Méthodes Numériques de base

Chapitre 5 – FACTORISATION DE CHOLESKY

.

Ensimag 1ère année MMIS – 2020-2021

Grenoble INP Ensimag

Rappel du plan du cours

Cours

1. différences finies pour pb lim 1D
2. interpolation polynomiale
3. méth. itératives linéaires
4. méth. itératives linéaires
5. factorisation LU A inv.
6. **factorisation Cholesky** A sym. déf > 0
7. équations différentielles
8. équations différentielles
9. équations non linéaires
10. équations non linéaires
11. optimisation

TD

1. différences finies
2. interpolation polynomiale
3. méth. itératives linéaires
4. méth. itératives linéaires
5. factorisation LU
6. **factorisation Cholesky**
7. conditionnement matriciel
8. équations différentielles
9. moindres carrés linéaires
10. équations non linéaires
11. équations non linéaires

(1) Factorisation de Cholesky

$A \in M_N(\mathbb{R})$ symétrique, définie positive.

Rappel : A est symétrique déf > 0 ssi

1. $A^T = A$
2. $\mathbf{x}^T A \mathbf{x} \geq 0 \ \forall \mathbf{x} \in \mathbb{R}^n$
3. $\mathbf{x}^T A \mathbf{x} = 0 \Rightarrow \mathbf{x} = 0$

A symétrique déf > 0 ssi toutes les v. p. > 0 / A déf $> 0 \Rightarrow A$ inversible / A symétrique $\Rightarrow A$ est diagonalisable ($\exists P$ t.q. $A = P^T D P$).

Théorème 3 Factorisation de Cholesky

Soit $A \in M_n(R)$ symétrique, définie positive. Il existe (au moins) une matrice réelle triangulaire inférieure B t.q.

$$\boxed{A = \cancel{B^T B}.} \quad BB^T \qquad B =$$

De plus, si on impose que les éléments diagonaux de B soient tous > 0, alors la factorisation $A = B^T B$ est unique.

(2) Preuve du Théorème (1/3)

Preuve : *On va établir qu'une matrice symétrique déf > 0 vérifie les conditions du th. 1. Elle admet donc une unique factorisation LU. En plus, cette factorisation ne fait qu'intervenir qu'une seule matrice.*

<u>(Existence) :</u> Notons Δ_k la sous-matrice d'éléments a_{ij}, $1 \leq i, j \leq k$. Δ_k sont inversibles, car elles sont symétriques, déf > 0.

En effet, si $\chi = (x_1, \dots, x_k)$ et $X = (x_1, \dots, x_k, 0 \dots 0)^T$ on a $\chi^T \Delta_k \chi = X^T AX$. ≥ 0

Donc d'après le th. 1 du chapitre précédent, on a de façon unique

$$A = LU$$

on cherche

L est triang. inf. avec $L_{ii} = 1$ et U triang. sup., U inversible ($u_{ii} \neq 0$.)

Notons $D = diag(u_{ii})$. On a

$$A = A^T = U^T L^T = \underbrace{U^T D^{-1}}_{L} \underbrace{DL^T}_{U}.$$

$U =$

$D = diag(u_{ii})$

$D^{-1} = diag(\frac{1}{u_{ii}})$

$U^T =$

$U^T D^{-1} =$

$L =$

$L^T =$

$D L^T = (u_{11} \dots u_{nn})$

Par unicité de la décomposition LU il vient $DL^T = U$ et donc

$$\Longrightarrow \quad A = LDL^T.$$

$\Rightarrow$ A et D ont la même signature. (le m̂ #vp > 0/m̂ #vp $= 0$/m̂ #vp < 0)

(2) Preuve du Théorème (2/3)

$(AB)^T = B^T A^T$

$\Delta = \begin{pmatrix} \ddots & & 0 \\ & \sqrt{u_{ii}} & \\ 0 & & \ddots \end{pmatrix}$

Comme A est sym. déf >0

$\Rightarrow$ toutes les v.p. de $A > 0$

$\Rightarrow$ toutes les v.p. de $D > 0 \quad \Rightarrow \quad u_{ii} \neq 0 \quad \forall i.$

Soit $\delta_i = \sqrt{u_{ii}}$ et $\Delta = diag(\delta_i) = diag(\sqrt{u_{ii}})$.

$A = L\underline{D}L^T$

$D = \Delta^2 = \Delta\Delta^T$

$\Rightarrow A = L\Delta\Delta^T L^T = \underbrace{(L\Delta)}_{B}\underbrace{(L\Delta)^T}_{B^T}$. Posons $B := (L\Delta)$.

$A = BB^T$

Donc $A = BB^T$, c.a.d. il existe une factorisation de Cholesky.

<u>(Unicité) :</u> Supposons qu'il existe B_1, B_2 triangulaires inférieures t.q.

$A = B_1 B_1^T = B_2 B_2^T \quad | \; B_2^{-1} \quad | \; (B_1^T)^{-1}$

B_2 est inversible, car A est inversible (Th. 1 : A inv. $\Rightarrow A = LU$ avec U inv.)

$\underbrace{B_2^{-1} B_1}_{(1)} = \underbrace{B_2^T (B_1^T)^{-1}}_{(2)} =: D \qquad \Rightarrow$ c'est une matrice diagonale$D$

$B_1, B_2 = \quad B_2^{-1} = \quad B_2^T = \quad B_1^T = \quad (B_1^T)^{-1} =$

$= \quad = D$

(2) Preuve du Théorème (3/3)

$$\begin{aligned} D &\overset{(2)}{=} B_2^T (B_1^T)^{-1} = (B_1^{-1} B_2)^T \quad \text{on utilise } (M^T)^{-1} = (M^{-1})^T \\ &= (\underbrace{(B_2^{-1} B_1)^{-1}}_{D^{-1}})^T \quad \text{on utilise } M^{-1}N = (N^{-1}M)^{-1} \\ &\overset{(1)}{=} (D^{-1})^T = D^{-1} \end{aligned}$$

Alors $D = D^{-1} \quad \Rightarrow \quad D^2 = I \Rightarrow (B_1)_{ii}^2 = (B_2)_{ii}^2 \quad \forall i$

Choisir $\left.\begin{array}{l}(B_1)_{ii} > 0 \\ (B_2)_{ii} > 0\end{array}\right\} \Rightarrow \boxed{B_1 = B_2.}$ ⇒ unicité de la fact. Cholesky.

Parce que $D \overset{(2)}{=} B_2^{-1} B_1 \quad \Leftrightarrow \quad B_2 D = B_1$

or, $D_{ii} = \pm 1$ donc $B_1 = B_2$ à un signe dans la diagonale près !

$\Rightarrow \quad (B_1)_{ii} = \pm (B_2)_{ii}.$

(3) Résolution de $A\mathbf{x} = \mathbf{b}$

Pour résoudre $A\mathbf{x} = \mathbf{b}$, c.à.d. $B\underbrace{B^T\mathbf{x}} = \mathbf{b}$, $(\mathbf{x}, \mathbf{b} \in I\!R^N)$ on résoud successivement

$$\begin{cases} B\mathbf{y} = \mathbf{b} & \text{étape de ''descente''} \\ B^T\mathbf{x} = \mathbf{y} & \text{étape de ''remontée''} \end{cases}$$

y

Remarque :

Même remarque que pour la factorisation LU : Cette méthode est intéressante lorsque l'on doit résoudre plusieurs systèmes linéaires avec des second membres **b** *différents, mais la même matrice A.*

Dans ce cas il n'y a qu'une seule factorisation à calculer, et pour chaque second membre une étape de déscente et remontée.

Ces dernières sont beaucoup moins coûteux en temps de calcul.

(4) Calcul de la matrice B (1/2)

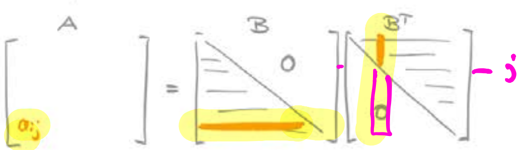
Factorisation de Cholesky $A = BB^T$. B peut être calculé à partir d'une factorisation LU, mais on expose ici une méthode moins coûteuse en temps de calcul.

La factorisation peut se faire par identification des coefficients situés dans la partie triang. inf. de A, i.e. pour les éléments a_{ij}, $i \geq j$, par

$$\boxed{a_{ij} = \sum_{k=1}^{j} b_{ik} b_{jk}} \qquad \text{pour } i \geq j \qquad (A \text{ sym.})$$

Explication :

Multiplication de matrices :



$$a_{ij} = \sum_{k=0}^{N} b_{ik} \underbrace{b^T_{kj}}_{=0 \text{ pour } k>j}$$

$$= \sum_{k=1}^{j} b_{ik} b^T_{kj} = \sum_{k=1}^{j} b_{ik} b_{jk} \qquad (b^T_{kj} = b_{jk})$$

(4) Calcul de la matrice B (2/2)

$$a_{ij} = \sum_{k=1}^{j} b_{ik} b_{kj}$$

Supposons connu les colonnes 1 à $j-1$ de B et calculons

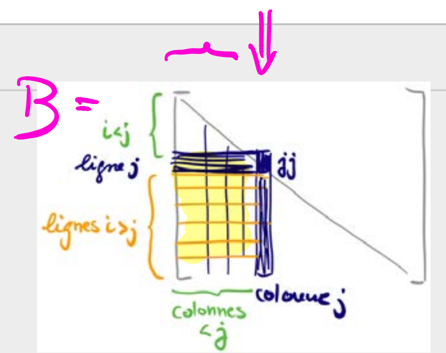
$$\begin{cases} i = j : & a_{jj} = \sum_{k=1}^{j-1} b_{jk}^2 + b_{jj}^2 \\ i > j : & a_{ij} = \sum_{k=1}^{j-1} b_{ik} b_{jk} + b_{ij} b_{jj} \end{cases}$$

Calcul de B

colonne j , $j = 1 \ldots N$

$$\Longrightarrow \begin{cases} i = j : & b_{jj} = \sqrt{a_{jj} - \sum_{k=1}^{j-1} b_{jk}^2} \\ & \quad - \quad + \quad * \\ i > j : & b_{ij} = \frac{1}{b_{jj}} \left(a_{ij} - \sum_{k=1}^{j-1} b_{ik} b_{jk} \right) \end{cases}$$

$$\underbrace{- \quad + \quad *}_{k=1 \,..\, j}$$



(5) Algorithme

$A = BB^T$

Algorithme avec une progression en **"colonne par colonne"**

Cet algorithme travaille en mémoire constant dans la place alloué à la matrice A donnée en entrée, cad les valeurs de b_{ij} écrasent les valeurs de a_{ij}.

FOR $j = 1, \ldots, N$

FOR $i = j, \ldots, N$

FOR $k = 1, \ldots, j-1$

$b_{ij} = a_{ij} := a_{ij} - \underbrace{a_{ik}a_{jk}}_{\hat{=} b_{ik}b_{jk}}$

la valeur de a_{ij} n'est plus utilisée ensuite. Sa place mémoire peut donc être re-écrite.

ENDFOR

IF i=j

THEN $a_{ii} := \sqrt{a_{ii}}$

ELSE $a_{ij} := \frac{a_{ij}}{a_{jj}}$

ENDFOR

ENDFOR

(6) Coût (1/2)

Evaluons le coût du calcul de B et de la résolution $BB^T\mathbf{x} = \mathbf{b}$.

On associe l'extraction de $\sqrt{(\ )}$ à une opération arithmétqiue élémentaire (ce qui est une approximation, car cette opération est plus coûteuse que les opérations élémentaires $+, -, *, /$).

Calul de la colonne j de B $\quad j = 1 — N$

calcul de b_{jj} (1 fois) : $(j-1)$ soustr., $(j-1)$ multi., 1 $\sqrt{(\ )}$

calcul de b_{ij} ($\underbrace{N-j}_{j+1 \leq i \leq N}$ fois) : $(j-1)$ soustr., $(j-1)$ multi., 1 division

Donc la j^{me} étape (colonne) nécessite :

$C_1 = \sum$ $2(N-j+1)(j-1)$ soustr. et multi.

$C_2 = \sum$ $(N-j)$ divisions

$C_2 = \sum$ $1\ \sqrt{(\ )}$

(6) Coût (2/2)

Au total (pour les N colonnes) :

$$C_1 = 2\sum_{j=1}^{N}(N-j+1)(j-1) = 2\left[(N+2)\sum_{j=1}^{N} j - \sum_{j=1}^{N} j^2 - N^2 - N\right]$$

$$\approx 2\left[(N+2)\frac{N(N-1)}{2} - \frac{N(N-1)(N-0.5)}{3} - N^2 - N\right]$$

$$\approx 2(\tfrac{1}{2}N^3) - 2(\tfrac{1}{3}N^3) + O(N^2)$$

$$\approx \tfrac{1}{3}N^3 + O(N^2) \text{ soustractions et multiplications}$$

$$\sum_{j=1}^{N} j = \frac{N(N+1)}{2} = O(N^2)$$

$$\sum_{j=1}^{N} j^2 = \frac{1}{3}N^3 + O(N^2)$$

$$C_2 = \sum_{j=1}^{N}(N-j) \approx O(N^2) \text{ divisions}$$

$$C_3 = N\sqrt{()} \quad \approx O(N)$$

Le coût de la factorisation de Cholesky

Le nombre opérations s'élève à $\frac{1}{3}N^3 + O(N^2)$

S'y rajoutent $O(N^2)$ opérations pour les étapes de descente et remontée.

Remarque : pour mémoire, la factorisation LU coûte $\frac{2}{3}N^3 + O(N^2)$ *opérations.*