

1 Matrices symétriques définies positives et leur inversion

Notation 1.1 (Matrice transposée) Si A est une matrice, alors A^T est sa matrice *transposée* : $(A^T)_{ij} = a_{ji}$.

Lemme 1.2 Si A, B, C, \dots, Y, Z sont des matrices, alors $(ABC \dots YZ)^T = Z^T Y^T \dots C^T B^T A^T$.

Preuve. Il faut vérifier d'abord que $(AB)^T = B^T A^T$; puis utiliser une induction sur le nombre des matrices à multiplier.

Définition 1.3 (Matrice symétrique) Une matrice A est *symétrique* si $A^T = A$.

Exercice 1.4 Si la matrice Y est symétrique, et la matrice X quelconque, alors les matrices XYX^T et $X^T Y X$ sont symétriques. Entre autres, les matrices XX^T et $X^T X$ sont aussi symétriques.

Preuve. Une simple vérification en utilisant le lemme 1.2

Lemme 1.5 Si une matrice A est non dégénérée et symétrique, alors la matrice A^{-1} est symétrique.

Preuve. On peut utiliser la formule

$$(A^{-1})_{ij} = (-1)^{i+j} \frac{\det A^{ji}}{\det A}$$

où A^{ji} est la matrice A avec la ligne j et la colonne i enlevées. Attention, la matrice A^{ji} n'est pas symétrique même si A l'est. Il faut vérifier que, pour enlever la ligne j et la colonne i d'une matrice, on peut faire la transposition, puis enlever la ligne i et la colonne j , puis transposer encore une fois. La première opération ne change pas la matrice (si A est symétrique) ; la troisième ne change pas le déterminant.

Définition 1.6 (Matrice définie positive) Une matrice symétrique A dont les éléments sont des nombres réels, est *définie positive* si pour tout vecteur $x \in \mathbb{R}^n$ non nul on a $x^T A x > 0$.

Proposition 1.7 Toute matrice symétrique et définie positive est non dégénérée.

Preuve. Si A est dégénérée, alors il existe un $x \neq 0$ tel que $Ax = 0$; pour ce vecteur-là on a $x^T Ax = 0$.

Lemme 1.8 *Si une matrice A est non dégénérée, alors la matrice $B = A^T A$ est symétrique (voir l'exercice 1.4) et définie positive.*

Preuve. On a $x^T Bx = x^T (A^T A)x = (x^T A^T)(Ax) = (Ax)^T (Ax) = |Ax|^2$. Si A est non dégénérée, alors $x \neq 0$ implique $Ax \neq 0$ et donc $|Ax|^2 > 0$, d'où $x^T Bx > 0$.

Proposition 1.9 *Si une matrice A est non dégénérée, alors*

$$A^{-1} = (A^T A)^{-1} A^T.$$

Preuve. Vérification :

$$(A^T A)^{-1} A^T \cdot A = (A^T A)^{-1} \cdot (A^T A) = I.$$

On peut conclure que le problème d'inversion de matrices se réduit au même problème pour les matrices définies positives.

Soit A une matrice $n \times n$ symétrique définie positive. Considérons sa représentation sous forme de blocs :

$$A = \begin{pmatrix} B & C^T \\ C & D \end{pmatrix} \quad (1)$$

où B et D sont des matrices carrées de taille $k \times k$ et $(n - k) \times (n - k)$ respectivement, et C une matrice rectangulaire $(n - k) \times k$. Il est évident que les matrices B et D sont symétriques.

Lemme 1.10 *Les matrices B et D sont définies positives.*

Preuve. Soit y un vecteur non nul de dimension k , $y^T = (y_1, \dots, y_k)$, et soit $x^T = (y_1, \dots, y_k, 0, \dots, 0)$. Alors $y^T B y = x^T A x > 0$. La preuve pour la matrice D est analogue.

Imposer la condition d'avoir les matrices B et D définies positives ne suffit pas pour obtenir une matrice A définie positive elle aussi. Cela devient presque évident si on remarque que ces deux conditions n'impliquent rien

pour la matrice C . Pour mieux comprendre qu'est-ce qu'il faut y ajouter, regardons d'abord le cas de dimension 2. Soit

$$A = \begin{pmatrix} b & c \\ c & d \end{pmatrix}, \quad x = \begin{pmatrix} y \\ z \end{pmatrix}.$$

Alors

$$\begin{aligned} x^T Ax &= (y \ z) \begin{pmatrix} b & c \\ c & d \end{pmatrix} \begin{pmatrix} y \\ z \end{pmatrix} = (y \ z) \begin{pmatrix} by + cz \\ cy + dz \end{pmatrix} = \\ & y(by + cz) + z(cy + dz) = by^2 + 2cyz + dz^2 = \\ & b \left(y^2 + 2y \cdot \frac{cz}{b} + \frac{c^2 z^2}{b^2} \right) + \left(d - \frac{c^2}{b} \right) z^2 = b \left(y + \frac{cz}{b} \right)^2 + \left(d - \frac{c^2}{b} \right) z^2. \end{aligned}$$

On peut maintenant prendre n'importe quel $z \neq 0$; calculer $y = -\frac{cz}{b}$, pour que $y + \frac{cz}{b}$ devienne nulle; et on obtient alors que la partie droite est égale à $(d - \frac{c^2}{b})z^2$. Pour qu'elle soit strictement positive, il faut imposer la condition

$$d - \frac{c^2}{b} > 0. \quad (2)$$

Vous avez sans doute reconnu la démarche : c'est l'extraction d'un carré parfait, une classique des classiques. On va maintenant répéter exactement la même chose mais dans le cas général, c'est-à-dire pour les matrices. Voici les difficultés qui nous attendent :

- Ce qui était un nombre c , devient une matrice, tantôt C , tantôt C^T . De même, y et z désormais dénotent des vecteurs (de dimension k et $n - k$ respectivement) et peuvent figurer sous forme y^T et z^T .
- Ce qui était division par b , devient multiplication par B^{-1} .
- Le plus important : la multiplication des matrices n'étant pas commutative, il faut être très attentif quant à l'ordre dans lequel on écrit les matrices (et les vecteurs) que l'on multiplie.

Alors :

$$\begin{aligned} x^T Ax &= (y^T \ z^T) \begin{pmatrix} B & C^T \\ C & D \end{pmatrix} \begin{pmatrix} y \\ z \end{pmatrix} = (y^T \ z^T) \begin{pmatrix} By + C^T z \\ Cy + Dz \end{pmatrix} = \\ & y^T By + y^T C^T z + z^T Cy + z^T Dz = \\ & y^T By + y^T BB^{-1}C^T z + z^T CB^{-1}By + z^T CB^{-1}BB^{-1}C^T z + \\ & z^T Dz - z^T CB^{-1}C^T z = \end{aligned}$$

$$(y + B^{-1}C^T z)^T B(y + B^{-1}C^T z) + z^T(D - CB^{-1}C^T)z.$$

On reconnaît dans la dernière formule quelques ingrédients déjà vus dans les calculs précédents : $y + B^{-1}C^T z$ remplace $y + \frac{cz}{b}$, et $D - CB^{-1}C^T$ remplace $d - \frac{c^2}{b}$.

Définition 1.11 (Complément de Schur) La matrice

$$S = D - CB^{-1}C^T$$

s'appelle le *complément de Schur* de la décomposition (1) de la matrice A en blocs.

Lemme 1.12 *La matrice S est symétrique définie positive.*

Preuve. La matrice B^{-1} est symétrique selon le lemme 1.5; la matrice $CB^{-1}C^T$, selon l'exercice 1.4. Étant donné que D est symétrique elle aussi, on obtient que S est symétrique.

La preuve du fait que S soit définie positive répète la preuve de la formule (2). Prenons n'importe quel vecteur $z \neq 0$; calculons $y = -B^{-1}C^T z$, pour que $y + B^{-1}C^T z$ soit nul; alors $x^T Ax$ devient égal à $z^T Sz$ et doit être strictement positif.

Théorème 1.13 *La matrice A^{-1} se calcule selon la formule suivante :*

$$A^{-1} = \begin{pmatrix} B^{-1} + B^{-1}C^T S^{-1}CB^{-1} & -B^{-1}C^T S^{-1} \\ -S^{-1}CB^{-1} & S^{-1} \end{pmatrix}. \quad (3)$$

Démonstration est directe, bien que pas très jolie : il faut multiplier la matrice ci-dessus par la “matrice de blocs” (1).

Avant de proposer un algorithme basé sur la formule (3), il est conseillé de contempler un peu la matrice, pour y voir des éléments qui se répètent. Par exemple, la sous-matrice du coin supérieur droit est la transposée de celle du coin inférieur gauche.

Algorithme. Appliquer récursivement la méthode suivante.

1. Découper la matrice A en quatre blocs selon la formule (1).
2. Calculer B^{-1} .
3. Calculer $C \cdot B^{-1}$; on remarque que $B^{-1}C^T = (CB^{-1})^T$.
4. Calculer $(CB^{-1}) \cdot C^T$, puis $S = D - CB^{-1}C^T$.

5. Calculer S^{-1} .
6. Calculer $S^{-1} \cdot (CB^{-1})$; c'est, au signe près, la sous-matrice du bas à gauche; celle du haut à droite est sa transposée.
7. Calculer $(B^{-1}C^T) \cdot (S^{-1}CB^{-1})$, puis la sous-matrice du haut à gauche.

Le bilan. Pour inverser une matrice de taille $n \times n$, il faut faire 2 inversions et 4 multiplications de matrices de taille deux fois plus petite, plus un certain nombre d'opérations quadratiques (addition, soustraction, transposition...). Ce qui nous donne l'inégalité suivante :

$$I(n) \leq 2I(n/2) + 4M(n/2) + an^2, \quad (4)$$

où $I(n)$ dénote la complexité de l'inversion et $M(n)$, celle de multiplication.

Théorème 1.14 $I(n) = O(M(n))$.

Démonstration. Nous supposons que la complexité de la multiplication est $\Omega(n^2)$ et que par conséquent $M(n) \geq 4M(n/2)$.

Supposons, par récurrence, que $I(n/2) \leq KM(n/2)$. Alors

$$I(n) \leq 2KM(n/2) + 4M(n/2) + an^2 \leq \left(\frac{K}{2} + 1\right) M(n) + an^2.$$

Pour que la partie droite soit inférieure à $KM(n)$ il suffit de demander que

$$an^2 \leq \left(\frac{K}{2} - 1\right) M(n),$$

ce qui est vrai pour une constante K suffisamment grande.

Proposition 1.15 $M(n) = O(I(n))$.

Preuve. Par rapport aux théorèmes précédents, cette preuve est remarquablement simple : il suffit de remarquer que

$$X = \begin{pmatrix} I & A & 0 \\ 0 & I & B \\ 0 & 0 & I \end{pmatrix} \quad \implies \quad X^{-1} = \begin{pmatrix} I & -A & AB \\ 0 & I & -B \\ 0 & 0 & I \end{pmatrix}.$$

Proposition 1.16 (Sans preuve) *La complexité de calcul de déterminant est $O(M(n))$.*