Décomposition de Cholesky

Riffaut Antonin

2013-2014

Théorème 1 (Décomposition de Cholesky)**.** *Soit $A \in \mathcal{S}_n^{++}(\mathbb{R})$. Il existe une unique matrice $T \in \mathcal{M}_n(\mathbb{R})$ triangulaire supérieure à coefficients diagonaux strictement positifs telle que $A = T^T T$.*

Démonstration. **Existence.** Les mineurs principaux de A sont non nuls, et même strictement positifs : en effet, pour $r \in \{1, \dots, n\}$, notons A_r la sous-matrice principale d'ordre r de A. A_r est symétrique. Par ailleurs, comme $A \in \mathcal{S}_n^{++}(\mathbb{R})$, alors $x^T A x > 0$, pour tout $x \in \mathbb{R}^n \setminus \{0\}$. Ainsi, pour tout $x' \in \mathbb{R}^r \setminus \{0\}$, en complétant x' par des zéros en un vecteur x de $\mathbb{R}^n \setminus \{0\}$, on obtient $x'^T A_r x' = x^T A x > 0$, donc A_r est définie positive : son déterminant est par conséquent strictement positif.

Il s'ensuit que la matrice A possède une décomposition LU : il existe d'uniques matrices $L, D, V \in \mathcal{M}_n(\mathbb{R})$, avec L triangulaire inférieure à diagonale unité, D diagonale à coefficients diagonaux strictement positifs, et V triangulaire supérieure à diagonale unité, telles que $A = LDV$ [1]. Comme A est symétrique, alors $A = A^T = V^T D L^T$, et par unicité de cette décomposition, on déduit que $V^T = L$. En posant $T = \sqrt{D} V$, on obtient une décomposition de Cholesky de A.

Unicité. Soit $S \in \mathcal{M}_n(\mathbb{R})$ une autre matrice triangulaire supérieure à coefficients diagonaux strictement positifs telle que $A = S^T S$. Alors $ST^{-1} = (TS^{-1})^T$. Le terme de gauche est une matrice triangulaire supérieure, tandis que le terme de droite est une matrice triangulaire inférieure : donc ST^{-1} est une matrice diagonale. De plus, $(ST^{-1})^{-1} = TS^{-1} = (TS^{-1})^T = ST^{-1}$, et tous les coefficients diagonaux de ST^{-1} sont strictement positifs, donc sont égaux à 1, c'est-à-dire $S = T$. ■

Remarque. La réciproque est vraie, au sens suivant : si $A \in \mathcal{M}_n(\mathbb{R})$ et s'il existe une matrice $T \in \mathcal{M}_n(\mathbb{R})$ triangulaire supérieure à coefficients diagonaux strictement positifs tel que $A = T^T T$, alors $A \in \mathcal{S}_n^{++}(\mathbb{R})$.

Calcul en pratique de la décomposition de Cholesky Notons $A = (a_{i,j})_{1 \le i,j \le n}$ et $T = (t_{i,j})_{1 \le i,j \le n}$ ($t_{i,j} = 0$ si $i > j$). Pour tous $i, j \in \{1, \dots, n\}$,

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

Étape 1. On a $a_{1,1} = t_{1,1}^2$, donc $t_{1,1} = \sqrt{a_{1,1}}$, ce qui requiert un calcul de racine carrée (on va considérer le calcul de la racine carrée comme une opération arithmétique élémentaire). Ensuite, pour $j \in \{2, \dots, n\}$, on a $a_{1,j} = t_{1,1} t_{1,j}$, donc $t_{1,j} = \frac{t_{1,1}}{a_{1,j}}$, ce qui requiert une division.

1. Cette factorisation s'obtient en factorisant la matrice U de la décomposition LU de A en $U = DV$, où D est la matrice diagonale des coefficients diagonaux de U.

Étape i, $2 \leq i \leq n$. À l'étape i, on suppose qu'on a calculé tous les coefficients $t_{i',j}$ pour $i' < i$. On a $a_{i,i} = \sum_{k=1}^{i} t_{k,i}^2$, donc $t_{i,i} = \sqrt{a_{i,i} - \sum_{k=1}^{i-1} t_{k,i}^2}$, ce qui requiert $i-1$ multiplications, $i-1$ additions, et un calcul de racine carrée. Ensuite, pour $j \in \{i+1, \dots, n\}$, on a $a_{i,j} = \sum_{k=1}^{i} t_{k,i} t_{k,j}$, donc $t_{i,j} = \frac{1}{t_{i,i}} \left(a_{i,j} - \sum_{k=1}^{i-1} t_{k,i} t_{k,j} \right)$, ce qui requiert $i-1$ multiplications, $i-1$ additions, et 1 division. Au total, l'étape i requiert $(n-i+1)(2i-1)$ opérations arithmétiques élémentaires.

Finalement, le coût total du calcul de la décomposition de Cholesky, en nombre d'opérations arithmétiques élémentaires, est

$$\sum_{i=1}^{n} (n-i+1)(2i-1) = \frac{1}{6} n(2n^2 + 3n + 1) \underset{n \to +\infty}{\sim} \frac{1}{3} n^3.$$

La décomposition de Cholesky est par conséquent légèrement plus efficace que la décomposition LU, dont le coût en nombre d'opérations arithmétiques élémentaires est asymptotiquement équivalent à $\frac{2}{3} n^3$.

Remarque. Soient $A \in \mathrm{GL}_n(\mathbb{R})$ et $b \in \mathbb{R}^n$. On cherche à résoudre le système linéaire $Ax = b$. Si A est symétrique définie positive, on peut calculer sa décomposition de Cholesky et procéder à une méthode de descente-remontée afin de déterminer x. Sinon, le système linéaire est équivalent à $A^T A x = A^T b$, où $A^T A$ est bien une matrice symétrique définie positive : on peut donc de même calculer la décomposition de Cholesky de $A^T A$, et procéder à une méthode de descente-remontée ; cette méthode nécessite les calculs supplémentaires de $A^T A$ et de $A^T b$, tous deux asymptotiquement négligeables devant le coût de la décomposition de Cholesky (pourvu qu'on dispose d'une méthode efficace pour calculer le produit matriciel $A^T A$).

Une application : l'inégalité de Hadamard

Corollaire 2. *Soit $A \in \mathcal{M}_n(\mathbb{R})$. Notons $X_1, \dots, X_n$ ses vecteurs colonnes. Alors*

$$|\det(A)| \leq \prod_{i=1}^{n} \|X_i\|_2 .$$

De plus, il y a égalité si et seulement si la famille $(X_1, \dots, X_n)$ est orthogonale.

Démonstration. Si A n'est pas inversible, le résultat est immédiat. Supposons donc que $A \in \mathrm{GL}_n(\mathbb{R})$. $A^T A$ est symétrique définie positive : on peut donc considérer sa décomposition de Cholesky $A^T A = T^T T$, avec $T \in \mathcal{M}_n(\mathbb{R})$ triangulaire supérieure à coefficients diagonaux strictement positifs. En notant $T = (t_{i,j})_{1 \leq i,j \leq n}$, on a $\det(A)^2 = \det(T)^2 = \prod_{i=1}^{n} t_{i,i}^2$. Or, en notant $A^T A = (b_{i,j})_{1 \leq i,j \leq n}$, on a $t_{i,i}^2 = b_{i,i} - \sum_{k=1}^{i-1} t_{k,i}^2 \leq b_{i,i} = \sum_{k=1}^{n} a_{k,i}^2 = \|X_i\|_2^2$. On en déduit l'inégalité de Hadamard.

En outre, il y a égalité si et seulement si $t_{k,i} = 0$, pour tout $i \in \{2, \dots, n\}$, pour tout $k \in \{1, \dots, i-1\}$, autrement dit si et seulement si T est diagonale, et a fortiori $A^T A$, ce qui signifie que les colonnes de A forment une famille orthogonale. ∎