Séance 5 : résolution de systèmes linéaire par une méthode directe : méthodes de factorisation

(Cette séance suit la séance 4 portant sur la méthode de Gauss)

Objectif : résoudre le système cramérien

$$\boxed{Ax = b} \quad (A \in GL_n(\mathbb{R}),\ b \in \mathbb{R}^n)$$

de manière :

→ exacte (aux arrondis numériques près)

→ rapide (en $O(n^3)$ opérations)

→ robuste (par rapport aux arrondis effectués)

La méthode de Gauss remplit ces 3 critères. Cependant, elle doit être répétée à chaque changement de second membre b.

Le principe des factorisations consiste à écrire A sous la forme :

$$A = A_1 A_2$$

(en $O(n^3)$ opérations) et où A_1 et A_2 sont faciles à "inverser" (en $\ll n^3$ opérations), par exemple : triangulaire, orthogonale, etc...

Pour résoudre $Ax=b$, on résoudra successivement : $A_1 u = b$

puis : $A_2 x = w$

I) Factorisation LU

Cette factorisation permet de décomposer A sous la forme :

$$A = \underset{\substack{\uparrow \\ \text{lower} \\ \text{(triangulaire)}}}{L}\ \underset{\substack{\uparrow \\ \text{upper}}}{U}$$

en reprenant la méthode d'Euler et en supposant qu'il n'y a pas eu d'échange de ligne.

Dans ce cas, on a vu :

$$\underset{\substack{\uparrow \\ \text{triangulaire} \\ \text{supérieure}}}{A_{n-1}} = E_{n-1}\, E_{n-2} \cdots E_1\, A$$

transvection

On a donc :

$$A = E_1^{-1} E_2^{-1} \cdots E_{n-1}^{-1} A_{n-1}$$

Si $E_i = \begin{array}{c} \begin{array}{ccccc} & & i & & \\ & & \downarrow & & \end{array} \\ \begin{pmatrix} 1 & & & & \\ & \ddots & & & 0 \\ & & 1 & & \\ 0 & & \alpha_{i+1} & \ddots & \\ & & \vdots & & \\ & & \alpha_n & & 1 \end{pmatrix} \end{array}$, on a

$$E_i^{-1} = \begin{pmatrix} 1 & & & & \\ & \ddots & & & 0 \\ & & 1 & & \\ & & -\alpha_{i+1} & \ddots & \\ & & \vdots & & \\ & & -\alpha_n & & 1 \end{pmatrix}$$

De plus, on remarque que

$$\begin{pmatrix}1&&&&\\&\ddots&&&\\&&1&&\\&&-\beta_{n-1}&1&\\&&-\beta_n&0&1\end{pmatrix}\begin{pmatrix}1&&&\\&\ddots&&\\&&1&\\&&-\alpha_n&1\end{pmatrix} =$$

$$\begin{pmatrix}1&&&&\\&\ddots&&&\\&&1&&\\&&-\beta_{n-1}&1&\\&&-\beta_n&-\alpha_n&1\end{pmatrix}$$

Au final, on a donc

$$A = \begin{pmatrix}1&&&&\\ \vdots&\ddots&&&\\ \vdots&&1&&\\ \vdots&&-\beta_{n-1}&1&\\ \cdots&\cdots&-\beta_n&-\alpha_n&1\end{pmatrix} A_{n-1}$$

c'est à dire :

$A = LU$ sans calcul supplémentaire.

On rassemble les résultats dans le théorème suivant :

Théorème : soit $A \in GL_n(\mathbb{R})$.

On suppose que les mineurs principaux :

$$\Delta_k = \begin{pmatrix}a_{11}&\cdots&a_{1,k}\\ \vdots&&\vdots\\ a_{k,1}&\cdots&a_{k,k}\end{pmatrix}$$

sont tous inversibles $(1 \leq k \leq n)$

Dans ce cas, il existe une unique factorisation :

$$\boxed{A = LU}$$

où

* L : triangulaire inférieure, telle que $\forall i \in \{1, \dots, n\}$, $l_{ii} = 1$
* U : triangulaire supérieure inversible

preuve :

→ existence : on raisonne par récurrence sur k, en montrant qu'il n'est pas nécessaire de faire un changement de ligne.

* $k = 1$: $a_{11} \neq 0$: ok
* $k \to k+1$: en reprenant la preuve de la justification de Gauss, on a, après l'itération k :

$$A_{k+1} = \begin{pmatrix} a_{11} & & & * & \\ 0 & \ddots & & & \\ & 0 & \tilde{a}_{k+1,k+1} & & \\ \vdots & & & & \\ 0 & \cdots & 0 & & \end{pmatrix} = E_k \cdots E_1 A$$

En tronquant sur les $(k+1)$ lignes/colonnes, et en prenant le déterminant, on a

$$\underset{\neq 0}{a_{11}} \cdots \underset{\neq 0}{\tilde{a}_{k,k}} \, \tilde{a}_{k+1,k+1} = 1 \cdot 1 \underset{\neq 0}{\det(\Delta_{k+1})}$$

On en déduit que $\tilde{a}_{k+1,k+1}$ et donc qu'un changement de ligne dans Gauss n'est pas utile.

* <u>Unicité</u> : par l'absurde, on suppose que

$$A = L_1 U_1 = L_2 U_2$$

On a donc

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

Or, l'inverse d'une triangulaire inférieure est une triangulaire inférieure (en effet, on peut calculer l'application linéaire réciproque par un principe de descente, elle est bien triangulaire inférieure)

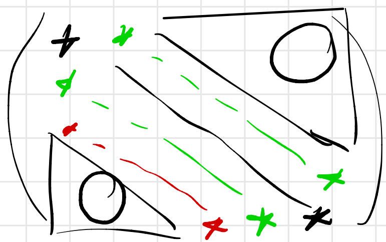
De plus $L_2^{-1} = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ * & & 1 \end{pmatrix}$

On en déduit que $L_2^{-1} L_1$ et $U_2^{-1} U_1$ sont deux matrices diagonales. Or, elles possèdent des "1" sur la diagonale. On a donc $L_2^{-1} L_1 = U_2^{-1} U_1 = Id$ //

Remarques

* En termes d'algorithmes, LU se confond avec Gauss
* Si A est quelconque, on peut montrer qu'il existe P, matrice de permutation ($P = P_\sigma$ où $(P_\sigma)_{i,j} = \delta_{\sigma(i),j}$) telle que $PA = L.U$.
* Attention au conditionnement de L et U

* La factorisation L.U. préserve la structure bande d'une matrice (ex. TD5)

$A =$ 

II) Factorisation de Cholesky

On s'intéresse à la factorisation des matrices symétriques, définies positives ($A \in \mathcal{S}_n^{++}(\mathbb{R})$).

On sait déjà qu'une factorisation LU est possible (critère de Sylvester).

On va plus loin avec le th. suivant :

Théorème : soit $A \in \mathcal{S}_n^{++}(\mathbb{R})$.

Il existe une unique matrice B telle que :

$$\boxed{A = B\,({}^tB)}$$

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

De plus, le calcul de B nécessite de l'ordre de $n^3/3$ opérations.

preuve :

→ existence : on part de

$$A = L.U.$$

On peut montrer que

$$\forall i \in \{1, \dots n\}, \ u_{ii} > 0$$

En effet, en reprenant l'expression :

$$a_{11} \cdots\cdots a_{k,k} \, a_{k+1,k+1} = 1 \cdot 1 \det(\Delta_{k+1})$$

>0 >0 $\Downarrow$ >0 >0

(avec Sylvester)

On note alors $D = \operatorname{diag}(\sqrt{u_{11}}, \dots \sqrt{u_{n,n}})$

et on écrit

$$A = \underbrace{(LD)}_{\substack{B \\ (\text{tri. inf})}} \underbrace{(D^{-1}U)}_{\substack{C \\ (\text{tri. sup.})}}$$

Or $A = {}^tA$, d'où

$${}^tC\,{}^tB = B\,C \text{ puis}$$

$$\underbrace{(B^{-1})({}^tC)}_{\text{tri. inf}} = \underbrace{C({}^tB)^{-1}}_{\text{tri. sup.}}$$

Cette matrice est diagonale, à coefficients diagonaux unité. D'où $B = {}^tC$

et donc $A = B({}^tB)$ //

→ Unicité : on se ramène à l'unicité de LU. Soit $\Delta = \operatorname{diag}(u_{11}, \dots u_{n,n})$

on a $A = \underbrace{(B\Delta^{-1})}_{L}\underbrace{(\Delta\,{}^tB)}_{U}$

Ainsi $\Delta^t B$ est unique.
En particulier, u_{ii}^2 est unique donc u_{ii} aussi (car $u_{ii} > 0$).
Δ est donc unique, B l'est aussi.

→ nbre d'opérations : avec Gauss, cette factorisation nécessite de l'ordre de $2n^3/3$ opérations. On repart ici du résultat

$$A = B\,{}^tB$$

pour déterminer les coefficients de B.
On écrit le système :

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

qu'on résout colonne par colonne :

$$\left.\begin{aligned} b_{1,1} &= \sqrt{a_{11}} \\ b_{2,1} &= \frac{a_{2,1}}{b_{1,1}} \\ &\vdots \\ b_{n,1} &= \frac{a_{n,1}}{b_{1,1}} \end{aligned}\right\} \text{(colonne 1)}$$

puis

$$\left.\begin{aligned} b_{i,i} &= \sqrt{a_{ii} - \sum_{k=1}^{i-1} b_{i,k}^2} \\ b_{i+1,i} &= \dots \\ &\vdots \\ b_{n,i} &= \dots \end{aligned}\right\} \text{(colonne i)}$$

justifié !

Cette méthode, appelée méthode des coefficients indéterminés, nécessite :

* n racines carrées
* $\frac{n(n-1)}{2}$ divisions
* $\frac{n(n-1)(n+1)}{6}$ add.
* // mult

soit de l'ordre de $n^3/3$ op.

III Factorisation QR

Cette 3ème factorisation à connaître s'écrit $A = Q R$

orthogonale — tri. sup.

A noter que pour résoudre un système du type $Qu = b$, on écrit $u = {}^tQ b$ (soit $O(n^2)$ opérations)

On a le résultat suivant :

Théorème : soit $A \in GL_n(\mathbb{R})$.

Il existe une unique factorisation

$$\boxed{A = QR}$$ avec

* Q : matrice orthogonale
* R : matrice triangulaire supérieure avec $\forall i \in \{1, \dots, n\}$, $r_{ii} > 0$

preuve :

→ existence : comme A est inversible, les colonnes $\{C_1, \dots, C_n\}$ de A forment une base de $\mathbb{R}^n$.

On l'orthonormalise par le principe de Gram-Schmidt en une b.o.n. $\{q_1, \dots, q_n\}$, telle que :

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

Matriciellement, ces relations peuvent s'écrire :

$$\underbrace{\begin{pmatrix} C_1 & \cdots & C_n \\ | & & | \\ | & & | \end{pmatrix}}_{A} = \underbrace{\begin{pmatrix} q_1 & \cdots & q_n \\ | & & | \\ | & & | \end{pmatrix}}_{Q} \underbrace{\begin{pmatrix} \overset{>0}{u_{11}} & * & \\ 0 & u_{22} & * \\ | & 0 & \ddots \\ 0 & | & \ddots \\ & 0 & \end{pmatrix}}_{R}$$

(opérations sur les colonnes ≡ mult. à droite)

→ unicité : on part de $Q_1 R_1 = Q_2 R_2$ puis

$$T = {}^tQ_2 Q_1 = \underbrace{R_2 R_1^{-1}}_{\text{tri. sup}}$$

On a ${}^tT \cdot T = {}^tQ_1 Q_2 {}^tQ_2 Q_1 = Id$

De plus $r_{ii} > 0$. On se trouve dans

le cas d'une décomposition de Cholesky par l'identité. Par unicité de Cholesky, on en déduit que $T = Id$.

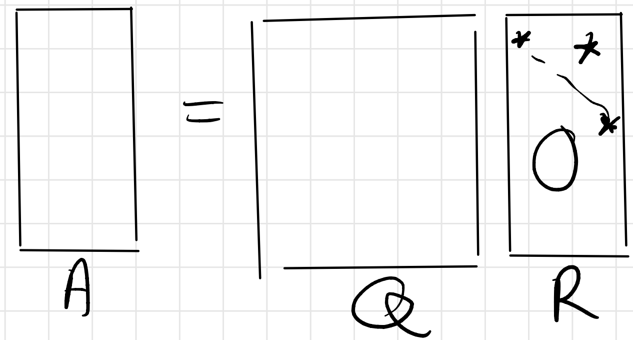
Remarques

* En termes de robustesse, l'algorithme de Gram Schmidt est instable (division par un petit nombre éventuel).

* Il existe un algorithme robuste, à l'aide de matrices de symétrie orthogonales hyperplanes, dit de Householder (voir TD5).

* Cet algorithme nécessite de l'ordre de $\frac{4n^3}{3}$ opérations.

* On peut étendre la factorisation QR aux matrices non carrées :



(utile par les problèmes de type moindres carrés)

* La méthode QR est à la base d'une méthode de calcul de valeurs propres.

TD5, exercice 1 :

* Factorisation LU de

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

On trouve :

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

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

En effet :

$$A_1 = \begin{pmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & -6 & -11 \end{pmatrix} \quad \begin{array}{l} L_2 \to L_2 - 4L_1 \\ L_3 \to L_3 - 7L_1 \end{array}$$

$$\underbrace{A_2 = \begin{pmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & 0 & 1 \end{pmatrix}}_{U} \quad L_3 \to L_3 - 2L_2$$

Ex 4 : Factorisation de Cholesky

de :

$$A = \begin{pmatrix} 4 & 0 & 12 & -6 \\ 0 & 1 & 2 & 1 \\ 12 & 2 & 49 & -4 \\ -6 & 1 & -4 & 51 \end{pmatrix}$$

$$B = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 6 & 2 & 3 & 0 \\ -3 & 1 & 4 & 5 \end{pmatrix}$$

Sol...

En effet :

$A = B^T B$ soit :

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

On détermine B colonne par colonne :

$$\begin{cases} b_{11} = \sqrt{4} = 2 \\ a_{2,1} = b_{2,1} b_{11} \Rightarrow b_{2,1} = 0 \\ a_{3,1} = b_{3,1} b_{11} \Rightarrow b_{3,1} = 6 \\ b_{4,1} = -3 \end{cases}$$

$$\begin{cases} b_{22} = \sqrt{a_{22} - b_{2,1}^2} = 1 \\ a_{3,2} = b_{3,1} b_{2,1} + b_{32} b_{22} \\ \Rightarrow b_{3,2} = 2 - 0 = 2 \\ b_{4,2} = a_{4,2} - b_{4,1} b_{2,1} = 1 \end{cases}$$

etc...