

Séance 5 : résolution de systèmes linéaires par une méthode directe : factorisations LU, Cholesky et QR

Cette séance fait suite à la séance 4 sur la méthode du pivot de Gauss.

Elle va présenter une extension de la méthode de Gauss par aboutir à une factorisation de la matrice A bien adaptée par la résolution de systèmes en grand nombre : si

$$\underline{A = A_1 A_2 \text{ avec}}$$

$A_1 x = b$ facile à résoudre ¹
(par exemple : A_1 triangulaire ou orthogonale)
et $A_2 x = b$ facile à résoudre,
la résolution $Ax = b$ l'est également ($A_1 \beta = b$ puis $A_2 x = \beta$)

I) Factorisation LU

On cherche à déterminer une factorisation du type $A = L \cdot U$
avec :

- L : triangulaire inférieure ("lower")
- U : " supérieure ("upper")

Cette factorisation est quasiment incluse dans la méthode de Gauss, sans stratégie de pivot. En effet, si aucun échange de ligne n'a été effectué dans Gauss, on a :

$$T = E_{n-1}(\alpha_n^{n-1}) E_{n-2}(\alpha_{n-1}^{n-2}, \alpha_n^{n-2}) \dots E(\alpha_2^1, \alpha_n^1) A$$

matrices de transvection

Un calcul matriciel simple montre que

$$\left(E_1(\alpha_2, \dots, \alpha_n) \right)^{-1} = \begin{pmatrix} 1 & & & \\ \alpha_2 & \diagdown & & 0 \\ & & \ddots & \\ \alpha_n & 0 & & 1 \end{pmatrix}^{-1}$$

On a bien

$$= \begin{pmatrix} 1 & & & \\ -\alpha_2 & \diagdown & & \\ & & \ddots & \\ -\alpha_n & & & 1 \end{pmatrix} \quad \text{2}$$

$$\text{puis } (E_1)^{-1} \dots (E_{n-2})^{-1} (E_{n-1})^{-1}$$

$$\parallel \begin{pmatrix} 1 & & & \\ -\alpha_2^1 & \diagdown & & 0 \\ & & \ddots & \\ -\alpha_n^1 & & & \diagdown & \\ & & & & -\alpha_n^{n-1} & 1 \end{pmatrix}$$

$$\parallel L$$

$$A = \underset{\text{lower}}{\uparrow} L \underset{\text{upper}}{\leftarrow} T$$

Plus précisément, on a :

Théorème : soit $A \in GL_n(\mathbb{R})$. On suppose que tous les mineurs principaux :

$$\Delta_k = \begin{pmatrix} a_{11} & & a_{1k} \\ & \ddots & \\ a_{k1} & & a_{kk} \end{pmatrix} \text{ sont inversibles}$$

Alors, il existe un unique couple (L, U) tel que

$$A = LU \text{ avec}$$

L : triangulaire inférieure et $l_{ii} = 1$ si $i \in \{1, \dots, n\}$
 U : triangulaire supérieure
L'algorithme de calcul de L et U

nécessite $2n^3/3$ opérations / 3
preuve :

* Existence : on montre qu'aucun échange de ligne n'est nécessaire dans le pivot de Gauss avec l'hypothèse :
* 1^{er} pivot : comme Δ_1 est inversible $a_{11} \neq 0$ et peut être choisi comme pivot.

* On suppose qu'aucun échange de lignes n'a été nécessaire après k pivots :
on a alors :

$$A_k = \begin{pmatrix} a_{11} & & & \\ \circ & & & \\ | & & & \\ | & & & \\ \circ & \circ & \circ & \circ \end{pmatrix} \quad \text{?} \quad \text{A}_{k+1}$$

($k \geq 2$)

On a $A_k = E_{k-1}(\dots) \dots E_1(\dots) A$
 Cette relation est également valable
 pour les matrices extraites $k \times k$:

$$\begin{pmatrix} a_{11} & \star \\ \circ & \sim a_{k,k} \end{pmatrix} = \begin{pmatrix} E_{k-1} \end{pmatrix}_{k \times k} \dots \Delta_k$$

Avec le déterminant, on a : 4
 $\left(\prod_{i=1}^{k-1} a_{ii} \right) \widetilde{a_{k,k}} = \det(\Delta_k)$

Comme $\det(\Delta_k) \neq 0$, on a bien $\widetilde{a_{k,k}} \neq 0$
 ce qui permet de le prendre comme
 nouveau pivot. //

* Unicité : par l'absurde, on suppose
 qu'il existe (L_1, U_1) et (L_2, U_2) tq
 $\left. \begin{matrix} L_1 U_1 = A \\ L_2 U_2 = A \end{matrix} \right\}$ on a donc

$$L_2^{-1} L_1 = U_2 U_1^{-1}$$

Gr, l'inverse d'une matrice

triangulaire inférieure (resp. sup)
 est triangulaire inférieure (resp. sup)
 tout comme le produit, ceci implique

que $\underbrace{L_2^{-1} L_1}_{\text{triang. sup}} = \underbrace{U_2 U_1^{-1}}_{\text{triang. inf}} = D$ diagonale

or $L_2^{-1} L_1 = \begin{pmatrix} 1 & & \\ * & \ddots & \\ & & 1 \end{pmatrix} \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix}$
 $= \begin{pmatrix} 1 & & \\ * & \ddots & \\ & & 1 \end{pmatrix}$

On en déduit :

$$L_2^{-1} L_1 = U_2 U_1^{-1} = Id$$

soit $L_1 = L_2$ et $U_1 = U_2$.

* Algorithme : en reprenant la formulation explicite de L et U issue du pivot de Gauss, on constate que les coefficients de L sont issus des

coefficients des transvections utilisées et $U = T$ obtenu à la fin de Gauss

(*) Pour inverser

($u_{ii} \neq 0$)

$$\begin{pmatrix} u_{11} & & * \\ & \ddots & \\ & & u_{nn} \end{pmatrix}$$

on peut faire une remontée avec un second membre forcé : $Tx = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$

$$\rightarrow x = T^{-1} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$$

$$\begin{pmatrix} * & & \\ & \ddots & \\ 0 & & \end{pmatrix}$$

on peut aussi utiliser Cramer par la m-observata

Exemple : factoriser sous forme LU
(si cela est possible) la matrice

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 10 \end{pmatrix}$$

$$\Downarrow \begin{pmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & -6 & -11 \end{pmatrix} \begin{matrix} L_2 \rightarrow L_2 - 4L_1 \\ L_3 \rightarrow L_3 - 7L_1 \end{matrix}$$

$$\Downarrow \begin{pmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & 0 & 1 \end{pmatrix} \begin{matrix} L_3 \rightarrow L_3 - 2L_2 \end{matrix}$$

On a donc

$$L = \begin{pmatrix} 1 & 0 & 0 \\ 4 & 1 & 0 \\ 7 & 2 & 1 \end{pmatrix}$$

$$\text{et } U = \begin{pmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & 0 & 1 \end{pmatrix}$$

Remarques :

* De manière générale, par tout $A \in GL_n(\mathbb{R})$, on peut montrer qu'il existe (P, L, U) tq

$$PA = LU \text{ avec } L, U$$

comme précédemment et P matrice de permutation ($P_{ij} = \delta_{i, \sigma(j)}$) ($\sigma \in \mathcal{S}_n$).

* L'absence de stratégie de pivot peut dégrader le conditionnement de L et U .

* La factorisation LU préserve la structure bande d'une matrice :

si $A = \begin{pmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \end{pmatrix}$, alors

$$L = \begin{pmatrix} 1 & & & & \\ * & 1 & & & \\ 0 & * & 1 & & \\ 0 & * & * & 1 & \\ 0 & * & * & * & 1 \end{pmatrix} \text{ et } U = \begin{pmatrix} \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \end{pmatrix}$$

$$U = \begin{pmatrix} \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \end{pmatrix}$$

Exercice :

* Factorisation $LDSt$ voir TD3 :

$$A = L D S$$

$\uparrow \quad \quad \uparrow \quad \quad \nwarrow$
 $l_{ij} = 1 \quad \quad \text{diagonale} \quad \quad u_{ij} = 1$

→ existence : $A = LU$ et

$$U = \begin{pmatrix} \lambda_1 & u_{2,1} & \dots \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

$$= \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \begin{pmatrix} 1 & \frac{u_{2,1}}{\lambda_1} & \dots \\ & \ddots & \\ 0 & & 1 \end{pmatrix}$$

($\lambda_i \neq 0$ car U inversible)

d'où $A = L D S$

* Unité: $A = L \underbrace{D S}_{U_1} = L \underbrace{D S_2}_{U_2}$

avec l'unité

de LU , on a :

$$\begin{cases} L_1 = L_2 \\ D_1 S_1 = D_2 S_2 \end{cases}$$

soit $D^{-1} D = \begin{pmatrix} S & S^{-1} \\ 2 & 1 \end{pmatrix}$ 8

diagonale tri. sup et terme diagonaux = 1

d'où $D^{-1} D = \begin{pmatrix} S & S^{-1} \\ 2 & 1 \end{pmatrix} = Id$

et $D_1 = D_2, S_1 = S_2$.

II) Factorisation de Cholesky

A présent, on suppose que A est symétrique définie positive, $A \in S_n^+(\mathbb{R})$ (symétrique et $\forall X, XAX > 0$ si $X \neq 0$) (aussi $\sigma_{\mathbb{C}}(A) \subset \mathbb{R}_+^*$, spectre).

Dans ce cas la factorisation LU se réécrit autrement et se simplifie :

Théorème : soit $A \in S_n^{++}(\mathbb{R})$

Il existe une unique matrice B telle que $A = B(\begin{smallmatrix} t \\ B \end{smallmatrix})$ avec

- * B triangulaire inférieure
- * $\forall i \in \{1, \dots, n\}, b_{ii} > 0$

De plus B s'obtient en $\frac{n^3}{3}$ opérations.

preuve :

Comme $A \in S_n^{++}(\mathbb{R})$, le critère de Sylvester indique que $\det(\Delta_k) > 0 \forall k \in \{1, \dots, n\}$. On en déduit que

$$A = \underbrace{L}_{\substack{\text{lower} \\ l_{ii}=1}} \underbrace{U}_{\text{upper}}$$

De plus, toujours avec Sylvester, $u_{ii} > 0 \forall i \in \{1, \dots, n\}$. On a alors

$$A = \underbrace{L} \underbrace{D} \underbrace{D^{-1}} \underbrace{U}$$

$$D = \text{diag}(\sqrt{u_1}, \dots, \sqrt{u_n})$$

En notant $B = LD$ et $C = D^{-1}U$

Comme $A = {}^t A$, on a :

$$({}^t C) {}^t B = B C \Rightarrow$$

$$\underbrace{({}^t C)^{-1}}_{\text{low}} \underbrace{B}_{\text{bw}} = \underbrace{({}^t B)}_{\text{up}} \underbrace{C^{-1}}_{\text{up}} = \text{Id}$$

diag $\uparrow$ $\frac{1}{\sqrt{u_{ii}}}$

D'où $({}^t C)^{-1} B = \text{Id}$ et

$C = {}^t B$, soit $A = B {}^t B$ //

coeff $\uparrow$ $\sqrt{u_{ii}} > 0$

* Unité : on suppose

$$A = B {}^t B = \underbrace{(B \Delta^{-1})}_L \underbrace{(\Delta {}^t B)}_U$$

où $\Delta = \text{diag}(u_{11}, \dots, u_{nn})$ 10

soit " $A = L \cdot U$ "

d'où $\Delta {}^t B$ est unique (unité de U)

En regardant les termes sur la diagonale, on en déduit u_{ii}^2 est unique. Comme $u_{ii} > 0$, u_{ii} est unique ; ainsi Δ est unique et B aussi.

* Avec Gauss, la factorisation Cholesky nécessite $\frac{2n^3}{3}$ opérations.

Il existe en fait un algorithme deux fois plus rapide, appelé

méthode des coefficients indéterminés :

On part de $A = B^t B$ et
en notant $B = [b_{ij}]$, on écrit
le système général :

$$a_{ij} = \sum_{k=1}^{\min(i,j)} b_{i,k} b_{j,k}$$

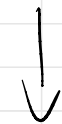
On résout ce système colonne par colonne :

$$\begin{cases} a_{11} = b_{11}^2 \\ a_{2,1} = b_{2,1} b_{11} \\ \vdots \\ a_{n,1} = b_{n,1} b_{11} \end{cases} \Rightarrow \begin{cases} b_{11} = \sqrt{a_{11}} \\ b_{2,1} = \frac{a_{2,1}}{b_{11}} \\ \vdots \\ b_{n,1} = \frac{a_{n,1}}{b_{1,1}} \end{cases}$$

$$\begin{cases} a_{ii} = b_{i,1}^2 + \dots + b_{i,i}^2 \\ a_{i+1,i} = b_{i+1,1} b_{1,i} + \dots + b_{i+1,i} b_{i,i} \end{cases}$$

justifié
par existence d'
une seule

$$\Rightarrow \begin{cases} b_{ii} = \sqrt{a_{ii} - \sum_{j=1}^{i-1} b_{ij}^2} \\ b_{i+1,i} = \frac{1}{b_{ii}} \left(\dots \right) \\ b_{n,i} = \dots \end{cases}$$



$$b_{n,n} = \sqrt{a_{n,n} - \sum_{j=1}^{n-1} b_{nj}^2}$$

Comme le coût des n racines carrées est en $O(n)$ (algorithme phénicien) on peut estimer le coût global de cette méthode :

$$\left\{ \begin{array}{l} \frac{n(n-1)}{2} \text{ divisions} \\ \frac{n(n-1)(n+1)}{6} \text{ add.} \\ \text{mult.} \end{array} \right. \Rightarrow \frac{n^3}{3} \text{ opérations}$$

Exercice : implémenter l'algorithme des coefficients indéterminés avec Python. (à envoyer)

Remarque :

* Il existe de nombreux cas de modélisation conduisant à la résolution de systèmes $Ax = b$ avec $A \in \mathcal{S}_n^{++}(\mathbb{R})$. Par exemple, on aboutit à la matrice dit du Laplacien discrétisé :

$$A = \begin{pmatrix} 2 & -1 & & \\ -1 & 2 & & \\ & & \ddots & \\ 0 & & -1 & 2 \end{pmatrix}$$

dans de nombreuses situations

$$\|X - AX\|_2^2 = \frac{1}{2} \sum_{i=0}^n (x_{i+1} - x_i)^2 \quad (\text{avec } x_0 = x_n = 0)$$

Exercice (à faire) : factorisation
LDS avec $A \in \mathcal{S}_n(\mathbb{R})$ (TD3)
Exemple : décomposition Cholesky
d'une matrice $n=4$ (TD3)

III Factorisation QR

Gn a le résultat suivant :

Théorème : soit $A \in \text{GL}_n(\mathbb{R})$.
Il existe un unique couple (Q, R)
tel que $A = Q \cdot R$ avec

* Q : orthogonale
* R : triangulaire supérieure avec
 $r_{ii} > 0$

Cette factorisation peut s'effectuer
en $\frac{4n^3}{3}$ opérations

preuve :

* Existence : orthonormalisation
de Gram-Schmidt :

$$A = \begin{pmatrix} c_1 & \dots & c_n \\ | & & | \\ q_1 & \dots & q_n \end{pmatrix}$$

Gn note $(q_1, \dots, q_n)$ la base
orthonormalisée par Gram Schmidt

de $(C_1, \dots, C_n)$,

$$\begin{cases} q_1 = \frac{C_1}{\|C_1\|} \\ q_2 = \frac{C_2 - \langle C_2, q_1 \rangle q_1}{\| \quad \|} \\ \vdots \\ q_n = \frac{C_n - \sum_{i=1}^{n-1} \langle C_n, q_i \rangle q_i}{\| \quad \|} \end{cases}$$

soit matriciellement :

$$\begin{pmatrix} C_1 & \dots & C_n \\ | & & | \end{pmatrix} = \begin{pmatrix} q_1 & \dots & q_n \\ | & & | \end{pmatrix} \begin{pmatrix} \|C_1\| & & \\ 0 & \ddots & * \\ 0 & & \end{pmatrix} \quad \text{14}$$

(red arrows pointing to the diagonal elements and the asterisk, with a red circle and arrow pointing to the bottom right)

(opération sur les colonnes)

multiplication à droite

Soit $A = Q R$

$\begin{matrix} \uparrow & \uparrow \\ O_n(\mathbb{R}) & \text{upper, } r_{ii} > 0 \end{matrix}$

* unicité : $Q, R = Q_2 R_2$

$$\Rightarrow \underbrace{R_2 Q_2}_{\substack{\text{red } h_{ii} > 0 \\ T}} = \underbrace{R_2 (R_1^{-1})}_{\text{upper}}$$

$$T T^T = R_1 Q_1 R_1^T Q_1^T = I_d$$

Par unicité de Cholesky par Id,
on a $T = Id \Rightarrow R_2 = R_1$ puis
 $Q_2 = Q_1$.

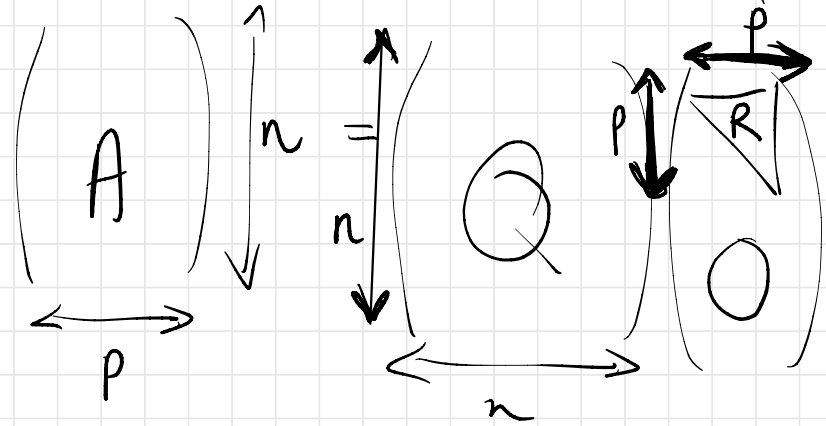
* Algorithme : la méthode précédente
nécessite $4n^3/3$ opérations, mais elle
est instable numériquement.

Une méthode plus stable est la
méthode de Householder (voir TD3)

Remarques

* La méthode QR s'étend à
des matrices non carrées (résolution

de systèmes non carrés)



* Cette factorisation est aussi utilisée
par la recherche de valeurs propres
(méthode QR).

Par finir, une remarque générale
sur la méthode de Gauss :
cette méthode sert aussi :

* à calculer un déterminant:

$$\det(A) = \pm \det(L)$$

↖ nbre d'échanges
de lignes

* à calculer l'inverse d'une matrice:

$$Ax = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \rightarrow x = A^{-1} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

(ou transformation de A en Id)

* à calculer le rang d'une matrice:

$$A \sim \begin{pmatrix} * & & 0 \\ 0 & * & 0 \\ 0 & 0 & 0 \end{pmatrix}$$