

CHAPITRE 1

INTERPOLATION POLYNÔMIALE

CHAPITRE 2

INTÉGRATION NUMÉRIQUE

CHAPITRE 3

RÉSOLUTION DE SYSTÈMES D'ÉQUATIONS LINÉAIRES

3.1 Généralités

Soit à résoudre le système linéaire

$$Ax = b, \quad (3.1)$$

où $A \in \mathbb{R}^{n \times n}$ est une matrice réelle inversible de taille n et $x, b \in \mathbb{R}^n$. Dans la suite, $x = A^{-1}b$ désignera la solution exacte de ce système.

On rappelle que théoriquement, la méthode de Cramer permet de donner chaque composante x_i de la solution x sous la forme d'un quotient de deux déterminants de taille n chacun, i.e.,

$$x_i = \frac{\Delta_i}{\Delta} = \frac{\det(A^{(1)}, \dots, A^{(i-1)}, b, A^{(i+1)}, \dots, A^{(n)})}{\det(A)}, \text{ pour } i = 1, 2, \dots, n. \quad (3.2)$$

Où $\Delta = \det(A)$ est le déterminant de la matrice A et Δ_i est le déterminant obtenu à partir de Δ en remplaçant $A^{(i)}$ la i -ème colonne de A par b le vecteur second membre.

Faisons remarquer que les formules de Cramer ci-dessus sont très peu utilisées (voire ne sont pas utilisées) dans la pratique car leur coût opératoire est de l'ordre de $(n+1)!$ *flops*. Ainsi, en supposant qu'on dispose d'un ordinateur capable d'effectuer 10^9 *flops* par seconde, alors il nous faudrait 9.6×10^{47} années pour résoudre un système linéaire ayant 50 inconnues comme le montre le tableau ci-dessous !

Taille n	Nombre de <i>flops</i>	Temps
10	$11! \simeq 3,99 \times 10^7$	0,04 seconde
20	$21! \simeq 5,1 \times 10^{19}$	1620 années
50	$51! \simeq 1,55 \times 10^{66}$	$9,6 \times 10^{47}$ années

Ajoutons à cela que même pour des systèmes linéaire de tailles "modestes", le calcul, sur des ordinateur, de la solution x par la méthode de Cramer est entaché par la propagation des

3.1. GÉNÉRALITÉS

erreurs d'arrondi et de troncature.

Pour toutes ses raisons, de nombreuses méthodes qui permettent de résoudre (3.1) ont été développées. Ces méthodes sont généralement subdivisées en :

- **Méthodes directes** : Ces méthodes fournissent la solution du système en un nombre fini d'étapes. Plus précisément, elles sont basées sur une décomposition de la matrice A sous forme d'un produit de matrices plus "simples" à manipuler : $A = LU$, $A = QR$, $A = LDL^T$, $A = SS^T$, ...
- **Méthodes itératives** : Ces méthodes déterminent, en théorie, la solution x du système (3.1) après un nombre infini d'itérations. Cette famille de méthodes consistent à construire une suite de vecteurs $x^{(0)}, x^{(1)}, \dots, x^{(k)}$ telle que

$$\lim_{k \rightarrow +\infty} \|x - x^{(k)}\|_{\infty} = 0.$$

Où $\|\cdot\|_{\infty}$ désigne la norme infinie dans \mathbb{R}^n définie par :

$$\|a\|_{\infty} = \max_{1 \leq i \leq n} |a_i| \text{ avec } a = (a_1, a_2, \dots, a_n)^T.$$

Définition 3.1. On dit qu'une méthode itérative est convergente si et seulement si pour tout vecteur de départ $x^{(0)}$ on a :

$$\lim_{k \rightarrow +\infty} \|x - x^{(k)}\|_{\infty} = 0.$$

Remarque 3.1. La convergence des méthodes itératives se teste généralement en utilisant un paramètre de tolérance ε . On arrête ainsi les calculs dès que l'erreur relative (respectivement absolue) est assez petite :

$$\frac{\|x^{(k+1)} - x^{(k)}\|_{\infty}}{\|x^{(k)}\|_{\infty}} < \varepsilon \quad (\|x^{(k+1)} - x^{(k)}\|_{\infty} < \varepsilon).$$

Une autre alternative est de tester le résidu relatif (respectivement absolu) :

$$\frac{\|r^{(k)}\|_{\infty}}{\|r^{(0)}\|_{\infty}} < \varepsilon \quad (\|r^{(k)}\|_{\infty} < \varepsilon),$$

où $r^{(k)} := b - Ax^{(k)}$.

Avant de passer à la description des méthodes citées ci-dessus, signalons qu'un des points les plus essentiels dans l'efficacité des méthodes de résolution de systèmes linéaires concerne la taille des systèmes à résoudre. Entre 1980 et 2000, la taille de la mémoire des ordinateurs a augmenté. La taille des systèmes qu'on peut résoudre sur ordinateur a donc également augmenté, selon l'ordre de grandeur suivant :

	1980	2000
Matrice "pleine" (tous les termes sont non nuls)	$n = 100$	$n = 10^6$
Matrice "creuse" (peu d'éléments non nuls)	$n = 10^6$	$n = 10^8$.

3.2 Rappels et compléments d'algèbre linéaire

Définition 3.2. Soit $A = (a_{ij})_{i,j=1,\dots,n}$ une matrice réelle de taille n . On dit que A est fortement inversible si les n sous matrices principales de A sont toutes inversibles, i.e., si $\det(A_k) \neq 0$, pour $k = 1, \dots, n$ et où $A_k = (a_{ij})_{i,j=1,\dots,k} \in \mathbb{R}^{k \times k}$.

Définition 3.3. Soit $A = (a_{ij})_{i,j=1,\dots,n}$ une matrice réelle de taille n . On dit que A est à diagonale dominante (respectivement strictement dominante) si et seulement si

$$|a_{ii}| \geq \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}| \quad (\text{respectivement } |a_{ii}| > \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}|) \quad \forall i = 1, \dots, n.$$

Définition 3.4. Soit A une matrice réelle de taille n . On dit que A est définie positive si pour tout $x \in \mathbb{R}^n$, $x \neq 0 \Rightarrow x^T A x > 0$.

Définition 3.5. Soit A une matrice réelle de taille n . On appelle rayon spectral de A qu'on note $\rho(A)$ la plus grande valeur propre en module, i.e.,

$$\rho(A) = \max_{i=1,\dots,n} \{|\lambda_i| : \lambda_i \text{ valeur propre de } A\}.$$

Proposition 3.1. Soit $A \in \mathbb{R}^{n \times n}$ une matrice réelle de taille n alors :

- Si A est à diagonale strictement dominante alors A est inversible.
- Si A est symétrique alors toutes ses valeurs propres sont réelles.
- Si A est symétrique définie positive alors toutes ses valeurs propres sont strictement positives.

3.3 Méthodes directes

3.3.1 Cas des systèmes triangulaires

Les systèmes linéaires dont la matrice A est triangulaire inférieure ou triangulaire supérieure sont parmi les systèmes les plus faciles à résoudre. En effet, la matrice A étant supposée inversible, alors ses coefficients diagonaux sont non nuls et donc pour un

- système triangulaire inférieur tel que celui décrit par le système (\mathcal{T}_i) ci-dessous,

$$(\mathcal{T}_i) : \begin{cases} a_{11}x_1 & & & = & b_1 \\ a_{21}x_1 & +a_{22}x_2 & & = & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1}x_1 & +a_{n2}x_2 & + \dots + & a_{nn}x_n & = & b_n, \end{cases}$$

la résolution se fait via les formules de descente (ou encore substitution directe) résumées dans l'algorithme ci-dessous

3.3. MÉTHODES DIRECTES

Algorithme : Résolution d'un système triangulaire inférieur

- $x_1 = \frac{b_1}{a_{1,1}}$;
- pour $i = 2, 3, \dots, n$ faire
 $x_i = \frac{1}{a_{i,i}} (b_i - \sum_{j=1}^{i-1} a_{i,j} x_j)$
fin pour.

- système triangulaire supérieur tel que celui décrit par la système (\mathcal{T}_s) ci-dessous,

$$(\mathcal{T}_s) : \begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \qquad \qquad \qquad \ddots \qquad \qquad \vdots \qquad \vdots \\ a_{nn}x_n = b_n, \end{cases}$$

la résolution se fait via les formules de remontée (ou encore substitution rétrograde) résumées dans l'algorithme ci-dessous

Algorithme : Résolution d'un système triangulaire supérieur

- $x_n = \frac{b_n}{a_{n,n}}$;
- pour $i = n-1, n-2, \dots, 1$ faire
 $x_i = \frac{1}{a_{i,i}} (b_i - \sum_{j=i+1}^n a_{i,j} x_j)$
fin pour.

Notons que pour les deux systèmes le coût opératoire est de l'ordre de n^2 flops.

3.3.2 Méthode de Gauss

Cette méthode est basée essentiellement sur l'utilisation de combinaisons linéaires entre lignes (et/ou de permutations de lignes ou de colonnes) du système dans le but de transformer progressivement le système de départ en un système triangulaire supérieure plus facile à résoudre. En effet, rappelons que la solution d'un système linéaire ne change pas quand on ajoute à une équation donnée une combinaison linéaire des autres équations.

La simplicité et l'efficacité de l'algorithme de cette méthode font que la méthode de Gauss est très utilisée dans la pratique et qu'elle est à la base de plusieurs algorithmes de résolution de systèmes linéaires. Dans la suite, nous abordons le principe de cette méthode à l'aide de quelques exemples.

Exemple 3.1. Soit le système

$$(\mathcal{S}) : \begin{cases} 2x_1 + x_2 + x_3 & = -3 \\ 4x_1 + 3x_2 + 3x_3 + x_4 & = -5 \\ 8x_1 + 7x_2 + 9x_3 + 5x_4 & = -7 \\ 6x_1 + 7x_2 + 9x_3 + 8x_4 & = 1 \end{cases}$$

Soit respectivement $A^{(0)} = A$ et $b^{(0)} = b$ la matrice et le vecteur second membre du système. Appliquons alors la méthode d'élimination de Gauss au système précédent en utilisant le tableau 4 lignes 5 colonnes ci-dessous

$$(A^{(0)} | b^{(0)}) = \left(\begin{array}{cccc|c} 2 & 1 & 1 & 0 & -3 \\ 4 & 3 & 3 & 1 & -5 \\ 8 & 7 & 9 & 5 & -7 \\ 6 & 7 & 9 & 8 & 1 \end{array} \right) \begin{array}{l} \ell_1 \\ \ell_2 \\ \ell_3 \\ \ell_4 \end{array}$$

$$(A^{(1)} | b^{(1)}) = \left(\begin{array}{cccc|c} 2 & 1 & 1 & 0 & -3 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 3 & 5 & 5 & 5 \\ 0 & 4 & 6 & 8 & 10 \end{array} \right) \begin{array}{l} \ell_1 \\ \ell_2 - 2\ell_1 = \ell_2 - \ell_{2,1}\ell_1 \\ \ell_3 - 4\ell_1 = \ell_3 - \ell_{3,1}\ell_1 \\ \ell_4 - 3\ell_1 = \ell_4 - \ell_{4,1}\ell_1 \end{array}$$

$$(A^{(2)} | b^{(2)}) = \left(\begin{array}{cccc|c} 2 & 1 & 1 & 0 & -3 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 2 & 2 & 2 \\ 0 & 0 & 2 & 4 & 6 \end{array} \right) \begin{array}{l} \ell_1 \\ \ell_2 \\ \ell_3 - 3\ell_2 = \ell_3 - \ell_{3,2}\ell_2 \\ \ell_4 - 4\ell_2 = \ell_4 - \ell_{4,2}\ell_2 \end{array}$$

$$(A^{(3)} | b^{(3)}) = \left(\begin{array}{cccc|c} 2 & 1 & 1 & 0 & -3 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 2 & 2 & 2 \\ 0 & 0 & 0 & 2 & 4 \end{array} \right) \begin{array}{l} \ell_1 \\ \ell_2 \\ \ell_3 \\ \ell_4 - \ell_3 = \ell_4 - \ell_{4,3}\ell_3, \end{array}$$

où $\ell_{i,k} = \frac{\alpha_{i,k}^{(k-1)}}{\alpha_{k,k}^{(k-1)}}$ pour $k = 1, 2, 3$ et $i = k + 1, \dots, 4$. Ainsi,

$$(\mathcal{S}) \Leftrightarrow \begin{cases} 2x_1 + x_2 + x_3 & = -3 \\ x_2 + x_3 + x_4 & = 1 \\ 2x_3 + 2x_4 & = 2 \\ 2x_4 & = 4 \end{cases}$$

et on obtient alors $x_1 = -1$, $x_2 = 0$, $x_3 = -1$ et $x_4 = 2$.

Exemple 3.2. Soit les systèmes

$$(\mathcal{S}_1) : \begin{cases} x_1 + x_2 + 2x_3 = 2 \\ 2x_1 + x_2 - x_3 = 3 \\ 3x_1 + 2x_2 + x_3 = 7 \end{cases} \quad \text{et} \quad (\mathcal{S}_2) : \begin{cases} x_1 + x_2 + 2x_3 = 1 \\ 2x_1 + x_2 - x_3 = 4 \\ 3x_1 + 2x_2 + x_3 = 5 \end{cases}$$

Les deux systèmes précédents ayant la même matrice A , nous allons appliquer la méthode d'élimination de Gauss en utilisant le tableau 3 lignes 5 colonnes formé par les éléments de la matrice $A^{(0)} = A$ et des vecteurs second membres $b_1^{(0)} = b_1$ et $b_2^{(0)} = b_2$. Nous avons alors

$$(A^{(0)} | b_1^{(0)} | b_2^{(0)}) = \left(\begin{array}{ccc|c|c} 1 & 1 & 2 & 2 & 1 \\ 2 & 1 & -1 & 3 & 4 \\ 3 & 2 & 1 & 7 & 5 \end{array} \right) \begin{matrix} \ell_1 \\ \ell_2 \\ \ell_3 \end{matrix}$$

$$(A^{(1)} | b_1^{(1)} | b_2^{(1)}) = \left(\begin{array}{ccc|c|c} 1 & 1 & 2 & 2 & 1 \\ 0 & -1 & -5 & -1 & 2 \\ 0 & -1 & -5 & 1 & 2 \end{array} \right) \begin{matrix} \ell_1 \\ \ell_2 - 2\ell_1 = \ell_2 - \ell_{2,1}\ell_1 \\ \ell_3 - 3\ell_1 = \ell_3 - \ell_{3,1}\ell_1 \end{matrix}$$

$$(A^{(2)} | b_1^{(2)} | b_2^{(2)}) = \left(\begin{array}{ccc|c|c} 1 & 1 & 2 & 2 & 1 \\ 0 & -1 & -5 & -1 & 2 \\ 0 & 0 & 0 & 2 & 0 \end{array} \right) \begin{matrix} \ell_1 \\ \ell_2 \\ \ell_3 - \ell_2 = \ell_3 - \ell_{3,2}\ell_2, \end{matrix}$$

avec $\ell_{i,k} = \frac{a_{i,k}^{(k-1)}}{a_{k,k}^{(k-1)}}$ pour $k = 1, 2$ et $i = k + 1, \dots, 3$. Ainsi,

$$(\mathcal{S}_1) \Leftrightarrow \begin{cases} x_1 + x_2 + 2x_3 = 2 \\ -x_2 - 5x_3 = -1 \\ 0 = 2 \end{cases} \quad \text{et} \quad (\mathcal{S}_2) \Leftrightarrow \begin{cases} x_1 + x_2 + 2x_3 = 1 \\ -x_2 - 5x_3 = 4 \\ 0 = 5 \end{cases}$$

On en déduit que le système (\mathcal{S}_1) n'a pas de solutions et n'est donc pas compatible. Quant au système (\mathcal{S}_2) , il possède une infinité de solutions de la forme $x_3 = \alpha$, $x_2 = -2 - 5\alpha$ et $x_1 = 3 + 3\alpha$ où $\alpha \in \mathbb{R}$.

Mise en Œuvre de l'algorithme

Soit le système linéaire $Ax = b$ où $A \in \mathbb{R}^{n \times n}$ est telle que le premier terme diagonal a_{11} soit non nul. Posons $A^{(0)} = A$ et $b^{(0)} = b$, alors comme $a_{11}^{(0)} \neq 0$, on peut éliminer l'inconnue x_1 de la ligne $\ell_i^{(0)}$, $i = 2, 3, \dots, n$, en retranchant à cette ligne la quantité $\ell_{i,1}\ell_1^{(0)}$ où $\ell_{i,1} = \frac{a_{i,1}^{(0)}}{a_{1,1}^{(0)}}$. Ainsi, en définissant

$$\begin{cases} a_{i,j}^{(1)} = a_{i,j}^{(0)} - \ell_{i,1}a_{1,j}^{(0)} & i, j = 2, 3, \dots, n, \\ b_i^{(1)} = b_i^{(0)} - \ell_{i,1}b_1^{(0)} & i, j = 2, 3, \dots, n, \end{cases}$$

on obtient le système $A^{(1)}x = b^{(1)}$ équivalent suivant

$$\underbrace{\begin{pmatrix} a_{11}^{(0)} & a_{12}^{(0)} & \dots & a_{1n}^{(0)} \\ 0 & a_{22}^{(1)} & \dots & a_{2n}^{(1)} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2}^{(1)} & \dots & a_{nn}^{(1)} \end{pmatrix}}_{:=A^{(1)}} \underbrace{\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}}_{:=x} = \underbrace{\begin{pmatrix} b_1^{(0)} \\ b_2^{(1)} \\ \vdots \\ b_n^{(1)} \end{pmatrix}}_{:=b^{(1)}}$$

De la même manière, pour $k = 2, 3, \dots, n$ et si $a_{k,k}^{(k-1)} \neq 0$, on peut éliminer x_k de la ligne $\ell_i^{(k-1)}$, $i = k+1, k+2, \dots, n$, en retranchant à cette ligne la quantité $\ell_{i,k} \ell_k^{(k-1)}$ où $\ell_{i,k} = \frac{a_{i,k}^{(k-1)}}{a_{k,k}^{(k-1)}}$. Ainsi, en définissant les quantités :

$$\begin{cases} a_{i,j}^{(k)} = a_{i,j}^{(k-1)} - \ell_{i,k} a_{k,j}^{(k-1)} & i, j = k+1, k+2, \dots, n, \\ b_i^{(k)} = b_i^{(k-1)} - \ell_{i,k} b_k^{(k-1)} & i = k+1, k+2, \dots, n, \end{cases}$$

on construit une suite finie de systèmes

$$A^{(k)}x = b^{(k)}, \text{ pour } k = 1, 2, \dots, n-1,$$

où la matrice $A^{(k)}$ pour $1 \leq k \leq n-1$ à la forme suivante

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

et où on a supposé que $a_{i,i}^{(i-1)} \neq 0$ pour $i = 1, 2, \dots, k$.

Finalement, notons que pour $k = n-1$, on obtient le système triangulaire supérieur $Ux = y$ suivant :

$$\underbrace{\begin{pmatrix} a_{11}^{(0)} & a_{12}^{(0)} & \dots & a_{1n}^{(0)} \\ 0 & a_{22}^{(1)} & \dots & a_{2n}^{(1)} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn}^{(n-1)} \end{pmatrix}}_{:=A^{(n-1)}} \underbrace{\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}}_{:=x} = \underbrace{\begin{pmatrix} b_1^{(0)} \\ b_2^{(1)} \\ \vdots \\ b_n^{(n-1)} \end{pmatrix}}_{:=b^{(n-1)}}$$

où $U := A^{(n-1)}$ et $y := b^{(n-1)}$, et qu'on a établi que la résolution du système $Ax = b$ équivaut à celle du système $Ux = y$. Finalement, la méthode d'élimination de Gauss, dont le coût opératoire est de l'ordre de $\frac{2}{3}n^3$, est résumé par l'algorithme ci-dessous :

Algorithme : Méthode de Gauss sans pivotage

```

1. Phase de triangularisation.
  Pour  $k = 1, 2, \dots, n-1$ 
    Pour  $i = k+1, k+2, \dots, n$ 
      Pour  $j = k+1, k+2, \dots, n$ 
         $a_{i,j} = a_{i,j} - \frac{a_{i,k}}{a_{k,k}} a_{k,j}$ ;
      Fin pour
       $b_i = b_i - \frac{a_{i,k}}{a_{k,k}} b_k$ ;
       $a_{i,k} = 0$ ;
    Fin pour
  Fin pour
2. Phase de résolution du système triangulaire supérieur.
 $x_n = \frac{b_n}{a_{n,n}}$ ;
  Pour  $i = n-1, n-2, \dots, 1$ 
     $x_i = \frac{1}{a_{i,i}} (b_i - \sum_{j=i+1}^n a_{i,j} x_j)$ ;
  Fin pour

```

Stratégie de pivotage

Exemple 3.3. Soit à résoudre le système $Ax = b$, où $b \in \mathbb{R}^3$ est quelconque et $A \in \mathbb{R}^{3 \times 3}$ la matrice

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

Après application de la première étape de la méthode de Gauss à la matrice A , on obtient

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

et on remarque qu'on ne peut pas continuer, et pourtant la matrice A est inversible ! Par contre si on permute les lignes 2 et 3 de la matrice A , et que l'on applique la méthode de Gauss à la nouvelle matrice obtenu

$$\tilde{A} = PA = \begin{pmatrix} 1 & 2 & 3 \\ 7 & 8 & 9 \\ 2 & 4 & 5 \end{pmatrix}, \text{ avec } P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

On obtient

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

qui est déjà une matrice triangulaire supérieure et donc le système $Ax = b$ peut être résolu.

Ce premier exemple nous montre que dans certains cas, la permutation de quelques équations d'un système linéaire donné et donc la permutation de certaines lignes de la matrice A est nécessaire afin de résoudre ce système.

Exemple 3.4. *Considérons le système*

$$(\mathcal{S}) : \begin{cases} \varepsilon x_1 + x_2 = 1 \\ x_1 + x_2 = 2, \end{cases}$$

où $\varepsilon \neq 0$. De plus, on suppose que $\varepsilon \neq 1$, et ainsi le système (\mathcal{S}) est inversible. La solution de ce système est alors donnée par $x = (x_1, x_2)$ avec $x_1 = \frac{1}{1-\varepsilon}$ et $x_2 = \frac{1-2\varepsilon}{1-\varepsilon}$.

Supposons maintenant que $\varepsilon \neq 1$ est "proche de 0" et appliquons l'élimination de Gauss décrite précédemment. La première étape transforme donc le système (\mathcal{S}) en $(\mathcal{S}^{(1)})$

$$(\mathcal{S}^{(1)}) : \begin{cases} \varepsilon x_1 + x_2 = 1 \\ (1 - \frac{1}{\varepsilon})x_2 = 2 - \frac{1}{\varepsilon}. \end{cases}$$

Comme les valeurs 1 et 2 sont très petites par rapport à $-\frac{1}{\varepsilon}$, on voit alors que la deuxième équation nous donne $-\frac{1}{\varepsilon}x_2 \simeq -\frac{1}{\varepsilon}$, d'où $x_2 \simeq 1$ et ensuite la première équation nous fournit $x_1 \simeq 0$. Mais ce résultat est faux!

L'erreur ne provient pas seulement du fait que ε est "très petit" car si on multiplie la première ligne par une puissance de 10 quelconque, on va trouver la même erreur!

L'anomalie provient du déséquilibre entre les coefficients de x_1 et x_2 de la ligne du pivot. Pour y remédier, échangeons maintenant les deux lignes de notre système et appliquons l'élimination de Gauss avec la valeur 1 comme premier pivot. On obtient alors

$$(\mathcal{S}^{(1)}) : \begin{cases} x_1 + x_2 = 1 \\ (1 - \varepsilon)x_2 = 1 - 2\varepsilon, \end{cases}$$

ce qui entraîne alors que $x_1 \simeq 1$ et $x_2 \simeq 1$. Et ce résultat est correct!

De même, ce deuxième exemple montre qu'il ne faut pas utiliser des pivots "trop petits" car les erreurs d'arrondi et/ou de troncature peuvent entraîner des solutions fausses. Et donc, il est nécessaire de recourir à des permutations de lignes voire des colonnes de la matrice du système.

Ainsi, pour éviter les problèmes rencontrés dans les deux exemples précédents, on peut appliquer l'une des deux stratégies suivantes :

- **Stratégie de pivotage partiel :** Cette stratégie, appelée encore stratégie du pivot partiel, consiste à chercher dans la sous colonne k , le plus grand pivot en valeur absolue, i.e., on cherche l'indice de ligne ℓ tel que $a_{\ell,k} = \max_{k \leq i \leq n} |a_{i,k}|$. Ensuite, on permute les lignes ℓ et k .
- **Stratégie de pivotage total :** Dans cette stratégie, on cherche le maximum en valeur absolue dans toute la sous matrice $(a_{i,j})_{i,j=k,\dots,n}$, i.e., on cherche les indices de ligne ℓ et de colonne m vérifiant $a_{\ell,m} = \max_{k \leq i,j \leq n} |a_{i,j}|$. Ensuite, on permute les lignes ℓ et k ainsi que les colonnes m et ℓ . A noter que cette stratégie est appelée encore stratégie du pivot total.

Remarque 3.2. - *Les pivots ayant une valeur petite ne disparaissent pas, que l'on utilise la stratégie du pivot total ou celle du pivot partiel. En fait, l'apparition des petits pivots est juste rejetée à la fin de la méthode de Gauss et dans ce cas l'influence de ces petits pivots est moins importante.*

- *Généralement, dans la pratique, la méthode de Gauss est implémentée avec une stratégie de pivot partiel.*

3.3.3 Décomposition LU

Nous commençons cette section en donnant quelques résultats théoriques concernant l'existence et l'unicité de la décomposition LU .

Théorème 3.1. *Une matrice $A \in \mathbb{R}^{n \times n}$ possède une décomposition LU , où la matrice L est triangulaire inférieure à diagonale unité et U triangulaire supérieure si et seulement si A est fortement inversible. Dans ce cas la décomposition LU est unique.*

Ce théorème permet d'affirmer que si la matrice A est fortement inversible, alors les n étapes de la méthode de Gauss peuvent être appliquées à la matrice A sans rencontrer un pivot nul.

Théorème 3.2. *Soit $A \in \mathbb{R}^{n \times n}$ une matrice inversible, alors il existe au moins une matrice de permutation P telle que la matrice PA admette une décomposition LU .*

Dans la suite, nous introduisons la décomposition LU d'une matrice A en donnant une interprétation matricielle des différentes étapes de la méthode d'élimination de Gauss.

Décomposition $A = LU$

Reprenons le système $Ax = b$ vu dans l'exemple 3.3 du paragraphe précédent. Soient alors la matrice A et le vecteur second membre b

$$A^{(0)} = A = \begin{pmatrix} 2 & 1 & 1 & 0 \\ 4 & 3 & 3 & 1 \\ 8 & 7 & 9 & 5 \\ 6 & 7 & 9 & 8 \end{pmatrix}, \quad b^{(0)} = b = \begin{pmatrix} -3 \\ -5 \\ -7 \\ 1 \end{pmatrix},$$

et notons ℓ_1, ℓ_2, ℓ_3 et ℓ_4 les vecteurs lignes des différentes matrices $A^{(k)}$ et des vecteurs $b^{(k)}$ obtenus durant la méthode d'élimination de Gauss.

- **Étape 1** : Par des combinaisons linéaires ci-dessous :

$$\begin{aligned}\ell_2 &\leftarrow \ell_2 - \ell_{2,1}\ell_1, \text{ avec } \ell_{2,1} = 2, \\ \ell_3 &\leftarrow \ell_3 - \ell_{3,1}\ell_1, \text{ avec } \ell_{3,1} = 4, \\ \ell_4 &\leftarrow \ell_4 - \ell_{4,1}\ell_1, \text{ avec } \ell_{4,1} = 3,\end{aligned}$$

nous avons transformé la matrice $A^{(0)}$ et le vecteur $b^{(0)}$ respectivement en la matrice $A^{(1)} = L_1A^{(0)}$ et en le vecteur $b^{(1)} = L_1b^{(0)}$, où

$$L_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ -4 & 0 & 1 & 0 \\ -3 & 0 & 0 & 1 \end{pmatrix}, \quad A^{(1)} = \begin{pmatrix} 2 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 3 & 5 & 5 \\ 0 & 4 & 6 & 8 \end{pmatrix} \quad \text{et} \quad b^{(1)} = \begin{pmatrix} -3 \\ 1 \\ 5 \\ 10 \end{pmatrix}.$$

Notons que la matrice L_1 n'est autre que la matrice identité où les 0 de la première colonne, au niveau des positions 2, 3 et 4, ont été remplacés par les coefficients $-\ell_{2,1}$, $-\ell_{3,1}$ et $-\ell_{4,1}$.

- **Étape 2** : Comme lors de la précédente étape, les combinaisons linéaires entre les lignes ci-dessous :

$$\begin{aligned}\ell_3 &\leftarrow \ell_3 - \ell_{3,2}\ell_2, \text{ avec } \ell_{3,2} = 3, \\ \ell_4 &\leftarrow \ell_4 - \ell_{4,2}\ell_2, \text{ avec } \ell_{4,2} = 4,\end{aligned}$$

ont permis de transformer $A^{(1)}$ en $A^{(2)} = L_2A^{(1)}$ et $b^{(1)}$ en $b^{(2)} = L_2b^{(1)}$, où

$$L_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -3 & 1 & 0 \\ 0 & -4 & 0 & 1 \end{pmatrix}, \quad A^{(2)} = \begin{pmatrix} 2 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & 2 & 4 \end{pmatrix} \quad \text{et} \quad b^{(2)} = \begin{pmatrix} -3 \\ 1 \\ 2 \\ 6 \end{pmatrix}.$$

Notons également que la matrice L_2 n'est autre que la matrice identité où les 0, au niveau des positions 3 et 4, de la deuxième colonne ont été remplacés par les coefficients $-\ell_{3,2}$ et $-\ell_{4,2}$.

- **Étape 3** : A l'aide de la combinaison linéaire

$$\ell_4 \leftarrow \ell_4 - \ell_{4,3}\ell_3, \text{ avec } \ell_{4,3} = 1,$$

on transforme $A^{(2)}$ en $A^{(3)} = L_3A^{(2)}$ et $b^{(2)}$ en $b^{(3)} = L_3b^{(2)}$, où

$$L_3 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{pmatrix}, \quad A^{(3)} = \begin{pmatrix} 2 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 2 \end{pmatrix} \quad \text{et} \quad b^{(3)} = \begin{pmatrix} -3 \\ 1 \\ 2 \\ 4 \end{pmatrix}.$$

Remarquons que la matrice L_3 n'est autre que la matrice identité où le 0, au niveau de la position 4, de la troisième colonne a été remplacé par le coefficient $-\ell_{4,3}$.

3.3. MÉTHODES DIRECTES

Ainsi, en posant $U = A^{(3)}$, on voit que $U = L_3 L_2 L_1 A$. D'où

$$A = LU \text{ avec } L = (L_3 L_2 L_1)^{-1} = L_1^{-1} L_2^{-1} L_3^{-1}.$$

Notons qu'il est facile de vérifier que

$$L_1^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 4 & 0 & 1 & 0 \\ 3 & 0 & 0 & 1 \end{pmatrix}, \quad L_2^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 3 & 1 & 0 \\ 0 & 4 & 0 & 1 \end{pmatrix} \text{ et } L_3^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix},$$

et que de plus la matrice L est une matrice triangulaire inférieure à diagonale unité dont les éléments non nuls ne sont autre que les coefficients $\ell_{i,j}$ utilisés précédemment, i.e.,

$$L = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 4 & 3 & 1 & 0 \\ 3 & 4 & 1 & 1 \end{pmatrix}.$$

Finalement, on a obtenu

$$\underbrace{\begin{pmatrix} 2 & 1 & 1 & 0 \\ 4 & 3 & 3 & 1 \\ 8 & 7 & 9 & 5 \\ 6 & 7 & 9 & 8 \end{pmatrix}}_A = \underbrace{\begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 4 & 3 & 1 & 0 \\ 3 & 4 & 1 & 1 \end{pmatrix}}_L \underbrace{\begin{pmatrix} 2 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 2 \end{pmatrix}}_U.$$

Décomposition $PA = LU$

Dans certains cas, la décomposition LU d'une matrice A donnée n'existe pas. Par contre, il est possible de trouver une matrice de permutation P telle que la décomposition LU de PA existe.

L'exemple ci-dessous montre comment trouver P , L et U en utilisant la stratégie de pivotage partiel. Soit donc

$$A^{(0)} = A = \begin{pmatrix} 2 & 1 & 1 & 0 \\ 4 & 3 & 3 & 1 \\ 8 & 7 & 9 & 5 \\ 6 & 7 & 9 & 8 \end{pmatrix}.$$

- **Etape 1** : Examinons les composantes de la première colonne de A pour déterminer la composante maximale en valeur absolue. On voit alors que $|a_{3,1}^{(0)}| = \max_{i=1,\dots,4} |a_{i,1}^{(0)}|$. Permutons alors, la 1^{ère} et la 3^{ème} ligne de $A^{(0)}$. Matriciellement, cette opération se traduit par une multiplication à gauche de $A^{(0)}$ par la matrice de permutation P_1 , i.e.,

$$\tilde{A}^{(0)} = P_1 A^{(0)} = \begin{pmatrix} 8 & 7 & 9 & 5 \\ 4 & 3 & 3 & 1 \\ 2 & 1 & 1 & 0 \\ 6 & 7 & 9 & 8 \end{pmatrix} \text{ avec } P_1 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

En utilisant les combinaisons linéaires

$$\begin{aligned}\ell_2 &\leftarrow \ell_2 - \ell_{2,1}\ell_1, \text{ avec } \ell_{2,1} = \frac{1}{2}, \\ \ell_3 &\leftarrow \ell_3 - \ell_{3,1}\ell_1, \text{ avec } \ell_{3,1} = \frac{1}{4}, \\ \ell_4 &\leftarrow \ell_4 - \ell_{4,1}\ell_1, \text{ avec } \ell_{4,1} = \frac{3}{4},\end{aligned}$$

transformons $\tilde{A}^{(0)}$ en $A^{(1)} = L_1\tilde{A}^{(0)} = L_1P_1A^{(0)}$, où

$$A^{(1)} = \begin{pmatrix} 8 & 7 & 9 & 5 \\ 0 & -\frac{1}{2} & -\frac{3}{2} & -\frac{3}{2} \\ 0 & -\frac{3}{4} & -\frac{5}{4} & -\frac{5}{4} \\ 0 & \frac{7}{4} & \frac{9}{4} & \frac{17}{4} \end{pmatrix} \text{ et } L_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 & 0 \\ -\frac{1}{4} & 0 & 1 & 0 \\ -\frac{3}{4} & 0 & 0 & 1 \end{pmatrix}.$$

- **Etape 2 :** Comme lors de la précédente étape, déterminons la composante maximale, en valeur absolue, de la deuxième colonne de $A^{(1)}$. On voit alors que $|a_{4,2}^{(1)}| = \max_{i=2,\dots,4} |a_{i,2}^{(1)}|$. Permutons alors, la 2^{ème} et la 4^{ème} ligne de $A^{(1)}$. On a alors les relations :

$$\tilde{A}^{(1)} = P_2A^{(1)} = \begin{pmatrix} 8 & 7 & 9 & 5 \\ 0 & \frac{7}{4} & \frac{9}{4} & \frac{17}{4} \\ 0 & -\frac{3}{4} & -\frac{5}{4} & -\frac{5}{4} \\ 0 & -\frac{1}{2} & -\frac{3}{2} & -\frac{3}{2} \end{pmatrix} \text{ avec } P_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}.$$

Maintenant, si on utilise les combinaisons linéaires suivantes :

$$\begin{aligned}\ell_3 &\leftarrow \ell_3 - \ell_{3,2}\ell_2, \text{ avec } \ell_{3,2} = -\frac{3}{7}, \\ \ell_4 &\leftarrow \ell_4 - \ell_{4,2}\ell_2, \text{ avec } \ell_{4,2} = -\frac{2}{7},\end{aligned}$$

alors $\tilde{A}^{(1)}$ se transforme en $A^{(2)} = L_2\tilde{A}^{(1)}$, où

$$A^{(2)} = \begin{pmatrix} 8 & 7 & 9 & 5 \\ 0 & \frac{7}{4} & \frac{9}{4} & \frac{17}{4} \\ 0 & 0 & -\frac{2}{7} & \frac{4}{7} \\ 0 & 0 & -\frac{6}{7} & -\frac{2}{7} \end{pmatrix} \text{ et } L_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & \frac{3}{7} & 1 & 0 \\ 0 & \frac{2}{7} & 0 & 1 \end{pmatrix}.$$

- **Etape 3 :** Maintenant, déterminons la composante maximale, en valeur absolue, de la troisième colonne de $A^{(2)}$. On voit alors que $|a_{4,3}^{(2)}| = \max_{i=3,\dots,4} |a_{i,3}^{(2)}|$. Permutons alors, la 3^{ème} et la 4^{ème} ligne de $A^{(2)}$. Ce qui se traduit par les relations matricielles :

$$\tilde{A}^{(2)} = P_3A^{(2)} = \begin{pmatrix} 8 & 7 & 9 & 5 \\ 0 & \frac{7}{4} & \frac{9}{4} & \frac{17}{4} \\ 0 & 0 & -\frac{6}{7} & -\frac{2}{7} \\ 0 & 0 & -\frac{2}{7} & \frac{4}{7} \end{pmatrix} \text{ avec } P_3 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

3.3. MÉTHODES DIRECTES

Finalement, en utilisant la combinaison linéaire

$$\ell_4 \leftarrow \ell_4 - \ell_{4,3}\ell_3, \text{ avec } \ell_{4,3} = \frac{1}{3},$$

la matrice $\tilde{A}^{(2)}$ est transformée en $A^{(3)} = L_3\tilde{A}^{(2)}$, où

$$A^{(2)} = \begin{pmatrix} 8 & 7 & 9 & 5 \\ 0 & \frac{7}{4} & \frac{9}{4} & \frac{17}{4} \\ 0 & 0 & -\frac{6}{7} & -\frac{2}{7} \\ 0 & 0 & 0 & \frac{2}{3} \end{pmatrix} \text{ et } L_3 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -\frac{1}{3} & 1 \end{pmatrix}.$$

Ainsi, en posant $U = A^{(3)}$, on voit que $U = L_3P_3L_2P_2L_1P_1A$. Cette dernière relation peut être reformulée de la manière suivante

$$U = (L_3P_3L_2P_2L_1P_1)A = (L'_3L'_2L'_1)(P_3P_2P_1)A,$$

où L'_k est égale à L_k à des permutations près. Plus précisément, on définit

$$L'_3 = L_3, \quad L'_2 = P_3L_2P_3, \quad L'_1 = P_3P_2L_1P_2P_3.$$

Finalement, on a

$$\underbrace{(P_3P_2P_1)}_{:=P} A = \underbrace{(L'_3L'_2L'_1)^{-1}}_{:=L} U,$$

avec

$$L'_3 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -\frac{1}{3} & 1 \end{pmatrix}, \quad L'_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & \frac{2}{7} & 1 & 0 \\ 0 & \frac{3}{7} & 0 & 1 \end{pmatrix} \text{ et } L'_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -\frac{3}{4} & 1 & 0 & 0 \\ -\frac{1}{2} & 0 & 1 & 0 \\ -\frac{1}{4} & 0 & 0 & 1 \end{pmatrix}.$$

C'est à dire $PA = LU$, où

$$P = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \quad L = \begin{pmatrix} 1 & 0 & 0 & 0 \\ \frac{3}{4} & 1 & 0 & 0 \\ \frac{1}{2} & -\frac{2}{7} & 1 & 0 \\ \frac{1}{4} & -\frac{3}{7} & \frac{1}{3} & 1 \end{pmatrix} \text{ et } U = \begin{pmatrix} 8 & 7 & 9 & 5 \\ 0 & \frac{7}{4} & \frac{9}{4} & \frac{17}{4} \\ 0 & 0 & -\frac{6}{7} & -\frac{2}{7} \\ 0 & 0 & 0 & \frac{2}{3} \end{pmatrix}.$$

Remarque 3.3. 1. Au cours de la recherche du pivot, il se peut que le maximum soit atteint en deux (ou plus) lignes. Dans ce cas, on retiendra comme ligne du pivot la première ligne où est atteint l'élément maximal.

2. La décomposition $PA = LU$ peut être utilisée pour résoudre un système linéaire $Ax = b$. En effet :

$$\begin{aligned} Ax = b &\Leftrightarrow PAx = \tilde{b} \quad (\text{avec } \tilde{b} = Pb), \\ &\Leftrightarrow L Ux = \tilde{b}, \\ &\Leftrightarrow Ly = \tilde{b} \quad (\text{avec } y = Ux). \end{aligned}$$

On voit alors que pour obtenir x la solution de $Ax = b$, il suffit de résoudre respectivement les deux systèmes triangulaires (inférieur et supérieur) suivants :

$$Ly = \tilde{b} \text{ (où } y \text{ est l'inconnu) avec } \tilde{b} = Pb \text{ et } Ux = y.$$

Exemple 3.5. Résoudre

$$(\mathcal{S}) : \begin{cases} 2x_1 + x_2 + 4x_4 = 1 \\ -4x_1 - 2x_2 + 3x_3 - 7x_4 = -3 \\ 4x_1 + x_2 - 2x_3 + 8x_4 = 1 \\ -3x_2 - 12x_3 - x_4 = -2. \end{cases}$$

Matriciellement, (\mathcal{S}) s'écrit

$$\underbrace{\begin{pmatrix} 2 & 1 & 0 & 4 \\ -4 & -2 & 3 & -7 \\ 4 & 1 & -2 & 8 \\ 0 & -3 & -12 & -1 \end{pmatrix}}_A \underbrace{\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}}_x = \underbrace{\begin{pmatrix} 1 \\ -3 \\ 1 \\ -2 \end{pmatrix}}_b.$$

En effectuant la factorisation LU de A avec stratégie de pivotage partiel, on trouve que $PA = LU$, avec

$$P = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \quad L = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & \frac{1}{3} & 1 & 0 \\ -\frac{1}{2} & 0 & \frac{3}{10} & 1 \end{pmatrix} \quad \text{et} \quad U = \begin{pmatrix} -4 & -2 & 3 & -7 \\ 0 & -3 & -12 & -1 \\ 0 & 0 & 5 & \frac{4}{3} \\ 0 & 0 & 0 & \frac{1}{10} \end{pmatrix}.$$

Ainsi,

$$\begin{aligned} Ax = b &\Leftrightarrow PAx = \tilde{b} \quad (\text{avec } \tilde{b} = Pb = (-3, -2, 1, 1)^T), \\ &\Leftrightarrow LUx = \tilde{b}, \\ &\Leftrightarrow Ly = \tilde{b} \quad (\text{avec } y = Ux). \end{aligned}$$

En résolvant le système triangulaire inférieur $Ly = \tilde{b}$, on trouve $y = (-3, -2, -\frac{4}{3}, -\frac{1}{10})^T$. Finalement, en résolvant le système triangulaire supérieur $Ux = y$, on trouve $x = (2, 1, 0, -1)^T$.

Mise en oeuvre de la décomposition LU

En supposant que la décomposition LU de la matrice A existe et en identifiant le terme général $a_{i,j}$ de la matrice A avec les éléments du produit scalaire $\ell_i u^j$, où $\ell_i = (\ell_{i,1}, \dots, \ell_{i,i-1}, 1, 0, \dots, 0)$ est la i -ème ligne de la matrice L et $u^j = (u_{1,j}, \dots, u_{j,j}, 0, \dots, 0)^T$ est la j -ème colonne de la matrice U , on a

$$a_{i,j} = \sum_{k=1}^n \ell_{i,k} u_{k,j} = \sum_{k=1}^{\min(i,j)} \ell_{i,k} u_{k,j},$$

3.3. MÉTHODES DIRECTES

car $\ell_{i,k} = 0$ lorsque $k > i$ et $u_{k,j} = 0$ pour $k > j$. Ainsi, on obtient :

$$\text{- Si } i \leq j, \text{ alors } a_{i,j} = \sum_{k=1}^i \ell_{i,k} u_{k,j} \text{ et donc } u_{i,j} = a_{i,j} - \sum_{k=1}^{i-1} \ell_{i,k} u_{k,j}, \quad (3.3)$$

$$\text{- Si } i > j, \text{ alors } a_{i,j} = \sum_{k=1}^j \ell_{i,k} u_{k,j} \text{ et donc } \ell_{i,j} = \frac{1}{u_{j,j}} \left(a_{i,j} - \sum_{k=1}^{j-1} \ell_{i,k} u_{k,j} \right). \quad (3.4)$$

Les deux égalités (3.3) et (3.4) permettent de former les éléments non nuls de la i -ème ligne de la matrice U ainsi que les éléments non nuls de la i -ème colonne de la matrice L . En effet, en prenant $j = 1$ dans l'égalité (3.3), nous avons $u_{1,i} = a_{1,i}$ pour $i = 1, \dots, n$ et puis l'égalité (3.4) nous donne $\ell_{i,1} = \frac{a_{i,1}}{u_{1,1}}$.

En supposant avoir déterminé les colonnes (respectivement les lignes) $1, \dots, i-1$ de la matrice U (respectivement L), nous obtenons à partir de (3.3) (respectivement à partir de (3.4)) les éléments $j = i, \dots, n$ de la i -ème ligne de U (respectivement colonne de L), i.e., nous avons

$$\begin{cases} u_{i,j} = a_{i,j} - \sum_{k=1}^{i-1} \ell_{i,k} u_{k,j} & \text{pour } j = i, \dots, n \\ \ell_{j,i} = \frac{1}{u_{i,i}} \left(a_{j,i} - \sum_{k=1}^{i-1} \ell_{i,k} u_{k,j} \right) & \text{pour } j = i+1, \dots, n. \end{cases}$$

Remarquons que la division par $u_{i,i}$ est possible puisque par hypothèse on suppose que la factorisation LU existe. De plus, toute nouveau coefficient $u_{i,j}$ ou $\ell_{j,i}$ est obtenu en fonction des éléments de L et U déjà calculés. Ainsi, les relations précédentes et les propriétés vérifiées par les matrices L et U nous permettent de donner l'algorithme suivant :

Algorithme. Décomposition LU .

1. Initialisation des matrices L et U .

$$L = I_n; U = 0_n;$$

2. Calcul de la première colonne de L et de la première ligne de U .

Pour $i = 1, \dots, n$

$$u_{1,i} = a_{1,i};$$

$$\ell_{i,1} = \frac{a_{i,1}}{a_{1,1}};$$

Fin pour

3. Calcul des éléments non nuls de la i -ème ligne de U et de la i -ème colonne de L .

Pour $i = 2, \dots, n$

Pour $j = i, \dots, n$

$$u_{i,j} = a_{i,j} - \sum_{k=1}^{i-1} l_{i,k} u_{k,j};$$

Fin pour

Pour $j = i + 1, \dots, n$

$$l_{j,i} = \frac{1}{u_{i,i}} (a_{j,i} - \sum_{k=1}^{i-1} l_{j,k} u_{k,i});$$

Fin pour

Fin pour

4. Phase de résolution du système triangulaire inférieur $Ly = b$.

5. Phase de résolution du système triangulaire supérieur $Ux = y$.

3.4 Méthodes itératives

Lorsque la taille des systèmes à résoudre est grande, et vu que la résolution d'un système linéaire $Ax = b$ de taille n requiert en général $\mathcal{O}(n^3)$ opérations, les méthodes directes sont coûteuses en terme de nombre d'opérations et donc aussi en terme de temps d'exécution sur ordinateur. Dans ce cas, il est préférable d'utiliser des méthodes itératives qui construisent des solutions approchées à la solution exacte du système.

Dans cette section, nous allons décrire une classe de méthodes itératives basées sur la décomposition de la matrice A sous la forme de matrices plus faciles à manipuler! Plus précisément, étant donnée A est une matrice carrée inversible, on suppose qu'on puisse écrire $A = M - N$ où les matrices M et N sont "convenablement choisies".

Afin de résoudre le système $Ax = b$, on remplace l'équation du système par la nouvelle équation :

$$Mx = Nx + b,$$

et en supposant que la matrice M est facilement inversible, nous obtenons

$$x = Bx + c,$$

où $B = M^{-1}N$ et $c = M^{-1}b$. On voit alors qu'on peut définir le schéma itératif suivant :

$$\begin{cases} x^{(k+1)} &= Bx^{(k)} + c \\ x^{(0)} &\text{donné ou à choisir.} \end{cases}$$

La convergence ou la non convergence de la suite $x^{(k)}$ vers la solution exacte $x = A^{-1}b$ du système dépend du rayon spectral de la matrice B . Plus précisément, nous avons le résultat suivant

Théorème 3.3. *La suite $x^{(k)}$ définie par $x^{(k+1)} = Bx^{(k)} + c$ converge pour tout $x^{(0)} \in \mathbb{R}^n$ si et seulement si $\rho(B) < 1$.*

Dans la suite, nous décomposons la matrice A sous la forme

$$A = \begin{pmatrix} \ddots & & -F \\ & D & \\ -E & & \ddots \end{pmatrix},$$

c'est à dire $A = D - E - F$ où

- $D = \text{diag}(a_{1,1}, \dots, a_{n,n})$ est la matrice diagonale formée par les éléments diagonaux de A .
- $-E = \text{tril}(A)$ est la matrice triangulaire inférieure formée par la partie triangulaire inférieure stricte de A .
- $-F = \text{triu}(A)$ est la matrice triangulaire supérieure formée par la partie triangulaire supérieure stricte de A .

De plus, nous supposons que les éléments diagonaux de A sont tous non nuls et ainsi D est inversible.

3.4.1 Méthode de Jacobi

Le schéma itératif de la méthode de Jacobi s'obtient en prenant $M = D$ et $N = E + F$. Ainsi, ce schéma s'écrit

$$\begin{cases} x^{(k+1)} &= B_J x^{(k)} + c_J \\ x^{(0)} &\text{donné ou à choisir,} \end{cases} \quad (\text{J})$$

avec $B_J = M^{-1}N = D^{-1}(E + F)$ et $c_J = D^{-1}b$.

Notons que l'égalité (J) entraîne que $Dx^{(k+1)} = (E + F)x^{(k)} + b$, et donc on a :

$$a_{i,i}x_i^{(k+1)} = - \sum_{\substack{j=1 \\ j \neq i}}^n a_{i,j}x_j^{(k)} + b_i.$$

Ainsi, le schéma itératif de la méthode de Jacobi s'écrit :

$$\begin{cases} x_i^{(k+1)} &= \frac{1}{a_{i,i}} \left(b_i - \sum_{\substack{j=1 \\ j \neq i}}^n a_{i,j}x_j^{(k)} \right), \quad \text{pour } i = 1, \dots, n \\ x^{(0)} &\text{donné ou à choisir,} \end{cases}$$

et la matrice B_J est telle que

$$B_J = \begin{pmatrix} 0 & -\frac{a_{1,2}}{a_{1,1}} & \cdots & -\frac{a_{1,n-1}}{a_{1,1}} & -\frac{a_{1,n}}{a_{1,1}} \\ -\frac{a_{2,1}}{a_{2,2}} & 0 & \cdots & -\frac{a_{2,n-1}}{a_{2,2}} & -\frac{a_{2,n}}{a_{2,2}} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ -\frac{a_{n-1,1}}{a_{n-1,n-1}} & -\frac{a_{n-1,2}}{a_{n-1,n-1}} & \cdots & 0 & -\frac{a_{n-1,n}}{a_{n-1,n-1}} \\ -\frac{a_{n,1}}{a_{n,n}} & -\frac{a_{n,2}}{a_{n,n}} & \cdots & -\frac{a_{n,n-1}}{a_{n,n}} & 0 \end{pmatrix}.$$

Exemple 3.6. Soit à résoudre le système

$$(\mathcal{S}_1) : \begin{cases} 4x_1 + x_2 = 3 \\ x_1 - 2x_2 = -15. \end{cases}$$

Pour obtenir le schéma itératif de la méthode de Jacobi associé au système (\mathcal{S}_1) ainsi que la matrice B_J et le vecteur c_J correspondants, il suffit d'écrire :

$$\begin{cases} 4x_1^{(k+1)} + x_2^{(k)} = 3 \\ x_1^{(k)} - 2x_2^{(k+1)} = -15. \end{cases}$$

Ce qui donne

$$\begin{cases} x_1^{(k+1)} = \frac{1}{4}(3 - x_2^{(k)}) \\ x_2^{(k+1)} = \frac{1}{2}(15 + x_1^{(k)}), \end{cases}$$

et ainsi $x^{(k+1)} = B_J x^{(k)} + c_J$, avec

$$B_J = \begin{pmatrix} 0 & -\frac{1}{4} \\ \frac{1}{2} & 0 \end{pmatrix} \quad \text{et} \quad c_J = \begin{pmatrix} \frac{3}{4} \\ \frac{15}{2} \end{pmatrix}.$$

Pour savoir si la méthode de Jacobi appliquée au système (\mathcal{S}_1) est convergente ou non, il suffit de déterminer le rayon spectrale de la matrice B_J . Et pour cela, nous pouvons calculer le polynôme caractéristique de B_J :

$$P_{B_J}(\lambda) = \det(B_J - \lambda I) = \begin{vmatrix} -\lambda & -\frac{1}{4} \\ \frac{1}{2} & -\lambda \end{vmatrix} = \lambda^2 + \frac{1}{8}.$$

Ce qui donne que les valeurs propres de B_J sont $\lambda_1 = \lambda_2 = \frac{i}{\sqrt{8}}$ et donc $\rho(B_J) = \frac{1}{\sqrt{8}} < 1$. Ainsi, la méthode de Jacobi appliquée au système (\mathcal{S}_1) est convergente pour tout $x^{(0)} \in \mathbb{R}^2$. En partant de $x^{(0)} = (0, 0)^T$, nous obtenons $x^{(1)} = (\frac{3}{4}, \frac{15}{2})^T$, $x^{(2)} = (-1.125, 7.875)^T$, ..., $x^{(10)} = (-1.000, 7.000)^T$. Et, on remarque alors que $x^{(k)} \rightarrow_{k \rightarrow +\infty} (-1, 7)^T$.

Exemple 3.7. Dans cet exemple, nous nous intéressons encore au système (\mathcal{S}_1) de l'exemple précédent que nous réécrivons comme suit

$$(\mathcal{S}_2) : \begin{cases} y_1 + 4y_2 = 3 \\ -2y_1 + y_2 = -15. \end{cases}$$

C'est à dire qu'on a $y_1 = x_2$ et $y_2 = x_1$.

En suivant la même démarche suivie pour l'exemple précédent, nous obtenons le schéma itératif associé au système (\mathcal{S}_2) ainsi que la matrice B_J et le vecteur c_J correspondants en écrivant

$$\begin{cases} y_1^{(k+1)} + 4y_2^{(k)} = 3 \\ -2y_1^{(k)} + y_2^{(k+1)} = -15, \end{cases}$$

et ainsi,

$$\begin{cases} y_1^{(k+1)} = 3 - 4y_2^{(k)} \\ y_2^{(k+1)} = -15 + 2y_1^{(k)}. \end{cases}$$

Ce qui donne $y^{(k+1)} = B'_J y^{(k)} + c'_J$, avec

$$B'_J = \begin{pmatrix} 0 & -4 \\ 2 & 0 \end{pmatrix} \quad \text{et} \quad c'_J = \begin{pmatrix} 3 \\ -15 \end{pmatrix}.$$

Et, en calculant le polynôme caractéristique de la matrice B'_J , nous trouvons que $P_{B'_J}(\lambda) = \lambda^2 - 8$, c'est à dire que $\rho(B'_J) = \sqrt{8} > 1$.

Ainsi, la méthode de Jacobi appliquée au système (\mathcal{S}_2) n'est pas convergente. Ce résultat est confirmé numériquement, puisqu'on constate que $x^{(k)} \rightarrow (\infty, \infty)^T$ lorsque $k \rightarrow +\infty$.

Nous terminons ce paragraphe en donnant l'algorithme de la méthode de Jacobi

Algorithme. Méthode de Jacobi.

1. *Initialisation :*

Choisir $x^{(0)}$ une solution de départ.

2. *Itération :*

Pour $k = 0, 1, \dots$ jusqu'à convergence

Pour $i = 1, \dots, n$

$$x_i^{(k+1)} = \frac{1}{a_{i,i}} \left(b_i - \sum_{\substack{j=1 \\ j \neq i}}^n a_{i,j} x_j^{(k)} \right).$$

Fin pour

Fin pour

3.4.2 Méthode de Gauss-Seidel

Le choix $M = D - E$ et $N = F$ permet de définir la méthode de Gauss-Seidel et dans ce cas le schéma itératif s'écrit

$$\begin{cases} x^{(k+1)} &= B_{GS}x^{(k)} + c_{GS} \\ x^{(0)} &\text{donné ou à choisir,} \end{cases} \quad (\text{GS})$$

avec $B_{GS} = M^{-1}N = (D - E)^{-1}F$ et $c_{GS} = (D - E)^{-1}b$.

Notons que dans la pratique, on ne calcule pas $(D - E)^{-1}$, mais on résout le système $(D - E)x^{(k+1)} = Fx^{(k)} + Fb$. Ce dernier système est un système triangulaire inférieure, donc facile à résoudre numériquement. On obtient ainsi,

$$\sum_{j=1}^i a_{i,j}x_j^{(k+1)} + \sum_{j=i+1}^n a_{i,j}x_j^{(k)} = b_i, \text{ pour } i = 1, \dots, n,$$

ou encore :

$$a_{i,i}x_i^{(k+1)} = -\sum_{j=1}^{i-1} a_{i,j}x_j^{(k+1)} - \sum_{j=i+1}^n a_{i,j}x_j^{(k)} + b_i, \text{ pour } i = 1, \dots, n.$$

Ceci permet d'obtenir le schéma itératif de la méthode de Gauss-Seidel qui s'écrit

$$\begin{cases} x_i^{(k+1)} &= \frac{1}{a_{i,i}} \left(b_i - \sum_{j=1}^{i-1} a_{i,j}x_j^{(k+1)} - \sum_{j=i+1}^n a_{i,j}x_j^{(k)} \right), \text{ pour } i = 1, \dots, n \\ x^{(0)} &\text{donné ou à choisir.} \end{cases}$$

Il est à remarquer que le calcul de $x_i^{(k+1)}$ la i -ème composante de la nouvelle solution approchée $x^{(k+1)}$ fait intervenir les valeurs des $x_j^{(k)}$ pour $j > i$ et des $x_j^{(k+1)}$ pour $j < i$. Il est donc nécessaire d'avoir calculé $x_1^{(k+1)}, \dots, x_{i-1}^{(k+1)}$ pour pouvoir calculer $x_i^{(k+1)}$. Ceci n'est pas nécessaire pour la méthode de Jacobi où le calcul de $x_i^{(k+1)}$ ne fait intervenir que les composantes de la solution approchée précédente $x^{(k)}$.

Exemple 3.8. Soit à résoudre le système

$$(\mathcal{S}_1) : \begin{cases} 4x_1 + x_2 = 3 \\ x_1 - 2x_2 = -15. \end{cases}$$

Pour obtenir le schéma itératif de Gauss-Seidel associé au système (\mathcal{S}_1) ainsi que la matrice B_{GS} et le vecteur c_{GS} correspondants, il suffit d'écrire :

$$\begin{cases} 4x_1^{(k+1)} + x_2^{(k)} = 3 \\ x_1^{(k+1)} - 2x_2^{(k+1)} = -15. \end{cases}$$

3.4. MÉTHODES ITÉRATIVES

Ce qui donne

$$\begin{cases} x_1^{(k+1)} &= \frac{1}{4}(3 - x_2^{(k)}) \\ x_2^{(k+1)} &= \frac{1}{2}(15 + x_1^{(k+1)}) = \frac{1}{2}(15 + \frac{1}{4}(3 - x_2^{(k)})) = \dots = \frac{63}{8} - \frac{1}{8}x_2^{(k)}, \end{cases}$$

et donc $x^{(k+1)} = B_{GS}x^{(k)} + c_{GS}$, avec

$$B_{GS} = \begin{pmatrix} 0 & -\frac{1}{4} \\ 0 & -\frac{1}{8} \end{pmatrix} \quad \text{et} \quad c_{GS} = \begin{pmatrix} \frac{3}{4} \\ \frac{63}{8} \end{pmatrix}.$$

Maintenant, calculons le polynôme caractéristique de B_{GS} pour déterminer son rayon spectrale.

$$P_{B_{GS}}(\lambda) = \det(B_{GS} - \lambda I) = \begin{vmatrix} -\lambda & -\frac{1}{4} \\ 0 & -\frac{1}{8} - \lambda \end{vmatrix} = \lambda\left(\frac{1}{8} + \lambda\right).$$

Ce qui donne que les valeurs propres de B_{GS} sont $\lambda_1 = -\frac{1}{8}$ et $\lambda_2 = 0$ et donc $\rho(B_{GS}) = \frac{1}{8} < 1$. Ainsi, la méthode de Gauss-Seidel appliquée au système (\mathcal{S}_1) est convergente pour tout $x^{(0)} \in \mathbb{R}^2$.

En partant de $x^{(0)} = (0, 0)^T$, nous obtenons $x^{(1)} = (3, -15)^T$, $x^{(2)} = (-1.218, 6.890)^T$, ..., $x^{(7)} = (-1, 7)^T$. Et, comme pour la méthode de Jacobi on remarque alors que $x^{(k)} \xrightarrow{k \rightarrow +\infty} (-1, 7)^T$. De plus, pour cet exemple, la méthode de Gauss-Seidel converge plus vite que la méthode de Jacobi puisque $\rho(B_{GS}) < \rho(B_J)$.

Exemple 3.9. Si on applique la méthode de Gauss-Seidel au système (\mathcal{S}_2) ci-dessous

$$(\mathcal{S}_2) : \begin{cases} y_1 + 4y_2 = 3 \\ -2y_1 + y_2 = -15. \end{cases}$$

on trouve que $y^{(k+1)} = B'_{GS}y^{(k)} + c'_{GS}$, avec

$$B'_{GS} = \begin{pmatrix} 0 & -4 \\ 0 & -8 \end{pmatrix} \quad \text{et} \quad c'_{GS} = \begin{pmatrix} -15 \\ -9 \end{pmatrix}.$$

De plus, on peut vérifier que $\rho(B'_{GS}) = 8 > 1$. Et donc la méthode de Gauss-Seidel (comme la méthode de Jacobi) appliquée au système (\mathcal{S}_2) n'est pas convergente.

Nous terminons ce paragraphe en donnant l'algorithme de la méthode de Gauss-Seidel.

Algorithme. Méthode de Gauss-Seidel.

1. *Initialisation :*

Choisir $x^{(0)}$ une solution de départ.

2. *Itération :*

Pour $k = 0, 1, \dots$ jusqu'à convergence

Pour $i = 1, \dots, n$

$$x_i^{(k+1)} = \frac{1}{a_{i,i}} \left(b_i - \sum_{j=1}^{i-1} a_{i,j} x_j^{(k+1)} - \sum_{j=i+1}^n a_{i,j} x_j^{(k)} \right).$$

Fin pour

Fin pour

3.4.3 Convergence des méthodes de Jacobi et de Gauss-Seidel

Dans cette section, nous allons donner des résultats concernant la convergence des méthodes de Jacobi et de Gauss-Seidel dans le cas où la matrice A est une matrice symétrique définie positive, ainsi que dans le cas où la matrice A est à diagonale strictement dominante.

Théorème 3.4. *Si la matrice A est symétrique définie positive, alors la méthode de Gauss-Seidel est convergente. (Attention : La méthode de Jacobi, ne converge pas forcément!)*

Théorème 3.5. *Si la matrice A est une matrice à diagonale strictement dominante, alors A est inversible et de plus les méthodes de Jacobi et de Gauss-Seidel sont convergentes.*

Enfin, on peut également établir le résultat suivant :

Théorème 3.6. *Soit A une matrice tridiagonale. Alors*

$$\rho(B_{GS}) = \rho(B_J)^2,$$

et donc les méthodes de Jacobi et Gauss-Seidel convergent ou divergent simultanément. Si elles convergent, la méthode de Gauss-Seidel est la plus rapide.