_{=\lambda_i v_i} \underbrace{AA^\top A v_i}_{\text{par def de } v_i} \\
&= \sqrt{\lambda_i} A v_i \\
&= \lambda_i \underbrace{\left(\frac{1}{\sqrt{\lambda_i}} A v_i \right)}_{U_i} \\
&= \lambda_i U_i
\end{aligned}$$

D'où U_i vecteur propre de $AA^\top$ associé à λ_i

- On pose

$$\begin{aligned}
\Sigma &= U^\top A V \\
&= U^\top \left(\underbrace{Av_1, \dots, Av_r}_{\sqrt{\lambda_i} u_i}, \underbrace{Av_{r+1}, \dots, Av_n}_{0 \text{ car } v_i \in \ker A} \right) \\
&= \begin{pmatrix} u_1^\top \\ \vdots \\ u_m^\top \end{pmatrix} \left(\frac{\sqrt{\lambda_j} \langle u_i, u_j \rangle | (0)}{\sqrt{\lambda_j} \langle u_i, u_j \rangle | (0)} \right) \\
&= \left(\begin{array}{ccc|c} \sqrt{\lambda_1} & & (0) & \\ & \ddots & & (0) \\ (0) & & \sqrt{\lambda_r} & \\ \hline & (0) & & (0) \end{array} \right)
\end{aligned}$$

$$\implies A = U \sigma V^\top$$

Définition: SVD de A Décomposition de la forme

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

Avec U, Σ, V comme définies précédemment

On appelle valeurs singulières notées $(\sigma_i)_i$ les valeurs propres racines carrées

Il manque un passage, j'étais en retard de 7 minutes au cours :/

2.3 Matrice pseudo-inverse

Ou pseudo-inverse généralisée

Soit $A \in \mathcal{M}_{n,n}(\mathbb{R})$ telle que $\operatorname{rg} A = r \geq 1$

Définition: Matrice pseudo-inverse Soit $A = U\Sigma U^\top$ la SVD de A

On définit la pseudo-inverse

$$A^+ = V\Sigma^+U^\top$$

avec $\Sigma^+ = \left(\begin{array}{c|c} \Sigma_r^{-1} & 0 \\ \hline 0 & 0 \end{array}\right)$ avec le bloc non-nul de taille $r \times r$.

Proposition

1. $AA^+ = \Pi_{\operatorname{Im} A}$
2. $A^+A = \Pi_{\ker A^\top}$

Avec Π_A la projection orthogonale sur A

Soit $Q \in \mathcal{O}_{n,r}(\mathbb{R})$.

ALors $QQ^\top$ est une projection othogonale sur le sous-espace engendré par les colonnes de Q.

Preuve *()* cf. CTD factorisation Q_r

1.

$$\begin{aligned} AA^+ &= U\Sigma \underbrace{V^\top V}_{\mathrm{Im}} \Sigma^+ U^\top \\ &= \begin{pmatrix} u_1 & u_2 \end{pmatrix} \begin{pmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \Sigma_1^{-1} & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} u_1^\top \\ u_2^\top \end{pmatrix} \\ &= u_1 u_1^\top \end{aligned}$$

Avec $U_1 = \begin{pmatrix} u_1 & \dots & u_r \end{pmatrix} \in \mathcal{M}_{m,r}$ telle que $u_1^\top u_2 = I_r$.

Donc

$$\begin{aligned} AA^+ &= \underbrace{\Pi_{(u_1,\dots,u_r)}}_{\operatorname{Im} A} \\ &= \Pi_{\operatorname{Im} A} \end{aligned}$$

2.

$$\begin{aligned} AA^+ &= V_1 V_1^\top \\ &= \underbrace{\Pi_{(v_1,\dots,v_r)}}_{\ker A^\top} \\ &= \Pi_{\ker A^\top} \end{aligned}$$

Théorème Caractérisation de Moore-Penrose A^+ est l'unique solution du système

$$(MP) \begin{cases} AXA & = A \\ XAX & = X \\ (AX)^\top & = AX \\ (XA)^\top & = XA \end{cases}$$

Avec $X \in \mathcal{M}_{n,n}(\mathbb{R})$

Preuve *(Etude de* (MP)*)*

Existance de solution

1.

$$\begin{aligned} AA^+A &= U\Sigma \underbrace{\mathrm{Im}}_{V^\top V} \Sigma^+ \underbrace{U^\top U}_{\mathrm{Im}} \Sigma V^\top \\ &= U\Sigma\Sigma^+\Sigma V^\top \end{aligned}$$

Or

$$\begin{aligned}
\Sigma\Sigma^+\Sigma &= \begin{pmatrix} \Sigma_r & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \Sigma_r^{-1} & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \Sigma_r & 0 \\ 0 & 0 \end{pmatrix} \\
&= \begin{pmatrix} \Sigma_r & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} \\
&= \begin{pmatrix} \Sigma_r & 0 \\ 0 & 0 \end{pmatrix} \\
&= \Sigma
\end{aligned}$$

Donc

$$AA^+A = U\Sigma V^\top = A$$

$$\begin{aligned}
A^+AA^+ &= V \underbrace{\Sigma^+\Sigma\Sigma^+}_{\Sigma^+} U^\top \\
&= V\Sigma^+U^\top \\
&= A^+
\end{aligned}$$

$$\begin{aligned}
(AA^+)^\top &= (U_1U_1^\top)^\top && \text{cf. prop. sur } AA^\top \text{ et } A^+A \\
&= U_1U_1^\top \\
&= AA^+
\end{aligned}$$

$$\begin{aligned}
(A^+A)^\top &= (V_1V_1^\top)^\top \\
&= V_1V_1^\top \\
&= A^+A
\end{aligned}$$

d(où A^+ solution de (MP) associée à A

Unicité de la solution Soient X_1, X_2 deux solutions de (MP) associées à A.

- Montrons $AX_1 = AX_2$

$$\begin{aligned}
AX_1 &= (AX_1)^\top && \text{par 3. appliquée à } X_1 \\
&= X_1^\top A^\top \\
&= X_1^\top (AX_2A)^\top \\
&= X_1^\top A^\top X_2^\top A^\top \\
&= (AX_1)^\top (AX_2)^\top \\
&= (AX_1)(AX_2) && \text{par 3.} \\
&= \underbrace{AX_1A}_{A} X_2 && \text{par 1. appliquée à } X_1 \\
&= AX_2
\end{aligned}$$

- On montre de même: $X_1A = X_2A$.
- Montrons $X_1AX_2 = A$

$$\begin{aligned}
X_1AX_2 &= X_2AX_2 \\
&= X_2 && \text{par 2. appliquée à } X_2 \\
&= X_2 && \text{car } AX_1 = AX_2 \\
\text{i.e.} \quad X_1 &= X_2 && \text{par 2. appliquée à } X_1
\end{aligned}$$

Donc on peut permuter $\cdot^\top$ et $\cdot^+$.

Théorème $^\top =^+$

$$A^{\top+} = A^{+\top}$$

Preuve *()* On vérifie que $(A^+)^\top$ est solution de (MP) associée à A.

1.

$$\begin{aligned}
A^\top (A^+)^\top A^\top &= (AA^+A)^\top \\
&= A^\top && \text{par 1. appliquée à } A^+
\end{aligned}$$

2.

$$\begin{aligned}(A^+)^\top A^\top (A^+)^\top &= \underbrace{(A^+AA^+)^\top}_{A+} && \text{par 2 appliquée à } A^+\\ &= (A^+)^\top\end{aligned}$$

3.

$$\begin{aligned}(A^\top (A^+)^\top)^\top &= (\underbrace{(A^+A)^\top}_{A^+A})^\top && \text{par 4. appliquée à } A^+\\ &= (A^+A)^\top\\ &= A^\top (A^+)^\top\end{aligned}$$

4.

$$\begin{aligned}((A^+)^\top A^\top)^\top &= [\underbrace{(AA^+)^\top}_{AA^+}]^\top && \text{par 3. appliquée à } A^+\\ &= (A^+)^\top A^\top\end{aligned}$$

d'où $A^{+\top}$ solution de (MP) associée à $A^\top$.

Par caractérisation de Moore-Penrose:

$$A^{\top +} = A^{+\top}$$

Théorème Caractérisation de A^+ par l'image d'un vecteur de $\mathbb{R}^n$ Soit $b \in \mathbb{R}^n$.

Alors A^+b est la solution de norme $\|\cdot\|_2$ minimale de

$$(\mathcal{P}) \min_{x\in\mathbb{R}^n} f(x) = \frac{1}{2}\|Ax - b\|_2^2$$

Preuve *()*

- $$\forall x \in \mathbb{R}^n, \begin{cases} \nabla f(x) & = A^\top (Ax - b) \\ \nabla^2 f(x) & = A^\top A \text{ semi-def pos.} \end{cases}$$

  Donc f convexe sur $\mathbb{R}^n$ et les points critiques ont des minima globaux.

- $$\begin{aligned} \nabla f &= A^\top (AA^+ b - b) \\ &= V_1 \Sigma_1 U_1^\top (U_1 \Sigma_r \underbrace{V_1^\top V_1}_{I_2} \Sigma_r^{-1} U_1^\top b - b) \\ &= V_1 \Sigma_r \underbrace{(\underbrace{U_1^\top U_1}_{I_2} U_1^\top - U_1^\top)}_{0} b \\ &= 0 \end{aligned}$$

  car

  $$\begin{aligned} A &= U \Sigma V^\top = U_1 \Sigma_r V_1^\top \\ A^+ &= V \Sigma^+ U^\top = V_1 \Sigma_r^{-1} U_1^\top \end{aligned}$$

  et $A^+ b$ est solution de $(\mathcal{P})$

- Montrons $\forall v \in \mathbb{R}^n,\ v$ solution de $(\mathcal{P}), \exists x_0 \in \ker A,\ x = x_0 + A^+ b$
  Soit $x \in \mathbb{R}^n$ solution de $(\mathcal{P})$.
  Or $\mathbb{R}^n = \ker A \oplus \ker A^\perp$.
  D'où $\exists!(x_0, x_\perp) \in \ker A \times \ker A^\perp, x = x_0 + x_\perp$.
  Or x solution de $(\mathcal{P})$ donc $\nabla f(x) = 0$.
  D'où

  $$\begin{aligned} & A^\top A x = A^\top b \\ \iff & A^\top A(x_0 + x_\perp) = A^\top b \\ \iff & A^\top \underbrace{A x_0}_{0 \text{ car } x_0 \in \ker A} + A^\top A x_\perp = A^\top b \\ \iff & A^\top A x_\perp = A^\top b \end{aligned}$$

Or $A^\top A(A^+b) = A^\top b$ car $A^\top b$ solution de $(\mathcal{P})$
D'où $A^\top Ax_\perp = A^\top A(A^+b)$ i.e. $A^\top A(x_\perp - A^+b) = 0$.
D'où $x_\perp - A^+b \in \ker(A^\top A) = \ker A$ (cf. prop. de $A^\top A$)

$$\begin{aligned} A^+b &= X_1\Sigma_r^{-1}U_1^\top b \\ &\in \operatorname{Im} V_1 \\ \text{or } V_1 &= \begin{pmatrix} v_1 & \dots & v_r \end{pmatrix} \quad \text{avec les } (u_i)_i \text{ b.o.n. de } \ker A^\top \end{aligned}$$

D'où $A^+b \in \ker A^\perp$
Or $x_\perp \in \ker A^\perp$ donc $x_\perp - A^+b \in \ker A^\perp$
D'où $x_\perp - A^+b \in \ker A \cap \ker A^\perp = \{0\}$ car $\ker A \oplus \ker A^\perp = \mathbb{R}^n$
et $x_\perp = A^+b$
D'où $\exists x_0 \in \ker A, x = x_0 + A^+b$

- **Bilan**

$$\forall x \in \mathbb{R}^n, \quad \text{solution de } (\mathcal{P}) \implies \|x\|_2 \geq \|A^+b\|_2$$

Corollaire Equation normale $\forall b \in \mathbb{R}^m, (A^\top A)A^+b = A^\top b,$

Proposition Cas particuliers

1. A inversible: $\operatorname{rg} A = m = n$

$$A^+ = A^{-1}$$

2. $\operatorname{rg} A = m$ $(m \geq n$)

$$A^+ = (A^\top A)^{-1}A^\top$$

3. $\operatorname{rg} A = m \quad (n \geq m)$

$$A^+ = A^\top(AA^\top)^{-1}$$

Preuve *()*

1. A^{-1} vérifie (MP)

2. $\operatorname{rg} A = n$ donc $A^\top A$ inversible. (cf. prop. de $A^\top A$).
   D'après les équations normales,

$$\begin{aligned} \forall b \in \mathbb{R}^m, \quad (A^\top A)A^+ b &= A^\top b \\ \text{donc} \quad (A^\top A)A^+ &= A^\top \\ \text{donc} \quad A^+ &= (A^\top A)^{-1} A^\top \end{aligned}$$

3. On pose $B = A^\top$. On a $B^\top B$ inversible.

$$\operatorname{rg} B = m$$

   d'où par 2.

$$\begin{aligned} B^+ &= (B^\top B)^{-1} B^\top \\ \text{i.e. } (A^\top)^+ &= (AA^\top)^{-1} A \\ \text{i.e. } (A^+)^\top &= (AA^\top)^{-1} A && \text{car } \cdot^+ \text{ et } \cdot^\top \text{ permuttent} \\ \text{i.e. } A^+ &= A^\top (AA^\top)^{-1} \end{aligned}$$

[-¿] (-1, 0) – (10, 0) nodeker A ; [-¿] (0, -1) – (0, 10) nodeker $A^\top$; [-¿] (1, 1) node$U|_{|\ker A}^{|\operatorname{Im} A^\top}$ – (20, 1) node$A = [U]$;

Figure 1: