Chapitre 3

Les valeurs singulières

On a vu que pour certaines matrices carrées on pouvait faire une décomposition en valeurs propres.

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

mais pour effectuer cette transformation, la matrice A doit être carrée et diagonalisable.

L'idée de la décomposition en valeurs singulières est similaire à la décomposition en valeurs propres, mais fonctionne pour n'importe quelle matrice de taille $m \times n$.

La décomposition en valeurs singulières est l'un des factorisations d'une matrice les plus importantes, car il n'y a pas aucune condition sur la matrice (rang, dimension).

Ce chapitre consiste à définir les valeurs singulières et la décomposition d'une matrice en valeurs singulières, avec quelques exemples illustratifs, puis dans la deuxième section, consiste à définir la notion d'inverse généralisé (pseudo-inverse).

3.1 Décomposition en valeurs singulières

L'importance de la décomposition en valeurs singulières est l'un des factorisations d'une matrice, fonctionne pour n'importe quelle matrice de taille $m \times n$.

Définition 3.1.1 *Soit $A \in M_{m\times n}(\mathbb{R})$, la décomposition en valeurs singulières de A s'écrit de la façon suivante :*

$$A = U\Sigma V^t \tag{3.1.1}$$

avec U et V deux matrices orthogonales de taille $m \times m$ et $n \times n$ respectivement et Σ une matrice diagonale de taille $m \times n$ contenant les valeurs singulière de A notées $\sigma_1, \sigma_{2,}, .., \sigma_p$ et $p = \min(m, n)$ telles que

$$\sigma_1 \geq \sigma_2 \geq ... \geq \sigma_r > \sigma_{r+1} = ...\sigma_p = 0,$$

avec

$$r = rang(A)$$

$$\Sigma = \begin{pmatrix} \sigma_1 & & & & 0 & .. & 0 \\ & \sigma_2 & & & 0 & ... & 0 \\ & & . & & & & 0 \\ & & & \sigma_n & 0 & ... & 0 \end{pmatrix} \text{ si } m \leq n$$

$$\Sigma = \begin{pmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & . & \\ & & & \sigma_n \\ 0 & & & 0 \\ . & & & \\ 0 & .. & .. & 0 \end{pmatrix} \text{ si } m > n$$

Les r premières colonnes de V et U sont dénommées respectivement vecteurs singulières droit et gauche de A.

Si $A = U\Sigma V^t$ on a

$$A^t A = \left(U\Sigma V^t\right)\left(U\Sigma V^t\right) = V\Sigma^2 V^t \tag{3.1.2}$$

$A^t A$ est une matrice réelle symétrique, elle est donc diagonalisable dans une base orthonormée de vecteurs propres

$$A^t A = PDP^{-1} = PDP^t \tag{3.1.3}$$

De plus $A^t A$ est définie positive car

$$< A^t Ax, x >= (A^t Ax)^t x = x^t A^t Ax = (Ax)^t (Ax) = \|Ax\|^2 > 0$$

Donc les valeurs propres λ_i sont positives ou nulles.

Par comparaison des équations (3.1.2) et (3.1.3) il est donc possible de dire que $\Sigma^2 = D$ et $P = V$.

Donc (V, Σ^2) représentent la décomposition aux valeurs propres de A^tA, i.e.

$$(A^tA = V\Sigma^2V^t).$$

On appliquant la même démarche à la matrice AA^t on obtient

$$(AA^t = U\Sigma^2U^t) = PDP^t.$$

D'où (U, Σ^2) représentent la décomposition aux valeurs propres de AA^t, i.e.

$$(AA^t = U\Sigma^2U^t).$$

On retrouve ici une propriété des valeurs singulières de A

$$\Sigma^2 = D \Leftrightarrow \sigma_i = \sqrt{\lambda_i}$$

Les valeurs singulières σ_i de A sont les racines carrées des valeurs propres de A^tA ou AA^t

Théorème 3.1.1 *Soit*

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

Alors

$$\| A \|_2 = \sigma_1 \quad et \qquad \| A \|_F = \sqrt{\sum_{i=1}^{n} \sigma_i^2}.$$

Preuve. On a U et V sont orthogonale, nous avons

$$\| A \|_2 = \| U\Sigma V^t \|_2 = \| \Sigma \|_2$$

.Maintenant

$$\| \Sigma \|_2^2 = \max \| \Sigma x \|_2^2 = \max_{\|x\|_2=1} (\sigma_1^2 x_1^2 + ... + \sigma_n^2 x_n^2) \leq \sigma_1^2(x_1^2 + ... + x_n^2) = \sigma_1^2,$$

et le maximum est vérifié pour $x = e_1$ alors $\| A \|_2 = \sigma_1$.

Pour la norme de Frobenuis, nous avons

$$\| A \|_F = \sqrt{\sum_{i,j}^{n} a_{ij}^2} = \sqrt{tr(A^tA)} = \sqrt{\sum_{i=1}^{n} \sigma_i^2},$$

Comme la trace d'une matrice est la somme de ses valeurs propres. ■

3.1.1 L'existance et l'unicité

Théorème 3.1.2 *Soit* $A \in R_{m\times n}$ *une matrice de rang* r. *Il existe deux matrices orthogonales* $U \in R^{m\times m}$*,*$(U^tU = UU^t = I_m)$ *et* $V \in R^{n\times n}$*,* $(V^tV = VV^t = I_n)$ *telles que :*

$$A = U\Sigma V^t, \qquad \Sigma = \begin{pmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{pmatrix} \tag{3.1.4}$$

ou $\Sigma \in R^{m\times n}, \Sigma_1 = diag(\sigma_1, \sigma_2, ..., \sigma_r)$, et

$$\sigma_1 \geq \sigma_2 \geq ...... \geq \sigma_r > 0$$

Composante par composante, l'identité matricielle (2.1.4) devient :

$$Av_j = \sigma_j u_j; \quad A^t u_j = \sigma_j v_j, \qquad pour\ j = 1, ..., n.$$

$$A^t u_j = 0, \qquad pour\ j = n+1, ..., m.$$

Si l'on note $U = (u_1, u_2, ..., u_m)$, $V = (v_1, v_2, ..., v_n)$ les colonnes des matrices U et V, les vecteurs u_j et v_j sont respectivement, les vecteurs singulières droits et gauches associés à les valeurs singulières σ_j.

Preuve. La preuve se fait par récurrence sur n.

Par définition de ce qu'est une norme matricielle subordonnée, il existe un vecteur $v_1 \in \mathbb{R}^n$ tel que :

$||v_1||_2 = 1, \qquad ||Av_1||_2 = ||A|| =_{déf} \sigma.$

Où σ est strictement positif (si $\sigma = 0$, alors $A = 0$, et il n'y a rien à démontrer). Posons $u_1 = 1/\sigma A v_1 \in R^m$.

Complétons le vecteur v_1 en une base orthogonale de $\mathbb{R}^n$, et notons $V = (v_1, V_1) \in \mathbb{R}^{n\times n}$ la matrice formée par les vecteurs de base.

Faisons de même pour u_1 et $\mathbb{R}^m$, notant $U = (u_1, U_1) \in R^{m\times m}$.

Remarquons que les matrices U et V sont orthogonales par construction.

D'après notre choix de U_1, $U_1^t A v_1 = \sigma U_1^t u_1 = 0$, et donc le produit $U^t AV$ a la structure par blocs suivante :

$$A_{1_{déf}} = U^t AV = \begin{pmatrix} \sigma & w^t \\ 0 & B \end{pmatrix}$$

avec $w^t = u_1^t AV^1$ et $B = U_1^t AV_1 \in \mathbb{R}^{(m-1)\times(n-1)}$.

Comme U et V sont orthogonales, $||A_1||_2 = ||A||_2 = \sigma$. Mais la double inégalité

$$||A_1||_2 \geq (\sigma^2 + w^t w)^{\frac{1}{2}} \geq \left\| A_1 \begin{pmatrix} \sigma \\ w \end{pmatrix} \right\|_2 = \left\| \begin{pmatrix} \sigma^2 + w^t w \\ Bw \end{pmatrix} \right\|_2 \geq \sigma^2 + w^t w,$$

montre que $||A_1|| \geq (s2 + w^t w)^{1/2}$. On doit donc avoir $w = 0$. On peut alors terminer la démonstration en appliquant l'hypothèse de récurrence à B. ■

Remarque 3.1.1 Σ *est unique mais* U *et* V *ne sont pas unique. Les colonnes* U *sont appelées les vecteurs singulières à gauche, et les colonnes de* V *sont vecteurs singulières à droite.*

3.1.2 Calcul de la SVD

On peut calculer la SVD de A de la façon suivante :

- Calculer $A^t A$,
- Calculer une décomposition en éléments propres $A^t A = VDV^t$,
- Soit la Σ la matrice diagonale de taille $m \times n$ ayant dans la diagonale les matrices carrée de D, $A = U\Sigma V^t \iff AV = U\Sigma$, U orthogonale,

$$A = U\Sigma V^t \iff AV = A[v_1, v_2, ..., v_r] = [u_1, u_2, ..., u_r] \begin{pmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & . & \\ & & & \sigma_r \\ 0 & & & 0 \\ . & & & \\ 0 & .. & .. & 0 \end{pmatrix},$$

d'ou

$$Av_i = \sigma_i u_i, \quad i = 1, ..., n$$

d'où

$$A = \Sigma \sigma_i u_i v_i^t \;\; \text{et} \;\; A^t = \Sigma \sigma_i v_i u_i^t$$

tel que $r = rang(A)$, u_i et v_i sont respectivement les $i^{éme}$ colonne de U et V.

Remarque 3.1.2 *Le rang de A est le nombre de SVD.*

Exemple 3.1.1 *Nous allons calculer la SVD de A telle que :* $A=\begin{pmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{pmatrix}$.

Pas 1 : la diagonalisation orthogonal A^tA *:*

$$A^tA = \begin{pmatrix} 80 & 100 & 40 \\ 100 & 170 & 140 \\ 40 & 140 & 200 \end{pmatrix}$$

Ses valeurs propres et vecteurs propres sont :

$$\begin{aligned} \lambda_1 &= 360, \; v_1 = (\; \tfrac{1}{3}, \;\; \tfrac{2}{3}, \;\; \tfrac{2}{3} \;) \\ \lambda_2 &= 90, \quad v_2 = (\; -\tfrac{2}{3}, \;\; -\tfrac{1}{3}, \;\; \tfrac{2}{3} \;) \\ \lambda_3 &= 0, \qquad v_3 = (\; \tfrac{2}{3}, \;\; -\tfrac{2}{3}, \;\; \tfrac{1}{3} \;) \end{aligned}$$

Pas 2 : Construire V *et* Σ *:*

$$\begin{aligned} V &= (v_1, v_2, v_3) = \begin{pmatrix} \frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{1}{3} & -\frac{2}{3} \\ \frac{2}{3} & \frac{2}{3} & \frac{1}{3} \end{pmatrix} \\ \Sigma &= \begin{pmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \end{pmatrix} = \begin{pmatrix} \sqrt{\lambda_1} & 0 & 0 \\ 0 & \sqrt{\lambda_2} & 0 \end{pmatrix} = \begin{pmatrix} 6\sqrt{10} & 0 & 0 \\ 0 & 3\sqrt{10} & 0 \end{pmatrix} \end{aligned}$$

Pas 3 : *Construire* U :

$$\begin{aligned} u_1 &= \frac{Av_1}{\sigma_1} = \begin{pmatrix} \frac{3}{\sqrt{10}} & \frac{1}{\sqrt{10}} \end{pmatrix} \\ u_2 &= \frac{Av_2}{\sigma_2} = \begin{pmatrix} \frac{1}{\sqrt{10}} & -\frac{3}{\sqrt{10}} \end{pmatrix} \\ &\Rightarrow U = \begin{pmatrix} \frac{3}{\sqrt{10}} & \frac{1}{\sqrt{10}} \\ \frac{1}{\sqrt{10}} & -\frac{3}{\sqrt{10}} \end{pmatrix} \end{aligned}$$

L'ensemble $\{u_1, u_2\}$ *est déjà une base de* $\mathbb{R}^2$*, donc Il n'est pas nécessaire de l'étendre.*

Finalement :

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

Donc

$$\begin{pmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{pmatrix} = \begin{pmatrix} \frac{3}{\sqrt{10}} & \frac{1}{\sqrt{10}} \\ \frac{1}{\sqrt{10}} & -\frac{3}{\sqrt{10}} \end{pmatrix} \begin{pmatrix} 6\sqrt{10} & 0 & 0 \\ 0 & 3\sqrt{10} & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{1}{3} & -\frac{2}{3} \\ \frac{2}{3} & \frac{2}{3} & \frac{1}{3} \end{pmatrix}$$

Le programme en MATLAB : $[U, S, V] = svd([4\ 11\ 14; 8\ 7\ -2])$

>> A=[4 11 14;8 7 -2]

$$A = \begin{pmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{pmatrix}$$

>> [U,S,V]=svd(A)

$$U = \begin{pmatrix} -0.9487 & -0.3162 \\ -0.3162 & 0.9487 \end{pmatrix}$$

$$S = \begin{pmatrix} 18.9737 & 0 & 0 \\ 0 & 9.4868 & 0 \end{pmatrix}$$

$$V = \begin{pmatrix} -0.3333 & 0.6667 & -0.6667 \\ -0.6667 & 0.3333 & 0.6667 \\ -0.6667 & -0.6667 & -0.3333 \end{pmatrix}$$

Exemple 3.1.2 *Nous allons calculer la SVD de A tel que :* $A = \begin{pmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{pmatrix}$.

Pas 1 : *la diagonalisation orthogonal A^tA :*

$$A^tA = \begin{pmatrix} 9 & -9 \\ -9 & 9 \end{pmatrix}$$

Ses valeurs propres et vecteur propres sont :

$$\begin{aligned} \lambda_1 &= 18,\ v_1 = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} \\ \lambda_2 &= 0, \quad v_2 = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \end{aligned}$$

Pas 2 : *construire V et Σ :*

$$\begin{aligned} V &= (v_1, v_2) = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \\ \Sigma &= \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} \sqrt{\lambda_1} & 0 \\ 0 & \sqrt{\lambda_2} \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 3\sqrt{2} & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \end{aligned}$$

Pas 3 : *Construire U :*

$$\begin{aligned} u_1 &= \frac{Av_1}{\sigma_1} = \begin{pmatrix} \frac{1}{3}, & -\frac{2}{3}, & \frac{2}{3} \end{pmatrix} \\ &\Rightarrow \ U = \begin{pmatrix} \frac{3}{\sqrt{10}} & \frac{1}{\sqrt{10}} \\ \frac{1}{\sqrt{10}} & -\frac{3}{\sqrt{10}} \end{pmatrix} \end{aligned}$$

L'ensemble $\{u_1\}$ce n'est pas une base de $\mathbb{R}^3$, donc Nous devons l'étendre, avec des vecteurs orthogonales. Tous les vecteurs orthogonaux à u_1 vérifiés :

$$u_1 \cdot u = 0 = \frac{1}{3}x_1 - \frac{2}{3}x_2 + \frac{2}{3}x_3 \Rightarrow x_1 = 2x_2 - 2x_3$$

Une base pour ce espace est $w_2 = (2, 1, 0)$ and $w_3 = (-2, 0, 1)$. Mais cette base n'est pas une base orthogonale.

Nous allons les mettre on forme orthogonale par le principe de Gram-Schmidt suivant :

$$\begin{aligned} u_2 &= \frac{w_2}{\parallel w_2 \parallel} = (\frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}}, 0) \\ w_3' &= w_3 - < w_3, v_2 > v_2 = (-\frac{2}{5}, \frac{4}{5}, 1) \\ u_3 &= \frac{w_3}{\parallel w_3 \parallel} = (-\frac{2}{3\sqrt{5}}, \frac{4}{3\sqrt{5}}, \frac{\sqrt{5}}{3}) \end{aligned}$$

Finalement :

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

Donc

$$\begin{pmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{pmatrix} = \begin{pmatrix} \frac{1}{3} & & \\ -\frac{2}{3} & & \\ & & \end{pmatrix} \begin{pmatrix} 3\sqrt{2} & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$$

Le programme en MATLAB : $[U, S, V] = svd([1\ \ \text{-1}; -2\ \ 2; 2\ \ \text{-2}])$

Exemple 3.1.3

$$A = \begin{pmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{pmatrix}$$

>> [U,S,V]=svd(A)

$$U = \begin{pmatrix} -0.3333 & 0.6667 & -0.6667 \\ 0.6667 & 0.6667 & 0.3333 \\ -0.6667 & 0.3333 & 0.6667 \end{pmatrix}$$

$$S = \begin{pmatrix} 4.2426 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$$

$$V = \begin{pmatrix} -0.7071 & 0.7071 \\ 0.7071 & 0.7071 \end{pmatrix}$$

Proposition 3.1.1 *On a les relations suivantes :*

1. *Le rang de A est égal au nombre des valeurs singulières non-nulles.*

2. $KerA = vect(v_{r+1}, ..., v_n)$, $\ ImA = vect(u_1, ..., u_r)$.

3. $\ker A^t = vect(v_1, ..., v_r)$, $\quad ImA^t = vect(u_{r+1}, ..., u_m)$.

4. $||A||_2 = \sigma_1$.

3.2 L'inverse généralisé

La notion d'inversible est une notion très importante dans tous les domaines des mathématiques.

Deux matrices A et B sont inverse, si leurs produits est égale à la matrice identité : $AB = I, BA = I$ alors $B = A^{-1}$.

Les matrices inverses, et plus généralement en utilisons la notion de l'inverse généralisé ou bien les pseudo-inverse, trouvent leurs applications à la résolution des systèmes

$$Ax = b$$

d'équation linéaires quelles que soient leurs dimentions : $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$ est le vecteur cherché $x \in \mathbb{R}^n$.

Le système linéaire possède une solution unique, lorsque la matrice A est inversible.

Malheureusement, on ne se trouve pas toujours dans ce cas, en pratique, on obtient souvent des données sous forme d'une matrice rectangulaire qui est un opérateur non inversible.

L'inverse généralisé d'un tel système est noté A^+.

L'inverse généralisé A^+ satisfait les conditions suivantes :

$$\begin{aligned} AA^+A &= A \\ A^+AA^+ &= A^+ \\ (AA^+)^t &= AA^+ \\ (A^+A)^t &= A^+A \end{aligned} \qquad \text{Condition de système}$$

La solution d'un système linéaire à partir de la pseudo-inverse A^+ s'écrit alors :

$$x = A^+b.$$

On distingue trois cas :

a) Si $m= n$, cas d'une matrice carré,

Si A une non singulière $(det(A) \neq 0)$ solution unique et

$$A^+ = A.$$

b) $m < n$, matrice rectangulaire, on a une infinité des solutions et

$$A^+ = A^t(AA^t)^{-1}.$$

c) $m > n$, pas de solution, il existe une solution approchée aux moindres carrés et

$$A^+ = (A^tA)^{-1}A^t.$$

3.2.1 L'inverse généralisé et SVD

Définition 3.2.1 *Soit $A \in \mathbb{R}^{m\times n}$ une matrice de rang r, ainsi que sa décomposition en valeurs singulières*

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

La matrice

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

est appelée matrice pseudo-inverse ou inverse généralisée de A, avec :

$$\Sigma^+ = diag(1/\sigma_1, \ldots, 1/\sigma_r, 0, \ldots, 0)$$

- $A^+A = I_r$ (matrice identité de $rang\ r$),
- Si $rang(A) = n < m$, alors $A^+ = (A^tA)^{-1}A^t$,
- Si $rang(A) = n = m$, alors $A^+ = A^{-1}$.