

Chapitre 1

Méthodes directes de résolution des systèmes linéaires

On considère la résolution du système linéaire

$$Ax = b, \tag{1.1}$$

avec A une matrice d'ordre n à coefficients réels inversible et b une matrice colonne de $M_{n,1}(\mathbb{R})$, par des méthodes dites *directes*, c'est-à-dire fournissant, en l'absence d'erreurs d'arrondi, la solution *exacte* en un nombre *fini*¹ d'opérations élémentaires. Ces méthodes consistent en la construction d'une matrice inversible M telle que MA soit une matrice triangulaire, le système linéaire équivalent (au sens où il possède la même solution) obtenu,

$$MAx = Mb,$$

étant alors « facile » à résoudre (on verra ce que l'on entend précisément par là). Une telle idée est par exemple à la base de la célèbre *méthode d'élimination de Gauss*, qui permet de ramener la résolution d'un système linéaire quelconque à celle d'un système triangulaire supérieur.

Après avoir présenté quelques cas pratiques d'application de ces méthodes et donné des éléments sur la résolution numérique des systèmes triangulaires, nous introduisons dans le détail la méthode d'élimination de Gauss. Ce procédé est ensuite réinterprété en termes d'opérations matricielles, donnant lieu à une méthode de *factorisation* (*factorization* ou *decomposition* en anglais) des matrices. Les propriétés d'une telle décomposition sont explorées, notamment dans le cas de matrices particulières. Le chapitre se conclut sur la présentation de quelques autres méthodes de factorisation.

1.1 Remarques sur la résolution des systèmes triangulaires

Observons tout d'abord que la solution du système linéaire $Ax = b$, avec A une matrice inversible, *ne s'obtient pas* en inversant A , puis en calculant le vecteur $A^{-1}b$, mais en réalisant plutôt des combinaisons linéaires sur les lignes du système et des substitutions. En effet, on peut facilement voir que le calcul de la matrice A^{-1} équivaut à résoudre n systèmes linéaires², ce qui s'avère bien plus coûteux que la résolution du *seul* système dont on cherche la solution.

Considérons à présent un système linéaire (1.1) dont la matrice est triangulaire inférieure, c'est-à-dire de la forme

$$\begin{array}{rccccccc} a_{11}x_1 & & & & & & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & & & & = & b_2 \\ \vdots & & \vdots & & \ddots & & & \vdots \\ a_{n1}x_1 & + & a_{n2}x_2 & + & \dots & + & a_{nn}x_n & = & b_n \end{array}$$

1. On oppose ici ce type de méthodes avec les méthodes dites *itératives*, qui nécessitent *a priori* un nombre infini d'opérations pour fournir la solution. Celles-ci sont l'objet du chapitre 3.

2. Ces systèmes sont

$$Ax_i = e_i, \quad 1 \leq i \leq n,$$

où e_i désigne le i^{e} vecteur de la base canonique de $M_{n,1}(\mathbb{R})$.

Si la matrice A est inversible, ses termes diagonaux a_{ii} , $i = 1, \dots, n$, sont tous non nuls³ et la résolution du système est alors extrêmement simple : on calcule x_1 par une division, que l'on substitue ensuite dans la deuxième équation pour obtenir x_2 , et ainsi de suite... Cette méthode, dite de « descente » (“*forward substitution*” en anglais), s’écrit

$$\begin{aligned} x_1 &= \frac{b_1}{a_{11}} \\ x_i &= \frac{1}{a_{ii}} \left(b_i - \sum_{j=1}^{i-1} a_{ij} x_j \right), \quad i = 2, \dots, n. \end{aligned} \quad (1.2)$$

L’algorithme mis en œuvre pour cette résolution effectue $\frac{1}{2}n(n-1)$ soustractions, $\frac{1}{2}n(n-1)$ multiplications et n divisions pour calculer la solution, soit un nombre d’opérations global de l’ordre de n^2 .

Exemple de résolution d’un système triangulaire inférieur. Appliquons une approche orientée colonne pour la résolution du système

$$\begin{pmatrix} 2 & 0 & 0 \\ 1 & 5 & 0 \\ 7 & 9 & 8 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 6 \\ 2 \\ 5 \end{pmatrix}.$$

On trouve que $x_1 = 3$ et l’on considère ensuite le système à deux équations et deux inconnues

$$\begin{pmatrix} 5 & 0 \\ 9 & 8 \end{pmatrix} \begin{pmatrix} x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2 \\ 5 \end{pmatrix} - 3 \begin{pmatrix} 1 \\ 7 \end{pmatrix},$$

pour lequel on trouve $x_2 = -\frac{1}{5}$. On a enfin

$$8x_3 = -16 + \frac{9}{5},$$

soit $x_3 = -\frac{71}{40}$.

Le cas d’un système linéaire dont la matrice est inversible et triangulaire supérieure se traite de manière analogue, par la méthode dite de « remontée » (“*back substitution*” en anglais) suivante

$$\begin{aligned} x_n &= \frac{b_n}{a_{nn}} \\ x_i &= \frac{1}{a_{ii}} \left(b_i - \sum_{j=i+1}^n a_{ij} x_j \right), \quad i = n-1, \dots, 1, \end{aligned} \quad (1.3)$$

et dont la complexité est également de l’ordre de n^2 opérations.

1.2 Méthode d’élimination de Gauss

Une technique de choix pour ramener la résolution d’un système linéaire quelconque à celle d’un système triangulaire est la *méthode d’élimination de Gauss* (*Gaussian elimination* en anglais). Celle-ci consiste en premier lieu à transformer, par des opérations simples sur les équations, le système en un système équivalent, c’est-à-dire ayant la (ou les) même(s) solution(s), $MA\mathbf{x} = M\mathbf{b}$, dans lequel MA est une matrice triangulaire supérieure⁴ (on dit encore que la matrice du système est sous forme *échelonnée*). Cette étape de mise à zéro d’une partie des coefficients de la matrice est qualifiée d’*élimination* et utilise de manière essentielle le fait qu’on ne modifie pas la solution d’un système linéaire en ajoutant à une équation donnée une combinaison linéaire des autres équations. Lorsque la matrice du système est inversible, la solution du système peut ensuite être obtenue par une méthode de remontée, mais le procédé d’élimination est très général et amène toute matrice sous forme échelonnée.

3. On a en effet $a_{11}a_{22}\dots a_{nn} = \det(A) \neq 0$.

4. Il faut bien noter qu’on ne calcule en pratique jamais explicitement la matrice d’élimination M , mais seulement les produits MA et $M\mathbf{b}$.

1.2.1 Élimination sans échange

Commençons par décrire étape par étape la méthode dans sa forme de base, dite *sans échange*, en considérant le système linéaire (1.1), avec A étant une matrice inversible d'ordre n . Supposons de plus que le terme a_{11} de la matrice A est non nul. Nous pouvons alors éliminer l'inconnue x_1 de la deuxième à la n^{e} ligne du système en leur retranchant respectivement la première ligne multipliée par le coefficient $\frac{a_{i1}}{a_{11}}$, $i = 2, \dots, n$. En notant $A^{(1)}$ et $\mathbf{b}^{(1)}$ la matrice et le vecteur second membre résultant de ces opérations⁵, on a alors

$$a_{ij}^{(1)} = a_{ij} - \frac{a_{i1}}{a_{11}} a_{1j} \text{ et } b_i^{(1)} = b_i - \frac{a_{i1}}{a_{11}} b_1, \quad i = 2, \dots, n, j = 1, \dots, n,$$

les coefficients de la première ligne restant les mêmes, et le système $A^{(1)}\mathbf{x} = \mathbf{b}^{(1)}$ est équivalent au système de départ. En supposant non nul le coefficient diagonal $a_{22}^{(1)}$ de $A^{(1)}$, on peut ensuite procéder à l'élimination de l'inconnue x_2 de la troisième à la n^{e} ligne de ce nouveau système, et ainsi de suite. On obtient, sous l'hypothèse $a_{k+1,k+1}^{(k)} \neq 0$, $k = 0, \dots, n-2$, une suite finie de matrices $A^{(k)}$, $1 \leq k \leq n-1$, de la forme

$$A^{(k)} = \begin{pmatrix} a_{11}^{(k)} & a_{12}^{(k)} & \dots & \dots & \dots & a_{1n}^{(k)} \\ 0 & a_{22}^{(k)} & & & & a_{2n}^{(k)} \\ \vdots & \ddots & \ddots & & & \vdots \\ 0 & \dots & 0 & a_{k+1,k+1}^{(k)} & \dots & a_{k+1,n}^{(k)} \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & a_{n,k+1}^{(k)} & \dots & a_{nn}^{(k)} \end{pmatrix},$$

telle que le système $A^{(n-1)}\mathbf{x} = \mathbf{b}^{(n-1)}$ est triangulaire supérieur. Les quantités $a_{k+1,k+1}^{(k)}$, $k = 0, \dots, n-2$, sont appelées les *pivots* et l'on a supposé qu'elles étaient non nulles à chaque étape, les formules permettant de passer du système linéaire de matrice $A^{(k)}$ et de second membre $\mathbf{b}^{(k)}$ au suivant se résumant à

$$a_{ij}^{(k+1)} = a_{ij}^{(k)} - \frac{a_{ik+1}^{(k)}}{a_{k+1,k+1}^{(k)}} a_{k+1,j}^{(k)} \text{ et } b_i^{(k+1)} = b_i^{(k)} - \frac{a_{ik+1}^{(k)}}{a_{k+1,k+1}^{(k)}} b_{k+1}^{(k)}, \quad i = k+2, \dots, n, j = k+1, \dots, n,$$

les autres coefficients étant inchangés. En pratique, pour une résolution « à la main » d'un système linéaire $A\mathbf{x} = \mathbf{b}$ par cette méthode, il est commode pour les calculs d'appliquer l'élimination à la matrice « augmentée » $(A \quad \mathbf{b})$.

Exemple d'application de la méthode d'élimination de Gauss sans échange. Considérons la résolution par la méthode d'élimination de Gauss sans échange du système linéaire suivant

$$\begin{cases} x_1 + 2x_2 + 3x_3 + 4x_4 = 11 \\ 2x_1 + 3x_2 + 4x_3 + x_4 = 12 \\ 3x_1 + 4x_2 + x_3 + 2x_4 = 13 \\ 4x_1 + x_2 + 2x_3 + 3x_4 = 14 \end{cases}.$$

À la première étape, le pivot vaut 1 et on soustrait de la deuxième (resp. troisième (resp. quatrième)) équation la première équation multipliée par 2 (resp. 3 (resp. 4)) pour obtenir

$$\begin{cases} x_1 + 2x_2 + 3x_3 + 4x_4 = 11 \\ -x_2 - 2x_3 - 7x_4 = -10 \\ -2x_2 - 8x_3 - 10x_4 = -20 \\ -7x_2 - 10x_3 - 13x_4 = -3 \end{cases}.$$

Le pivot vaut -1 à la deuxième étape. On retranche alors à la troisième (resp. quatrième) équation la deuxième équation multipliée par -2 (resp. -7), d'où le système

$$\begin{cases} x_1 + 2x_2 + 3x_3 + 4x_4 = 11 \\ -x_2 - 2x_3 - 7x_4 = -10 \\ -4x_3 + 4x_4 = 0 \\ 4x_3 + 36x_4 = 40 \end{cases}.$$

5. On pose $A^{(0)} = A$ et $\mathbf{b}^{(0)} = \mathbf{b}$ pour être consistant.

À la dernière étape, le pivot est égal à -4 et on soustrait à la dernière équation l'avant-dernière multipliée par -1 pour arriver à

$$\begin{cases} x_1 + 2x_2 + 3x_3 + 4x_4 = 11 \\ -x_2 - 2x_3 - 7x_4 = -10 \\ -4x_3 + 4x_4 = 0 \\ 40x_4 = 40 \end{cases}.$$

Ce système triangulaire, équivalent au système d'origine, est enfin résolu par remontée :

$$\begin{cases} x_4 = 1 \\ x_3 = x_4 = 1 \\ x_2 = 10 - 2 - 7 = 1 \\ x_1 = 11 - 2 - 3 - 4 = 2 \end{cases}.$$

Comme on l'a vu, la méthode d'élimination de Gauss, dans sa forme sans échange, ne peut s'appliquer que si tous les pivots $a_{k+1,k+1}^{(k)}$, $k = 0, \dots, n-2$, sont non nuls, ce qui rejette de fait des matrices inversibles aussi simples que

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

De plus, le fait que la matrice soit inversible n'empêche aucunement l'apparition d'un pivot nul durant l'élimination, comme le montre l'exemple ci-dessous.

Exemple de mise en échec de la méthode d'élimination de Gauss sans échange. Considérons la matrice inversible

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 7 & 8 & 9 \end{pmatrix} = A^{(0)}.$$

On a alors

$$A^{(1)} = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 0 & -1 \\ 0 & -6 & -12 \end{pmatrix},$$

et l'élimination s'interrompt à l'issue de la première étape, le pivot $a_{22}^{(1)}$ étant nul.

Il apparaît donc que des conditions plus restrictives que l'inversibilité de la matrice sont nécessaires pour assurer la bonne exécution de cette méthode. Celles-ci sont fournies par le théorème 2. Indiquons qu'il existe des catégories de matrices pour lesquelles la méthode de Gauss sans échange peut-être utilisée sans aucun risque. Parmi celles-ci, on trouve les matrices à *diagonale dominante par lignes ou par colonnes* et les matrices *symétriques définies positives*.

1.2.2 Élimination de Gauss avec échange

Dans sa forme générale, la méthode d'élimination de Gauss permet de transformer un système linéaire dont la matrice est carrée (inversible ou non) ou même rectangulaire en un système échelonné équivalent. En considérant le cas d'une matrice A carrée inversible, nous allons maintenant décrire les modifications à apporter à la méthode déjà présentée pour mener l'élimination à son terme. Dans tout ce qui suit, les notations de la section 1.2.1 sont conservées.

À la première étape, au moins l'un des coefficients de la première colonne de la matrice $A^{(0)} (= A)$ est non nul, faute de quoi la matrice A ne serait pas inversible. On choisit⁶ un de ces éléments comme premier pivot d'élimination et l'on échange alors la première ligne du système avec celle du pivot avant de procéder à l'élimination de la première colonne de la matrice résultante, c'est-à-dire l'annulation de tous les éléments de la première colonne de la matrice (permutée) du système situés sous la diagonale. On note $A^{(1)}$ et $\mathbf{b}^{(1)}$ la matrice et le second membre du système obtenu et l'on réitère ce procédé. Par la suite, à l'étape $k+1$, $1 \leq k \leq n-2$, la matrice $A^{(k)}$ est inversible⁷, et donc l'un au moins des éléments $a_{i,k+1}^{(k)}$, $k+1 \leq i \leq n$, est différent de zéro. Après avoir choisi

6. Pour l'instant, on ne s'intéresse pas au choix *effectif* du pivot, qui est cependant d'une importance cruciale pour la stabilité numérique de la méthode. Ce point est abordé dans la section 1.2.3.

7. On a en effet que $\det(A^{(k)}) = \pm \det(A)$, $k = 0, \dots, n-1$. On renvoie à la section 1.3.1 pour une justification de ce fait.

comme pivot l'un de ces coefficients non nuls, on effectue l'échange de la ligne de ce pivot avec la $k + 1^{\text{e}}$ ligne de la matrice $A^{(k)}$, puis l'élimination conduisant à la matrice $A^{(k+1)}$. Ainsi, on arrive après $n - 1$ étapes à la matrice $A^{(n-1)}$, dont le coefficient $a_{nn}^{(n-1)}$ est non nul.

En raison de l'échange de lignes qui a éventuellement lieu avant chaque étape d'élimination, on parle de méthode d'élimination de Gauss *avec échange*.

Exemple d'application de la méthode d'élimination de Gauss avec échange. Considérons la résolution du système linéaire $Ax = b$, avec

$$A = \begin{pmatrix} 2 & 4 & -4 & 1 \\ 3 & 6 & 1 & -2 \\ -1 & 1 & 2 & 3 \\ 1 & 1 & -4 & 1 \end{pmatrix} \text{ et } b = \begin{pmatrix} 0 \\ -7 \\ 4 \\ 2 \end{pmatrix},$$

par application de la méthode d'élimination de Gauss avec échange. On trouve successivement

$$A^{(1)} = \begin{pmatrix} 2 & 4 & -4 & 1 \\ 0 & 0 & 7 & -\frac{7}{2} \\ 0 & 3 & 0 & \frac{7}{2} \\ 0 & -1 & -2 & \frac{1}{2} \end{pmatrix} \text{ et } b^{(1)} = \begin{pmatrix} 0 \\ -7 \\ 4 \\ 2 \end{pmatrix},$$

$$A^{(2)} = \begin{pmatrix} 2 & 4 & -4 & 1 \\ 0 & 3 & 0 & \frac{7}{2} \\ 0 & 0 & 7 & -\frac{7}{2} \\ 0 & 0 & -2 & \frac{5}{3} \end{pmatrix} \text{ et } b^{(2)} = \begin{pmatrix} 0 \\ 4 \\ -7 \\ \frac{10}{3} \end{pmatrix},$$

et

$$A^{(3)} = \begin{pmatrix} 2 & 4 & -4 & 1 \\ 0 & 3 & 0 & \frac{7}{2} \\ 0 & 0 & 7 & -\frac{7}{2} \\ 0 & 0 & 0 & \frac{2}{3} \end{pmatrix} \text{ et } b^{(3)} = \begin{pmatrix} 0 \\ 4 \\ -7 \\ \frac{4}{3} \end{pmatrix},$$

d'où la solution $x = (1 \quad -1 \quad 0 \quad 2)^T$. On note que l'on a procédé au cours de la deuxième étape à l'échange des deuxième et troisième lignes.

On pourra remarquer que si la matrice A est non inversible, alors tous les éléments $a_{ik+1}^{(k)}$, $k + 1 \leq i \leq n$, seront nuls pour au moins une valeur de k entre 0 et $n - 1$. Si cela vient à se produire alors que $k < n - 1$, on n'a pas besoin de réaliser l'élimination dans la $k + 1^{\text{e}}$ colonne (puisque celle-ci est déjà nulle) et l'on passe simplement à l'étape suivante en posant $A^{(k+1)} = A^{(k)}$ et $b^{(k+1)} = b^{(k)}$. L'élimination est donc bien possible pour une matrice carrée non inversible et l'on a démontré le résultat suivant.

Théorème 1 Soit A une matrice carrée, inversible ou non. Il existe au moins une matrice inversible M telle que la matrice MA soit triangulaire supérieure.

Il reste à compter le nombre d'opérations élémentaires que requiert l'application de la méthode d'élimination de Gauss pour la résolution d'un système linéaire de n équations à n inconnues. Tout d'abord, pour passer de la matrice $A^{(k)}$ à la matrice $A^{(k+1)}$, $0 \leq k \leq n - 2$, on effectue $(n - k - 1)^2$ soustractions, $(n - k - 1)^2$ multiplications et $n - k - 1$ divisions, ce qui correspond à un total de $\frac{1}{6}(2n - 1)n(n - 1)$ soustractions, $\frac{1}{6}(2n - 1)n(n - 1)$ multiplications et $\frac{1}{2}n(n - 1)$ divisions pour l'élimination complète. Pour la mise à jour du second membre à l'étape $k + 1$, on a besoin de $n - k - 1$ soustractions et autant de multiplications, soit en tout $\frac{1}{2}n(n - 1)$ soustractions et $\frac{1}{2}n(n - 1)$ multiplications. Enfin, il faut faire $\frac{1}{2}n(n - 1)$ soustractions, autant de multiplications et n divisions pour résoudre le système final par une méthode de remontée.

En tout, la résolution du système par la méthode d'élimination de Gauss nécessite donc environ $\frac{n^3}{3}$ soustractions, $\frac{n^3}{3}$ multiplications et $\frac{n^2}{2}$ divisions. À titre de comparaison, le calcul de la solution du système par la règle de Cramer (voir la proposition 6) requiert, en utilisant un développement « brutal » par ligne ou colonne pour le calcul des déterminants, environ $(n + 1)!$ additions ou soustractions, $(n + 2)!$ multiplications et n divisions. Ainsi, pour $n = 10$ par exemple, on obtient un compte d'approximativement 700 opérations pour la méthode d'élimination de Gauss contre près de 479000000 opérations pour la règle de Cramer !

1.2.3 Choix du pivot

Revenons à présent sur le choix des pivots lors de l'élimination. À la $k + 1^{\text{e}}$ étape du procédé, si l'élément $a_{k+1,k+1}^{(k)}$ est non nul, il semble naturel de l'utiliser comme pivot (c'est d'ailleurs ce que l'on fait dans la méthode de Gauss sans échange). Cependant, à cause de la présence d'erreurs d'arrondi en pratique, cette manière de procéder est en général à proscrire, comme l'illustre l'exemple d'instabilité numérique suivant.

Exemple d'application numérique. Supposons que les calculs soient effectués en virgule flottante dans le système décimal, avec une mantisse à trois chiffres, et considérons le système

$$\begin{pmatrix} 10^{-4} & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix},$$

dont la solution « exacte » est $x_1 = 1,0001$ et $x_2 = 0,9998$. En choisissant le nombre 10^{-4} comme pivot à la première étape de l'élimination de Gauss, on obtient le système triangulaire

$$\begin{pmatrix} 10^{-4} & 1 \\ 0 & -9990 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ -9990 \end{pmatrix},$$

car les nombres $-10^4 + 1 = -9999$ et $-10^4 + 2 = -9998$ sont tous deux arrondis au nombre -9990 . La solution numérique calculée est alors

$$x_1 = 0 \text{ et } x_2 = 1,$$

ce qui très différent de la véritable solution du système. Si, par contre, on commence par échanger les deux équations du système pour utiliser le nombre 1 comme pivot, on trouve

$$\begin{pmatrix} 1 & 1 \\ 0 & 0,999 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 0,999 \end{pmatrix},$$

puisque les nombres $-10^{-4} + 1 = 0,9999$ et $-2 \cdot 10^{-4} + 1 = 0,9998$ sont chacun arrondis au nombre 0,999. La solution calculée vaut alors

$$x_1 = 1 \text{ et } x_2 = 1,$$

ce qui est cette fois très satisfaisant.

En général, le changement de pivot n'a pas un effet aussi spectaculaire que dans cet exemple, mais il n'en demeure pas moins essentiel lorsque les calculs sont effectués en arithmétique à virgule flottante. De fait, pour éviter la propagation d'erreurs et obtenir une meilleure stabilité numérique de la méthode, il faut chercher, même lorsque le pivot « naturel » est non nul, à choisir le plus grand pivot en valeur absolue. On peut pour cela suivre, au début de la $k + 1^{\text{e}}$ étape, $0 \leq k \leq n - 2$, de l'élimination,

- soit une stratégie de *pivot partiel* (*partial pivoting* en anglais) dans laquelle le pivot est l'élément $a_{rk+1}^{(k)}$ de la $k + 1^{\text{e}}$ colonne de la matrice $A^{(k)}$ situé sous la diagonale ayant la plus grande valeur absolue,

$$|a_{rk+1}^{(k)}| = \max_{k+1 \leq i \leq n} |a_{ik+1}^{(k)}|,$$

- soit une stratégie de *pivot total* (*complete pivoting* en anglais), plus coûteuse, dans laquelle le pivot est l'élément $a_{rs}^{(k)}$ de la sous-matrice $(a_{ij}^{(k)})_{k+1 \leq i, j \leq n}$ le plus grand en valeur absolue,

$$|a_{rs}^{(k)}| = \max_{k+1 \leq i, j \leq n} |a_{ij}^{(k)}|,$$

Dans le derniers cas, si le pivot n'est pas dans la $k + 1^{\text{e}}$ colonne, il faut procéder à un échange de colonnes en plus d'un éventuel échange de lignes.

Quelle que soit la stratégie adoptée, la recherche des pivots doit également être prise en compte dans l'évaluation du coût global de la méthode d'élimination de Gauss. Celle-ci demande de l'ordre de n^2 comparaisons au total pour la stratégie de pivot partiel et de l'ordre de n^3 comparaisons pour celle de pivot total, la première étant privilégiée en raison de sa complexité algorithmique moindre et de performances généralement très bonnes.

1.3 Interprétation matricielle de l'élimination de Gauss : la factorisation LU

Nous allons maintenant montrer que la méthode de Gauss dans sa forme sans échange est équivalente à la décomposition de la matrice A sous la forme d'un produit de deux matrices, $A = LU$, avec L une matrice triangulaire inférieure (*lower triangular* en anglais), qui est l'inverse de la matrice M des transformations successives appliquées à la matrice A lors de l'élimination de Gauss sans échange, et U une matrice triangulaire supérieure (*upper triangular* en anglais), avec $U = A^{(n-1)}$ en reprenant la notation utilisée dans la section 1.2.1.

1.3.1 Formalisme matriciel

Chacune des opérations que nous avons effectuées pour transformer le système linéaire lors de l'élimination de Gauss, que ce soit l'échange de deux lignes ou l'annulation d'une partie des coefficients d'une colonne de la matrice $A^{(k)}$, $0 \leq k \leq n-2$, peut se traduire matriciellement par la multiplication de la matrice et du second membre du système linéaire courant par une matrice inversible particulière. L'introduction de ces matrices va permettre de traduire le procédé d'élimination dans un formalisme matriciel débouchant sur une factorisation remarquable de la matrice A .

Matrices des transformations élémentaires

Soit $(m, n) \in (\mathbb{N} \setminus \{0, 1\})^2$ et $A = (a_{ij})_{1 \leq i \leq m, 1 \leq j \leq n} \in M_{m,n}(\mathbb{R})$. On appelle *opérations élémentaires sur les lignes de* A les transformations suivantes :

- l'échange (entre elles) des i^{e} et j^{e} lignes de A ;
- la multiplication de la i^{e} ligne de A par un scalaire $\lambda \in \mathbb{R} \setminus \{0\}$;
- le remplacement de la i^{e} ligne de A par la somme de cette même ligne avec la j^{e} ligne de A multipliée par un scalaire λ , où $\lambda \in \mathbb{R} \setminus \{0\}$.

Explicitons à présent les opérations matricielles correspondant à chacune de ces opérations. Tout d'abord, échanger les i^{e} et j^{e} lignes, $(i, j) \in \{1 \dots, m\}^2$, de la matrice A revient à multiplier à gauche cette matrice par la *matrice de permutation* (*permutation matrix* en anglais)

$$P_{ij} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & & & & & & & & \vdots \\ \vdots & \ddots & 1 & \ddots & & & & & & & \vdots \\ \vdots & & \ddots & 0 & 0 & \dots & 0 & 1 & & & \vdots \\ \vdots & & & 0 & 1 & \ddots & & 0 & & & \vdots \\ \vdots & & & \vdots & \ddots & \ddots & \ddots & \vdots & & & \vdots \\ \vdots & & & 0 & & \ddots & 1 & 0 & & & \vdots \\ \vdots & & & 1 & 0 & \dots & 0 & 0 & & & \vdots \\ \vdots & & & & & & & & 1 & & \vdots \\ \vdots & & & & & & & & & \ddots & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix} = I_m + (E_{ij} + E_{ji} - E_{ii} - E_{jj}) \in M_m(\mathbb{R}).$$

Cette matrice est symétrique et orthogonale, de déterminant valant -1 .

La multiplication de la i^{e} ligne de la matrice A par un scalaire non nul λ s'effectue en multipliant à gauche

cette matrice par la *matrice de dilatation* (*directional scaling matrix* en anglais)

$$D_i(\lambda) = \begin{pmatrix} 1 & 0 & \dots & \dots & \dots & \dots & 0 \\ 0 & \ddots & \ddots & & & & \vdots \\ \vdots & \ddots & 1 & \ddots & & & \vdots \\ \vdots & & \ddots & \lambda & \ddots & & \vdots \\ \vdots & & & \ddots & 1 & \ddots & \vdots \\ \vdots & & & & \ddots & \ddots & 0 \\ 0 & \dots & \dots & \dots & \dots & 0 & 1 \end{pmatrix} = I_m + (\lambda - 1)E_{ii} \in M_m(\mathbb{R}).$$

Cette matrice est inversible et $D_i(\lambda)^{-1} = D_i(\frac{1}{\lambda})$.

Enfin, le remplacement de la i^e ligne de A par la somme de la i^e ligne et de la j^e ligne, $i \neq j$, multipliée par un scalaire non nul λ est obtenu en multipliant à gauche la matrice A par la *matrice de transvection* (*shear matrix* en anglais)

$$T_{ij}(\lambda) = \begin{pmatrix} 1 & 0 & \dots & \dots & \dots & \dots & 0 \\ 0 & \ddots & \ddots & & & & \vdots \\ \vdots & \ddots & 1 & \dots & 0 & & \vdots \\ \vdots & & \vdots & \ddots & \vdots & & \vdots \\ \vdots & & \lambda & \dots & 1 & \ddots & \vdots \\ \vdots & & & & \ddots & \ddots & 0 \\ 0 & \dots & \dots & \dots & \dots & 0 & 1 \end{pmatrix} = I_m + \lambda E_{ij} \in M_m(\mathbb{R})$$

(on a supposé dans l'exemple que $j < i$). Cette matrice a pour inverse⁸ $T_{ij}(-\lambda)$. Pour la suite, il est important de remarquer que, étant donné trois entiers naturels i, k et l , tels que $1 \leq k < i < l \leq m$, et deux scalaires non nuls λ et μ , le produit des matrices de transvection $T_{ik}(\lambda)$ et $T_{lk}(\mu)$ de $M_m(\mathbb{R})$ est commutatif et vaut

$$T_{ik}(\lambda)T_{lk}(\mu) = I_m + \lambda E_{ik} + \mu E_{lk}.$$

On effectue de manière analogue des opérations élémentaires *sur les colonnes* de la matrice A en multipliant à *droite* cette dernière par les matrices d'ordre n correspondantes.

Exemple d'action d'une matrice de permutation. Soit les matrices $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$ et $P_{23} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$. On a

$$P_{23}A = \begin{pmatrix} 1 & 2 & 3 \\ 7 & 8 & 9 \\ 4 & 5 & 6 \end{pmatrix} \text{ et } AP_{23} = \begin{pmatrix} 1 & 3 & 2 \\ 4 & 6 & 5 \\ 7 & 9 & 8 \end{pmatrix}.$$

Factorisation LU

Si l'élimination arrive à son terme sans qu'il y ait besoin d'échanger des lignes du système linéaire, la matrice inversible M du théorème 1 est unique et égale au produit

$$M = E^{(n-1)} \dots E^{(2)} E^{(1)}$$

8. L'ensemble des matrices de transvection de $M_m(\mathbb{R})$ engendre le *groupe spécial linéaire* (*special linear group* en anglais) $SL_m(\mathbb{R})$, constitué des matrices d'ordre m dont le déterminant vaut 1, et, avec l'ensemble des matrices de dilatation de $M_m(\mathbb{R})$, le groupe général linéaire $GL_m(\mathbb{R})$.

de $n - 1$ matrices d'élimination définies par

$$E^{(k)} = \prod_{i=k+1}^n T_{ik} \left(-\frac{a_{ik}^{(k-1)}}{a_{kk}^{(k-1)}} \right) = \begin{pmatrix} 1 & 0 & \dots & \dots & \dots & \dots & \dots & 0 \\ 0 & 1 & \ddots & & & & & \vdots \\ \vdots & \ddots & \ddots & \ddots & & & & \vdots \\ \vdots & & & 0 & 1 & \ddots & & \vdots \\ \vdots & & & \vdots & -\frac{a_{k+1k}^{(k-1)}}{a_{kk}^{(k-1)}} & 1 & \ddots & \vdots \\ \vdots & & & \vdots & -\frac{a_{k+2k}^{(k-1)}}{a_{kk}^{(k-1)}} & 0 & \ddots & \vdots \\ \vdots & & & \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & \dots & 0 & -\frac{a_{nk}^{(k-1)}}{a_{kk}^{(k-1)}} & 0 & \dots & 0 & 1 \end{pmatrix}, \quad 1 \leq k \leq n-1. \quad (1.4)$$

Par construction, la matrice M est triangulaire inférieure et son inverse est donc également une matrice triangulaire inférieure. Il en résulte que la matrice A s'écrit comme le produit

$$A = LU, \quad (1.5)$$

dans lequel $L = M^{-1}$ et $U = MA = A^{(n-1)}$ est une matrice triangulaire supérieure. Fait remarquable, la matrice L se calcule de manière immédiate à partir des matrices $E^{(k)}$, $1 \leq k \leq n-1$, alors qu'il n'existe pas d'expression simple pour M . En effet, chacune des matrices d'élimination définies par (1.4) étant produit de matrices de transvection, il est facile de vérifier que son inverse vaut

$$(E^{(k)})^{-1} = \prod_{i=k+1}^n T_{ik} \left(\frac{a_{ik}^{(k-1)}}{a_{kk}^{(k-1)}} \right) = \begin{pmatrix} 1 & 0 & \dots & \dots & \dots & \dots & \dots & 0 \\ 0 & 1 & \ddots & & & & & \vdots \\ \vdots & \ddots & \ddots & \ddots & & & & \vdots \\ \vdots & & & 0 & 1 & \ddots & & \vdots \\ \vdots & & & \vdots & \frac{a_{k+1k}^{(k-1)}}{a_{kk}^{(k-1)}} & 1 & \ddots & \vdots \\ \vdots & & & \vdots & \frac{a_{k+2k}^{(k-1)}}{a_{kk}^{(k-1)}} & 0 & \ddots & \vdots \\ \vdots & & & \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & \dots & 0 & \frac{a_{nk}^{(k-1)}}{a_{kk}^{(k-1)}} & 0 & \dots & 0 & 1 \end{pmatrix}, \quad 1 \leq k \leq n-1,$$

et l'on a⁹

$$L = (E^{(1)})^{-1}(E^{(2)})^{-1} \dots (E^{(n-1)})^{-1} = \begin{pmatrix} 1 & 0 & \dots & \dots & \dots & \dots & \dots & 0 \\ \frac{a_{21}^{(0)}}{a_{11}^{(0)}} & 1 & \ddots & & & & & \vdots \\ \vdots & \ddots & \ddots & \ddots & & & & \vdots \\ \vdots & & & \frac{a_{k+1k}^{(k-1)}}{a_{kk}^{(k-1)}} & 1 & \ddots & & \vdots \\ \vdots & & & \vdots & \vdots & \ddots & \ddots & 0 \\ \frac{a_{n1}^{(0)}}{a_{11}^{(0)}} & \dots & \frac{a_{nk}^{(k-1)}}{a_{kk}^{(k-1)}} & \dots & \frac{a_{n,n-1}^{(n-2)}}{a_{n-1,n-1}^{(n-2)}} & \dots & 1 & \vdots \end{pmatrix}.$$

Si des échanges de lignes ont eu lieu lors de l'élimination, la¹⁰ matrice M s'écrit

$$M = E^{(n-1)}P^{(n-1)} \dots E^{(2)}P^{(2)}E^{(1)}P^{(1)},$$

9. La vérification est laissée en exercice.

10. Il n'y a pas forcément unicité de la matrice dans ce cas, en raison de possibles multiples choix de pivots.

où la matrice $P^{(k)}$, $1 \leq k \leq n-1$, est soit la matrice de permutation correspondant à l'échange de lignes effectué à la k^{e} étape, soit la matrice identité si le pivot « naturel » est utilisé. En écrivant que

$$M = E^{(n-1)}(P^{(n-1)}E^{(n-2)}P^{(n-1)}) \dots (P^{(n-1)} \dots P^{(2)}E^{(1)}P^{(2)} \dots P^{(n-1)})(P^{(n-1)} \dots P^{(2)}P^{(1)}),$$

et en posant $P = P^{(n-1)} \dots P^{(2)}P^{(1)}$, on obtient $L = PM^{-1}$ et $U = (MP^{-1})PA$, d'où

$$PA = LU.$$

Enfin, si la stratégie de choix de pivot employée conduit à des échanges de colonnes en plus des éventuels échanges de lignes, la factorisation prend la forme

$$PAQ = LU,$$

les matrices P et Q rendant respectivement compte des échanges de lignes et de colonnes.

Exemple d'application de la factorisation $PA = LU$. Revenons à l'exemple de mise en échec de la méthode d'élimination de Gauss, pour lequel le pivot « naturel » est nul à la seconde étape. La recherche d'un pivot partiel conduit à l'échange de la deuxième ligne avec la troisième et l'on arrive à

$$A^{(2)} = \begin{pmatrix} 1 & 2 & 3 \\ 0 & -6 & -12 \\ 0 & 0 & -1 \end{pmatrix} = U.$$

Les matrices d'élimination aux deux étapes effectuées sont respectivement

$$E^{(1)} = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -7 & 0 & 1 \end{pmatrix} \text{ et } E^{(2)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

et les matrices d'échange sont respectivement

$$P^{(1)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ et } P^{(2)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix},$$

d'où

$$L = \begin{pmatrix} 1 & 0 & 0 \\ 7 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix}.$$

et la matrice de permutation

$$P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

On a

$$LU = \begin{pmatrix} 1 & 2 & 3 \\ 7 & 8 & 9 \\ 2 & 4 & 5 \end{pmatrix} = PA.$$

La résolution du système linéaire (1.1) connaissant la factorisation LU de la matrice A se ramène à celle du système triangulaire inférieur

$$Ly = \mathbf{b} \tag{1.6}$$

par une méthode de descente, suivie de celle du système triangulaire supérieur

$$Ux = \mathbf{y} \tag{1.7}$$

par une méthode de remontée, ce que l'on accomplit en effectuant $n(n-1)$ additions, $n(n-1)$ multiplications et n divisions. Dans le cas d'une factorisation de type $PA = LU$ (resp. $PAQ = LU$), il faudra appliquer la matrice de permutation P au vecteur \mathbf{b} de manière à tout d'abord résoudre le système $Ly = P\mathbf{b}$ avant de considérer le système $Ux = \mathbf{y}$ (resp. les systèmes $Uz = \mathbf{y}$ et $Q^{-1}\mathbf{x} = \mathbf{z}$).

Exemple d'application de la factorisation LU pour la résolution d'un système linéaire. Considérons la matrice

$$A = \begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 10 \end{pmatrix}.$$

En appliquant de l'algorithme de factorisation, on arrive à

$$A = LU = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 4 & 7 \\ 0 & -3 & -6 \\ 0 & 0 & 1 \end{pmatrix}.$$

Si $\mathbf{b} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, la solution de $Ly = \mathbf{b}$ est $\mathbf{y} = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$ et celle de $Ux = \mathbf{y}$ est $\mathbf{x} = \frac{1}{3} \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$.

On observera que ce type de factorisation est particulièrement avantageux lorsque l'on doit résoudre plusieurs systèmes linéaires ayant tous la même matrice et des seconds membres différents. En effet, une fois obtenue la factorisation LU de la matrice, la résolution de chaque système ne nécessite que de l'ordre de n^2 opérations.

Terminons cette section en indiquant que la méthode de factorisation LU fournit une manière rapide de calculer le déterminant de la matrice A , qui n'est autre, au signe près, que le produit des pivots, puisque¹¹

$$\det(PA) = \det(LU) = \det(L)\det(U) = \det(U) = \left(\prod_{i=1}^n u_{ii} \right),$$

et

$$\det(A) = \frac{\det(PA)}{\det(P)} = \begin{cases} \det(PA) & \text{si on a effectué un nombre pair d'échanges de lignes,} \\ -\det(PA) & \text{si on a effectué un nombre impair d'échanges de lignes,} \end{cases}$$

le déterminant d'une matrice de permutation étant égal à -1 .

1.3.2 Condition d'existence de la factorisation LU

Commençons par donner une condition suffisante assurant qu'il n'y aura pas d'échange de lignes durant l'élimination de Gauss, ce qui conduira bien à une factorisation de la forme (1.5) de la matrice. On va à cette occasion aussi établir que cette décomposition est unique si l'on impose la valeur 1 aux éléments diagonaux de L (c'est précisément la valeur obtenue avec la construction par élimination de Gauss).

Théorème 2 (condition suffisante d'existence et d'unicité de la factorisation LU) Soit A une matrice d'ordre n . La factorisation LU de A , avec $l_{ii} = 1$ pour $i = 1, \dots, n$, existe et est unique si toutes les sous-matrices principales

$$A_k = \begin{pmatrix} a_{11} & \dots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \dots & a_{kk} \end{pmatrix}, \quad 1 \leq k \leq n, \quad (1.8)$$

extraites de A sont inversibles.

DÉMONSTRATION. Il est possible de montrer l'existence de la factorisation LU de manière constructive, en utilisant le procédé d'élimination de Gauss. En supposant que les n sous-matrices principales extraites de A sont inversibles, on va ici prouver en même temps l'existence et l'unicité par un raisonnement par récurrence¹².

Pour $k = 1$, on a

$$A_1 = a_{11} \neq 0,$$

et il suffit de poser $L_1 = 1$ et $U_1 = a_{11}$. Montrons à présent que s'il existe une unique factorisation de la sous-matrice A_{k-1} , $2 \leq k \leq n$, de la forme $A_{k-1} = L_{k-1}U_{k-1}$, avec $(L_{k-1})_{ii} = 1$, $i = 1, \dots, k-1$, alors il existe une unique factorisation de ce type pour A_k . Pour cela, décomposons A_k en blocs

$$A_k = \begin{pmatrix} A_{k-1} & \mathbf{b} \\ \mathbf{c}^\top & d \end{pmatrix},$$

11. On laisse le lecteur traiter le cas d'une factorisation de la forme $PAQ = LU$.

12. Notons que ce procédé de démonstration permet aussi de prouver *directement* (c'est-à-dire sans faire appel à un résultat sur la factorisation LU) l'existence et l'unicité de la factorisation de Cholesky d'une matrice symétrique définie positive (voir le théorème 5).

avec \mathbf{b} et \mathbf{c} des matrices de $M_{k-1,1}(\mathbb{R})$ et d une matrice d'ordre 1, et cherchons une factorisation de A_k de la forme

$$\begin{pmatrix} A_{k-1} & \mathbf{b} \\ \mathbf{c}^\top & d \end{pmatrix} = \begin{pmatrix} L_{k-1} & \mathbf{0} \\ \mathbf{l}^\top & 1 \end{pmatrix} \begin{pmatrix} U_{k-1} & \mathbf{u} \\ \mathbf{0}^\top & \mu \end{pmatrix}$$

où $\mathbf{0}$ désigne le vecteur nul de $M_{k-1,1}(\mathbb{R})$, \mathbf{l} et \mathbf{u} sont des matrices de $M_{k-1,1}(\mathbb{R})$ et μ est une matrice d'ordre 1. En effectuant le produit de matrices et en identifiant par blocs avec A_k , on obtient

$$L_{k-1}U_{k-1} = A_{k-1}, \quad L_{k-1}\mathbf{u} = \mathbf{b}, \quad \mathbf{l}^\top U_{k-1} = \mathbf{c}^\top \quad \text{et} \quad \mathbf{l}^\top \mathbf{u} + \mu = d.$$

Si la première de ces égalités n'apporte aucune nouvelle information, les trois suivantes permettent de déterminer les matrices \mathbf{l} , \mathbf{u} et μ . En effet, on a par hypothèse $0 \neq \det(A_{k-1}) = \det(L_{k-1})\det(U_{k-1})$, les matrices L_{k-1} et U_{k-1} sont donc inversibles. Par conséquent, les matrices \mathbf{l} et \mathbf{u} existent et sont uniques et $\mu = d - \mathbf{l}^\top \mathbf{u}$. Ceci achève la preuve par récurrence. \square

Dans cette preuve, on utilise de manière fondamentale le fait les termes diagonaux de la matrice L sont tous égaux à 1. On aurait tout aussi bien pu choisir d'imposer d'autres valeurs (non nulles) ou encore décider de fixer les valeurs des éléments diagonaux de la matrice U . Ceci implique que plusieurs factorisations LU existent, chacune pouvant être déduite d'une autre par multiplication par une matrice diagonale convenable (voir la section ??).

On remarque également que la condition du théorème n'est que *suffisante*. Il n'est en effet pas nécessaire que la matrice A soit inversible pour que sa factorisation LU existe et soit unique (ce cas étant cependant le seul ayant vraiment un intérêt pratique). Nous laissons au lecteur le soin d'adapter et de compléter la démonstration précédente pour obtenir le résultat ci-après.

Théorème 3 (condition nécessaire et suffisante d'existence et d'unicité de la factorisation LU) Soit A une matrice d'ordre n et de rang non nul r . Une factorisation LU de A , avec $l_{ii} = 1$ pour $i = 1, \dots, n$, existe si et seulement si les sous-matrices principales A_k d'ordre $k = 1, \dots, r$ extraites de A sont inversibles. Cette factorisation est de plus unique sous la même condition si et seulement si r est supérieur ou égal à $n - 1$.

Pour toute matrice A inversible, il est possible de se ramener à la condition suffisante du théorème 2 après des échanges préalable de lignes de la matrice (comme on l'a vu lors de la description de l'élimination de Gauss avec échange). En ce sens, la factorisation LU des matrices inversibles est toujours possible. Si une stratégie de pivot partiel ou de pivot total est appliquée à l'élimination de Gauss, on a plus précisément le résultat suivant.

Théorème 4 Soit A une matrice d'ordre n inversible. Alors, il existe une matrice P (resp. des matrices P et Q) tenant compte d'une stratégie de pivot partiel (resp. de pivot total), une matrice triangulaire inférieure L , dont les éléments sont inférieurs ou égaux à 1 en valeur absolue, et une matrice triangulaire supérieure U telles que

$$PA = LU \quad (\text{resp. } PAQ = LU).$$

1.4 Factorisation de Cholesky

Nous présentons dans cette dernière section une factorisation adaptée aux matrices réelles *symétriques définies positives*. On peut en effet remarquer que de telles matrices vérifient les hypothèses du théorème 2 en vertu du critère de Sylvester (voir le théorème 7), en observant que les termes diagonaux de la matrice U issue de leur factorisation LU sont alors strictement positifs. Ceci conduit à l'introduction d'une factorisation tirant partie de la propriété de symétrie en ne faisant intervenir qu'une seule matrice triangulaire inférieure. On a plus précisément le résultat suivant.

Théorème 5 (« factorisation de Cholesky ») Soit A une matrice réelle symétrique définie positive. Alors, il existe une unique matrice triangulaire inférieure B , dont les éléments diagonaux sont strictement positifs, telle que

$$A = BB^\top.$$

DÉMONSTRATION. Supposons que la matrice A est d'ordre n . On sait, par le théorème 7, que les déterminants des sous-matrices principales extraites A_k , $1 \leq k \leq n$, de A (définies par (1.8)), sont strictement positifs et les conditions du théorème 2 sont vérifiées. La matrice A admet donc une unique factorisation LU. Les éléments diagonaux de la matrice U sont de plus strictement positifs, car on a

$$\prod_{i=1}^k u_{ii} = \det(A_k) > 0, \quad 1 \leq k \leq n.$$

En introduisant la matrice diagonale Δ définie par $(\Delta)_{ii} = \sqrt{u_{ii}}$, $1 \leq i \leq n$, la factorisation se réécrit

$$A = L\Delta\Delta^{-1}U.$$

En posant $B = L\Delta$ et $C = \Delta^{-1}U$, la symétrie de A entraîne que $BC = C^T B^T$, d'où $C(B^T)^{-1} = B^{-1}C^T = I_n$ (une matrice étant triangulaire supérieure, l'autre triangulaire inférieure et toutes deux à coefficients diagonaux égaux à 1) et donc $C = B^T$. On a donc montré l'existence d'au moins une factorisation de Cholesky. Pour montrer l'unicité de cette décomposition, on suppose qu'il existe deux matrices triangulaires inférieures B_1 et B_2 telles que

$$A = B_1 B_1^T = B_2 B_2^T,$$

d'où $B_2^{-1} B_1 = B_2^T (B_1^T)^{-1}$. Il existe donc une matrice diagonale D telle que $B_2^{-1} B_1 = D$ et, par conséquent, $B_1 = B_2 D$. Finalement, on a

$$B_2 B_2^T = B_1 B_1^T = B_2 D D^T B_2^T,$$

et donc $D^2 = I_n$. Les coefficients diagonaux d'une matrice de factorisation de Cholesky étant par hypothèse positifs, on a nécessairement $D = I_n$ et donc $B_1 = B_2$. \square

Pour la mise en œuvre de cette factorisation, on procède de la manière suivante. On pose $B = (b_{ij})_{1 \leq i, j \leq n}$ avec $b_{ij} = 0$ si $i < j$ et l'on déduit alors de l'égalité $A = BB^T$ que

$$a_{ij} = \sum_{k=1}^n b_{ik} b_{jk} = \sum_{k=1}^{\min(i,j)} b_{ik} b_{jk}, \quad 1 \leq i, j \leq n.$$

La matrice A étant réelle symétrique, il suffit, par exemple, que les relations ci-dessus soient vérifiées pour $j \leq i$ et l'on construit alors les colonnes de la matrice B à partir de celles de A . En fixant l'indice j à 1 et en faisant varier l'indice i de 1 à n , on trouve

$$\begin{aligned} a_{11} &= (b_{11})^2, & \text{d'où } b_{11} &= \sqrt{a_{11}}, \\ a_{21} &= b_{11} b_{21}, & \text{d'où } b_{21} &= \frac{a_{21}}{b_{11}}, \\ &\vdots & &\vdots \\ a_{n1} &= b_{11} b_{n1}, & \text{d'où } b_{n1} &= \frac{a_{n1}}{b_{11}}, \end{aligned}$$

ce qui permet la détermination de la première colonne de B . Les coefficients de la j^{e} colonne de B , $2 \leq j \leq n$, s'obtiennent en utilisant les relations

$$\begin{aligned} a_{jj} &= (b_{j1})^2 + (b_{j2})^2 + \cdots + (b_{jj})^2, & \text{d'où } b_{jj} &= \sqrt{a_{jj} - \sum_{k=1}^{j-1} (b_{jk})^2}, \\ a_{j+1j} &= b_{j1} b_{j+11} + b_{j2} b_{j+12} + \cdots + b_{jj} b_{j+1j}, & \text{d'où } b_{j+1j} &= \frac{a_{j+1j} - \sum_{k=1}^{j-1} b_{jk} b_{j+1k}}{b_{jj}}, \\ &\vdots & &\vdots \\ a_{nj} &= b_{j1} b_{n1} + b_{j2} b_{n2} + \cdots + b_{jj} b_{nj}, & \text{d'où } b_{nj} &= \frac{a_{nj} - \sum_{k=1}^{j-1} b_{jk} b_{nk}}{b_{jj}}, \end{aligned}$$

après avoir préalablement déterminé les $j-1$ premières colonnes, le théorème 5 assurant que les quantités sous les racines carrées sont strictement positives. En pratique, on ne vérifie d'ailleurs pas que la matrice réelle A est définie positive, mais simplement qu'elle est symétrique, avant de débiter la factorisation. En effet, si la valeur trouvée à la k^{e} étape, $1 \leq k \leq n$, pour la quantité $(b_{kk})^2$ est négative ou nulle, c'est que A n'est pas définie positive. Au contraire, si l'algorithme de factorisation arrive à son terme, cela prouve que A est définie positive, car, pour toute matrice inversible B et tout vecteur \mathbf{v} non nul, on a

$$(BB^T \mathbf{v}, \mathbf{v}) = \|B^T \mathbf{v}\|_2^2 > 0.$$

Il est à noter que le calcul du déterminant d'une matrice dont on connaît la factorisation de Cholesky est immédiat, puisque

$$\det(A) = \det(BB^T) = (\det(B))^2 = \left(\prod_{i=1}^n b_{ii} \right)^2.$$

Le nombre d'opérations nécessaires pour effectuer la factorisation de Cholesky d'une matrice réelle A symétrique définie positive d'ordre n par les formules ci-dessus est de $\frac{1}{6}(n^2 - 1)n$ additions et soustractions, $\frac{1}{6}(n^2 - 1)n$ multiplications, $\frac{1}{2}n(n - 1)$ divisions et n extractions de racines carrées, soit une complexité favorable par rapport à la factorisation LU de la même matrice. Si l'on souhaite résoudre un système linéaire $Ax = b$ associé, il faut ajouter $n(n - 1)$ additions et soustractions, $n(n - 1)$ multiplications et $2n$ divisions pour la résolution des systèmes triangulaires $By = b$ et $B^T x = y$.

Exemple d'application de la factorisation de Cholesky. Considérons la matrice réelle symétrique définie positive

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 10 \\ 3 & 10 & 26 \end{pmatrix}.$$

En appliquant de l'algorithme de factorisation de Cholesky, on obtient

$$A = BB^T = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 4 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{pmatrix}.$$

1.A Annexe du chapitre

Proposition 6 (« règle de Cramer ») Soit \mathbb{K} le corps des nombres réels \mathbb{R} ou des nombres complexes \mathbb{C} . On désigne par les vecteurs \mathbf{a}_j , $j = 1, \dots, n$, de $M_{n,1}(\mathbb{K})$ les colonnes d'une matrice inversible A de $M_n(\mathbb{K})$. Les composantes de la solution du système $Ax = b$, avec b un vecteur de $M_{n,1}(\mathbb{K})$, sont données par

$$\forall i \in \{1, \dots, n\}, x_i = \frac{\det(\mathbf{a}_1, \dots, \mathbf{a}_{i-1}, b, \mathbf{a}_{i+1}, \dots, \mathbf{a}_n)}{\det(A)}.$$

DÉMONSTRATION. Le déterminant étant une forme multilinéaire alternée, on a

$$\forall (i, j) \in \{1, \dots, n\}^2, i \neq j, \forall (\lambda, \mu) \in \mathbb{K}^2, \det(\mathbf{a}_1, \dots, \mathbf{a}_{i-1}, \lambda \mathbf{a}_i + \mu \mathbf{a}_j, \mathbf{a}_{i+1}, \dots, \mathbf{a}_n) = \lambda \det(A).$$

Or, si le vecteur x est solution de $Ax = b$, ses composantes sont les composantes du vecteur b dans la base de $M_{n,1}(\mathbb{K})$ formée par les colonnes de A , c'est-à-dire

$$b = \sum_{j=1}^n x_j \mathbf{a}_j.$$

On en déduit que

$$\forall i \in \{1, \dots, n\}, \det(\mathbf{a}_1, \dots, \mathbf{a}_{i-1}, \sum_{j=1}^n x_j \mathbf{a}_j, \mathbf{a}_{i+1}, \dots, \mathbf{a}_n) = \det(\mathbf{a}_1, \dots, \mathbf{a}_{i-1}, b, \mathbf{a}_{i+1}, \dots, \mathbf{a}_n) = x_i \det(A),$$

d'où la formule. □

Théorème 7 (« critère de Sylvester ») Soit n un entier naturel non nul. Une matrice réelle symétrique M d'ordre n est définie positive si et seulement si tous ses mineurs principaux dominants (ou primaires) sont strictement positifs, c'est-à-dire

$$\forall k \in \{1, \dots, n\}, \det(M_k) > 0,$$

où

$$M_k = \begin{pmatrix} m_{11} & \dots & m_{1k} \\ \vdots & & \vdots \\ m_{k1} & \dots & m_{kk} \end{pmatrix}.$$

DÉMONSTRATION. Montrons tout d'abord que la condition est nécessaire. Pour cela, introduisons une forme quadratique q sur un espace vectoriel réel E de dimension égale à n , dont la matrice symétrique M est la matrice relativement à une base donnée de E . Si M est définie positive, alors q également et l'on sait qu'il existe une base $\{\varepsilon_i\}_{i=1, \dots, n}$ relativement à laquelle la matrice de q est la matrice identité, c'est-à-dire qu'il existe une matrice inversible P telle que

$$I_n = P^T M P.$$

On a ainsi

$$1 = \det(P)^2 \det(M),$$

dont on déduit que $\det(M)$ est strictement positif. En considérant la restriction de la forme q au sous-espace vectoriel engendré $\text{Vect}(\{\varepsilon_1, \dots, \varepsilon_k\})$ et en procédant de manière similaire pour k allant de 1 à $n-1$, on montre que tout mineur principal dominant de M est strictement positif.

Pour montrer que la condition est suffisante, on raisonne par récurrence sur l'ordre de la matrice, c'est-à-dire sur l'entier naturel non nul n . Pour $n = 1$, c'est évident, puisque qu'on peut identifier M_1 à un réel strictement positif. Supposons le résultat vrai pour toute matrice d'ordre $n-1$, avec n un entier supérieur ou égal à 2, et considérons la matrice symétrique M_n , que l'on peut écrire par blocs sous la forme

$$M_n = \begin{pmatrix} M_{n-1} & m \\ m^\top & \alpha \end{pmatrix},$$

avec m appartenant à $M_{n-1,1}(\mathbb{R})$ et α appartenant à \mathbb{R} . La matrice M_{n-1} étant symétrique définie positive par hypothèse de récurrence, elle est inversible et ses colonnes forment une base de $M_{n-1,1}(\mathbb{R})$ et l'on peut par conséquent écrire m comme une (unique) combinaison linéaire de ces dernières, dont on note $\alpha_1, \dots, \alpha_{n-1}$ les coefficients. En considérant alors la matrice inversible d'ordre n

$$P = \begin{pmatrix} 1 & 0 & \dots & 0 & -\alpha_1 \\ 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots \\ \vdots & & \ddots & \ddots & -\alpha_{n-1} \\ 0 & \dots & \dots & 0 & 1 \end{pmatrix},$$

on vérifie que

$$P^\top M_n P = \begin{pmatrix} M_{n-1} & 0 \\ 0^\top & \alpha \end{pmatrix},$$

d'où $\det(P^\top M_n P) = (\det(P))^2 \det(M_n) = \det(M_{n-1})\alpha$. Puisque les mineurs $\det(M_n)$ et $\det(M_{n-1})$ sont strictement positifs, on en déduit que α l'est aussi. Le caractère défini positif de M_n étant équivalent à celui de la matrice $P^\top M_n P$, on peut alors conclure en écrivant que

$$\forall X \in M_{n,1}(\mathbb{R}), X^\top P^\top M_n P X = \begin{pmatrix} x_1 \\ \vdots \\ x_{n-1} \end{pmatrix}^\top M_{n-1} \begin{pmatrix} x_1 \\ \vdots \\ x_{n-1} \end{pmatrix} + (x_n)^2 \alpha,$$

et en utilisant que la matrice M_{n-1} est définie positive et le réel α est strictement positif. □