

Mathématiques Numériques

Chapitre 3 : Systèmes linéaires et décompositions de matrice

Partie II : Factorisation de Cholesky, QR, SVD

Bruno Pinçon / J.-F. Scheid

Télécom Nancy 1A

2025-2026

7 Factorisation de Cholesky

8 Factorisation QR

9 Décomposition SVD

7. Factorisation de Cholesky

1. Introduction

André-Louis Cholesky (1875-1918), polytechnicien (affecté au service géographique de l'état-major de l'armée) ;

"Sur la résolution numérique des systèmes d'équations linéaires",
manuscrit (8 pages) écrit en 1910, déposé par sa famille aux Archives
de l'École Polytechnique en 2004.

Cholesky n'a jamais publié ses travaux. La méthode de Cholesky a
été exposée pour la première fois en 1924, soit plusieurs années après
sa mort, par le commandant Benoît. La méthode de Cholesky semble
avoir été connue et utilisée par les topographes. Mais son essor est dû
à John Todd qui l'exposa dans son cours d'analyse numérique au
King's College de Londres à partir de 1946.

Sur la résolution numérique
des systèmes d'équations linéaires.

Le procédé que nous allons indiquer s'applique aux systèmes d'équations différentielles auxquels on applique la méthode de Runge-Kutta; mais nous n'embrassons tout d'abord que la résolution d'un système de 2 équations linéaires à 2 inconnues. On peut étendre facilement ce raisonnement à la résolution d'un système ~~linéaire~~ de 2 équations linéaires symétriques à 2 inconnues.

[illegible]
$$\text{II} \quad \left\{ \begin{array}{l} \gamma_1 = \alpha_1^1 \lambda_1 + \alpha_1^2 \lambda_2 + \dots + \alpha_1^n \lambda_n \\ \gamma_2 = \alpha_2^1 \lambda_1 + \alpha_2^2 \lambda_2 + \dots + \alpha_2^n \lambda_n \\ \dots \\ \gamma_r = \alpha_r^1 \lambda_1 + \alpha_r^2 \lambda_2 + \dots + \alpha_r^n \lambda_n \end{array} \right.$$
$$\text{III} \left\{ \begin{array}{l} A_1^1 \lambda_1 + A_2^1 \lambda_2 + \dots + A_n^1 \lambda_n + C_1 \\ A_1^2 \lambda_1 + A_2^2 \lambda_2 + \dots + A_n^2 \lambda_n + C_2 = 0 \\ A_1^3 \lambda_1 + A_2^3 \lambda_2 + \dots + A_n^3 \lambda_n + C_3 \end{array} \right.$$

De plus, l'axe des ordonnées du rectangle le calcul des λ à partir du système de équations (V), car on a remplacé dans la signature III, la trame constante C par la trame de signature V, l'opération consistant à changer λ en $(1-\lambda)$, le calcul de inconnues, fait avec le V donne donc de valeurs λ' tels que $\lambda_p + \lambda'_p = 1$

La résolution d'un système de 10 équations à 10 inconnues
peut être faite ~~en 15 s~~ avec 5 chiffres exacts, en
4 à 5 heures : y compris la vérification des équations et le calcul
des résidus.

On a résolu par cette méthode plusieurs systèmes dépassant 30 équations, et en particulier un système de 56 équations. Ce dernier cas fait partie d'un calcul de compensation.

Viscous L 2 Decem. 1910

Richard

Définition

Soit A une matrice carrée, symétrique. On dit que A est *définie positive* (sdp) si

$$(Ax \mid x) > 0, \quad \forall x \neq 0.$$

- Une matrice symétrique définie positive vérifie les conditions du théorème de factorisation LU^1 , en particulier elle est inversible. Elle admet donc une factorisation LU unique.
- Cependant, on peut déterminer une factorisation analogue mais plus simple et moins coûteuse (en nb d'opérations et en mémoire) car ne faisant intervenir qu'une seule matrice.

1. On montre sans trop de difficulté (exercice) que A sdp est inversible et toutes les sous-matrices diagonales Δ_k sont aussi sdp donc inversibles.

2. Factorisation de Cholesky

Théorème 3. Factorisation de Cholesky

Si A est une matrice symétrique définie positive, alors il existe (au moins) une matrice triangulaire inférieure C telle que

$$A = CC^{\top}.$$

De plus, si on impose que tous les éléments diagonaux de la matrice C soient strictement positifs, alors la matrice C est unique.

Esquisse de la preuve. A admet une unique factorisation LU . On montre que $U = DL^{\top}$ où D est la matrice diagonale avec $d_{ii} = u_{ii} > 0$. On a donc

$$A = LDL^{\top} = LD^{1/2}D^{1/2}L^{\top} = \underbrace{LD^{1/2}}_{:=C}(LD^{1/2})^{\top} = CC^{\top}.$$

3. Algorithme de Cholesky

On veut déterminer la matrice C triangulaire inférieure telle que $A = CC^T$ (A sdp).

- On pose a priori

$$C = \begin{pmatrix} c_{11} & 0 & \dots & 0 \\ c_{21} & c_{22} & \ddots & \vdots \\ \vdots & & \ddots & 0 \\ c_{n1} & \dots & \dots & c_{nn} \end{pmatrix}$$

- On écrit $A = CC^T$ et on obtient alors le système d'équations :

$$\sum_{k=1}^j c_{ik}c_{jk} = a_{ij}, \quad \forall 1 \leq j \leq i \leq n \quad (1)$$

Il y a $n(n+1)/2$ inconnues pour $n(n+1)/2$ équations.

- La résolution de ce système peut se faire de proche en proche, colonne par colonne :

- ▶ *Première colonne ($j = 1$).*

- ★ avec $i = j = 1$ dans (1), on commence par déterminer

$$c_{11} = \sqrt{a_{11}} > 0$$

($a_{11} > 0$ car A est définie positive)

- ★ puis, pour $1 = j < i \leq n$ dans (1) :

$$c_{i1} = a_{i1}/c_{11}$$

- ▶ *Deuxième colonne ($j = 2$)* : on détermine ensuite c_{22} puis c_{i2} pour $2 = j < i \leq n$ dans (1).
- ▶ etc, ...

On obtient ainsi l'algorithme suivant pour la factorisation de Cholesky.

Algorithme de Cholesky

pour j de 1 à n :

$$c_{jj} = \sqrt{a_{jj} - \sum_{k=1}^{j-1} c_{jk}^2}$$

pour i de $j + 1$ à n :

$$c_{ij} = \left(a_{ij} - \sum_{k=1}^{j-1} c_{ik} c_{jk} \right) / c_{jj}$$

fin pour

fin pour

4. Coût de la factorisation de Cholesky

- Le nombre d'opérations de la factorisation de Cholesky est de l'ordre de $n^3/6$ additions, $n^3/6$ multiplications, $n^2/2$ divisions et n extractions de racines.

La factorisation est donc moins coûteuse que la méthode de Gauss qui requière de l'ordre de $n^3/3$ additions, $n^3/3$ multiplications et $n^2/2$ divisions.

- La méthode de Cholesky est donc fortement conseillée quand la matrice est symétrique définie positive.
- La résolution d'un système linéaire à partir d'une factorisation de Cholesky, se fera aussi par une méthode de *descente-remontée* (comme pour LU).

8. Factorisation QR

1. Matrice de Householder

Définition : matrice de Householder

Soit $v \in \mathbb{R}^n$, $v \neq 0$. La matrice de Householder de taille $n \times n$ associée à v , est définie par

$$H(v) = I - 2 \frac{v v^T}{v^T v}.$$

Remarque. $v v^T$ est le produit d'un vecteur colonne avec un vecteur ligne donc fournit une matrice de taille $n \times n$ avec

$$(v v^T)_{ij} = v_i v_j,$$

tandis que

$$v^T v = \|v\|^2 = \sum_{i=1}^n v_i^2$$

est un nombre (la norme euclidienne de v au carré).

Proposition 1.

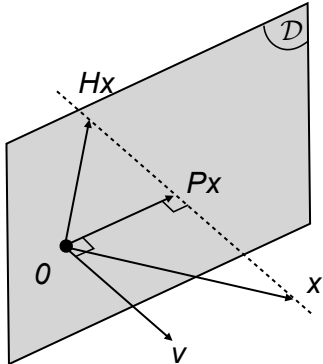
Soit $H = H(v)$ la matrice de Householder associée à $v \in \mathbb{R}^n$, $v \neq 0$.

- ① H est la matrice de la *symétrie* par rapport à l'hyperplan vectoriel

$$\mathcal{D} = \{y \in \mathbb{R}^n, (y | v) = 0\}$$

perpendiculaire à v .

- ② H est symétrique et orthogonale ($H^{-1} = H^T$) avec $HH = I$.
- ③ $\det(H) = -1$



Vérification de 1.

Pour $x \in \mathbb{R}^n$, le projeté orthogonal de x sur $\mathcal{D}$ est défini comme l'unique vecteur $Px \in \mathcal{D}$ vérifiant

$$(Px - x \mid y) = 0, \quad \forall y \in \mathcal{D}.$$

On note Hx le symétrique de x par rapport à $\mathcal{D}$.

On a $x = Px + \alpha v$ et $Hx = Px - \alpha v$ avec nécessairement

$$\alpha = (x \mid v) / \|v\|^2 \in \mathbb{R}.$$

Comme de plus $(x \mid v)v = v(v^\top x) = (vv^\top)x$, on obtient

$$Px = (I - vv^\top / \|v\|^2)x \quad \text{et} \quad Hx = (I - 2vv^\top / \|v\|^2)x.$$

Proposition 2.

Soient x et y deux vecteurs *non-colinéaires*, $x \neq 0$, $y \neq 0$. On peut trouver un vecteur $v \neq 0$ et $\alpha \in \mathbb{R}$ tels que

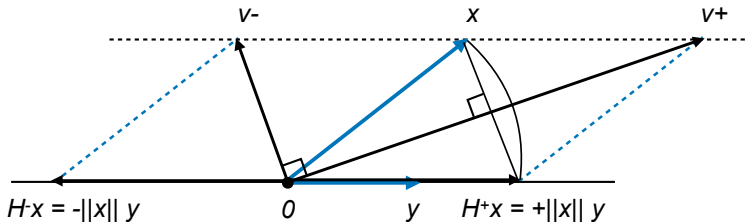
$$Hx = \alpha y \quad \text{où} \quad H = I - 2vv^T / \|v\|^2$$

Preuve. On peut toujours supposer que $\|y\| = 1$. Il y a deux vecteurs v possibles (exercice)

$$v^+ = x + \|x\|y \quad \text{et} \quad H^-x = -\|x\|y \quad (1)$$

ou bien

$$v^- = x - \|x\|y \quad \text{et} \quad H^+x = +\|x\|y \quad (2)$$



2. Factorisation QR

Théorème 4. Factorisation QR

Soit A une matrice carrée réelle. Il existe une matrice orthogonale Q et une matrice triangulaire supérieure R telles que

$$A = QR.$$

De plus, on peut toujours choisir les éléments diagonaux de R strictement positifs. Si la matrice A est inversible, la factorisation est unique.

Preuve (existence). Comme pour la factorisation $A = LU$, on essaye de mettre des 0 sous la diagonale de A , colonne par colonne.

- Pour la première étape, on choisit

$$v = A^1 \pm \|A^1\|e^1$$

où A^1 désigne la première colonne de A (e^1 est le premier vecteur de la base canonique) et on note $H^{(1)}$ la matrice de Householder associée à v .

D'après la Proposition 2, la matrice $A^{(1)} = H^{(1)}A$ aura alors sa première colonne égale à $\mp \|A^1\| e^1$ cad

$$(A^{(1)})^1 = \mp \|A^1\| e^1$$

et donc toutes ses composantes sauf la première sont nulles.

La i -ème colonne de $A^{(1)}$ vaut

$$(A^{(1)})^i = H^{(1)}A^i$$

- Pour continuer (deuxième étape), on applique la même démarche à la sous-matrice formée des $(n - 1)$ dernières lignes et colonnes ; ainsi de suite jusqu'à la dernière étape (étape $n - 1$) avec une sous-matrice formée des 2 dernières lignes et colonnes.

- A chaque étape $k \geq 1$, on détermine les matrices $A^{(k)}$ par

$$A^{(k)} = H^{(k)} A^{(k-1)}$$

avec

$$H^{(k)} = \begin{pmatrix} \overbrace{\begin{matrix} I_{k-1} & 0 \\ 0 & \tilde{H}^{(k)} \end{matrix}}^{k-1} \end{pmatrix}, \quad A^{(k-1)} = \begin{pmatrix} \overbrace{\begin{matrix} \dots & \dots \\ 0 & \dots \end{matrix}}^{k-1} \\ 0 & \tilde{A}^{(k-1)} \end{pmatrix}$$

où

- ▶ I_{k-1} est la matrice identité d'ordre $k - 1$
- ▶ $\tilde{H}^{(k)}$ est la matrice de Householder pour la sous-matrice carrée principale $\tilde{A}^{(k-1)}$ composée des $n - (k - 1)$ dernières lignes et colonnes de $A^{(k-1)}$:

$$\tilde{A}^{(k-1)} = \left(a_{ij}^{(k-1)} \right)_{k \leq i, j \leq n}$$

Pour construire $\tilde{H}^{(k)}$, on choisit le vecteur $v \in \mathbb{R}^{n-(k-1)}$ associé de façon à obtenir des 0 en dessous de la diagonale de la k -ème colonne² (cf. 1ère étape).

La matrice $A^{(k)}$ résultante est ainsi de la forme :

$$A^{(k)} = \begin{pmatrix} a_{11}^{(1)} & \times & \times & \times & \times & \times & \times \\ 0 & a_{22}^{(2)} & \times & \times & \times & \times & \times \\ \vdots & \ddots & \ddots & \times & \times & \times & \times \\ \vdots & \dots & 0 & a_{kk}^{(k)} & \times & \times & \times \\ \vdots & \dots & \vdots & 0 & \boxed{a_{k+1,k+1}^{(k)}} & \times & \times \\ \vdots & \dots & \vdots & \vdots & \boxed{\vdots} & \times & \times \\ 0 & \dots & 0 & 0 & \boxed{a_{n,k+1}^{(k)}} & \times & \times \end{pmatrix}$$

2. si $a_{ik}^{(k-1)} = 0$ pour tout $k+1 \leq i \leq n$, alors on prend $H^{(k)} = I$

- Au bout de $n - 1$ étapes, on obtient

$$A^{(n-1)} = R$$

matrice triangulaire supérieure. On obtient la décomposition annoncée³ avec

$$Q^{\top} = Q^{-1} = H^{(n-1)} \dots H^{(1)}$$



3. le produit de 2 matrices orthogonales est une matrice orthogonale

Annexe. Formules de la factorisation QR.

Soit A une matrice de taille (n, n) . A l'étape $k \in \llbracket 1, n-1 \rrbracket$, on détermine la matrice $A^{(k)}$ en calculant uniquement la sous-matrice principale $\mathcal{A}_{n-k+1}^{(k)}$ formée de $n-k+1$ dernières lignes et colonnes de $A^{(k)}$ (les $k-1$ premières lignes sont inchangées)⁴. On a

$$\mathcal{A}_{n-k+1}^{(k)} = \left(a_{ij}^{(k)} \right)_{k \leq i, j \leq n}$$

et

$$\mathcal{A}_{n-k+1}^{(k)} = \tilde{H}^{(k)} \tilde{A}^{(k-1)}$$

où $\tilde{H}^{(k)} = I - 2v_k v_k^\top / \|v_k\|^2$ est la matrice de Householder associée au vecteur

$$v_k = c_{k-1}^1 \pm \|c_{k-1}^1\| (1, 0, \dots, 0)^\top \in \mathbb{R}^{n-(k-1)},$$

où $c_{k-1}^1 = \left(\tilde{A}^{(k-1)} \right)^1 \in \mathbb{R}^{n-(k-1)}$ désigne la première colonne de $\tilde{A}^{(k-1)}$.

4. avec cette notation, on a $\mathcal{A}_{n-k}^{(k)} = \tilde{A}^{(k)}$

Exemple. Soit $A = \begin{pmatrix} -4 & -4 & -1 \\ 4 & 1 & 13 \\ 2 & -1 & 8 \end{pmatrix}$

1ère étape. On choisit $v = A^1 - \|A^1\|e^1 = (-4, 4, 2)^\top - 6(1, 0, 0)^\top = (-10, 4, 2)^\top$ avec $\|v\|^2 = 120$. On obtient

$$H^{(1)} = \frac{1}{3} \begin{pmatrix} -2 & 2 & 1 \\ 2 & 11/5 & -2/5 \\ 1 & -2/5 & 14/5 \end{pmatrix}$$

et

$$A^{(1)} = H^{(1)}A = \begin{pmatrix} 6 & 3 & 12 \\ 0 & -9/5 & 39/5 \\ 0 & -12/5 & 27/5 \end{pmatrix}.$$

2ème étape. On continue le procédé avec la sous-matrice principale

$$\tilde{A}^{(1)} = \begin{pmatrix} -9/5 & 39/5 \\ -12/5 & 27/5 \end{pmatrix}.$$

On choisit $v = (\tilde{A}^{(1)})^1 - \|(\tilde{A}^{(1)})^1\|\tilde{e}^1 = (-9/5, -12/5)^\top - 3(1, 0)^\top = (-24/5, -12/5)^\top$ avec $\|v\|^2 = 144/5$.

On obtient $\tilde{H}^{(2)} = \begin{pmatrix} -3/5 & -4/5 \\ -4/5 & 3/5 \end{pmatrix}$, $H^{(2)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -3/5 & -4/5 \\ 0 & -4/5 & 3/5 \end{pmatrix}$

et

$$R = A^{(2)} = H^{(2)} A^{(1)} = \begin{pmatrix} 6 & 3 & 12 \\ 0 & 3 & -9 \\ 0 & 0 & -3 \end{pmatrix},$$

$$Q = H^{(1)} H^{(2)} = \frac{1}{3} \begin{pmatrix} -2 & -2 & -1 \\ 2 & -1 & -2 \\ 1 & -2 & 2 \end{pmatrix}.$$

3. Choix du vecteur de Householder.

Dans la pratique, on choisit entre v^+ et v^- pour des raisons de stabilité numérique due aux erreurs d'arrondi. La norme $\|v\|^2$ intervient au dénumérateur dans H . On choisit donc v^+ ou v^- pour que $\|v\|^2$ ait *la plus grande valeur*. Comme

$$\|v\|^2 = (x \pm \|x\|y \mid x \pm \|x\|y) = \|x\|^2(1 + \|y\|^2) \pm 2\|x\|(x \mid y)$$

on choisit $v = v^+$ si $(x \mid y) > 0$ et $v = v^-$ sinon. Ainsi, on choisit

$$\boxed{v = x + \text{signe}((x \mid y))\|x\|y} \quad (3)$$

Remarques.

- Cas d'une matrice A rectangulaire de taille (m, n) avec $m \geq n$.

Il existe une matrice Q orthogonale de taille (m, m) et une matrice R de taille (m, n) triangulaire supérieure telle que $A = QR$ avec

$$R = \left(\begin{array}{c|c} \tilde{R} & \\ \hline 0 & \end{array} \right)$$

où $\tilde{R}$ est une matrice triangulaire supérieure de taille (n, n) .

- Permet de résoudre les problèmes de moindres carrés de la forme $\min_x \|Ax - b\|^2$ (cf. Chapitre 4). En particulier, si $A = QR$ alors $A^\top A = R^\top Q^\top QR = R^\top R = \tilde{R}^\top \tilde{R}$ qui est une factorisation de Cholesky de $A^\top A$.
- Le nombre d'opérations est en $\mathcal{O}(\frac{2}{3}n^3)$ soit environ deux fois plus que pour LU . Mais la factorisation QR possède des bonnes propriétés de stabilité numérique.

9. Décomposition SVD

1. Rappels sur les éléments propres d'une matrice

Soit A une matrice carrée de taille (n, n) à coefficients réels ou complexes.

a. Produit scalaire et matrice complexe.

- Soit $x, y \in \mathbb{C}^n$. Le produit scalaire dans $\mathbb{C}^n$ entre x et y est le nombre complexe défini par

$$(x | y) = \sum_{i=1}^n x_i \overline{y_i} = \overline{(y | x)}$$

Si A^* désigne la matrice *adjointe* définie par $A^* = \overline{A}^\top$ (la transposée de la matrice conjuguée), on a

$$(Ax | y) = (x | A^*y)$$

- Une matrice A à coefficients complexes est dite *unitaire* (*orthogonale* dans le cas réel) si elle est inversible et son inverse vaut $A^{-1} = A^*$ ($A^{-1} = A^\top$ dans le cas réel).

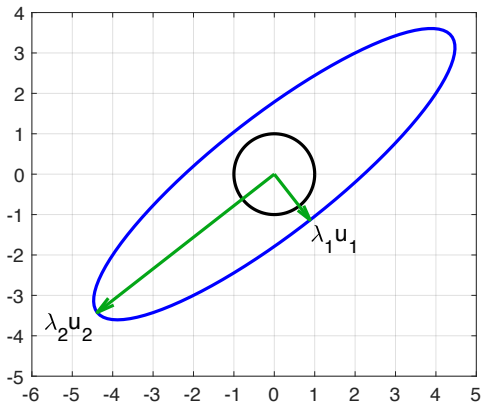
b. Valeurs propres.

- $\lambda \in \mathbb{C}$ est une *valeur propre* de A s'il existe un vecteur $u \in \mathbb{C}^n$, $u \neq 0$ tel que $Au = \lambda u$. Le vecteur u est appelé *vecteur propre* associé à λ .
- Les valeurs propres λ de A sont les zéros du polynôme caractéristique $p_A(\lambda) = \det(A - \lambda I)$. Le polynôme p_A est de degré n par rapport à λ donc il existe n valeurs propres, non nécessairement distinctes.
- On note $Sp(A)$ l'ensemble des valeurs propres de A (le spectre). On a $Sp(A) = Sp(A^\top)$ et $Sp(A^*) = Sp(\overline{A})$.
- Si A est réelle, p_A est un polynôme à coefficients réels et les valeurs propres complexes de A apparaissent par paires de complexes conjugués.
- Si A est *hermitienne* i.e. $A^* = A$ (ou réelle et symétrique) alors toutes ses valeurs propres sont réelles.
- Les *valeurs singulières* d'une matrice A sont les racines carrées des valeurs propres de la matrice A^*A .

Soit $A = \begin{pmatrix} 4 & 2 \\ 2 & 3 \end{pmatrix}$ qui admet les éléments propres

$$\lambda_1 \simeq 1.4384, u_1 \simeq \begin{pmatrix} 0.6154 \\ -0.7882 \end{pmatrix} \quad \text{et} \quad \lambda_2 \simeq 5.556, u_2 \simeq \begin{pmatrix} -0.7882 \\ -0.6154 \end{pmatrix}.$$

Le cercle unité est transformé par A en une ellipse dont les axes principales sont déterminés par les vecteurs u_1 et u_2 (de normes 1) et les longueurs des demi-axes par λ_1 et λ_2 .



2. Décomposition spectrale

On considère désormais des matrices à coefficients réels.

Proposition 3.

Soit A une matrice réelle symétrique, il existe une matrice orthogonale U et une matrice diagonale Λ telle

$$U^{-1}AU = \Lambda$$

où Λ est formée des valeurs propres de A et les colonnes de U sont les vecteurs propres associés.

Les vecteurs propres u_i de A (avec $Au_i = \lambda_i u_i$) sont donc 2 à 2 orthogonaux ; on peut toujours les choisir orthonormés i.e. de norme euclidienne 1. On a la décomposition spectrale (avec $U^{-1} = U^T$) :

$$A = U\Lambda U^{-1} = \sum_{i=1}^n \lambda_i u_i u_i^T \quad (4)$$

3. Décomposition SVD

Toute matrice peut être réduite à une matrice diagonale par multiplication à gauche et à droite par des matrices orthogonales. C'est la décomposition *SVD* (Singular Value Decomposition).

Proposition 4. Décomposition SVD

Soit $A \in \mathbb{R}^{m \times n}$. Il existe deux matrices orthogonales $U \in \mathbb{R}^{m \times m}$ et $V \in \mathbb{R}^{n \times n}$ telles que

$$U^T A V = \Sigma = \text{diag}(\sigma_1, \dots, \sigma_p) \in \mathbb{R}^{m \times n} \quad (5)$$

avec $p = \min(m, n)$ et $\sigma_1 \geq \dots \geq \sigma_p \geq 0$. La relation (5) est appelée "Décomposition en valeurs singulières"^a de A et les nombres σ_i sont les *valeurs singulières* de A avec

$$\sigma_i = \sigma_i(A) = \sqrt{\lambda_i(A^T A)} \quad (6)$$

a. Singular Value Decomposition, *SVD*

La matrice Σ a la forme suivante :

$$\Sigma = \begin{pmatrix} \sigma_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \sigma_n \\ 0 & \dots & 0 \\ \vdots & & \vdots \\ 0 & \ddots & 0 \end{pmatrix} \quad \text{ou} \quad \Sigma = \begin{pmatrix} \sigma_1 & 0 & 0 & 0 & \dots & 0 \\ 0 & \ddots & 0 & 0 & & 0 \\ 0 & 0 & \sigma_m & 0 & \dots & 0 \end{pmatrix}$$

$(\mathbf{m} < \mathbf{n})$

$(\mathbf{m} > \mathbf{n})$

On a

$$A = U\Sigma V^{\top} \quad (7)$$

Vecteurs singuliers

- Les colonnes de V , appelées *vecteurs singuliers à droite* sont les vecteurs propres de $A^{\top}A = V(\Sigma^{\top}\Sigma)V^{\top}$.
- Les colonnes de U , appelées *vecteurs singuliers à gauche* sont les vecteurs propres de $AA^{\top} = U(\Sigma\Sigma^{\top})U^{\top}$.

Vérification de (6). On a la décomposition en valeurs singulières :

$$A = U\Sigma V^\top = \sum_{i=1}^p \sigma_i U^i V^{i\top} \quad (8)$$

où U^i (resp. V^i) désigne la i -ème colonne de U (resp. V).

La matrice

$$A^\top A = V(\Sigma^\top \Sigma)V^\top \quad (9)$$

est symétrique, ses valeurs propres $\lambda_i(A^\top A)$ sont donc réelles et

$$\lambda_i(A^\top A) = \lambda_i(\Sigma^\top \Sigma) = (\sigma_i(A))^2 \geq 0.$$

Exemple. Soit $A = \begin{pmatrix} 1 & 0 & 1 \\ -2 & 1 & 0 \end{pmatrix} = U\Sigma V^\top$

- *Vecteur singulier droit* V . Les colonnes de V sont les vecteurs

propres de $A^\top A = \begin{pmatrix} 5 & -2 & 1 \\ -2 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$. On obtient

$$V = \begin{pmatrix} 0.9129 & 0 & -0.4082 \\ -0.3651 & 0.4472 & -0.8165 \\ 0.1826 & 0.8944 & 0.4082 \end{pmatrix} \text{ associées aux valeurs}$$

propres $\lambda_1 = 6$, $\lambda_2 = 1$, $\lambda_3 = 0$.

- *Valeurs singulières*. On a $\Sigma = \begin{pmatrix} \sqrt{6} & 0 & 0 \\ 0 & \sqrt{1} & 0 \end{pmatrix}$

- *Vecteur singulier gauche* U . Les colonnes de U sont les vecteurs

propres de $AA^\top = \begin{pmatrix} 2 & -2 \\ -2 & 5 \end{pmatrix}$. On obtient

$$U = \begin{pmatrix} -0.4472 & -0.8944 \\ 0.8944 & -0.4472 \end{pmatrix} \text{ associé aux valeurs propres de}$$

$\lambda_1 = 6$, $\lambda_2 = 1$.

Applications de la SVD.

- Problèmes de moindres carrés (cf. Chapitre 4)
- *Approximation de matrice.*

On a vu que $A = U\Sigma V^\top = \sum_{i=1}^p \sigma_i U^i V^{i\top}$ où U^i (resp. V^i) est la i -ème colonne de U (resp. V) et $p = \min(m, n)$.

On peut introduire l'approximation $\hat{A}(k)$ de A obtenue en ne retenant que k premiers termes de la somme :

$$\hat{A}(k) = \sum_{i=1}^k \sigma_i U^i V^{i\top}, \quad 1 \leq k \leq p \quad (10)$$

(on rappelle que les valeurs singulières sont ordonnées : $\sigma_1 \geq \dots \geq \sigma_p \geq 0$).

Exemple : compression d'image. On considère une image 1432×1910 en niveau de gris. On dispose donc d'une matrice $A \in \mathbb{R}^{1432 \times 1910}$ composée de valeurs entre 0 (noir) et 1 (blanc).

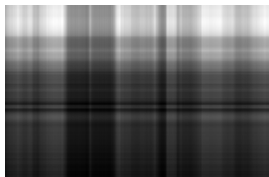


D'après M. Deisenroth, A. Faisal, C. Ong, *Mathematics for machine learnings*, 2019.

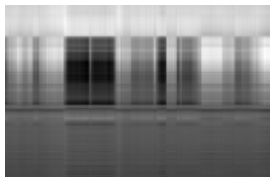
Avec les approximations $\hat{A}$ par SVD, on obtient les images suivantes :



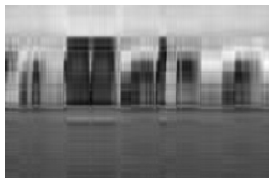
(a) Original image A .



(b) Rank-1 approximation $\hat{A}(1)$.



(c) Rank-2 approximation $\hat{A}(2)$.



(d) Rank-3 approximation $\hat{A}(3)$. (e) Rank-4 approximation $\hat{A}(4)$. (f) Rank-5 approximation $\hat{A}(5)$.

Si on stocke la matrice A complète, on doit stocker $1432 \times 1910 = 2.735120 \cdot 10^6$ nombres. Pour l'approximation $\hat{A}(5)$, on ne stocke que les 5 valeurs singulières et les 5 vecteurs singuliers gauches et droits, soit $5 \times 1432 + 5 \times 1910 + 5 = 16715$ nombres, soit 0.6% du stockage complet....