Chapitre I. Méthodes directes de résolution des systèmes linéaires

M. Granger

Table des matières

Dans ce chapitre, la section 1 est pour l'essentiel formée de rappels de première année d'Université Il faudra prêter attention aux notations de 1.1. qui seront utilisées dans tout le cours. En particulier on veillera à bien comprendre la remarque 4 qui montre comment voir le produit de deux matrices A et B colonne par colonne ou ligne par ligne.

1 Rappels d'algèbre linéaire élémentaire

1.1 Applications linéaires et matrices

On note E et F deux espaces vectoriels de dimensions finies sur le corps $\mathbb{K}$ avec $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. On note leurs dimensions respectives par : $dimE = n$ et $dimF = m$.
Soient $\mathcal{B} = (e_1, \dots, e_n)$ et $\mathcal{C} = (e'_1, \dots, e'_m)$ des bases respectives de E et F. Pour tout vecteur v (resp. w) de E (resp. F) on note ses coordonnées dans la base $\mathcal{B}$ (resp.$\mathcal{C}$) sous la forme d'une colonne :

$$V = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \text{ resp. } W = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}$$

ce qui signifie que :

$$v = \sum_{j=1}^{n} x_j e_j, \text{ resp. } w = \sum_{j=1}^{m} y_j e'_j$$

Lorsque $E = \mathbb{K}^n$ *et lorsque* $\mathcal{B}$ *est la base canonique*, on identifiera v et la colonne V, et on traitera dans les formules v comme un vecteur colonne.

Soit $\varphi : E \longrightarrow F$ une application linéaire. La matrice de φ dans la paire de bases $(\mathcal{B},\mathcal{C})$ est la matrice de taille $m \times n$:

$$Mat_{\mathcal{B},\mathcal{C}}(\varphi) = \begin{pmatrix} a_{1,1} & \dots & a_{1,n} \\ & \dots & \\ a_{m,1} & \dots & a_{m,n} \end{pmatrix}$$

déterminée par les égalités $\varphi(e_j) = \sum_{i=1}^m a_{i,j} e'_i$, autrement dit par le fait que la $j$ième colonne de cette matrice, $C_j = \begin{pmatrix} a_{1,j} \\ \vdots \\ a_{m,j} \end{pmatrix}$, est la colonne des coordonnées sur la base $\mathcal{C}$ du vecteur $c_j = \varphi(e_j)$.

On note $\mathcal{L}(E,F)$ l'espace vectoriel des applications linéaires $E \to F$, et pour $E = F$ on simplifie la notation : on note $\mathcal{L}(E) = \mathcal{L}(E,E)$, appelé espace vectoriel des endomorphismes de E. C' est un anneau non commutatif, avec pour multiplication la composition $\circ$ des endomorphismes.
On note $\mathcal{M}_{m,n}(K)$ l'espace vectoriel des matrices de taille $m \times n$ à coefficients dans $\mathbb{K}$.

Proposition 1.1 .- *L'application* $\mathcal{L}(E,F) \longrightarrow \mathcal{M}_{m,n}(\mathbb{K})$ *qui à* φ *associe* $Mat_{\mathcal{B},\mathcal{C}}(\varphi)$ *est un isomorphisme d'espaces vectoriels (qui dépend du choix des bases).*

Remarque 1.- Pour éclairer cet énoncé, explicitons à l'aide de $A = Mat_{\mathcal{B},\mathcal{C}}(\varphi)$, l'effet de φ sur les coordonnées des vecteurs dans les bases $\mathcal{B}$ et $\mathcal{C}$:

Soit $Mat_{\mathcal{B},\mathcal{C}}(\varphi) = A$, l'image de φ par l' isomorphisme de la proposition 1.1. Soient $v = \sum_{j=1}^n x_j e_j$ un vecteur de E, $w = \varphi(v) = \sum_{j=1}^m y_j e'_j$, son image par φ et V, W, les vecteurs colonnes des coordonnées x_i (resp. y_j) de v (resp.w) dans les bases respectives $\mathcal{B}$ et $\mathcal{C}$. Alors on trouve par un calcul direct les relations :

$$y_i = \sum_{j=1}^n a_{i,j} x_j \text{ pour } i = 1, \cdots, m.$$

Ceci peut être résumé par la formule $W = A \cdot V$, où v et W sont des matrices colonnes et $A \cdots V$ un cas particulier du produit matriciel dont nous rappelons plus bas la définition (Remarque 4).

Remarque 2.- L'isomorphisme de la proposition 1.1 dépend des bases $\mathcal{B}, \mathcal{C}$ utilisées. Cette dépendance est décrite par les formules de changement de bases dans les applications linéaires : si $\mathcal{B}', \mathcal{C}'$ sont deux autres bases de E et F respectivement, et si $P = P_{\mathcal{B}\to\mathcal{B}'}$, resp $Q = P_{\mathcal{C}\to\mathcal{C}'}$ sont les deux matrices de changements de bases [1], les matrices $A := Mat_{\mathcal{B},\mathcal{C}}(\varphi)$ et $A' := Mat_{\mathcal{B}',\mathcal{C}'}(\varphi)$ sont liées par la relation :

$$A' = Q^{-1}.A.P.$$

Remarque 3.- Lorsque $E = \mathbb{K}^n$ et $F = \mathbb{K}^m$, et lorsqu'on choisit pour $\mathcal{B}, \mathcal{C}$ les bases canoniques $\mathcal{C}a_n$ et $\mathcal{C}a_m$, la proposition 1.1 fournit un isomorphisme particulier et on se permettra donc d'identifier de cette façon $\mathcal{M}_{m,n}(\mathbb{K})$ à $\mathcal{L}(\mathbb{K}^n, \mathbb{K}^m)$.

En vertu de la remarque 1, cette identification fait correspondre à la matrice A, l'application linéaire $\varphi_A : \mathbb{K}^n \to \mathbb{K}^m$, définie par $\varphi_A(V) = A.V$.

Z.- Prendre garde au fait qu'on sera souvent amené à exprimer l'application linéaire φ_A, dans des bases autres que les bases canoniques. La matrice obtenue $Mat_{\mathcal{B},\mathcal{C}}(\varphi_A)$ *n'est pas* alors pas la matrice A.

Remarque 4.- Rappelons que le produit matriciel $(B, A) \to B.A$,

$$\mathcal{M}_{p,m}(\mathbb{K}) \times \mathcal{M}_{m,n}(\mathbb{K}) \longrightarrow \mathcal{M}_{p,n}(\mathbb{K})$$

[1] Rappelons que $P = P_{\mathcal{B}\to\mathcal{B}'}$, est la matrice dont les colonnes sont les coordonnées des vecteurs de $\mathcal{B}'$ dans la base $\mathcal{B}$

est défini par $(B.A)_{i,k} = \sum B_{i,j} A_{j,k}$. Par les isomorphismes de la proposition 1.1, le produit matriciel correspond à la composition des applications linéaires. Si $\psi : F \longrightarrow G$ est une deuxième application linéaire, composable avec φ, dans le sens indiqué par F, on a :

$$Mat_{\mathcal{B},\mathcal{D}}(\psi \circ \varphi) = Mat_{\mathcal{C},\mathcal{D}}(\psi) \cdot Mat_{\mathcal{B},\mathcal{C}}(\varphi)$$

En particulier : $\varphi_{B.A} = \varphi_B \circ \varphi_A$.

N.B.- On a utilisé dans cette définition la convention, très commode, consistant à désigner par $A_{i,j}$ le coefficient situé à la place (i,j) d'une matrice A.

Remarque 5.- La présentation qui suit du produit matriciel en termes de lignes ou de colonnes est indispensable à une bonne compréhension des algorithmes de Gauss étudiés au chapitre suivant.

On peut écrire une matrice A, selon les besoins, comme une liste de lignes L_i de longueur n, ou de colonnes C_j de longueur m respectivement, ce qui donne deux présentations possibles d'une matrice de taille $m \times n$:

$$A = \begin{bmatrix} L_1 \\ \vdots \\ L_m \end{bmatrix} = \begin{bmatrix} C_1, & \dots & , C_n \end{bmatrix}$$

Voir l'exercice 0) pour lever tout malentendu sur cette présentation par lignes ou par colonnes et apprendre à l'utiliser.

En sus du fait que le calcul pratique d'un produit de matrice est largement facilité par ce point de vue, il importe de rappeler ce que ces notions *signifient du point de vue de l'algèbre linéaire* : Les colonnes de A engendrent l'image de l'application linéaire φ_A. Dans le *système* d'équations linéaires ([2]) $Au = b$, de second membre $b \in \mathbb{K}^m$ et d'inconnue $u \in \mathbb{K}^n$, chaque ligne de A est la liste des coefficients du premier membre d'une équation linéaire de ce système.

- Le produit d'une ligne $L = (l_1, \dots, l_m)$ par une colonne (de même longueur !) : $\Gamma = \begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_m \end{pmatrix} \in \mathcal{M}_{m,1}(\mathbb{K})$ est la matrice scalaire de taille $(1,1)$: $(c) = \sum l_i \gamma_i$
- Le produit $L \cdot A$ de A, à gauche par la ligne $L = (l_1, \dots, l_m) \in \mathcal{M}_{1,m}(\mathbb{K})$ est la matrice ligne de $\mathcal{M}_{1,n}(\mathbb{K})$:
$$L \cdot A = l_1 L_1 + \ldots + l_m L_m$$
De même le produit $A \cdot C$ de A par la colonne $C = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} \in \mathcal{M}_{n,1}(\mathbb{K})$ est la matrice colonne de $\mathcal{M}_{m,1}(\mathbb{K})$ suivante :
$$A \cdot C = c_1 C_1 + \ldots + c_n C_n$$
Ces deux observations sont deux cas particuliers des produits par blocs, qu'on étudiera plus loin en toute généralité.
- Plus généralement, le produit $B.A$ s'exprime alors de la façon commode suivante : On note, pour préciser la dépendance par rapport à A, $L_i(A)$, $1 \le i \le m$ et $C_j(A)$, $1 \le i \le n$, les rangées de A et de même B est une matrice de taille (p,m), multipliable à droite par A, écrite en ligne ou en colonne :
$$B = \begin{bmatrix} L_1(B) \\ \vdots \\ L_p(B) \end{bmatrix} = \begin{bmatrix} C_1(B), & \dots & , C_m(B) \end{bmatrix}$$
Alors on peut donner du produit matriciel trois expressions, en lignes, en colonnes, ou enfin par le tableau de taille $p \times n$ des produits des lignes de B par les colonnes de A :

[2] objet des chapitres 2 à 4 de ce cours !

1.
$$B \cdot A = B \cdot \begin{bmatrix} L_1(A) \\ \vdots \\ L_m(A) \end{bmatrix} = \begin{bmatrix} L'_1 \\ \vdots \\ L'_p \end{bmatrix}$$

avec $L'_i = b_{i,1}L_1(A) + \ldots + b_{i,m}L_m(A) = L_i(B) \cdot A$, produit de A à gauche par la $i$ème ligne de B.

2.
$$B \cdot A = \begin{bmatrix} C_1(B), & \ldots, & C_m(B) \end{bmatrix} \cdot A = \begin{bmatrix} C'_1 & \ldots & C'_n \end{bmatrix}$$

avec $C'_j = a_{1,j}C_1(B) + \ldots + a_{m,j}C_m(B) = B \cdot C_j(A)$, produit de B à droite par la $j$ème colonne de A.

3.
$$B \cdot A = \begin{pmatrix} L_1(B) \cdot C_1(A) & \cdots & L_1(B) \cdot C_n(A) \\ \vdots & L_i(B) \cdot C_j(A) & \cdots \\ L_p(B) \cdot C_1(A) & \cdots & L_p(B) \cdot C_n(A) \end{pmatrix}$$

Si $\mathbb{K} = \mathbb{R}$, $B.A$ est le tableau de produits scalaires dans $\mathbb{R}^m$, $L_i(A) \cdot C_j(B) = < L_i(A), C_j(B) >$.

1.2 Rappels sur les déterminants

Soit A une matrice carrée. Son déterminant est :

$$detA = \sum_{\sigma \in \mathcal{S}_n} (-1)^{\epsilon(\sigma)} (\Pi_{i=1}^n a_{i,\sigma(i)}) \tag{1}$$

On peut aussi voir $detA = det(V_1, \ldots, V_n)$, comme l'unique forme multilinéaire alternée sur les vecteurs colonnes $V_1, \ldots, V_n$ de A qui prend la valeur $det(e_1, \ldots, e_n) = 1$ sur la base canonique de $\mathbb{K}^n$.

On suppose connues -voir dans un manuel de première année- les propriétés élémentaires du déterminant, notamment déterminant d'un produit, calcul pratique par modification de lignes ou de colonnes etc.

Dans l'exercice 1) on évalue le temps nécessaire pour calculer un déterminant 20×20 puis 100×100, avec un ordinateur effectuant 10^8 opérations par seconde.

Le résultat est dissuasif, s'exprimant déja en milliers d'années pour le cas 20×20 et ce n'est surement pas en utilisant des formules fondées sur les déterminants qu'on résoud les systèmes linéaires dans la pratique. Les méthodes de résolution des systèmes linéaires constituent l'objectif du chapitre II, ainsi que du chapitre IV (méthodes itératives).

1.3 Théorème fondamental

Théorème 1.2 *Soit $f : E \longrightarrow F$ une application linéaire entre espaces de dimensions finies. Alors :*

$$dim(Im(f)) + dim(ker(f)) = dim(E)$$

Si E et F sont de la même dimension (par exemple $E = F = \mathbb{K}^n$), cela donne l'équivalence :

$$f \text{ est injective } \iff f \text{ est surjective }.$$

Autre formulation possible pour $f : E \to F$ linéaire entre deux espaces de dimensions n, m, à priori quelconques.

$$f \text{ est bijective } \Leftrightarrow f \text{ est injective et } m = n \Leftrightarrow f \text{ est surjective et } m = n.$$

Rappelons pour cela que, pour une application *linéaire* f, l'injectivité est équivalente à la condition $ker(f) = 0$, qui équivaut aussi à :

Le système linéaire $A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = (0)$, n'a que la solution nulle.

Signalons pour se préparer au point de vue du chapitre II que la surjectivité s'exprime aussi en terme de systèmes linéaires par :

"Quel que soit le second membre $b \in \mathbb{K}^n$, le système linéaire $A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = (b)$ a au moins une solution.

2 Généralités sur les systèmes linéaire de Cramer.

On considère une matrice inversible A et on fixe un vecteur $b \in \mathbb{K}^n$, $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ ("le second membre"). L'objectif est d'étudier des méthodes de résolution du système linéaire :

$$Au = b$$

dont l'inconnue est un vecteur $u = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{K}^n$. On sait que sous l'hypothèse d'inversibilité un tel système de n équations à n inconnues a une solution unique.
L'objectif est d'étudier des méthodes de calcul approché de cette solution, qui en même temps contrôlent que A est bien inversible, et qui soient praticables pour des systèmes de grande taille (p.ex. au moins $n \geq 100$), et aussi de déterminer en fonction du système proposé la fiabilité de la solution obtenue, du point de vue de la sensibilité aux petites perturbations des coefficients, et aux erreurs d'arrondis.

On peut se rappeler qu'il existe une expression très élégante de la solution avec des quotients de déterminants :

$$u_i = \frac{det(C_1, \cdots, C_{i-1}, u, C_{i+1}, \cdots, C_n)}{det(A)}$$

Inutile d'insister car ces formules sont impraticables dès que n est un peu grand, disons ≥ 10. Essayez déjà d'écrire un déterminant 5×5. Voir aussi l'exercice 1 du chapitre I

Remarque 1. Il ne s'agit pas non plus dans la pratique de calculer A^{-1} pour ensuite "faire $u = A^{-1}.b$"
En effet le calcul de $A^{-1} = [u_1, \ldots, u_n]$, décrit par ses vecteurs colonnes revient déjà à résoudre n systèmes linéaires : $A.u_i = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} := e_i$, ce qui est bien sur trop coûteux pour résoudre un seul système.

Remarque 2. Un système triangulaire $(\mathcal{T})$ du type ci-dessous est aisée à résoudre, par la méthode dite de remontée.

$$(\mathcal{T}) \quad \left\{ \begin{array}{cccccc} a_{1,1}x_1+ & \ldots & & \ldots & +a_{1,n}x_n & = & b_1 \\ & & \ldots & \ldots & \ldots & \ldots & \ldots \\ & & & \ldots & & & \ldots \\ & & a_{i,i}x_i+ & \ldots & +a_{i,n}x_n & = & b_i \\ & & & \ldots & & & \\ & & & & \ldots & & \\ & & & & a_{n,n}x_n & = & b_n \end{array} \right.$$

La méthode de remontée consiste à déterminer successivement :

$$x_n = \frac{b_n}{a_{n,n}}$$

puis chaque x_i pour i variant de $n-1$ à 1, par ordre décroissant des indices, en fonction des termes déjà calculés selon la formule :

$$x_i = \frac{1}{a_{i,i}}[b_i - a_{i,i+1}x_{i+1} - \ldots - a_{i,n}x_n]$$

Remarquer que le fait que la matrice soit inversible équivaut ici à : $\forall i, a_{i,i} \neq 0$.

On verra dans l'exercice 1) de ce chapitre que le nombre d'opérations nécessaires pour résoudre un tel système est de l'ordre de n^2. Le coût est faible par rapport au coût de l'algorithme à mettre en oeuvre pour réduire un système quelconque à cette forme, algorithme du pivot de Gauss qui fait l'objet du prochain paragraphe.

3 La méthode de Gauss

3.1 Description de l'algorithme

La méthode consiste à construire une matrice inversible M telle que $MA = U$ soit une matrice triangulaire supérieure.
Le système initial $Au = b$ est alors équivalent parr multiplication à gauche par M inversible au système triangulaire $MAu = Mb$ c'est à dire au système :

$$U.u = b',$$

de second membre $b' = Mb$. Il suffit donc d'appliquer à ce dernier système la méthode de remontée.
En fait, on ne calcule pas directement la matrice M, mais on décrit une succession d'opérations dites opérations élémentaires, chacune d'elle s'interprétant comme la multiplication *à gauche* par une matrice inversible M_i, à la fois de A et du second membre b, le résultat de cette succession d'opérations étant $MA = U$ et $Mb = b'$ avec $M = M_q \ldots M_1$. Voici la succession de ces étapes :

Etape 1

Sous-étape 1.1. *Puisque A est inversible sa première colonne est non nulle : On peut donc choisir i tel que $a_{i,1} \neq 0$. On appelle $a_{i,1}$ le pivot pour la première étape. Si $i \neq 1$ on échange les lignes 1 et i de A* ***ainsi que*** *les seconds membres b_1 et b_i. Sinon on ne change rien.*

Ecrivons A comme une succession de n lignes de longueur n

$$A = \begin{bmatrix} L_1 \\ \vdots \\ L_n \end{bmatrix}$$

avec $L_i = [a_{i,1}, \dots, a_{i,n}]$. On rappelle que selon le chapitre I, $[\lambda_1, \dots, \lambda_n].A$ est la matrice ligne $\lambda_1.L_1 + \dots + \lambda_n.L_n$.

Par conséquent, si $i \neq 1$ l'échange de lignes indiqué n'est autre que la transformation $A \longrightarrow T_{1,i}.A$, $b \longrightarrow T_{1,i}.b$, où $T_{1,i}$ est la matrice de l'application linéaire associée à la transposition des vecteurs de base d'indices 1 et i (voir aussi l'exercice 2) :

$$T_{1,i} = \begin{array}{c} \begin{matrix} & & & & i & & & \end{matrix} \\ \begin{bmatrix} 0 & & \cdots & 0 & 1 & 0 & \cdots & 0 \\ & 1 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ & \vdots & \ddots & \vdots & \vdots & & & \vdots \\ 0 & 0 & \cdots & 1 & & & & \\ 1 & 0 & \cdots & & 0 & & \cdots & 0 \\ 0 & & \cdots & & 0 & 1 & 0 & 0 \\ \vdots & & & & & & \ddots & \\ 0 & & \cdots & & 0 & 0 & \cdots & 1 \end{bmatrix} \end{array}$$

Par convention $T_{1,1} = I$. C'est le cas $i = 1$ où le pivot est d'emblée "à la bonne place". Et la multiplication à gauche par $T_{1,1}$ consiste comme souhaité à ne rien changer.

<u>Sous-étape 1.2.</u>-

On effectue sur la matrice $A' = T_{i,1}.A = \left(a'_{i,j}\right) = \begin{bmatrix} L'_1 \\ \vdots \\ L'_n \end{bmatrix}$, *qu'on vient d'obtenir, des combinaisons linéaires de lignes du type*

$$A_2 = \begin{bmatrix} L'_1 \\ L'_2 - \gamma_2.L'_1 \\ \vdots \\ L'_n - \gamma_n.L'_1 \end{bmatrix},$$

de façon que la première colonne devienne : $\begin{bmatrix} a'_{1,1} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$.

On change simultanément le second membre en $b^{(2)} = \begin{bmatrix} b_1^{(2)} \\ b_2^{(2)} \\ \vdots \\ b_n^{(2)} \end{bmatrix} = \begin{bmatrix} b'_1 \\ b'_2 - \gamma_2.b'_1 \\ \vdots \\ b'_n - \gamma_n.b'_1 \end{bmatrix}$.

Précisément, le choix nécessaire et suffisant des γ_k pour que la matrice A_2 soit de la forme voulue

$$(1) \qquad A_2 = \begin{bmatrix} a_{1,1} & \alpha_2, & \cdots, & \alpha_n \\ 0 & & & \\ \vdots & & \widetilde{A_2} & \\ 0 & & & \end{bmatrix},$$

avec $\widetilde{A_2}$, une matrice $(n-1) \times (n-1)$, est $\gamma_k = \frac{a'_{k,1}}{a'_{1,1}}$.

Pour bien voir la valeur de γ_k, on peut écrire sous forme développée la combinaison de ligne $L'_k - \gamma_k.L'_1$

$$\begin{array}{ccccc} a'_{1,1}, & a'_{1,2}, & \cdots, & a'_{1,n} & | & \times(-\gamma_k) \\ a'_{k,1}, & a'_{k,2}, & \cdots, & a'_{k,n} & | & \times 1 \end{array}$$

dont il s'agit d'annuler le premier coefficient ce qui donne l'équation $a'_{1,1} \times (-\gamma_k) + a'_{k,1} = 0$, et donc l'unique valeur de γ_k annoncée.

Pour l'harmonie des notations on pose $A = A_1$, $b = b^{(1)}$.

La transformation de la sous-étape 1.2.

$$A' = \begin{bmatrix} L'_1 \\ \vdots \\ L'_n \end{bmatrix} \longrightarrow \begin{bmatrix} L'_1 \\ L'_2 - \gamma_2.L'_1 \\ \vdots \\ L'_n - \gamma_n.L'_1 \end{bmatrix}, \quad b' \longrightarrow b^{(2)},$$

n'est autre que la multiplication $A' \longrightarrow E.A'$ $b' \longrightarrow E.b'$ par la matrice

$$E = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ -\gamma_2 & 1 & 0 & \cdots & 0 \\ -\gamma_3 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \ddots & \\ -\gamma_n & 0 & 0 & \cdots & 1 \end{bmatrix}$$

comme on peut le voir facilement par la technique de multiplication des matrices vue dans le chapitre 1, remarque 5 de la première section. Ainsi :

$$\boxed{E.T_{i,1}.A = A_2, \quad E.T_{i,1}.b = b^{(2)} \quad \text{et } Au = b \Leftrightarrow A_2 u = b^{(2)}}.$$

Etapes 2 à $n-1$.

La poursuite de l'algorithme du pivot consiste alors à répéter le même processus à partir des lignes $2, \cdots, n-1$, en plaçant par un échange éventuel de lignes un pivot non nul aux positions successives $(2,2), \cdots, (n-1, n-1)$, de façon à rendre la matrice transformée triangulaire jusqu'à la colonne k incluse à l'issue de la k ième étape.

Cela revient par récurrence à appliquer, sur la matrice $\widetilde{A_2}$ de (1), l'algorithme de Gauss pour les matrices de taille $(n-1) \times (n-1)$ immédiatement inférieure à celle de A.

Pour être totalement explicite nous décrivons le passage de A_k, résultat de la $(k-1)$ième étape à A_{k+1}, résultat de la $k$ième étape, et la tranformation du second membre $b^{(k)}$ de la $(k-1)$ième étape en $b^{(k+1)}$. La matrice A_k, triangulaire jusqu'à la $(k-1)$ième colonne s'écrit donc :

$$A_k = \begin{bmatrix} a^{(1)}_{1,1} & & \cdots & & & a^{(1)}_{1,n} \\ 0 & a^{(2)}_{2,2} & & \cdots & & a^{(2)}_{2,n} \\ & & \ddots & & & \\ 0 & & \cdots & a^{(k-1)}_{k-1,k-1} & \cdots & a^{(k-1)}_{k-1,n} \\ 0 & & \cdots & 0 & & \\ \vdots & & & \vdots & \widetilde{A_k} & \\ 0 & & \cdots & 0 & & \end{bmatrix},$$

Les coefficients de $\widetilde{A_k}$ seront notés $a^{(k-1)}_{i,j}$, pour $k \leq i, j \leq n$.

Si $k \leq n-1$ on passe alors de A_k à A_{k+1} par une transformation du type : $A_{k+1} = E_k T_k A_k$. La matrice T_k, est soit la matrice unité soit une matrice de transposition $T_k = T_{k,i_k}$, avec $i_k > k$, s'il y a lieu d'échanger les lignes d'indices k et i_k, pour placer un coefficient non nul $a^{(k-1)}_{k,i_k}$ en position de pivot à la place (k,k). La matrice E_k est triangulaire inférieure de la forme :

$$E_k = \begin{bmatrix} 1 & \dots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & & \vdots \\ 0 & \dots & 1 & 0 & 0 & \cdots & 0 \\ 0 & \dots & 0 & 1 & 0 & \cdots & 0 \\ 0 & \dots & 0 & -\gamma_{k+1,k} & 1 & \cdots & 0 \\ \vdots & \dots & & \vdots & \vdots & \ddots & \\ 0 & \dots & 0 & -\gamma_{n,k} & 0 & \cdots & 1 \end{bmatrix}.$$

On appelle *transvection* toute matrice de ce type, avec une diagonale unité. Le second membre suit le mouvement et est changé en $b^{(k+1)} = E_k T_k b^{(k)}$.

Remarquons que chacune de ces transformations élémentaires conserve le déterminant au signe près puisque $\det(E_k) = 1$ et $\det(T_k) = -1$, sauf si $T_k = I$, donc $\det(A_{k+1}) = \det(T_k.A_k) = \pm\det(A_k)$. Par récurrence, A_k est inversible, donc $\widetilde{A_k}$ l'est aussi, ce qui permet de justifier, de la même façon que dans la sous étape 1.1., le fait que l'algorithme peut se poursuivre :

Pour être encore plus détaillé : la $k$ième opération consiste à traiter $\widetilde{A_k}$ exactement comme $A = A_1$ dans la première étape, en ne modifiant dans A_k que les $n-k$ dernières lignes et colonnes : cela revient à choisir un pivot non nul $a_{k,i_k}^{(k-1)}$ dans la première colonne de $\widetilde{A_k}$, à échanger les lignes k et i_k de $\widetilde{A_k}$, (et de A_k) puis à transformer $\widetilde{A_k}$ à l'aide du pivot.

Le processus complet se résume donc par l'égalité suivante où A_n est une matrice triangulaire.

$$\boxed{A_n = E_{n-1}.T_{n-1}\dots E_1.T_1.A\ , \quad b^{(n)} = E_{n-1}.T_{n-1}\dots E_1.T_1.b}.$$

3.2 Stratégie de choix du pivot

Que se passe il si on adopte la stratégie "bête" consistant à choisir le premier coefficient rencontré dans la colonne considérée, par exemple $a_{1,1}$ si $a_{1,1} \neq 0$? Cela peut donner lieu à de grandes erreurs d'arrondi, comme le montre l'exemple suivant :

$$(\mathcal{S}) \qquad \left\{ \begin{array}{cccc} 0,0001x & + & y & = 1 \\ x & + & y & = 2 \end{array} \right.$$

On suppose que la machine fonctionne avec 3 chiffres significatifs en virgules flottantes (les vraies machines travaillent au moins à 16 décimales). Par exemple 0.0001 , $\pi = 3.14159\dots$, $1/3 = 0.33333\dots$ sont convertis respectivement en :

$$1.00 \times 10^{-4}\ ,\ \ 3.14 \times 10^{0}\ ,\ \ 3.33 \times 10^{-1}$$

valeurs approchées sauf la première qui se trouve être exacte.

Remarques.- Une autre convention souvent utilisée est l'écriture commençant par 0. comme dans $\pi \sim 0.314 \times 10^{+1}$. Dans 1.00 il faut remarquer que les deux 0 sont significatifs et que 1.00 est l'écriture approchée de tout nombre strictement compris entre 0.995 et 1.005. Rappelons que les calculateurs utilisent le signe . plutot que la virgule de la tradition française pour séparer les entiers des décimales.

Dans l'exemple $(\mathcal{S})$ et selon "la stratégie bête" on prend pour pivot $a_{1,1} = 1.00 \times 10^{-4}$. On transforme donc la deuxième ligne par : $L_2 - 10^{+4} L_1$. Le calcul exact est

$$(\mathcal{S}_1) \qquad \left\{ \begin{array}{cccc} 0.0001x & + & y & = 1 \\ & & 9999y & = 9998 \end{array} \right.$$

mais la machine prend la version arrondie :

$$(\mathcal{S}_1') \quad \left\{ \begin{array}{cccc} 10^{-4}x & + & y & = 1 \\ & & 1.00 \times 10^{4} y & = 1.00 \times 10^{4} \end{array} \right.$$

et la résolution du système triangulaire obtenu est alors :

$$y = 1 \quad \text{puis} \quad 0.0001x = 1 - y = 0$$

donc la solution trouvée est

$$x = 0\,, \quad y = 1$$

ce qui est très éloigné de la solution exacte tirée de $(\mathcal{S}_1)$:

$$x = 1/0.9999 = 1.00010001... \sim 1.00.$$

Explication : Lorsqu'on calcule x dans la première équation $0.0001x = 1 - y = 0$, on divise par un petit pivot 0.0001, et l'erreur sur y dans le calcul approché de la deuxième équation se trouve multipliée par 10000 : l'erreur d'arrondi qui peut aller initialement jusqu'à $0.5 \times 10{-}3$ devient du même ordre de grandeur que les solutions.

Si au contraire, on échange les lignes et on choisit pour pivot $a_{2,1} = 1$, la résolution donne le système triangulaire : $\left\{\begin{array}{llll} x & + & y & = 2 \\ & & 0.9999y & = 0.9998 \sim 0.999 \end{array}\right.$

D'où $y = 1$ et cette fois $x = 2 - y = 1$, ce qui est correct à 10^{-3} près.

Conclusion : *Pivot total, pivot partiel.* Même si à l'issue de la $(k-1)$ième étape on trouve $a_{k,k}^{(k-1)} \neq 0$, on vient de voir qu'il est préférable de choisir un autre pivot.

La stratégie du *pivot partiel* consiste à choisir $a_{i,k}^{(k-1)}$, avec $i \geq k$ *de module maximum* dans la $k$ième colonne de $\widetilde{A_k}$.

La méthode du *pivot total*, consiste à utiliser le plus grand des coefficients $a_{i,j}^{(k-1)}$, parmi toutes les paires d'indices (i, j), ce qui mène à des échanges de lignes et de colonnes. La méthode est plus précise mais plus coûteuse en temps. De plus, il faut prendre garde au fait que la solution du système transformé n'est pas u, mais un vecteur obtenu en permutant les coordonnées de u, de la même façon qu'on a permuté les colonnes.

Notre exemple de mauvais choix de pivot $a_{1,1}$ décrit un cas particulièrement défavorable où le rapport $a_{2,1}/a_{1,1}$ est le plus élevé possible. Pour une matrice assez générale, le choix bête de pivot à chaque étape est souvent numériquement acceptable et donne alors lieu à une factorisation de la matrice A en produit de matrices triangulaires qui est l'objet de la section 4 suivante.

3.3 Calcul du nombre d'opérations.

Proposition 3.1 .-

La mise en oeuvre de la méthode du pivot nécessite $\frac{n^3-n}{3}$ *additions-soustractions, autant de multiplications et* $\frac{n(n-1)}{2}$ *divisions, soit un total d'opérations de l'ordre de* $\frac{2n^3}{3}$

La vérification fera l'objet de l'exercice 3). On comparera le résultat obtenu à la "méthode" de Cramer consistant à calculer des déterminants.

4 La factorisation LU d'une matrice carrée.

4.1

Supposons qu'on puisse trouver deux matrices L triangulaire inférieure (L pour lower), et U triangulaire supérieure (U pour upper) telles que

$$A = LU.$$

Alors résoudre le système $Ax = b$, revient à résoudre successivement les deux systèmes $Ly = b$ puis $Ux = y$, le vecteur intermédiaire y s'obtenant par une méthode de descente, analogue à celle de remontée mais où cette fois on calcule successivement $x_1, x_2, \ldots, x_n$ dans l'ordre croissant des indices, puisque la matrice L est triangulaire *inférieure.*

Le nombre d'opérations de l'ordre de $2n^2$. En ajoutant la mise sous forme LU, dont on va voir qu'elle revient à un pivot de Gauss, avec un nombre d'opérations de l'ordre de $\frac{2n^3}{3}$ cela donne pour résoudre simultanément p systèmes $Ax = b_j$, avec des seconds membres $b_1, \cdots, b_p$, un nombre d'opérations de l'ordre de $2n^2.p + \frac{2n^3}{3}$.

Il n'existe pas toujours une décomposition LU comme le montre l'exemple suivant :

Exemple : La matrice suivante : $\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$ n'admet pas de décomposition LU.

En effet une décomposition LU s'écrirait $\left(\begin{smallmatrix} 0 & 1 \\ 1 & 1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} a & 0 \\ b & c \end{smallmatrix}\right)\left(\begin{smallmatrix} d & e \\ 0 & f \end{smallmatrix}\right)$. Ceci impliquerait $ad = 0,\ ae = 1,\ bd = 1$ qui est impossible.

Nous nous plaçons maintenant dans le cas particulier où ce que nous avons appelé la stratégie bête fonctionne jusqu'au bout : si à chaque étape de l'algorithme, on constate que $a_{k,k}^{(k-1)} \neq 0$ on choisit comme pivot, ce premier coefficient "en haut à gauche" $a_{k,k}^{(k-1)}$ de $\widetilde{A_k}$. C'est la stratégie qui consiste à éviter toute permutation de lignes.

Dans les notations de l'algorithme de Gauss que nous avons utilisées dans la section 3, on a alors $T_k = I$, et les matrices intermédiaires A_k sont donc telles que $A_{k+1} = E_k T_k A_k = E_k A_k$, ce qui donne :

$$A_{k+1} = E_k T_k A_k = E_k \cdots E_1 A_1.$$

On aboutit à la formule $MA = A_n = U$, triangulaire supérieure avec $M = E_{n-1}. \cdots E_1$, triangulaire inférieure .

Le fait que chaque E_i est triangulaire inférieure avec une diagonale unité, entraîne qu'il en est de même de chaque E_i^{-1}, puis de $L = M^{-1} = E_1^{-1} \cdots E_{n-1}^{-1}$. C'est l'objet de l'exercice 4).

On a donc obtenu une décomposition $A = LU$ avec :

$$L = M^{-1} = E_1^{-1} \cdots E_{n-1}^{-1}$$

matrice triangulaire à diagonale unité. Une telle matrice est aussi appelée matrice unipotente. Ceci démontre la première partie (existence) de la proposition 4.1 ci dessous.

La justification du fait que la prise de l'inverse et le produit préservent la structure triangulaire inférieure, que ce soit avec ou sans la condition de diagonale unité fait l'objet d'un exercice, comme déjà indiqué. Le deuxième question de ce même exercice 4) contient aussi la démonstration de la remarque suivante.

Remarque : L'objet de cette remarque est de montrer que le calcul de $L = M^{-1}$ ne coûte RIEN en temps de calcul, sinon de recopier des données. On vérifie en effet les faits suivants :

$$E_k = \begin{bmatrix} 1 & \ldots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & & \vdots \\ 0 & \ldots & 1 & 0 & 0 & \cdots & 0 \\ 0 & \ldots & 0 & 1 & 0 & \cdots & 0 \\ 0 & \ldots & 0 & -\gamma_{k+1,k} & 1 & \cdots & 0 \\ \vdots & \ldots & & \vdots & \vdots & \ddots & \\ 0 & \ldots & 0 & -\gamma_{n,k} & 0 & \cdots & 1 \end{bmatrix}, \text{ a pour inverse } \quad E_k^{-1} = \begin{bmatrix} 1 & \ldots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & & \vdots \\ 0 & \ldots & 1 & 0 & 0 & \cdots & 0 \\ 0 & \ldots & 0 & 1 & 0 & \cdots & 0 \\ 0 & \ldots & 0 & +\gamma_{k+1,k} & 1 & \cdots & 0 \\ \vdots & \ldots & & \vdots & \vdots & \ddots & \\ 0 & \ldots & 0 & +\gamma_{n,k} & 0 & \cdots & 1 \end{bmatrix}.$$

puis

$$L = E_1^{-1} \cdots E_{n-1}^{-1} = \begin{bmatrix} 1 & \cdots & \cdots & 0 & \cdots & 0 & 0 \\ +\gamma_{2,1} & \ddots & & 0 & \cdots & 0 & 0 \\ \vdots & \ddots & \ddots & \vdots & & \vdots & \vdots \\ \vdots & \ddots & \ddots & 1 & & 0 & 0 \\ +\gamma_{k+1,1} & \cdots & \cdots & +\gamma_{k+1,k} & 1 & \cdots & 0 \\ \vdots & & & \vdots & & \ddots & \vdots \\ +\gamma_{n,1} & \cdots & \cdots & +\gamma_{n,k} & \cdots & +\gamma_{n,n-1} & 1 \end{bmatrix}.$$

On remarquera l'importance de l'ordre des multiplications de matrices : le calcul de L s'effectue par simple recopiage, alors que celui de $E_{n-1}^{-1} \cdots E_1^{-1}$ serait bien plus compliqué.

Faire l'essai sur des matrices 3×3 en comparant :

$$\begin{bmatrix} 1 & 0 & 0 \\ a & 1 & 0 \\ b & 0 & 1 \end{bmatrix} \times \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & c & 1 \end{bmatrix} \quad \text{et} \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & c & 1 \end{bmatrix} \times \begin{bmatrix} 1 & 0 & 0 \\ a & 1 & 0 \\ b & 0 & 1 \end{bmatrix}$$

Proposition 4.1 .- *Si le choix de pivot bête, sans permutation de lignes, est possible à chaque étape, la matrice A admet une décomposition LU, avec L unipotente (i.e.à diagonale unitaire). De plus cette décomposition est unique.*

Démonstration.- L'existence vient d'être établie. Voyons l'unicité :

Supposons que $A = LU = L'U'$, avec L, L' (resp U, U') triangulaire inférieure unipotente (resp supérieure.) Alors :

$$L^{-1}L' = UU'^{-1}$$

est une matrice à la fois triangulaire inférieure et supérieure donc diagonale, et c'est la matrice unité I, puisque selon l'écriture $L^{-1}L'$, elle est à diagonale unité.

Ainsi $L^{-1}L' = I$, donc $L = L'$ et de même $U = U'$. □

Voici une condition nécessaire et suffisante pour l'existence d'une factorisation LU :

Théorème 4.2 .- *Soit $A = (a_{i,j})$ une matrice carrée. On considère les déterminants $\Delta_k = det(M_k)$ des sous-matrices dites principales :*

$$M_k = \begin{bmatrix} a_{1,1} & \cdots & a_{1,k} \\ \vdots & \cdots & \vdots \\ a_{k,1} & \cdots & a_{k,k} \end{bmatrix}$$

(Ce sont les "coins supérieurs gauche" de toutes les tailles possibles). Alors A admet une factorisation LU, si et seulement si pour tout $k \in \{1, \cdots, n\}$, on a $\Delta_k \neq 0$, et cette condition équivaut aussi au fait que l'algorithme de Gauss fonctionne sans permutation de lignes ("l'algorithme bête fonctionne").

Démonstration.- On se place à l'issue de l'étape $k-1$ de l'algorithme de la section 3 en supposant que le pivot "bête" a pu être choisi, à chaque étape. Ainsi $A_k = E.A$ avec $E = E_{k-1} \cdots E_1$. Appliquons la mutiplication par blocs de tailles k et $n-k$. Comme E et A sont de la forme :

$$E = \begin{bmatrix} \mathcal{T}_k & \mathbf{0} \\ E_{II,I} & I \end{bmatrix} \text{ et } A = \begin{bmatrix} M_k & A_{I,II} \\ A_{II,I} & A_{II,II} \end{bmatrix}$$

avec $\mathcal{T}_k$ triangulaire de taille k, à diagonale unité, cela donne, pour le premier bloc en haut à gauche de A_k :

$$\mathcal{T}_k.M_k + 0 \times A_{II,I} = \mathcal{T}_k.M_k = \begin{bmatrix} a_{1,1}^{(1)} & \cdots & \cdots & \cdots & a_{1,k}^{(1)} \\ 0 & \ddots & \cdots & \cdots & \cdots \\ & \mathbf{0} & \ddots & & \cdots \\ 0 & \cdots & 0 & a_{k-1,k-1}^{(k-1)} & a_{k-1,k}^{(k-1)} \\ 0 & \cdots & \cdots & 0 & a_{k,k}^{(k-1)} \end{bmatrix}$$

On en déduit, puisque $\det(\mathcal{T}_k) = 1$, que $\Delta_k = \det(M_k) = a_{1,1}^{(1)} \cdots a_{k-1,k-1}^{(k-1)} a_{k,k}^{(k-1)}$.

On a aussi pour la même raison $\Delta_{k-1} = \det(M_{k-1}) = a_{1,1}^{(1)} \cdots a_{k-1,k-1}^{(k-1)}$. Donc sachant que $\Delta_{k-1} \neq 0$, on trouve que $\Delta_k \neq 0$ implique la non nullité du pivot $a_{k,k}^{(k-1)}$ rencontré dans la k ième étape.

Par récurrence sur k on a donc obtenu l'équivalence annoncée entre l'ensemble des conditions $\Delta_k \neq 0$, et des conditions $a_{k,k}^{(k-1)} \neq 0$ de non nulité des pivots bêtes.
On a déja vu que ces dernières conditions impliquent l'existence d'une factorisation LU.

Réciproquement le même argument de multiplication par blocs montre que l'égalité $A = LU$, implique la non nullité de Δ_k, si L et U sont triangulaires inversibles.

□

N.B. Observer que $a_{k,k}^{(k-1)} = a_{k,k}^{(k)}$ sous les conditions de ce théorème.

4.2 Deux exemples importants

1) Les matrices tridiagonales.

Ce sont les matrices carrées telles $a_{i,j} = 0$, si $|i - j| \geq 2$, c'est à dire de la forme

$$A = \begin{pmatrix} a_1 & b_1 & 0 & \dots & 0 \\ c_2 & a_2 & b_2 & \dots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & & & & b_{n-1} \\ 0 & \dots & 0 & c_n & a_n \end{pmatrix}$$

On montre que si la matrice A admet une décomposition LU, les matrices L et U sont bidiagonales et que les pivots $d_1 = a_1, d_2, \dots, d_n$ qui constituent la diagonale du U, sont donnés par la relation de récurrence :

$$d_{p+1} = a_{p+1} - \frac{b_{p+1} c_p}{d_p}$$

2) Les matrices à diagonale dominante. **Cet exemple important fait l'objet de l'exercice 6).**

Définition 4.3 *On dit que la matrice carrée $A = (a_{i,j})$ est à diagonale dominante si*

$$\forall i \in \{1, \cdots, n\} \quad |a_{i,i}| > \sum_{1 \leq j \leq j \neq i} |a_{i,j}|$$

Dans cette définition $|a_{i,i}|$ domine sa ligne. On pourrait adapter sur les colonnes avec des résultats analogues.

Lemme 4.4 *Si A est à diagonale dominante alors c'est une matrice inversible.*

Corollaire 4.5 *Si A est à diagonale dominante alors elle admet une factorisation LU.*

En effet les sous-matrices principales M_k sont aussi clairement à diagonale dominante donc inversibles et le théorème précédent s'applique.

Corollaire 4.6 *Soit A une matrice carrée et $\lambda \in \mathbb{C}$ une valeur propre de A. Alors :*

$$|\lambda - a_{i,i}| \leq \sum_{1 \leq j \leq j \neq i} |a_{i,j}|$$

En d'autres termes le spectre de A est contenu dans une réunion de disques de centres respectifs $a_{i,i}$ et rayons $\rho_i = \sum_{1 \leq j \leq j \neq i} |a_{i,j}|$. Les cercles correspondants sont appelés cercles de Gerschgörin. Un exemple concret est donné à la fin de l'exercice 6).

5 Factorisation de Cholesky

La décomposition de Cholesky n'est qu'une variante de la décomposition LU pour des matrices particulières : Dans toute cette section $A \in \mathcal{M}_n(\mathbb{R})$ est une matrice carrée *réelle symétrique définie positive.* c'est à dire telle que :

$$A = {}^tA \quad \text{et} \quad \forall x \in \mathbb{R}^n \setminus \{0\}, \quad {}^txAx > 0$$

Une telle matrice est forcément inversible puisque $Ax = 0 \Rightarrow^t xAx = 0$.

Théorème 5.1 *Il existe une matrice B triangulaire inférieure telle que :*

$$A = B({}^tB)$$

On peut exiger en plus $b_{i,i} > 0$ pour tout i. Et sous cette condition sur ses coefficients diagonaux, la matrice B est unique.

Démonstration.- Pour tout k la sous-matrice $M_k = (a_{i,j})_{1 \leq i,j \leq k}$ introduite dans le théorème 4.2 est inversible, car c'est la matrice de la restriction à $\mathbb{R}^k \times \{0\}$, de la forme quadratique $x \mapsto {}^txAx$, et le fait d'être définie positive passe aux sous-espaces. Il existe donc d'après la section 4 une décomposition

$$A = LU = \begin{bmatrix} 1 & \cdots & \cdots & 0 \\ \vdots & \ddots & \mathbf{0} & \vdots \\ \vdots & l_{i,j} & \ddots & \vdots \\ l_{n,1} & \cdots & \cdots & 1 \end{bmatrix} . \begin{bmatrix} u_{1,1} & \cdots & \cdots & u_{1,n} \\ \vdots & \ddots & u_{i,j} & \vdots \\ \vdots & \mathbf{0} & \ddots & \vdots \\ 0 & \cdots & \cdots & u_{n,n} \end{bmatrix}$$

La matrice L ayant une diagonale de 1 et on note $u_{1,1}, \ldots, u_{n,n}$, les termes diagonaux de U.

On a d'après un calcul déjà fait dans la démonstration du théorème 4.2 : $\Delta_k = \det(M_k) = u_{1,1} \ldots u_{k,k}$. Le fait que pour tout k, $\Delta_k > 0$ implique par une récurrence immédiate que $\forall k,\ u_{k,k} > 0$. Notons alors Λ la matrice diagonale de diagonale $(\sqrt{u_{1,1}}, \ldots, \sqrt{u_{n,n}})$. On obtient alors une nouvelle décomposition en matrices triangulaires *ayant la même diagonale :*

$$A = BC = (L\Lambda)(\Lambda^{-1}U)$$

$$B = L\Lambda = \begin{bmatrix} \sqrt{u_{1,1}} & \cdots & \cdots & 0 \\ \vdots & \ddots & \mathbf{0} & \vdots \\ \vdots & b_{i,j} & \ddots & \vdots \\ b_{n,1} & \cdots & \cdots & \sqrt{u_{n,n}} \end{bmatrix} \quad \text{et} \quad C = \Lambda^{-1}U = \begin{bmatrix} \sqrt{u_{1,1}} & \cdots & \cdots & c_{1,n} \\ \vdots & \ddots & c_{i,j} & \vdots \\ \vdots & \mathbf{0} & \ddots & \vdots \\ 0 & \cdots & \cdots & \sqrt{u_{n,n}} \end{bmatrix}$$

La relation ${}^tA = A$ implique $BC = {}^tC.{}^tB$, donc

$$(*) \qquad C({}^tB)^{-1} = B^{-1}.{}^tC$$

On constate, en appliquant les résultats de l'exercice 4) sur les produits et inverses des matrices triangulaires que cette matrice est triangulaire supérieure d'après le membre de gauche de l'égalité et triangulaire inférieure d'après le membre de droite : c'est donc une matrice diagonale. D'autre part les termes diagonaux des matrices triangulaires supérieures C et ${}^tB$ étant égaux ceux de $C({}^tB)^{-1}$ qui sont leurs quotients respectifs valent 1, et la matrice commune aux deux membres de $(*)$ est donc $C({}^tB)^{-1} = I$. Ceci donne $C = {}^tB$, ce qu'il fallait démontrer.

□

Observons que la partie existence de la démonstration a donné directement la solution à diagonale de coefficients positifs. Sans cette condition de positivité il y a d'autres solutions : on peut changer B en $B\Delta$ avec Δ diagonale à coefficients dans $\{-1, +1\}$. En effet on a alors $\Delta^2 = I$, matrice unité, donc $B\Delta^t(B\Delta) = B(\Delta)^{2t}B = B^tB = A$.

Calcul de B : on peut naturellement appliquer l'algorithme de factorisation LU, puis en tirer $B = L\Lambda$, en utilisant les coefficients diagonaux de U.

Toutefois on peut aussi calculer B directement et on constate que le nombre d'opérations est alors divisé par deux de l'ordre de $\frac{n^3}{3}$. C'est l'objet de l'exercice 8) précédé d'un exemple dans l'exercice 7).

Rappel et remarque.- La forme quadratique associée à une matrice symétrique A est $q(x) = {}^txAx = \sum_{i,j} a_{i,j}x_ix_j$.

Selon la décomposotion de Cholesky, on trouve :

$$q(x) = {}^txB^tBx = {}^t({}^tBx)({}^tBx) = \|{}^tBx\|^2$$

ce qui souligne le caractère positif de la matrice A et de la forme quadratique.

Remarque Il existe une décomposition de Cholesky, que nous ne développons pas ici par souci de simplicité, pour les matrices à coefficients dans $\mathbb{C}$, hermitiennes définies positives.