Factorisation de Cholesky

On considère un système linéaire $A\,x = b$ avec $b \in \mathrm{IR}^N$ et $A \in M_N(\mathbb{R})$ symétrique définie positive ($y^T A\,y > 0\ \forall y \in \mathrm{IR}^N \setminus \{0\}$) et donc A inversible. La méthode de Cholesky pour résoudre ce type de système est basée sur le résultat suivant:

Théorème (factorisation de Cholesky)

Soit $A \in M_N(\mathrm{IR})$ symétrique définie positive. Alors il existe une unique matrice $T \in M_N(\mathrm{IR})$ triangulaire inférieure telle que :
$t_{ii} > 0$ et $A = T\,T^T$.

Remarque :

Réciproquement, l'existence d'une factorisation de Cholesky implique A symétrique définie positive.

(car $x^T Ax = \left\|T^T x\right\|_2^2$)

En utilisant la factorisation de Cholesky, on a donc:

$$A\,x = b \Leftrightarrow \begin{cases} A = T\ T^T \\ T\,(T^T x) = b \end{cases} \Leftrightarrow \begin{cases} A = T\ T^T \\ T^T x = y \\ Ty = b \end{cases}$$

La méthode de Cholesky pour résoudre $A\,x = b$ consiste en 3 étapes:

* factorisation de Cholesky $A = T\ T^T$
* résolution du système triangulaire $Ty = b$ (descente)
* résolution du système triangulaire $T^T x = y$ (remontée)

Cette décomposition est très intéressante lorqu'on doit résoudre plusieurs systèmes de même matrice A sym def >0 avec des seconds membres b différents: la factorisation est effectuée une seule fois (c'est l'opération la plus coûteuse)

Nous allons donner un algorithme permettant de calculer la factorisation $A = T\ T^T$ en $\approx N^3/3$ opérations arithmétiques élémentaires lorsque $N \to \infty$ (pour une matrice A pleine)

Le coût total pour résoudre le système linéaire $Ax = b$ par la méthode de Cholesky est donc $\approx N^3/3$ lorsque $N \to \infty$:

* factorisation de Cholesky $A = T\ T^T$: $\approx N^3/3$
* résolution de $Ty = b$: $O(N^2)$ car système triangulaire
* résolution de $T^T x = y$: $O(N^2)$ car système triangulaire

$\Rightarrow$ moins coûteux que méthode de Gauss $\approx 2N^3/3$

Plan :

1- Preuve du théorème de factorisation
de Cholesky

2- Calcul de la factorisation :

- algorithme de Cholesky (identification)
- exemple
- calcul du coût lorsque $N \to \infty$
- remarques (erreurs d'arrondis,…)

Preuve de la factorisation de Cholesky à partir de A=LU

Soit $A \in M_N(\mathrm{IR})$ symétrique définie positive.

Alors $\boxed{A = LU}$ car $\Delta_k = \begin{pmatrix} a_{1,1} & \cdots & a_{1,k} \\ \vdots & & \vdots \\ a_{k,1} & \cdots & a_{k,k} \end{pmatrix}$ inversible $\forall k \geq 1$

En effet Δ_k est sym def >0: $\forall x \in \mathrm{IR}^k \setminus \{0\}$,

$$x^T \Delta_k x = \begin{pmatrix} x_1 & \cdots & x_k & 0 & \cdots & 0 \end{pmatrix} \begin{pmatrix} a_{1,1} & \cdots & a_{1,N} \\ \vdots & & \vdots \\ a_{N,1} & \cdots & a_{N,N} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_k \\ 0 \\ \vdots \\ 0 \end{pmatrix} > 0$$

et (Δ_k sym def >0) $\Rightarrow$ (Δ_k inversible)

Comme A est symétrique, nous allons montrer que:

$$\boxed{U = DL^T, \quad D = \mathrm{diag}(u_{11}, u_{22}, \ldots, u_{NN})}$$

et donc la factorisation $A = LU$ s'écrit:

$$\boxed{A = LDL^T}$$

Preuve: on remarque que $A = A^T = U^T L^T = U^T D^{-1} D L^T$

(U inversible donc $D = \mathrm{diag}(u_{11}, u_{22}, \ldots, u_{NN})$ l'est aussi)

Donc $A = LU = (U^T D^{-1})(DL^T)$

* $U^T D^{-1}$ est triangulaire inférieure de diagonale unité

* DL^T est triangulaire supérieure

$\Rightarrow$ par unicité de la factorisation LU : $U = DL^T \quad (L = U^T D^{-1})$

Puisque $A = L\,D\,L^T$ avec L inversible (matrices A, D congruentes) :

A sym def >0 $\Leftrightarrow$ D sym def >0 $\quad (x^T A x = y^T D y$ avec $y = L^T x)$

Puisque $D = \mathrm{diag}(u_{11}, u_{22}, \ldots, u_{NN})$, on a donc :

A sym def >0 $\Leftrightarrow$ $u_{ii} > 0 \;\forall i = 1, \ldots, N$

En notant $D^{1/2} = \mathrm{diag}(u_{11}^{1/2}, u_{22}^{1/2}, \ldots, u_{NN}^{1/2})$: $A = L\,D^{1/2}\,D^{1/2} L^T$, d'où :

$\boxed{A = T\,T^T \text{ avec } T = L\,D^{1/2} \text{ triangulaire inférieure, } t_{ii} = u_{ii}^{1/2} > 0}$

On a donc montré l'existence de la factorisation de Cholesky de A

Preuve de l'unicité de la factorisation de Cholesky:

supposons $A = T\,T^T$ avec T triangulaire inférieure et $t_{ii} > 0$
et montrons qu'on a nécessairement $T = L\,D^{1/2}$.
Pour cela on utilise l'unicité de la factorisation $A = LU$.

On a $A = T\,T^T = \left(T\Lambda^{-1}\right)\left(\Lambda T^T\right)$ avec $\Lambda = \text{diag}(t_{11}, t_{22}, \ldots, t_{NN})$,
$T\Lambda^{-1}$ est triang. inf de diagonale unité, ΛT^T est triang. sup,
donc par unicité de la factorisation $A = LU$:
$\Lambda T^T = U \;\Rightarrow\; t_{ii}^2 = u_{ii} \;\Rightarrow\; t_{ii} = \sqrt{u_{ii}}$ car $t_{ii} > 0$, i.e. $\Lambda = D^{1/2}$
et $T\Lambda^{-1} = L \;\Rightarrow\; T = L\Lambda = L\,D^{1/2}$

Algorithme de factorisation Cholesky :

Soit $A \in M_N(\mathrm{IR})$ symétrique définie positive.

La factorisation de Cholesky $A = T\,T^T$

($T \in M_N(\mathrm{IR})$ triangulaire inférieure, $t_{ii} > 0$) se calcule comme suit:

pour tout $j = 1,\ldots,N$:

$$\mathrm{t}_{j,j} = \left(a_{j,j} - \sum_{k<j} \mathrm{t}_{j,k}^2 \right)^{1/2}$$

pour tout $i = j+1,\ldots,N$:

$$\mathrm{t}_{i,j} = \frac{1}{\mathrm{t}_{j,j}} \left(a_{i,j} - \sum_{k<j} \mathrm{t}_{i,k}\mathrm{t}_{j,k} \right)$$

Justification de l'algorithme de Cholesky :

on calcule $(t_{jj},\ldots,t_{Nj})$ successivement pour $j = 1,\ldots,N$,

par identification des coefficients $(a_{jj},\ldots,a_{Nj})$ dans $A = T\,T^T$

Puisque $A = T\,T^T$ avec T triangulaire inférieure:

$$\boxed{a_{i,j} = \sum_{k=1}^{\min(i,j)} \mathrm{t}_{i,k}\mathrm{t}_{j,k}}$$

–Etape 1 : calcul de 1ère colonne de T à partir de celle de A

$a_{1,1} = \mathrm{t}_{1,1}^2 \Rightarrow \boxed{\mathrm{t}_{1,1} = \left(a_{1,1}\right)^{1/2}}$ $(a_{1,1} > 0$ car A sym def >0)

pour tout $i \geq 2$: $a_{i,1} = \mathrm{t}_{i,1}\mathrm{t}_{1,1} \Rightarrow \boxed{\mathrm{t}_{i,1} = \dfrac{a_{i,1}}{\mathrm{t}_{1,1}}}$

–Récurrence: on suppose connues les colonnes 1,..., p de T, calculons la (p+1)ième colonne de T

On rappelle que (identification dans $A = T\,T^T$) :

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

$$\Rightarrow a_{p+1,p+1} = \sum_{k=1}^{p+1} \mathrm{t}^2_{p+1,k} = \mathrm{t}^2_{p+1,p+1} + \sum_{k=1}^{p} \mathrm{t}^2_{p+1,k}$$

$$\Rightarrow \mathrm{t}^2_{p+1,p+1} = a_{p+1,p+1} - \sum_{k=1}^{p} \mathrm{t}^2_{p+1,k} \;\; > 0 \text{ (puisque } T \text{ existe)}$$

On en déduit (puisque $\mathrm{t}_{p+1,p+1} > 0$):

$$\boxed{\mathrm{t}_{p+1,p+1} = \left(a_{p+1,p+1} - \sum_{k=1}^{p} \mathrm{t}^2_{p+1,k} \right)^{1/2}}$$

Pour compléter le calcul de la colonne $p+1$ de T,

il reste à calculer $t_{i,p+1}$ pour tout $i \geq p+2$:

$$a_{i,p+1} = \sum_{k=1}^{p+1} t_{i,k} t_{p+1,k} = t_{i,p+1} t_{p+1,p+1} + \sum_{k=1}^{p} t_{i,k} t_{p+1,k}$$

$$\Rightarrow t_{i,p+1} t_{p+1,p+1} = a_{i,p+1} - \sum_{k=1}^{p} t_{i,k} t_{p+1,k} \text{ , où } t_{p+1,p+1} > 0 \text{ est connu,}$$

$$\Rightarrow \quad \boxed{t_{i,p+1} = \frac{1}{t_{p+1,p+1}} \left(a_{i,p+1} - \sum_{k=1}^{p} t_{i,k} t_{p+1,k} \right)}$$

Remarque : l'algorithme de Cholesky permet de savoir si $A \in M_N(\mathrm{IR})$ symétrique est définie positive. C'est le cas si et seulement si

l'algo peut être mené à son terme: $a_{j,j} - \sum_{k<j} t_{j,k}^2 > 0$ pour tout $j = 1, \ldots, N$

Exemple de factorisation de Cholesky :

$$A = \begin{pmatrix} 1 & -1 & 1 \\ -1 & 5 & 3 \\ 1 & 3 & 6 \end{pmatrix} = T\,T^T = \begin{pmatrix} t_{11} & 0 & 0 \\ t_{21} & t_{22} & 0 \\ t_{31} & t_{32} & t_{33} \end{pmatrix} \begin{pmatrix} t_{11} & t_{21} & t_{31} \\ 0 & t_{22} & t_{32} \\ 0 & 0 & t_{33} \end{pmatrix}$$

$t_{11}^2 = 1$ et $t_{11} > 0$ donc $t_{11} = 1$,

$t_{11}t_{21} = -1$ donc $t_{21} = -1$, $\quad t_{11}t_{31} = 1$ donc $t_{31} = 1$

$t_{21}^2 + t_{22}^2 = a_{22}$ et $t_{22} > 0$ donc $t_{22} = \left(a_{22} - t_{21}^2\right)^{1/2} = 2$

$t_{31}t_{21} + t_{32}t_{22} = a_{32}$ donc $t_{32} = (a_{32} - t_{31}t_{21}) / t_{22} = 2$

$t_{31}^2 + t_{32}^2 + t_{33}^2 = a_{33}$ et $t_{33} > 0$ donc $t_{33} = \left(a_{33} - t_{31}^2 - t_{32}^2\right)^{1/2} = 1$

$$\boxed{\Rightarrow T = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 2 & 0 \\ 1 & 2 & 1 \end{pmatrix}}$$

Coût de l'algorithme de Cholesky

équivalent quand $N \to \infty$ du nb d'op. arithmétiques (matrice A pleine)

Etape 1 : calcul de 1ère colonne de T $\Rightarrow$ N opérations

$\mathrm{t}_{11} = \sqrt{a_{11}}$, $\mathrm{t}_{i1} = a_{i1} / t_{11}$ pour i=2,...,N (N-1 divisions, une $\sqrt{\ }$)

Etape p+1 : calcul de (p+1)ième colonne de T (connaissant colonnes 1...p)

$$\mathrm{t}_{p+1,p+1} = \left(a_{p+1,p+1} - \sum_{k=1}^{p} \mathrm{t}^2_{p+1,k} \right)^{1/2}$$

$\Rightarrow$ 2p+1 opérations (p soustractions, p $\times$, une $\sqrt{\ }$)

$$\mathrm{t}_{i,p+1} = \frac{1}{\mathrm{t}_{p+1,p+1}} \left(a_{i,p+1} - \sum_{k=1}^{p} \mathrm{t}_{i,k} \mathrm{t}_{p+1,k} \right)$$

$\Rightarrow$ 2p+1 opérations (1 division, p soustractions, p $\times$)

$\Rightarrow$ le calcul de (p+1)ième colonne de T correspondant à $t_{i,p+1}$ pour $i = p+1,\ldots,N$ nécessite (2p+1)(N-p) opérations

calcul $j-$ième colonne de T : $\boxed{C_j = (2j-1)(\mathrm{N+1-j}) \text{ opérations}}$

Coût du calcul des colonnes 1,2,...,N de T :

$$\sum_{j=1}^{N} C_j = 2\sum_{j=1}^{N} j(\mathrm{N+1-j}) - \sum_{j=1}^{N}(\mathrm{N+1-j})$$

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

$$= 2(\mathrm{N+1})\left(\frac{N^2}{2} + O(N)\right) - 2\left(\frac{N^3}{3} + O(N^2)\right) + O(N^2)$$

$$= N^3 - (2/3)N^3 + O(N^2)$$

$$\boxed{\approx (1/3)N^3 \quad ((1/6)N^3 \text{ multiplications et soustractions}, N\sqrt{\ })}$$

Remarques :

$*$Nous verrons que le calcul d'une $\sqrt{\ }$ en double précision machine $\sim$ 4 itérations de la méthode de Newton (~ 12 opérations élémentaires), donc le coût du calcul des $N\sqrt{\ }$ est $O(N)$, négligeable devant $(1/3)N^3$

$*$Borne sur les coefficients de T : $\sum_{k=1}^{i} \mathrm{t}_{i,k}^2 = a_{i,i} \Rightarrow \mathrm{t}_{i,k} \leq \sqrt{a_{i,i}}$

$*$Estimation des erreurs d'arrondi (Wilkinson, 1968):

(u désigne l'unité d'arrondi, $\mathrm{u} \approx 10^{-16}$ en double précision,

cond(A) = conditionnement de A pour norme euclidienne)

L'algo de Cholesky en arithmétique flottante peut être mené à son terme si $q_N u\,\mathrm{cond(A)} \leq 1$ pour une constante $q_N = O(N^{3/2})$, et la solution numérique $\hat{x}$ de $Ax = b$ vérifie $(A+E)\hat{x} = b$ avec $\|E\|_2 = O(N^{3/2}\mathrm{u}\|A\|_2)$