

La décomposition de Cholesky

Marc Lorenzi

18 février 2025

1 Matrices symétriques définies positives

Dans tout l'article, n désigne un entier naturel non nul.

Définition 1. Soit $A \in \mathcal{M}_n(\mathbb{R})$ une matrice symétrique. La matrice A est *définie positive* lorsque A est symétrique et pour tout $X \in \mathcal{M}_{n,1}(\mathbb{R})$ non nul, $X^T A X > 0$.

Notation. Nous noterons $\mathcal{D}_n(\mathbb{R})$ l'ensemble des matrices réelles $n \times n$ symétriques définies positives.

Soit E un $\mathbb{R}$ -espace vectoriel de dimension n . Soit $\varphi : E^2 \longrightarrow \mathbb{R}$ une forme bilinéaire sur E . Soit $\mathcal{B} = (e_1, \dots, e_n)$ une base de E . La *matrice de φ dans la base $\mathcal{B}$* est

$$A = (\varphi(e_i, e_j))_{i,j \in \llbracket 1, n \rrbracket}$$

La forme bilinéaire φ est symétrique si et seulement si pour tous $i, j \in \llbracket 1, n \rrbracket$,

$$A_{ij} = \varphi(e_i, e_j) = \varphi(e_j, e_i) = A_{ji}$$

La symétrie de φ équivaut donc à celle de A .

Soient $x, y \in E$. Notons X et Y les matrices respectives de x et y dans la base $\mathcal{B}$. On vérifie facilement que par la bilinéarité de φ ,

$$\varphi(x, y) = X^T A Y$$

La forme bilinéaire φ est positive et non dégénérée si et seulement si pour tout $x \in E$ non nul (de matrice X dans la base $\mathcal{B}$),

$$\varphi(x, x) = X^T A X > 0$$

Ainsi, les matrices symétriques définies positives sont les matrices des formes bilinéaires symétriques positives non dégénérées, c'est à dire des *produits scalaires*.

Proposition 1. $\mathcal{D}_n(\mathbb{R}) \subset GL_n(\mathbb{R})$.

Démonstration. Soit $A \in \mathcal{D}_n(\mathbb{R})$. Soit $X \in \mathcal{M}_{n1}(\mathbb{R})$. Supposons $AX = 0$. A fortiori, $X^TAX = 0$. Comme A est définie positive, X n'est pas nul. Ainsi, $X = 0$. Le noyau de A est donc $\{0\}$, donc A est inversible. $\square$

Proposition 2. Si $M \in GL_n(\mathbb{R})$, alors $A = M^T M \in \mathcal{D}_n(\mathbb{R})$.

Démonstration. On a

$$A^T = (M^T M)^T = M^T M^{TT} = M^T M = A$$

La matrice A est donc symétrique. Soit $X \in \mathcal{M}_{n1}(\mathbb{R})$ non nul. On a

$$X^TAX = X^T M^T M X = (MX)^T(MX)$$

Posons $Y = MX$. Comme M est inversible, $Y \neq 0$. On a donc

$$X^TAX = \sum_{k=1}^n Y_k^2 > 0$$

$\square$

Exemple. Prenons

$$M = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 2 & 3 \\ 0 & 0 & 3 \end{pmatrix}$$

On a

$$A = M^T M = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 8 & 12 \\ 3 & 12 & 27 \end{pmatrix}$$

Le théorème de Cholesky montrera que, inversement, toute matrice symétrique définie positive est de cette forme. En fait, nous ferons mieux que cela puisque nous verrons que la matrice M peut être choisie triangulaire supérieure. En supposant les coefficients diagonaux de M strictement positifs, nous aurons même l'unicité de M .

Proposition 3. Soit $A \in \mathcal{M}_n(\mathbb{R})$ une matrice symétrique. La matrice A appartient à $\mathcal{D}_n(\mathbb{R})$ si et seulement si ses valeurs propres sont strictement positives.

Démonstration. La matrice A étant symétrique, elle est diagonalisable en base orthonormée. Soit $P \in \mathcal{O}_n(\mathbb{R})$ telle que

$$P^TAP = \text{diag}(\lambda_1, \dots, \lambda_n) = D$$

où $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ sont les valeurs propres de A . Pour tout $X \in \mathcal{M}_{n1}(\mathbb{R})$,

$$X^TAX = X^T P D P^T X = Y^T D Y$$

où $Y = P^T X$. On a donc

$$X^TAX = \sum_{k=1}^n \lambda_k Y_k^2$$

L'application $X \mapsto P^T X$ est une bijection linéaire de $\mathcal{M}_{n1}(\mathbb{R})$ sur lui-même. De là, A est définie positive si et seulement si pour tout $Y \in \mathcal{M}_{n1}(\mathbb{R})$ non nul, $\sum_{k=1}^n \lambda_k Y_k^2 > 0$. Cette condition équivaut clairement à la positivité stricte des λ_k . $\square$

Exemple. En reprenant l'exemple ci-dessus, les valeurs propres de A valent environ 33.20, 0.46 et 2.33.

2 Le théorème de Cholesky (unicité)

Proposition 4. Soit $A \in \mathcal{D}_n(\mathbb{R})$. Il existe au plus une matrice $R \in \mathcal{M}_n(\mathbb{R})$ vérifiant

- R est triangulaire supérieure.
- Les coefficients diagonaux de R sont strictement positifs.
- $A = R^T R$.

Démonstration. Supposons qu'une telle matrice R existe.

Pour tous $i, j \in \llbracket 1, n \rrbracket$ tels que $i \leq j$,

$$A_{ij} = \sum_{k=1}^n (R^T)_{ik} R_{kj} = \sum_{k=1}^n R_{ki} R_{kj} = \sum_{k=1}^i R_{ki} R_{kj}$$

ou encore

$$(\star) \quad A_{ij} = R_{ii} R_{ij} + \sum_{k=1}^{i-1} R_{ki} R_{kj}$$

Soit $i \in \llbracket 1, n \rrbracket$. Supposons déterminées les lignes 1, 2, ..., $i-1$ de la matrice R . En prenant $j = i$ dans $(\star)$, on obtient

$$A_{ii} = R_{ii}^2 + \sum_{k=1}^{i-1} R_{ki}^2$$

Comme $R_{ii} > 0$, on en déduit que

$$R_{ii} = \sqrt{A_{ii} - \sum_{k=1}^{i-1} R_{ki}^2}$$

Toujours grâce à $(\star)$, on a ensuite pour tout $j \in \llbracket i+1, n \rrbracket$,

$$R_{ij} = \frac{1}{R_{ii}} \left(A_{ij} - \sum_{k=1}^{i-1} R_{ki} R_{kj} \right)$$

La ligne i de la matrice R est donc déterminée de façon unique à partir de la ligne i de A et des lignes 1, ..., $i-1$ de R . On en déduit l'unicité de R par récurrence forte sur i . $\square$

L'existence de R serait immédiate si nous savions que les réels $A_{ii} - \sum_{k=1}^{i-1} R_{ki}^2$ sont tous strictement positifs. En effet, $(\star)$ définirait la matrice R par des relations de récurrence forte sur i . Cependant, ceci est loin d'être évident. Nous allons, pour montrer l'existence de R , suivre une voie différente.

3 Le théorème de Cholesky (existence)

Commençons par quelques résultats préliminaires.

Proposition 5. Soit $A \in \mathcal{D}_n(\mathbb{R})$. Pour tout $i \in \llbracket 1, n \rrbracket$, $A_{ii} > 0$.

Démonstration. Soit $i \in \llbracket 1, n \rrbracket$. Soit $X = (0 \dots 0 \ 1 \ 0 \dots 0)^T \in \mathcal{M}_{n1}(\mathbb{R})$, où le 1 est en position i . On a alors

$$X^T A X = A_{ii} > 0$$

□

Proposition 6. Soit $A \in \mathcal{D}_n(\mathbb{R})$. Considérons une décomposition en blocs de A de la forme

$$A = \begin{pmatrix} A' & B^T \\ B & A'' \end{pmatrix}$$

où $A' \in \mathcal{M}_p(\mathbb{R})$, $A'' \in \mathcal{M}_{n-p}(\mathbb{R})$, $B \in \mathcal{M}_{(n-p)p}(\mathbb{R})$ et $p \in \llbracket 1, n-1 \rrbracket$. Alors, les matrices A' et A'' sont symétriques définies positives.

Démonstration. La matrice A' est clairement symétrique. Soit $X \in \mathcal{M}_{p1}(\mathbb{R})$ non nul. Complétons X avec $n-p$ zéros pour former

$$Y = (X_1 \dots X_p \ 0 \dots 0)^T \in \mathcal{M}_{n1}(\mathbb{R}) \setminus \{0\}$$

On a

$$X^T A' X = Y^T A Y > 0$$

Ainsi, $A' \in \mathcal{D}_p(\mathbb{R})$. De même, $A'' \in \mathcal{D}_{n-p}(\mathbb{R})$. □

Proposition 7. Si $A \in \mathcal{D}_n(\mathbb{R})$ et $P \in \text{GL}_n(\mathbb{R})$, alors $P^T A P \in \mathcal{D}_n(\mathbb{R})$.

Démonstration. Posons $B = P^T A P$. On a

$$B^T = (P^T A P)^T = P^T A^T P^{TT} = P^T A P = B$$

La matrice B est donc symétrique. Soit $X \in \mathcal{M}_{n1}(\mathbb{R})$ non nul. On a

$$X^T B X = (P X)^T A X$$

Comme P est inversible, $Y = P X \neq 0$ et donc

$$X^T B X = Y^T A Y > 0$$

La matrice B est donc définie positive. □

Théorème 8. [CHOLESKY]

Soit $A \in \mathcal{D}_n(\mathbb{R})$. Il existe une unique matrice $R \in \mathcal{M}_n(\mathbb{R})$ vérifiant

- R est triangulaire supérieure.
- Les coefficients diagonaux de R sont strictement positifs.
- $A = R^T R$.

Démonstration. Nous avons déjà prouvé l'unicité. Montrons l'existence par récurrence sur n . C'est immédiat si $n = 1$. Soit $n \geq 2$. Supposons la propriété vérifiée pour toutes les matrices symétriques définies positives de taille $n - 1$. Soit $A \in \mathcal{D}_n(\mathbb{R})$. Écrivons la décomposition en blocs

$$A = \begin{pmatrix} a & X^T \\ X & \tilde{A} \end{pmatrix}$$

où $a \in \mathbb{R}$, $X \in \mathcal{M}_{(n-1)1}(\mathbb{R})$ et $\tilde{A} \in \mathcal{M}_{n-1}(\mathbb{R})$. Par la proposition 5, $a > 0$. Posons

$$\begin{aligned} r &= \sqrt{a} \\ Y &= \frac{1}{r}X \\ \hat{A} &= \tilde{A} - Y Y^T \end{aligned}$$

On a alors facilement

$$A = \begin{pmatrix} r & 0 \\ Y & I_{n-1} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \hat{A} \end{pmatrix} \begin{pmatrix} r & Y^T \\ 0 & I_{n-1} \end{pmatrix}$$

La matrice $\begin{pmatrix} r & Y^T \\ 0 & I_{n-1} \end{pmatrix}$ est triangulaire supérieure, à coefficients diagonaux non nuls. Elle est donc inversible. Notons P son inverse. Soit $B = \begin{pmatrix} 1 & 0 \\ 0 & \hat{A} \end{pmatrix}$. On a

$$A = (P^{-1})^T B P^{-1}$$

et donc

$$B = P^T A P$$

Par la proposition 7, $B \in \mathcal{D}_n(\mathbb{R})$. Par la proposition 6, on a donc $\hat{A} \in \mathcal{D}_{n-1}(\mathbb{R})$. Par l'hypothèse de récurrence, il existe une matrice $\hat{R} \in \mathcal{M}_{n-1}(\mathbb{R})$ triangulaire supérieure à coefficients diagonaux strictement positifs telle que

$$\hat{A} = \hat{R}^T \hat{R}$$

On a donc

$$\begin{aligned} A &= \begin{pmatrix} r & 0 \\ Y & I_{n-1} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \hat{R}^T \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \hat{R} \end{pmatrix} \begin{pmatrix} r & Y^T \\ 0 & I_{n-1} \end{pmatrix} \\ &= \begin{pmatrix} r & 0 \\ Y & \hat{R}^T \end{pmatrix} \begin{pmatrix} r & Y^T \\ 0 & \hat{R} \end{pmatrix} \end{aligned}$$

Posons

$$R = \begin{pmatrix} r & Y^T \\ 0 & \hat{R} \end{pmatrix}$$

La matrice R est triangulaire supérieure à coefficients diagonaux strictement positifs, et $A = R^T R$. Ceci complète la preuve. $\square$

4 L'algorithme de décomposition

Soit $A \in \mathcal{D}_n(\mathbb{R})$. Soit $R \in \mathcal{M}_n(\mathbb{R})$ triangulaire supérieure à coefficients diagonaux strictement positifs telle que $A = R^T R$. La preuve de l'unicité de la décomposition de Cholesky fournit un algorithme pour le calcul de R . Nous avons vu que pour tout $i \in \llbracket 1, n \rrbracket$,

$$R_{ii} = \sqrt{A_{ii} - \sum_{k=1}^{i-1} R_{ki}^2}$$

et pour tout $j \in \llbracket i+1, n \rrbracket$,

$$R_{ij} = \frac{1}{R_{ii}} \left(A_{ij} - \sum_{k=1}^{i-1} R_{ki} R_{kj} \right)$$

Pour tout $i \in \llbracket 1, n \rrbracket$, on effectue pour le calcul de R_{ii} , $i-1$ élévations au carré, $i-1$ soustractions et 1 calcul de racine carrée. On a donc au total $2i-1$ opérations.

Puis, pour tout $j \in \llbracket i+1, n \rrbracket$, le calcul de R_{ij} nécessite $i-1$ produits, $i-1$ soustractions et 1 division. On a donc encore au total $2i-1$ opérations pour chaque valeur de j .

Le coût total du calcul de la ligne i de la matrice R est donc

$$2i-1 + (n-i)(2i-1) = (2i-1)(n-i+1) \text{ flops}$$

Le coût total du calcul de R est donc

$$C_n = \sum_{i=1}^n (2i-1)(n-i+1) \text{ flops}$$

Un calcul simple montre que

$$C_n = \frac{1}{6}n(n+1)(2n+1) \sim \frac{1}{3}n^3 \text{ flops}$$