

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique



Université de Ghardaïa

Faculté des Sciences et Technologie
Département des Mathématiques et Informatique

Projet de fin d'étude présenté en vue de l'obtention du diplôme de

LICENCE

Domaine : Mathématiques et Informatique

Spécialité : Mathématiques

THEME:

*Quelques méthodes numériques pour la résolution des systèmes
linéaires et non linéaires*

PAR :

AMMI SAID Lakhdar

AKIF Zineb

Jury:

M: TELLAB Brahim

Maitre Assistant B Univ. Ghardaïa

Encadreur

M: ELHADJ MOUSSA Yacine

Maitre Assistant B Univ. Ghardaïa

Examineur

ANNEE UNIVERSITAIRE: 2013/2014

Remerciements

Nous souhaitons remercier dans un premier temps, toute l'équipe pédagogique du département des mathématiques et d'informatique et en particulier, le responsable de la formation : **Dr. Guerbati Kaddour**, et les membres de jury : l'encadreur **M. Tellab Brahim**, l'examineur **M. Elhadj Moussa Yacine** pour leur aide et le temps qu'ils nous ont consacrés.

Nous remercions également nos enseignants, et aussi les étudiants du troisième année mathématiques pour avoir animé l'année universitaire 2013/2014.

Nous n'oublions pas les membres de nos familles pour le soutien et la patience qu'ils nous ont témoigné.

Nous saluons tous les étudiants du département des mathématiques et d'informatique.

Merci

Ammi Said Lakhdar
Akif Zineb

Table des matières

Introduction générale	6
1 Notions de base sur l'algèbre linéaire	7
1.1 Espaces vectoriels et vecteurs	7
1.2 Matrices et vecteurs	8
1.2.1 Opérations sur les matrices	8
1.2.2 Trace, déterminant, rang d'une matrice	9
1.2.3 Matrices inversibles	10
1.2.4 Inverse d'une matrice	10
1.2.5 Transposée d'une matrice	11
1.2.6 Matrices symétriques, matrices anti-symétriques	11
1.2.7 Diagonalisation, trigonalisation	11
1.3 Produit scalaire et normes vectorielles	12
1.4 Normes matricielles	13
1.5 Matrices semblables	13
1.6 Matrices définies positives	13
2 Résolution des systèmes linéaires par des méthodes directes	15
2.1 Introduction	15
2.2 Méthode d'élimination de Gauss	15
2.3 Méthode de la décomposition LU	17
2.3.1 Existence	18
2.3.2 Algorithme	18
2.4 Factorisation de Cholesky	20
2.4.1 La complexité des méthodes	22
3 Résolution des systèmes linéaires par des méthodes itératives	23
3.1 Introduction	23
3.1.1 Convergence des méthodes itératives	23
3.2 Méthodes itératives	24
3.2.1 Méthode de Jaccobi	24
3.2.2 Méthode de Gauss-Seidel	25
3.2.3 Méthode de Relaxation	27
3.3 Étude de la convergence dans le cas d'une matrice symétrique définie positive	28
3.4 Étude de la convergence dans le cas d'une matrice à diagonale dominante	28

4	Résolution des équations et des systèmes non linéaires	31
4.1	Introduction	31
4.2	Quelques méthodes itératives classiques	33
4.2.1	Méthode de dichotomie	33
4.2.2	Méthode de la corde	33
4.2.3	Méthode de Lagrange	34
4.2.4	Méthode de Newton	34
4.3	La méthode du point fixe	35
4.4	Convergence de la méthode de Newton, de Lagrange et de la corde	35
4.4.1	Convergence de la méthode de Newton	35
4.4.2	Convergence de la méthode de Lagrange	37
4.4.3	Convergence de la méthode de la corde	38
4.5	Systèmes d'équations non linéaires	40
	Conclusion et perspectives	42
	Bibliographie	43

Introduction générale

Dans ce travail, nous s'intéressons à la résolution des systèmes linéaires, $Ax = b$, et des systèmes non linéaires de n équations à n inconnues à coefficients dans le corps \mathbb{K} . ($\mathbb{K} = \mathbb{R}$ ou \mathbb{C})

Nous distinguons deux catégories de méthodes de résolution d'un système linéaire, les méthodes directes et les méthodes itératives.

Si nous supposons une précision infinie, les méthodes directes conduisent à la solution exacte du système en un nombre fini d'étapes alors que les méthodes itératives donnent toujours une approximation de la solution, la solution exacte étant obtenue en un nombre infini d'étapes.

Les méthodes directes ont un caractère exact, mais pour les grands systèmes, la propagation des erreurs d'arrondis en diminue l'efficacité.

Les méthodes itératives sont bien adaptées au cas des matrices creuses (contient beaucoup de termes nuls), car ces méthodes ne transforment pas la matrice de départ, contrairement aux méthodes directes. Pour l'aspect programmation et pour plus d'information on pourra consulter [2] ou [4].

Ce mémoire est subdivisé en quatre chapitres

Premier chapitre : Dans ce chapitre, nous présentons quelques définitions, théorèmes et notions de base sur l'algèbre linéaire qui sont utilisés tout au long de ce mémoire.

Deuxième chapitre : Dans ce chapitre, nous intéressons à la résolutions des systèmes linéaires par des méthodes directes, et nous présentons particulièrement trois méthodes essentielles :

- Méthode d'élimination de Gauss.
- Méthode de la décomposition LU .
- Méthode d'élimination de Cholesky.

Troisième chapitre : Dans ce chapitre, nous présentons quelques méthodes itératives de résolutions des systèmes linéaires et nous examinons par la suite, la convergence de ces méthodes, nous étudions particulièrement trois méthodes :

- Méthode de Jaccobi.
- Méthode de Gauss-Seidel.
- Méthode de relaxation.

Quatrième chapitre : Dans ce chapitre, nous résumons quelques méthodes classiques de résolution des équations et des systèmes non linéaires, en particulier :

- Méthode de dichotomie.
- Méthode de la corde.
- Méthode de Lagrange.
- Méthode de Newton.

Nous donnons par la suite quelques conditions de convergence pour ces méthodes, et enfin, nous terminons notre mémoire par l'exposition de la méthode de Newton pour les systèmes d'équations non linéaires.

Chapitre 1

Notions de base sur l'algèbre linéaire

Dans ce chapitre, nous introduisons quelques définitions, notations et opérations de base sur l'algèbre des matrices, qui seront utiles tout au long de ce mémoire.

1.1 Espaces vectoriels et vecteurs

Définition 1.1.1 *Un espace vectoriel V sur un corps \mathbb{K} ($\mathbb{K}=\mathbb{R}$ ou \mathbb{C}) est un ensemble non vide muni deux lois, une loi interne notée $+$ et une loi externe notée \cdot , telles que pour tout $v, w \in V$ et $\lambda, \mu \in \mathbb{K}$ on a :*

- $(V, +)$ est un groupe commutatif
- $\lambda \cdot (v + w) = \lambda \cdot v + \lambda \cdot w$
- $(\lambda + \mu) \cdot v = \lambda \cdot v + \mu \cdot v$
- $\lambda \cdot (\mu \cdot v) = (\lambda \mu) \cdot v$
- $1 \cdot v = v$.

Exemple 1.1.1 $\mathcal{M}_{2,2}(\mathbb{R})$, ensemble de matrices à coefficients réels, est un espace vectoriel sur \mathbb{R} pour les deux lois de compositions définies par :

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} + \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} = \begin{pmatrix} a + a' & b + b' \\ c + c' & d + d' \end{pmatrix}$$
$$\lambda \cdot \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} \lambda a & \lambda b \\ \lambda c & \lambda d \end{pmatrix}$$

Remarque 1.1.1 *Dans toute la suite, on note \mathbb{R}^n un espace vectoriel de dimension n .*

La base canonique de \mathbb{R}^n est (e_1, \dots, e_n) .

Un vecteur $x \in \mathbb{R}^n$, de composantes x_1, \dots, x_n , est noté $x = (x_i)$.

Définition 1.1.2 *On dit qu'une partie non vide F de E est un sous-espace vectoriel de E si et seulement si*

$$\forall (v, w) \in F^2, \forall (\alpha, \beta) \in \mathbb{K}^2 : \alpha v + \beta w \in F.$$

En particulier, l'ensemble F des combinaisons linéaires d'une famille de p vecteurs de E , $\{v_1, \dots, v_p\}$, est un sous-espace vectoriel de E , appelé sous-espace engendré par la famille de vecteurs. On la note

$$\begin{aligned} F &= \text{vect}\{v_1, \dots, v_p\} \\ &= \{v = \alpha_1 v_1 + \dots + \alpha_p v_p, \alpha_i \in \mathbb{K}, i = 1, \dots, p\}. \end{aligned}$$

La famille $\{v_1, \dots, v_p\}$ est appelé famille génératrice de F .

1.2 Matrices et vecteurs

Définition 1.2.1 Soient m et n deux entiers positifs. On appelle matrice à m lignes et n colonnes ou matrice $m \times n$, à coefficients dans \mathbb{K} , un ensemble de mn scalaires $a_{ij} \in \mathbb{K}$, avec $i = 1, \dots, m$, $j = 1, \dots, n$ représentés dans le tableau rectangulaire suivant :

$$A = \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{pmatrix}$$

ou encore $A = (a_{ij})$ avec $1 \leq i \leq m$ et $1 \leq j \leq n$.

- L'ensemble des matrices $m \times n$ à coefficients dans \mathbb{K} est noté $\mathcal{M}_{m,n}(\mathbb{K})$.
- Si $\mathbb{K} = \mathbb{R}$, on écrit $A \in \mathbb{R}^{m \times n}$.
- Si $\mathbb{K} = \mathbb{C}$, on écrit $A \in \mathbb{C}^{m \times n}$.
- Si $n = m$, la matrice A est dite matrice carrée ou d'ordre n , de diagonale principale (a_{11}, \dots, a_{nn}) .
- Un vecteur ligne est une matrice n'ayant qu'une seule ligne, c'est-à-dire que $m = 1$ et un vecteur colonne est une matrice n'ayant qu'une seule colonne, c'est-à-dire que $n = 1$.

1.2.1 Opérations sur les matrices

- **Somme de deux matrices :** On appelle somme de deux matrices $A = (a_{ij})$ et $B = (b_{ij})$, la matrice $C(m \times n)$ dont les coefficients sont $C_{ij} = a_{ij} + b_{ij}$ où $i = 1, \dots, m$ et $j = 1, \dots, n$.
- **Multiplication d'une matrice par un scalaire :** La multiplication de A par un scalaire $\lambda \in \mathbb{K}$, est la matrice $C(m \times n)$ dont les coefficients sont donnés par $C_{ij} = \lambda a_{ij}$ où $i = 1, \dots, m$ et $j = 1, \dots, n$.
- **Produit de deux matrices :** Le produit d'une matrice $A(m, p)$ par une matrice $B(p, n)$ est la matrice $C(m, n)$, dont les coefficients sont donnés par :

$$c_{ij} = \sum_{k=1}^p a_{ik} b_{kj},$$

où $i = 1, \dots, m$ et $j = 1, \dots, n$.

Le produit matriciel est associatif et distributif par rapport à la somme matricielle mais il n'est pas commutatif en général. Si $AB = BA$, on dira que les matrices carrées A et B commutent.

Dans le cas des matrices carrées, l'élément neutre pour le produit matriciel est la matrice carrée d'ordre n , appelée matrice identité, définie par $I_n = (\delta_{ij})$.

La matrice identité est la seule matrice $n \times n$ qui vérifie $AI_n = I_n A = A$, pour toutes les matrices carrées A .

Remarque 1.2.1 • L'ensemble des matrices $\mathbb{R}^{m \times n}$ est un espace vectoriel de dimension mn .

- Le produit de deux matrices diagonales est une matrice diagonale.
- Le produit de deux matrices triangulaires supérieures (inférieures) est une matrice triangulaire supérieure (inférieure).

1.2.2 Trace, déterminant, rang d'une matrice

Définition 1.2.2 *A est une matrice carrée d'ordre n . On définit la trace de A comme la somme des éléments de la diagonale principale :*

$$\text{tr}(A) = \sum_{i=1}^n a_{ij}.$$

Proposition 1.2.1 *Si A et B sont deux matrices carrées d'ordre n , alors*

- $\text{tr}(A^T) = \text{tr}(A)$.
- $\text{tr}(A + B) = \text{tr}(A) + \text{tr}(B)$.
- *Si A est une matrice $m \times n$ et B une matrice $n \times m$, alors*

$$\text{tr}(A \times B) = \text{tr}(B \times A).$$

Définition 1.2.3 *Soit A une matrice carrée d'ordre n .*

Etant donné un couple (i, j) d'entiers, $1 \leq i, j \leq n$, on note (A_{ij}) la matrice carrée d'ordre $n - 1$ obtenue en supprimant la i -ème ligne et la j -ème colonne de A .

On définit le déterminant de A , et on le note $\det(A)$, par récurrence sur l'ordre de la matrice A .

- *Si $n = 1$: le déterminant de A est le nombre*

$$\det(A) = a_{11},$$

- *Si $n > 1$: le déterminant de A est le nombre*

$$\det(A) = \sum_{j=1}^n (-1)^{i+j} a_{ij} \det(A_{ij}) \quad \forall i, \quad 1 \leq i \leq n,$$

ou, de manière équivalente, le nombre

$$\det(A) = \sum_{i=1}^n (-1)^{i+j} a_{ij} \det(A_{ij}) \quad \forall j, \quad 1 \leq j \leq n.$$

Remarque 1.2.2 *(déterminant d'une matrice d'ordre 2)*

Soit A la matrice carrée d'ordre $n = 2$, alors

$$\det(A) = \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$

Remarque 1.2.3 *(règle de Sarrus)*

Soit A la matrice carrée d'ordre $n = 3$, alors

$$\begin{aligned} \det(A) &= \det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \\ &= (a_{11}a_{22}a_{33} + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{23}) - (a_{13}a_{22}a_{31} + a_{23}a_{32}a_{11} + a_{33}a_{12}a_{21}). \end{aligned}$$

Remarque 1.2.4 *le déterminant d'une matrice triangulaire est égal au produit des éléments diagonaux.*

Proposition 1.2.2 • $\det(A^T) = \det(A)$,

- $\det(A^{-1}) = \frac{1}{\det(A)}$
- $\det(A \times B) = \det(A) \cdot \det(B)$.

Définition 1.2.4 Le rang d'une matrice quelconque A est égal au plus grand entier s tel que l'on puisse extraire de A une matrice carrée s inversible, c'est-à-dire de déterminant non nul.

Exemple 1.2.1 Soit la matrice $A = \begin{pmatrix} 1 & 3 & 2 \\ 1 & 3 & 1 \end{pmatrix}$ Le rang de A est 2.

A est d'ordre 2×3 donc $s \leq \min 2, 3$ donc $s = 0, 1$ ou 2.

Comme le déterminant de la sous-matrice composée de la première et de la deuxième colonne est nul, on ne peut pas conclure,

comme le déterminant de la sous-matrice composée de la première et de la troisième colonne est non nul, alors $s = 2$.

Exemple 1.2.2 Soit la matrice $A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 5 & -1 \\ -1 & 0 & -1 \end{pmatrix}$.

A est d'ordre 3×3 donc $s \leq 3$.

Le déterminant de A est 0 donc $s \neq 3$.

le déterminant de la sous-matrice $\begin{pmatrix} 1 & 0 \\ 0 & 5 \end{pmatrix}$ est 5, donc $s = 2$.

1.2.3 Matrices inversibles

Définition 1.2.5 Une matrice carrée A d'ordre n est dite inversible (ou régulière ou non singulière) s'il existe une matrice carrée B d'ordre n telle que $AB = BA = I$. Dans ce cas B est la matrice inverse de A , notée A^{-1} . Une matrice qui n'est pas inversible est dite singulière.

Remarque 1.2.5 • Si A est inversible, son inverse est aussi inversible et on a : $(A^{-1})^{-1} = A$.

- Si A et B sont deux matrices inversibles d'ordre n , leur produit AB est aussi inversible et on a : $(AB)^{-1} = B^{-1}A^{-1}$.

Proposition 1.2.3 Une matrice carrée est inversible si et seulement si ses vecteurs colonnes sont linéairement indépendants.

1.2.4 Inverse d'une matrice

Si A est une matrice inversible d'ordre n , alors

$$A^{-1} = \frac{1}{\det(A)} C,$$

où C est la matrice de coefficients $c_{ij} = (-1)^{i+j} \det(A_{ji})$, $i = 1, \dots, n$, $j = 1, \dots, n$.

Par conséquent, une matrice carrée est inversible si et seulement si son déterminant est non nul. Dans le cas d'une matrice diagonale inversible, l'inverse est aussi une matrice diagonale ayant pour éléments les inverses des éléments de la matrice.

Toute matrice orthogonale A est inversible, son inverse est A^T et $\det(A) = \pm 1$.

1.2.5 Transposée d'une matrice

Définition 1.2.6 On appelle transposée d'une matrice $A \in \mathbb{R}^{m \times n}$ la matrice $n \times m$, notée A^T , obtenue en échangeant les lignes et les colonnes.

Proposition 1.2.4

$$\begin{aligned} \bullet (A^T)^T &= A & \bullet (A+B)^T &= A^T + B^T & \bullet (AB)^T &= B^T A^T \\ \bullet (\alpha A)^T &= \alpha A^T & \bullet (A^T)^{-1} &= (A^{-1})^T & \text{si } A \text{ est inversible.} \end{aligned}$$

1.2.6 Matrices symétriques, matrices anti-symétriques

Définition 1.2.7 Une matrice A est dite symétrique si $A^T = A$.

Une matrice A est dite anti-symétrique si $A^T = -A$.

Exemple 1.2.3

$$\begin{pmatrix} 1 & 5 & -9 \\ 5 & 4 & 0 \\ -9 & 0 & 7 \end{pmatrix} \text{ est une matrice symétrique.}$$

$$\begin{pmatrix} 1 & 5 & -9 \\ -5 & 4 & 0 \\ 9 & 0 & 7 \end{pmatrix} \text{ est une matrice anti-symétrique.}$$

Définition 1.2.8 Soit $A \in \mathbb{C}^{m \times n}$. La matrice $B = A^* \in \mathbb{C}^{n \times m}$ est appelée adjointe de A si $b_{ij} = \bar{a}_{ji}$, où \bar{a}_{ji} le conjugué de a_{ji} et on a :

- $(A+B)^* = A^* + B^*$
- $(AB)^* = B^* A^*$
- $(\alpha A)^* = \bar{\alpha} A^*, \forall \alpha \in \mathbb{C}$.

1.2.7 Diagonalisation, trigonalisation

Définition 1.2.9 On dit qu'une matrice $A \in \mathbf{M}_n(\mathbb{K})$, ($K = \mathbb{R}$ ou \mathbb{C}) est diagonalisable s'il existe une matrice inversible $P \in \mathbf{M}_n(\mathbb{K})$, et une matrice diagonale D telles que $P^{-1}AP = D$

- les valeurs propres de A sont les éléments de la diagonale de D et les vecteurs propres sont les vecteurs colonnes de P .

Définition 1.2.10 On dit qu'une matrice $A \in \mathbf{M}_n(\mathbb{K})$, $\mathbb{K} = \mathbb{R}$ ou \mathbb{C} est trigonalisable s'il existe une matrice inversible $P \in \mathbf{M}_n(\mathbb{K})$, et une matrice triangulaire T inférieure (les coefficients $a_{i,j} = 0$ si $i > j$) ou supérieure (les coefficients $a_{i,j} = 0$ si $i < j$) telles que $P^{-1}AP = T$

- les valeurs propre de A sont les éléments de la diagonale de T .

1.3 Produit scalaire et normes vectorielles

- Le produit scalaire de deux vecteurs x et y est définie par :

$$x^T y = \sum_{i=1}^n x_i y_i.$$

- Dans le cas de vecteurs complexes, le produit scalaire hermitien de deux vecteurs x et y est définie par :

$$x^* y = \sum_{i=1}^n \bar{x}_i y_i.$$

- Il est possible de définir plusieurs normes dans l'espace vectoriel \mathbb{R}^n .

Définition 1.3.1 Une norme d'un espace vectoriel E est une application $\| \cdot \|$ de E dans \mathbb{R} qui vérifie les propriétés suivantes :

- 1) $\forall x \in E, \| x \| \geq 0$
- 2) $\forall \lambda \in \mathbb{R}, \forall x \in E, \| \lambda x \| = | \lambda | \| x \|$
- 3) $\forall x, y \in E, \| x + y \| \leq \| x \| + \| y \|$
- 4) $\forall x \in E, \| x \| = 0 \iff x = 0.$

Dans \mathbb{R}^n , les trois normes les plus courantes sont :

- Norme infinie : $\| x \|_\infty = \max_{i=1, \dots, n} | x_i |,$
- Norme 1 : $\| x \|_1 = \sum_{i=1}^n | x_i |,$
- Norme 2 ou norme euclidienne : $\| x \|_2 = \left(\sum_{i=1}^n x_i^2 \right)^{\frac{1}{2}},$ Où $x = (x_1, \dots, x_n)^T.$

Proposition 1.3.1 (Inégalité de Cauchy-Schwarz)

$$\forall (x, y) \in \mathbb{R}^n \times \mathbb{R}^n : | x^T y | \leq \| x \|_2 \| y \|_2.$$

L'égalité a lieu si et seulement si les vecteurs x et y sont liés.

Comme toutes les normes des espaces vectoriels de dimension finie, ces trois normes sont équivalentes.

Les constantes qui les relient sont données par la proposition suivante :

Proposition 1.3.2 Pour tout $x \in \mathbb{R}^n$, on a :

$$\begin{aligned} \| x \|_2 &\leq \| x \|_1 \leq \sqrt{n} \| x \|_2, \\ \| x \|_\infty &\leq \| x \|_2 \leq \sqrt{n} \| x \|_\infty, \\ \| x \|_\infty &\leq \| x \|_1 \leq n \| x \|_\infty. \end{aligned}$$

1.4 Normes matricielles

Définition 1.4.1 On suppose que l'on a choisi une norme dans chacun des deux espaces \mathbb{R}^n et \mathbb{R}^m . On définit alors la norme matricielle subordonnée dans l'espace des matrices $\mathbb{R}^{m \times n}$ par :

$$\forall A \in \mathbb{R}^{m \times n}, \|A\| = \max_{\|x\|=1} \|Ax\|.$$

Lorsque les normes 1, 2 ou infinie sont choisies respectivement pour les deux ensembles à la fois, on note les normes subordonnées correspondantes de la même manière.

Proposition 1.4.1 Soit $A = (a_{ij}) \in \mathbb{R}^{m \times n}$. Alors :

$$\|A\|_1 = \max_{j=1, \dots, n} \sum_{i=1}^m |a_{ij}|,$$

$$\|A\|_\infty = \max_{i=1, \dots, m} \sum_{j=1}^n |a_{ij}|,$$

Remarque 1.4.1 La norme matricielle euclidienne n'est pas facile à calculer dans le cas général, contrairement aux autres normes.

1.5 Matrices semblables

Définition 1.5.1 On dit que deux matrices A et B sont semblables si et seulement s'il existe une matrice régulière P telle que : $A = P^{-1}BP$

1.6 Matrices définies positives

Définition 1.6.1 (mineurs principaux) Soit $A \in M_n(\mathbb{R})$.

Une sous-matrice $k \times k$ formée à partir de A , en éliminant $n - k$ colonnes, disons les colonnes i_1, i_2, \dots, i_{n-k} et les mêmes $n - k$ lignes i_1, i_2, \dots, i_{n-k} est appelée une sous-matrice de A , d'ordre principal k . Le déterminant d'une sous-matrice principale $k \times k$ est appelée le mineur principal d'ordre k de la matrice A .

Exemple 1.6.1 Soit A une matrice 3×3

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$

- Les mineurs principaux de premier ordre sont : a_{11}, a_{22} et a_{33} .
- Les mineurs principaux de second ordre sont :

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \det \begin{pmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{pmatrix} \text{ et } \det \begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix}.$$

- Le seul mineur principal de troisième ordre est : $\det(A)$.

Définition 1.6.2 (*mineurs principaux dominants*) Soit $A \in M_n(\mathbb{R})$.

Une sous-matrice $k \times k$ formée à partir de A , en éliminant les $n - k$ dernières colonnes, et les mêmes $n - k$ dernières lignes, est appelée une sous-matrice de A , d'ordre principal dominant k . Le déterminant d'une sous-matrice principale dominante $k \times k$ est appelée le mineur principal dominant d'ordre k de la matrice A .

Exemple 1.6.2 Soit A une matrice 3×3

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$

- Le seul mineur principal dominant de premier ordre est : a_{11} .
- Le seul mineur principal dominant de second ordre est : $\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$
- Le seul mineur principal dominant de troisième ordre est : $\det(A)$.

Définition 1.6.3 Soit A une matrice symétrique définie positive d'ordre n . Elle est dite définie positive si elle vérifie la propriété suivante :

$$\forall x \in \mathbb{K}^n, \quad x^T A x > 0.$$

Théorème 1.6.1 (*critère de Sylvester*) Pour qu'une matrice $A = (a_{ij})_{1 \leq i, j \leq n}$, réelle symétrique ou complexe hermitienne, soit définie positive, il faut et il suffit que les n mineurs principaux dominants soient strictement positifs.

Remarque 1.6.1 (*intérêt des matrices définies positives*)

- Les problèmes de résolution de systèmes linéaires les plus faciles à traiter numériquement sont ceux dont les matrices sont symétriques définies positives.

Chapitre 2

Résolution des systèmes linéaires par des méthodes directes

2.1 Introduction

On se propose de résoudre l'équation matricielle $Ax = b$ où A est une matrice inversible de taille n . Lorsque A est triangulaire supérieure, la résolution de l'équation $Ax = b$ est immédiate. On calcule successivement X_n à partir de la dernière équation, puis X_{n-1} à partir de l'avant-dernière et ainsi de suite jusqu'à x_1 .

On procède d'une manière analogue si A est triangulaire inférieure.

2.2 Méthode d'élimination de Gauss

Supposons que A n'est pas triangulaire. L'idée de la méthode d'élimination de Gauss (ou pivot de Gauss ou trigoalisation de Gauss) est de se ramener à un système linéaire $A'x = b'$ dont la matrice A' est triangulaire supérieure, on obtient ensuite la solution par simple remontée.

Exemple 2.2.1 On cherche à résoudre le système suivant :

$$\begin{cases} 2x + 3y + 3z + t = 15 \\ -4x - 6y + 3z + 2t = 3 \\ -x + y + z + t = 5 \\ -2x - y + z + t = 1 \end{cases} \quad (2.1)$$

La méthode de Gauss, consiste à éliminer x des lignes 2, 3 et 4, puis y des lignes 3 et 4, puis z de la ligne 4. On obtient alors la valeur de t et on en déduit les autres valeurs en remontant. Le système (2.1) s'écrit sous la forme matricielle $Ax = b$:

$$\begin{pmatrix} 2 & 3 & 3 & 1 \\ -4 & -6 & 3 & 2 \\ -1 & 1 & 1 & 1 \\ -2 & -1 & 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} 15 \\ 3 \\ 5 \\ 1 \end{pmatrix}$$

Étape 1

On pose $A_1 = A$ et on note $a_{ij}^1, 1 \leq i, j \leq 4$ l'élément (i, j) de la matrice A . Lorsqu'il est non nul,

on appelle pivot l'élément a_{11}^1 . Dans notre exemple $a_{11}^1 = 2$. On effectue alors pour toute ligne L_i , $2 \leq i \leq 4$:

$$L_i \leftarrow L_i - \frac{a_{i1}^1}{a_{11}^1} L_1$$

on arrive à :

$$\begin{pmatrix} 2 & 3 & 3 & 1 \\ 0 & 0 & 9 & 4 \\ 0 & \frac{5}{2} & \frac{5}{2} & \frac{3}{2} \\ 0 & 2 & 4 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} 15 \\ 33 \\ \frac{25}{2} \\ 16 \end{pmatrix}.$$

Étape 2

On note A_2 la matrice obtenue à l'étape 1. Maintenant, le pivot a_{22}^2 de la nouvelle matrice est nul. Alors on échange la deuxième et la troisième ligne. On obtient :

$$\begin{pmatrix} 2 & 3 & 3 & 1 \\ 0 & \frac{5}{2} & \frac{5}{2} & \frac{3}{2} \\ 0 & 0 & 9 & 4 \\ 0 & 2 & 4 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} 15 \\ \frac{25}{2} \\ 33 \\ 16 \end{pmatrix}.$$

Le nouveau pivot maintenant est a_{22}^2 . On effectue alors :

$$L_4 \leftarrow L_4 - \frac{a_{42}^2}{a_{22}^2} L_2 = L_4 - \frac{4}{5} L_2.$$

Ce qui nous donne :

$$\begin{pmatrix} 2 & 3 & 3 & 1 \\ 0 & \frac{5}{2} & \frac{5}{2} & \frac{3}{2} \\ 0 & 0 & 9 & 4 \\ 0 & 0 & 2 & \frac{4}{5} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} 15 \\ \frac{25}{2} \\ 33 \\ 6 \end{pmatrix}.$$

Étape 3

Le pivot de la matrice A_3 est l'élément $a_{33}^3 = 9$. On effectue :

$$L_4 \leftarrow L_4 - \frac{a_{43}^3}{a_{33}^3} L_3 = L_4 - \frac{2}{9} L_3.$$

Ce qui nous donne :

$$\begin{pmatrix} 2 & 3 & 3 & 1 \\ 0 & \frac{5}{2} & \frac{5}{2} & \frac{3}{2} \\ 0 & 0 & 9 & 4 \\ 0 & 0 & 0 & -\frac{4}{25} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} 15 \\ \frac{25}{2} \\ 33 \\ -\frac{4}{3} \end{pmatrix}.$$

Étape 4

La matrice obtenue est triangulaire supérieure, et on peut résoudre le système par remontée. On trouve :

$$\begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} 6 \\ -1 \\ -3 \\ 15 \end{pmatrix}.$$

Algorithme :

De l'exemple précédent on tire l'algorithme suivant :

Étape 1

Si $a_{11} \neq 0$, alors :

La ligne 1 ne change pas. et pour $i=2,\dots,n$ on, effectue : $L_i \leftarrow L_i - \frac{a_{i1}}{a_{11}}L_1$ et on a :

$$\begin{cases} a_{ij} = a_{ij} - \frac{a_{i1}}{a_{11}}a_{1j}, & i = 2, \dots, n, \quad j = 1, \dots, n. \\ b_i = b_i - \frac{a_{i1}}{a_{11}}b_1, & i = 2, \dots, n \end{cases}$$

Étape 2

Si $a_{22} \neq 0$, alors :

La ligne 1 et 2 ne changent pas. et pour $i=3,\dots,n$, on effectue : $L_i \leftarrow L_i - \frac{a_{i2}}{a_{22}}L_2$ et on a :

$$\begin{cases} a_{ij} = a_{ij} - \frac{a_{i2}}{a_{22}}a_{2j}, & i = 3, \dots, n, \quad j = 1, \dots, n. \\ b_i = b_i - \frac{a_{i2}}{a_{22}}b_2, & i = 3, \dots, n \end{cases}$$

⋮

Étape k

Si $a_{kk} \neq 0$, alors :

$$\begin{cases} a_{ij} = a_{ij} - \frac{a_{ik}}{a_{kk}}a_{kj}, & i = k + 1, \dots, n, \quad j = 1, \dots, n. \\ b_i = b_i - \frac{a_{ik}}{a_{kk}}b_k, & i = k + 1, \dots, n \end{cases}$$

Après ces étapes on obtient un système $A'x = b'$ (avec A' est triangulaire supérieure) équivalent au système $Ax = b$.

Remarque 2.2.1 Si $a_{kk} = 0$, on utilise la stratégie de pivot (soit partiel ou total), qui est basée sur le changement des lignes.

Pour plus d'information, on pourra consulter [1].

2.3 Méthode de la décomposition LU

La méthode de décomposition LU consiste à factoriser la matrice A en produit de deux matrices triangulaires

$$A = LU$$

où L est triangulaire inférieure (lower) et U est triangulaire supérieure (upper). La résolution du système linéaire $Ax = b$ est alors ramenée à la résolution de deux systèmes $Ly = b$ et $Ux = y$.

Dans la méthode de Gauss pour éliminer les éléments de la colonne k ($k^{\text{ième}}$ étape), on multiplie la matrice A à droite par une matrice L_k nulle sauf la colonne k et on a :

$$l_{k1} = 1, \quad l_{kk} = -\frac{a_{ik}}{a_{kk}}.$$

Si on sauvegarde les matrice L_k on peut définir à la fin d'élimination une matrice triangulaire supérieure L .

Et donc on définit une nouvelle méthode directe de résolution des systèmes linéaires qui s'appelle méthode de décomposition LU.

Ceci est intéressant lorsqu'on a à résoudre plusieurs systèmes avec la même matrice et différents second membres car la décomposition est effectuée une fois pour toutes.

2.3.1 Existence

Théorème 2.3.1 Soit $A \in \mathbb{R}^{n \times n}$. La factorisation LU de A avec L est une matrice triangulaire inférieure dont les éléments diagonaux sont égaux à 1 et U est une matrice triangulaire supérieure est unique si et seulement si tous les mineurs principaux dominants de A sont non nuls.

Preuve : Pour la démonstration, on pourra consulter [1].

■

2.3.2 Algorithme

Soit la matrice :

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{n1} & & \cdots & a_{nn} \end{pmatrix} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ l_{21} & 1 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ l_{n1} & \cdots & l_{nn-1} & 1 \end{pmatrix} \cdot \begin{pmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ 0 & u_{22} & \cdots & u_{2n} \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & u_{nn} \end{pmatrix}$$

Calcul des matrices L et U .

• Étape 1

$$\begin{cases} u_{1i} = a_{1i} & i = 1, \dots, n \\ l_{i1} = \frac{a_{i1}}{u_{11}}, & i = 2, \dots, n \end{cases}$$

• Étape 2

$$\begin{cases} u_{2i} = a_{2i} - l_{21} \times u_{1i}, & i = 2, \dots, n \\ l_{i2} = \frac{a_{i2} - l_{i1} \times u_{12}}{u_{22}}, & i = 3, \dots, n \end{cases}$$

⋮

• Étape k

$$\begin{cases} u_{kj} = a_{kj} - \sum_{i=1}^{j-1} l_{ki} \times u_{ik}, & j = k, \dots, n \\ l_{ik} = \frac{1}{u_{kk}} \times (a_{ik} - \sum_{j=1}^{k-1} l_{ij} \times u_{jk}), & i = (k+1), \dots, n \end{cases}$$

L'étape k est vérifiée pour $k = 2, \dots, n-1$, et pour $k = n$ on trouve juste u_{nn} .

Exemple 2.3.1 On considère le système suivant :

$$\begin{cases} x + y + 3z + 3t = 15 \\ x + 2y + 4z + 2t = 14 \\ 6x + 9y + 7z = 28 \\ 3x + 5y + 8z + t = 22 \end{cases} \quad (2.2)$$

Le système (2.2) s'écrit sous la forme matricielle de la façon suivante :

$$\begin{pmatrix} 1 & 1 & 3 & 3 \\ 1 & 2 & 4 & 2 \\ 6 & 9 & 7 & 0 \\ 3 & 5 & 8 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} 15 \\ 14 \\ 28 \\ 22 \end{pmatrix}$$

c'est-à-dire de la forme

$$Ax = b$$

On vérifie facilement que les mineurs principaux de la matrice A sont non nuls, donc A vérifie les conditions du théorème 3, et par la suite on peut appliquer la méthode de factorisation LU à la résolution du système (2.2).

On a :

$$\begin{pmatrix} 1 & 1 & 3 & 3 \\ 1 & 2 & 4 & 2 \\ 6 & 9 & 7 & 0 \\ 3 & 5 & 8 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ l_{21} & 1 & 0 & 0 \\ l_{31} & l_{32} & 1 & 0 \\ l_{41} & l_{42} & l_{43} & 1 \end{pmatrix} \cdot \begin{pmatrix} u_{11} & u_{12} & u_{13} & u_{14} \\ 0 & u_{22} & u_{23} & u_{24} \\ 0 & 0 & u_{33} & u_{34} \\ 0 & 0 & 0 & u_{44} \end{pmatrix}$$

• $k = 1$

$$\begin{cases} u_{1i} = a_{1i}; & i = 1, \dots, 4 \\ l_{i1} = \frac{a_{i1}}{u_{11}}; & i = 2, \dots, 4 \end{cases}$$

Donc :

$$\begin{cases} l_{21} = \frac{a_{21}}{u_{11}} = \frac{1}{1} = 1 \\ l_{31} = \frac{a_{31}}{u_{11}} = \frac{6}{1} = 6 \\ l_{41} = \frac{a_{41}}{u_{11}} = \frac{3}{1} = 3 \end{cases}$$

Alors, on peut écrire

$$\begin{pmatrix} 1 & 1 & 3 & 3 \\ 1 & 2 & 4 & 2 \\ 6 & 9 & 7 & 0 \\ 3 & 5 & 8 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 6 & l_{32} & 1 & 0 \\ 3 & l_{42} & l_{43} & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 & 3 & 3 \\ 0 & u_{22} & u_{23} & u_{24} \\ 0 & 0 & u_{33} & u_{34} \\ 0 & 0 & 0 & u_{44} \end{pmatrix}$$

• $k = 2$

$$\begin{cases} u_{22} = a_{22} - l_{21} \times u_{12} = 1 \\ u_{23} = a_{23} - l_{21} \times u_{13} = 1 \\ u_{24} = a_{24} - l_{21} \times u_{14} = -1 \\ l_{32} = \frac{a_{32} - l_{31} \times u_{12}}{u_{22}} = 3 \\ l_{42} = \frac{a_{42} - l_{41} \times u_{12}}{u_{22}} = 2 \end{cases}$$

Alors la décomposition devient :

$$\begin{pmatrix} 1 & 1 & 3 & 3 \\ 1 & 2 & 4 & 2 \\ 6 & 9 & 7 & 0 \\ 3 & 5 & 8 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 6 & 3 & 1 & 0 \\ 3 & 2 & l_{43} & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 & 3 & 3 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & u_{33} & u_{34} \\ 0 & 0 & 0 & u_{44} \end{pmatrix}$$

- $k = 3$

$$\begin{cases} u_{33} = a_{33} - \sum_{i=1}^2 l_{3i} \times u_{i3} = -14 \\ u_{34} = a_{34} - \sum_{i=1}^2 l_{3i} \times u_{i4} = -15 \\ l_{43} = \frac{1}{u_{33}} \times (a_{43} - \sum_{j=1}^2 l_{4j} \times u_{j3}) = \frac{3}{14} \end{cases}$$

- $k = 4$

$$u_{44} = a_{44} - \sum_{i=1}^3 l_{4i} \times u_{i4} = -\frac{39}{14}.$$

Donc le système (2.2) devient

$$LUx = b, \tag{2.3}$$

avec :

$$L = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 6 & 3 & 1 & 0 \\ 3 & 2 & \frac{3}{14} & 1 \end{pmatrix} \quad \text{et} \quad U = \begin{pmatrix} 1 & 1 & 3 & 3 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & -14 & -15 \\ 0 & 0 & 0 & -\frac{39}{14} \end{pmatrix}.$$

Le système (2.3) est équivalent au système :

$$\begin{cases} Ly = b \\ Ux = y \end{cases} \tag{2.4}$$

La résolution du système (2.4), nous donne

$$x = \begin{pmatrix} 2 \\ 1 \\ 1 \\ 3 \end{pmatrix}.$$

Donc la solution du système (2.2) est

$$x = \begin{pmatrix} 2 \\ 1 \\ 1 \\ 3 \end{pmatrix}.$$

2.4 Factorisation de Cholesky

La méthode de Cholesky ne s'applique qu'aux matrices réelles symétriques définies positives. Elle consiste en une factorisation $A = BB^T$, où B est une matrice triangulaire inférieure. La résolution du système $Ax = b$ peut se ramener à la résolution du système

$$\begin{cases} By = b \\ B^T x = y \end{cases}$$

Ceci est intéressant lorsqu'on a à résoudre plusieurs systèmes avec la même matrice et différents second membres car la décomposition est effectuée une fois pour toutes.

Théorème 2.4.1 Soit $A \in \mathbb{R}^{n \times n}$ une matrice symétrique définie positive. Il existe une unique matrice réelle B triangulaire inférieure, telle que tous ses éléments diagonaux soient positifs et qui vérifie

$$A = BB^T.$$

Preuve : Pour la démonstration, on pourra consulter [5].

Algorithme :

Soit la matrice :

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn-1} & a_{nn} \end{pmatrix} = \begin{pmatrix} l_{11} & 0 & \cdots & 0 \\ l_{21} & l_{22} & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ l_{n1} & \cdots & l_{nn-1} & l_{nn} \end{pmatrix} \cdot \begin{pmatrix} l_{11} & l_{12} & \cdots & l_{1n} \\ 0 & l_{22} & \cdots & l_{2n} \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & l_{nn} \end{pmatrix}$$

on a :

$$\begin{cases} L_{11} = \sqrt{a_{11}} \\ L_{ij} = (a_{ij} - \sum_{k=1}^{j-1} L_{ik}L_{kj})L_{jj}, \quad j = 1, \dots, i-1 \\ L_{ii} = (a_{ij} - \sum_{k=1}^{j-1} L_{ik})^{\frac{1}{2}} \end{cases}$$

Exemple 2.4.1 On considère la matrice

$$A = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \end{pmatrix}$$

La factorisation de Cholesky de A est :

$$A = \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{2} & \frac{\sqrt{3}}{6} & 0 \\ \frac{1}{3} & \frac{\sqrt{3}}{6} & \frac{\sqrt{5}}{30} \end{pmatrix} \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ 0 & \frac{\sqrt{3}}{6} & \frac{\sqrt{3}}{6} \\ 0 & 0 & \frac{\sqrt{5}}{30} \end{pmatrix} = BB^T$$

$$\det(A) = \left(\frac{\sqrt{3}}{6} \cdot \frac{\sqrt{5}}{30} \right)^2 = \frac{1}{2160}.$$

Pour résoudre le système $Ax = b$, avec $b = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.

On a :

$$By = b \implies y = \begin{pmatrix} 1 \\ \sqrt{3} \\ \sqrt{5} \end{pmatrix},$$

Puis ,

$$B^T x = y \implies x = \begin{pmatrix} 3 \\ -24 \\ 30 \end{pmatrix}.$$

2.4.1 La complexité des méthodes

Pour une matrice d'ordre n , la méthode de Gauss nécessite $\frac{2}{3}n^3 + \frac{3}{2}n^2 - \frac{7}{6}n$ opérations élémentaires, mais la décomposition LU demande $\frac{1}{3}n^3 + n^2 - \frac{1}{3}n$ opérations élémentaires. Pour la méthode de Cholesky on utilise $\frac{1}{3}n^3 + \frac{1}{2}n^2 + \frac{1}{2}n$ opérations élémentaires.

Chapitre 3

Résolution des systèmes linéaires par des méthodes itératives

3.1 Introduction

Les méthodes directes sont très coûteuses en termes de nombre d'opérations, donc de temps de calcul lorsque la taille du système linéaire est assez élevée. On peut alors avoir recours à des méthodes où la solution est obtenue par itérations successives et où chaque itération consiste à résoudre un système linéaire moins coûteux.

On peut écrire de manière générique une classe de méthodes itératives de la manière suivante :
Considérons le système linéaire

$$Ax = b$$

où A est une matrice $n \times n$ inversible, et soit une décomposition de A sous la forme

$$A = M - N$$

où M est une matrice inversible, facile à inverser. En se donnant un vecteur $x^0 \in \mathbb{R}^n$, on construit une suite de vecteurs $(x^{(k)})$ par itérations

$$Mx^{(k+1)} = Nx^{(k)} + b, \quad k = 0, 1, \dots \quad (3.1)$$

Notons que si cette méthode converge vers un vecteur $x \in \mathbb{R}^n$ alors on a nécessairement

$$Mx = Nx + b,$$

c'est-à-dire

$$Ax = b.$$

3.1.1 Convergence des méthodes itératives

Soit $x \in \mathbb{R}^n$, On définit la norme :

$$\|x\| := \left(\sum_{i=1}^n x_i^2 \right)^{\frac{1}{2}}.$$

On dit que la méthode itérative (3.1) converge s'il existe un vecteur $x \in \mathbb{R}^n$ tel que

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

pour tout choix de vecteur initial $x^0 \in \mathbb{R}^n$.

Définition 3.1.1 Soit A une matrice carrée d'ordre n . On appelle rayon spectral de A et on le note $\rho(A)$, la plus grande valeur propre en module, i.e.

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

On a alors le théorème fondamental suivant :

Théorème 3.1.1 La méthode itérative (3.1) converge si et seulement si

$$\rho(M^{-1}N) < 1.$$

Preuve : Pour la démonstration, on pourra consulter [5].

■

3.2 Méthodes itératives

Dans cette section, on s'intéresse d'exposer trois méthodes itératives telles que : méthode de Jacobi, méthode de Gauss-seidel et méthode de Relaxation. Pour cela nous écrivons A sous la forme

$$A = L + D + U$$

où

$$d_{ij} = \begin{cases} a_{ii} & \text{si } i = j \\ 0 & \text{si non} \end{cases}, \quad l_{ij} = \begin{cases} a_{ij} & \text{si } i > j \\ 0 & \text{si non} \end{cases}, \quad u_{ij} = \begin{cases} a_{ij} & \text{si } i < j \\ 0 & \text{si non} \end{cases}.$$

On suppose en outre que la matrice diagonale D est inversible, autrement dit $a_{ii} \neq 0$ pour tout $1 \leq i \leq n$.

3.2.1 Méthode de Jacobi

Dans cette méthode nous choisissons

$$M = D, \quad N = -(L + U).$$

Alors la méthode s'écrit

$$Dx^{(k+1)} := -(L + U)x^{(k)} + b, \quad k = 0, 1, \dots$$

i.e.

$$x^{(k+1)} := -D^{-1}(L + U)x^{(k)} + D^{-1}b, \quad k = 0, 1, \dots$$

ou encore

$$\begin{cases} x_i^{(k+1)} := \frac{1}{a_{ii}} \left(b_i - \sum_{j \neq i} a_{ij} x_j^{(k)} \right) & 1 \leq j \leq n, k = 0, 1, \dots \\ x_i^{(0)} \text{ donné} & 1 \leq i \leq n. \end{cases}$$

Chaque itération consiste donc à résoudre un système linéaire diagonal.

Exemple 3.2.1 On considère le système suivant :

$$\begin{pmatrix} 4 & 2 & 1 \\ -1 & 2 & 0 \\ 2 & 1 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 4 \\ 2 \\ 9 \end{pmatrix}. \quad (3.2)$$

Le système peut se mettre sous la forme :

$$\begin{cases} 4x + 2y + z = 4 \\ -x + 2y = 2 \\ 2x + y + 4z = 9 \end{cases}$$

ce qui nous donne :

$$\begin{cases} x = 1 - \frac{y}{2} - \frac{z}{4} \\ y = 1 + \frac{x}{2} \\ z = \frac{9}{4} - \frac{x}{2} - \frac{y}{4} \end{cases}$$

Soit $x_0(0, 0, 0)$ le vecteur initial, en calculant les itérées on trouve :

$$\begin{aligned} x^1 &= \frac{1}{a_{11}} \left(b_1 - a_{12}y^0 - a_{13}z^0 \right) = \frac{4}{4} = 1 \\ y^1 &= \frac{1}{a_{22}} \left(b_2 - a_{21}x^0 - a_{23}z^0 \right) = \frac{2}{2} = 1 \\ z^1 &= \frac{1}{a_{33}} \left(b_3 - a_{31}x^0 - a_{32}y^0 \right) = \frac{9}{4}. \end{aligned}$$

Donc $x_1 = (1, 1, \frac{9}{4})$. D'une manière analogue on trouve :

$$\begin{aligned} x_2 &= \left(-\frac{1}{16}, \frac{3}{2}, \frac{3}{2} \right) \\ x_3 &= \left(-\frac{1}{8}, -\frac{1}{32}, \frac{61}{32} \right) \\ x_4 &= \left(\frac{5}{128}, \frac{15}{16}, \frac{265}{128} \right) \\ x_5 &= \left(\frac{7}{512}, \frac{261}{256}, \frac{511}{256} \right). \end{aligned}$$

La suite (x_k) converge donc vers la solution exacte $(0, 1, 2)$ du système (3.2).

3.2.2 Méthode de Gauss-Seidel

Dans la méthode de Gauss-Seidel, on choisit $M = D + L$ et $N = -U$, donc cette méthode peut être écrite sous la forme :

$$(D + L)x^{(k+1)} = -Ux^{(k)} + b, \quad k = 0, 1, \dots$$

ou encore

$$x^{(k+1)} = -(D + L)^{-1}Ux^{(k)} + (D + L)^{-1}b, \quad k = 0, 1, \dots$$

La méthode de Jacobi montre que, pour calculer les itérés $x_i^{(k+1)}$, les valeurs $x_j^{(k+1)}$, sont disponibles pour $1 \leq j \leq n$. En modifiant cette méthode, on trouve :

$$\begin{cases} x_i^{(k+1)} := \frac{1}{a_{ii}} \left(b_i - \sum_{j < i} a_{ij}x_j^{(k+1)} - \sum_{j > i} a_{ij}x_j^{(k)} \right) & 1 \leq j \leq n, \quad k = 0, 1, \dots \\ x_i^{(0)} \text{ donné} & 1 \leq i \leq n. \end{cases}$$

Exemple 3.2.2 *Considérons le système suivant :*

$$\begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{4} \\ -\frac{1}{2} & 1 & 0 \\ \frac{1}{2} & \frac{1}{4} & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ \frac{9}{4} \end{pmatrix}, \quad (3.3)$$

i.e.

$$Ax = b.$$

Le système (3.3) nous donne :

$$\begin{cases} x = 1 - \frac{y}{2} - \frac{z}{4} \\ y = 1 + \frac{x}{2} \\ z = \frac{9}{4} - \frac{x}{2} - \frac{y}{4} \end{cases}$$

en partant du vecteur initial $x_0(0, 0, 0)$ on calcule x_1 :

$$\begin{aligned} x^1 &= \frac{1}{a_{11}} \left(b_1 - a_{12}y^0 - a_{13}z^0 \right) = \frac{1}{1} = 1 \\ y^1 &= \frac{1}{a_{22}} \left(b_2 - a_{21}x^1 - a_{23}z^0 \right) = \frac{3}{2} \\ z^1 &= \frac{1}{a_{33}} \left(b_3 - a_{31}x^1 - a_{33}y^1 \right) = \frac{11}{8}. \end{aligned}$$

Donc

$$x_1 = \left(1, \frac{3}{2}, \frac{11}{8} \right)$$

On calcule de la même façon

$$\begin{aligned} x_2 &= \left(-\frac{3}{32}, \frac{61}{64}, \frac{527}{256} \right) \\ x_3 &= \left(\frac{9}{1024}, \frac{2047}{2048}, \frac{16349}{8192} \right). \end{aligned}$$

Cette suite de point (x_k) converge vers la solution exacte $(0, 1, 2)$.

Remarque 3.2.1 *La méthode de Gauss-Seidel et la méthode de Jacobi peuvent être aussi utilisées pour la résolution des systèmes non linéaires.*

Exemple 3.2.3 *Soit à résoudre le système*

$$\begin{cases} x = \sin(xy) - \frac{y}{2\pi} \\ y = 2\pi x - \left(\pi - \frac{1}{4}\right)(e^{2x-1} - 1) \end{cases} \quad (3.4)$$

partant du point $x_0 = \left(\frac{2}{5}, 3\right)$, on calcule successivement

$$\begin{aligned} x_1 &= (0.455, 3.03) \\ x_2 &= (0.499, 3.11) \\ x_3 &= (0.505, 3.14). \end{aligned}$$

C'est une suite (x_k) qui converge vers la solution exacte $\left(\frac{1}{2}, \pi\right)$ du système (3.4).

3.2.3 Méthode de Relaxation

Dans la méthode de relaxation, on choisit

$$M = \frac{1}{\omega}D + L, \quad N = \left(\frac{1}{\omega} - 1\right)D - U.$$

Où $\omega \in \mathbb{R}_+^*$. La méthode de relaxation s'écrit donc sous la forme :

$$\left(\frac{1}{\omega}D + L\right)x^{(k+1)} = \left(\frac{1-\omega}{\omega}D - U\right)x^{(k)} + b, \quad k = 0, 1, \dots$$

i.e.

$$x^{(k+1)} = \left(\frac{1}{\omega}D + L\right)^{-1} \left(\frac{1-\omega}{\omega}D - U\right)x^{(k)} + \left(\frac{1}{\omega}D + L\right)^{-1} b, \quad k = 0, 1, \dots$$

Remarque 3.2.2 Pour $\omega = 1$, nous retrouvons la méthode de Gauss-Seidel. Pour $\omega < 1$, cette méthode est appelée méthode de sous-relaxation et pour $\omega > 1$ est appelée méthode de sur-relaxation

Remarque 3.2.3 Le but de la méthode de relaxation est d'accélérer la méthode de Gauss-Seidel.

Exemple 3.2.4 On considère le système $Ax = b$ qui admet une solution exacte $x = (1, 2, 3)$; où

$$A = \begin{pmatrix} 3 & 1 & -1 \\ 1 & 5 & 2 \\ 2 & -1 & -6 \end{pmatrix} \text{ et } b = \begin{pmatrix} 2 \\ 17 \\ -18 \end{pmatrix}$$

pour $\omega = 1.1$ et $x_0 = (0, 0, 0)$:

$$\begin{aligned} x_1 &= (0.7333, 3.5787, 2.9128) \\ x_2 &= (0.4158, 2.0090, 2.7927) \\ x_3 &= (0.9792, 2.0948, 2.9957) \end{aligned}$$

pour $\omega = 1$ et $x_0 = (0, 0, 0)$:(méthode de Gauss-seidel)

$$\begin{aligned} x_1 &= (0.6667, 3.2667, 2.6778) \\ x_2 &= (0.4704, 2.2348, 2.7843) \\ x_3 &= (0.8498, 2.1163, 2.9306) \end{aligned}$$

Donc on remarque que la méthode de Relaxation pour le paramètre $\omega = 1.1$ plus converge que la méthode de Gauss-seidel.

3.3 Étude de la convergence dans le cas d'une matrice symétrique définie positive

Dans ce paragraphe on va examiner la convergence de ces méthodes itératives dans le cas d'une matrice symétrique définie positive.

Théorème 3.3.1 *Soit A une matrice symétrique définie positive et soit $A = M - N$ une décomposition de la matrice A où M est inversible. On suppose que la matrice $M^T + N$ est symétrique définie positive. Alors, la méthode itérative $Mx^{(k+1)} = Nx^{(k)} + b$ converge.*

De ce résultat, on peut déduire les corollaires suivants :

Corollaire 3.3.1 *Si A est une matrice symétrique définie positive, la méthode de relaxation converge pour $0 < \omega < 2$.*

Preuve : Pour la méthode de relaxation, on a vu que :

$$M = \frac{1}{\omega}D + L, \quad N = \frac{1-\omega}{\omega}D - U.$$

Comme A est symétrique, on a $L = U^T$. La diagonale de M est celle de D . On déduit que M est inversible puisque les éléments d_{ii} sont positifs. Soit

$$M^T + N = \frac{1}{\omega}D + L^T + \frac{1-\omega}{\omega}D - L^T = \frac{2-\omega}{\omega}D.$$

Cette matrice diagonale est définie positive si et seulement si $0 < \omega < 2$. ■

Corollaire 3.3.2 *Soit A une matrice symétrique définie positive telle que $2D - A$ soit définie positive. Alors, la méthode de Jacobi converge.*

Preuve :

$$\begin{aligned} M^T + N &= D^T - (L + U) \\ &= D - (A - D) \\ &= 2D - A. \end{aligned}$$
■

3.4 Étude de la convergence dans le cas d'une matrice à diagonale dominante

On va étudier maintenant la convergence de ces méthodes itératives pour un autre type de matrices, c'est le cas des matrices à diagonale dominante. Pour ceci on a besoin de la définition suivante.

Définition 3.4.1 On dit qu'une matrice A est à diagonale dominante si on a :

$$\sum_{j \neq i} |a_{ij}| \leq |a_{ii}| \quad \forall i = 1, \dots, n.$$

On dit qu'elle est à diagonale strictement dominante si

$$\sum_{j \neq i} |a_{ij}| < |a_{ii}| \quad \forall i = 1, \dots, n.$$

Théorème 3.4.1 Soit A une matrice à diagonale strictement dominante, alors A est inversible. De plus, les méthodes de Jacobi et de Gauss-seidel convergent.

Preuve : Montrons d'abord que A est inversible.

- Soit $x \in \mathbb{R}^n$ avec $x \neq 0$ et $Ax = 0$. Soit i un entier tel que $|x_i| \geq |x_j|$ pour tout $j = 1, \dots, n$. On a

$$a_{ii}x_i = - \sum_{j \neq i} a_{ij}x_j.$$

Alors

$$|a_{ii}| |x_i| = \left| \sum_{j \neq i} a_{ij}x_j \right| \leq \sum_{j \neq i} |a_{ij}| |x_j| \leq |x_i| \sum_{j \neq i} |a_{ij}|.$$

Ce qui est impossible puisque $x \neq 0$. Donc la matrice A est inversible.

- Montrons maintenant la convergence.

Posons $B = M^{-1}N$ et soit λ une valeur propre et x un vecteur propre de B , c'est-à-dire :

$$Bx = \lambda x, \quad x \neq 0.$$

Puisque on s'intéresse à la plus grande valeur propre en module, on suppose que $\lambda \neq 0$. Par la suite, on a

$$\left(M - \frac{1}{\lambda}N \right) x = 0.$$

- Pour la méthode de Jacobi, la dernière équation s'écrit

$$\left(D + \frac{1}{\lambda}L + \frac{1}{\lambda}U \right) x = 0.$$

Posons $C = D + \frac{1}{\lambda}L + \frac{1}{\lambda}U$. Supposons que $|\lambda| \geq 1$, dans ce cas, on a

$$|c_{ii}| = |a_{ii}| > \sum_{j \neq i} |a_{ij}| \geq \frac{1}{|\lambda|} \sum_{j \neq i} |a_{ij}| = \sum_{j \neq i} |c_{ij}|.$$

On en déduit que C est à diagonale strictement dominante, donc C est inversible. Alors $x = 0$. Ceci est une contradiction avec le fait que x est un vecteur propre de B .

- Pour la méthode de Gauss-Seidel, on pose

$$C = D + L + \frac{1}{\lambda}U.$$

D'une manière analogue que précédemment, on suppose que $|\lambda| \geq 1$, on déduit que :

$$\begin{aligned} |c_{ii}| &= |a_{ii}| \\ &> \sum_{j \neq i} |a_{ij}| \\ &\geq \sum_{j < i} |a_{ij}| + \frac{1}{|\lambda|} \sum_{j > i} |a_{ij}| \\ &= \sum_{j \neq i} |c_{ij}|. \end{aligned}$$

C'est une contradiction.

■

Chapitre 4

Résolution des équations et des systèmes non linéaires

4.1 Introduction

L'objet de ce chapitre est la résolution de l'équation

$$f(x) = 0 \tag{4.1}$$

où f est une fonction de $I \subset \mathbb{R}$ dans \mathbb{R} . Ce problème concernait déjà les mathématiciens de l'antiquité puisque les babyloniens et les grecs s'intéressaient déjà aux racines des polynômes de degré 2. Plus tard, c'est en s'intéressant aux racines des polynômes de degré 2 et 3 que les mathématiciens italiens du sixième siècle, Cardan et Bombelli découvrirent les nombres complexes. On sait aujourd'hui qu'il n'existe pas de formule générale exacte pour calculer les racines des polynômes de degré supérieur ou égal à cinq. Cette remarque permet de comprendre l'intérêt des méthodes numériques. Elles fournissent un outil d'approcher la solution d'un problème que l'on ne sait pas résoudre exactement. Elles permettent également le calcul d'une solution en un temps réduit. Cependant elles ne remplacent pas la théorie : théorie et simulations sont complémentaires. Les simulations donnent un aspect concret. La réflexion se nourrit des aspects concrets, augmentent la compréhension et guident les nouvelles simulations. Actuellement, la résolution de l'équation (4.1) préoccupe toujours les mathématiciens : ainsi, les points d'équilibre d'un système dynamique de dimension finie ou infinie, la minimisation d'une fonction ou d'une fonctionnelle peuvent conduire à la résolution d'un problème de type (4.1).

Les méthodes pour approcher une racine α de f sont en général itératives : elles consistent à construire une suite $\{x^{(k)}\}$ telle que

$$\lim_{k \rightarrow \infty} x^{(k)} = \alpha.$$

La convergence des itérations est caractérisée par la définition suivante :

Définition 4.1.1 *On dit qu'une suite $\{x^{(k)}\}$ construite par une méthode numérique converge vers α avec ordre $p \geq 1$ si*

$$\exists C > 0 : \frac{|x^{(k+1)} - \alpha|}{|x^{(k)} - \alpha|^p} \leq C, \quad \forall k \geq k_0, \tag{4.2}$$

Contrairement au cas des systèmes linéaires, la convergence des méthodes itératives pour la détermination des racines d'une équation non linéaire dépend en général du choix de la donnée initiale $x^{(0)}$. Le plus souvent, on ne sait établir que des résultats de convergence locale, c'est-à-dire valables seulement pour un $x^{(0)}$ appartenant à un certain voisinage de la racine α . Les méthodes qui convergent vers α pour tout choix de $x^{(0)}$ dans l'intervalle I sont dites globalement convergentes vers α .

Pour introduire ce chapitre, on commence par rappeler le théorème des valeurs intermédiaires qui permet d'assurer l'existence d'une solution à l'équation.

Théorème 4.1.1 (*théorème des valeurs intermédiaires*)

Soit f une fonction continue, à valeurs dans \mathbb{R} et définie sur un intervalle $[a, b]$. On suppose que

$$f(a)f(b) < 0.$$

Alors, il existe $l \in]a, b[$ tel que

$$f(l) = 0.$$

Preuve : On utilise la méthode de dichotomie. On pose $a_0 = a$, $b_0 = b$. Puis,

$$c_n = \frac{a_n + b_n}{2}$$

- Si $f(c_n) = 0$, on s'arrête.
- Si $f(c_n)f(a_n) > 0$, on pose $a_{n+1} = c_n$, $b_{n+1} = b_n$.
- Si $f(c_n)f(a_n) < 0$, on pose $b_{n+1} = c_n$, $a_{n+1} = a_n$.

Les suites (a_n) , (b_n) , (c_n) ainsi définies vérifient :

- (a_n) est croissante majorée, (b_n) est décroissante minorée.
- $(b_n - a_n) = \frac{(b-a)}{2^n}$.

On en déduit que (a_n) et (b_n) convergent vers une même limite l . Par ailleurs on a

$$f(a_n)f(b_n) \leq 0,$$

et par passage à la limite

$$(f(l))^2 \leq 0.$$

Ce qui nous donne

$$f(l) = 0.$$

■

Théorème d'existence d'un point fixe :

Un autre théorème joue un rôle important en analyse, il s'agit du théorème de point fixe. On donne ici ce résultat dans le cas d'une fonction de \mathbb{R} dans \mathbb{R} , mais il faut savoir que ce théorème admet des généralisation dans \mathbb{R}^n et dans des espaces plus abstraits.

Théorème 4.1.2 (*théorème d'existence d'un point fixe*)

Soit g une fonction continue définie sur un intervalle $[a, b]$ et vérifiant $g([a, b]) \subset [a, b]$. Alors il existe $l \in [a, b]$ tel que

$$g(l) = l$$

Preuve : On pose

$$f(x) = g(x) - x$$

et on applique le théorème précédent à la fonction f .

■

4.2 Quelques méthodes itératives classiques

Nous avons démontré en introduction l'existence d'une solution de l'équation (4.1). Il faut maintenant chercher à calculer cette solution effectivement. Nous savons résoudre les équations du premier et second degré, et avec un peu plus de technique les équations du troisième et du quatrième degré. Mais pour le reste, en général on ne sait pas calculer les solutions de manière exacte. Il est donc nécessaire de développer des méthodes qui vont approcher les solutions. Nous allons présenter dans ce chapitre quelques méthodes itératives classiques. Elles permettent de construire une suite qui converge vers une solution de (4.1). Plus précisément nous introduisons les méthodes de dichotomie, de la corde, de Lagrange et de Newton. toutes ces méthodes reposent sur la résolution d'une équation linéaire approchant (4.1). Plus précisément, on construit une suite x_n jusqu'à l'ordre n , on approche $f(x_{n+1})$ par une fonction affine :

$$f(x_{n+1}) \simeq f(x_n) + \alpha_n(x_{n+1} - x_n)$$

où α_n est une approximation de $f'(x_n)$.

Et puisque l'on aimerait que $f(x_{n+1})$ soit proche de 0, on résoud :

$$f(x_n) + \lambda_n(x_{n+1} - x_n) = 0.$$

Ce qui donne

$$x_{n+1} = x_n - \frac{f(x_n)}{\lambda_n}.$$

4.2.1 Méthode de dichotomie

Soit f une fonction continue et $f(a)f(b) < 0$. On pose :

$$a_0 = a, \quad b_0 = b.$$

Puis,

$$x_n = \frac{a_n + b_n}{2}$$

- Si $f(x_n) = 0$, on s'arrête.
- Si $f(x_n)f(a_n) > 0$, on pose $a_{n+1} = x_n, b_{n+1} = b_n$.
- Si $f(x_n)f(a_n) < 0$, on pose $b_{n+1} = x_n, a_{n+1} = a_n$.

Cette méthode assure la convergence de la suite (x_n) vers une solution de l'équation (4.1), dès que $f(a)f(b) < 0$.

Remarque 4.2.1 *Notre problème peut avoir plusieurs solutions, et la méthode converge donc vers une des solutions.*

4.2.2 Méthode de la corde

Cette méthode, consiste à choisir

$$\lambda_n = \frac{f(b) - f(a)}{b - a}.$$

Donc l'équation

$$f(x_{n+1}) = 0,$$

devient,

$$f(x_n) + \frac{f(b) - f(a)}{b - a}(x_{n+1} - x_n) = 0.$$

Et par la suite :

$$x_{n+1} = x_n - \frac{b - a}{f(b) - f(a)}f(x_n).$$

Remarque 4.2.2 Soient $A(a, f(a)), B(b, f(b)), X_n(x_n, f(x_n))$ et (Δ_n) la droite parallèle à (AB) passant par X_n . Géométriquement, x_{n+1} est l'abscisse du point d'intersection de (Δ_n) et l'axe des abscisses. Cette méthode n'assure pas la convergence vers une solution mais certaines hypothèses permettent d'assurer la convergence.

4.2.3 Méthode de Lagrange

La méthode de Lagrange consiste à choisir

$$\lambda_n = \frac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}}.$$

Donc l'équation

$$f(x_{n+1}) = 0,$$

devient,

$$f(x_n) + \frac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}}(x_{n+1} - x_n) = 0.$$

Et par la suite

$$x_{n+1} = x_n - \frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})}f(x_n).$$

Remarque 4.2.3 Soient $X_n(x_n, f(x_n)), X_{n-1}(x_{n-1}, f(x_{n-1}))$ et (Δ_n) la droite passant par X_{n-1} et X_n . Géométriquement, x_{n+1} est l'abscisse du point d'intersection de (Δ_n) et l'axe des abscisses.

Comme pour la méthode de la corde, cette méthode n'assure pas la convergence vers une solution.

4.2.4 Méthode de Newton

La méthode de Newton consiste à choisir

$$\lambda_n = f'(x_n).$$

L'équation $f(x_{n+1}) = 0$, nous conduit à l'équation,

$$f(x_n) + f'(x_n)(x_{n+1} - x_n).$$

Et par la suite :

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}.$$

Remarque 4.2.4 Soit (Δ_n) la tangente à la courbe représentative de f au point d'abscisse x_n . Alors x_{n+1} est l'abscisse du point d'intersection de (Δ_n) et l'axe des abscisses. L'algorithme de Newton n'assure pas la convergence vers une solution de (4.1)

4.3 La méthode du point fixe

Théorème 4.3.1 Soit g une fonction de classe C^1 sur un intervalle $[a, b] \subset \mathbb{R}$. On suppose que $g([a, b]) \subset [a, b]$ et qu'il existe k , $0 < k < 1$ tel que pour tout $x \in [a, b]$, $g'(x) < k$, alors :

- la fonction g admet un unique point fixe dans $[a, b]$.
- pour tout $x_0 \in [a, b]$, la suite $(x_n)_{n \in \mathbb{N}}$ définie par

$$x_{n+1} = g(x_n)$$

vérifie :

$$\forall n \in \mathbb{N}, x_n \in [a, b] \text{ et } \lim_{n \rightarrow +\infty} x_n = l.$$

Preuve : Puisque la fonction g est continue et $g([a, b]) \subset [a, b]$, le théorème 4.1.2 assure l'existence d'un point fixe. Soient l_1 et l_2 deux points fixes de g . D'après le théorème des accroissements finis, on a :

$$|l_1 - l_2| = |g(l_1) - g(l_2)| \leq k |l_1 - l_2| < |l_1 - l_2|$$

dès que $l_1 \neq l_2$. Donc, $l_1 = l_2$ et le point fixe est unique. Le fait que $x_n \in [a, b]$ résulte du fait que $g([a, b]) \subset [a, b]$. Par ailleurs pour tout $n \in \mathbb{N}^*$,

$$|x_n - l| = |g(x_{n-1}) - l| = |g(x_{n-1}) - g(l)| \leq k |x_{n-1} - l|.$$

On montre ainsi par récurrence que

$$|x_n - l| \leq k^n |x_0 - l|,$$

puisque $0 < k < 1$, cela implique que,

$$\lim_{n \rightarrow +\infty} |x_n - l| = 0.$$

■

4.4 Convergence de la méthode de Newton, de Lagrange et de la corde

4.4.1 Convergence de la méthode de Newton

Soit x_0 donné, la suite des itérés de Newton est définie par la relation de récurrence suivante :

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

Théorème 4.4.1 Soit f une fonction de classe C^2 sur l'intervalle $[a, b]$. On suppose que :

- $f(a)f(b) < 0$
- $f' \neq 0$ sur $[a, b]$
- $f'' > 0$ sur $[a, b]$.

Alors, si $f(x_0) > 0$, la suite des itérés de Newton converge vers l'unique solution l de (4.1) sur cet intervalle.

Preuve : On suppose que $f' > 0$ sur $[a, b]$. (le cas $f' < 0$ se traite d'une manière analogue) D'après les hypothèses f réalise une bijection de $[a, b]$ sur $[f(a), f(b)]$. Donc, il existe un unique nombre $l \in]a, b[$ solution de l'équation $f(x) = 0$. Le fait que x_{n+1} soit plus grand que l est dû au fait que f est convexe. En effet puisque f est convexe, sa courbe est au dessus des tangentes. Donc la tangente coupe l'axe des abscisses en un point situé à droite de l . D'un point de vue analytique, d'après la formule de Taylor-Lagrange à l'ordre 2, on a pour tout $x \in [a, b]$ il existe $\xi_n \in [a, b]$ tel que

$$f(x) = f(x_n) + f'(x_n)(x - x_n) + \frac{f''(\xi_n)}{2}(x - x_n)^2$$

Comme $f'' > 0$ pour tout $x \neq x_n$, alors on a :

$$f(x) > f(x_n) + f'(x_n)(x - x_n).$$

Ce qui donne pour $x = l$:

$$0 = f(l) > f(x_n) + f'(x_n)(l - x_n).$$

Comme x_{n+1} vérifie

$$0 = f(x_n) + f'(x_n)(x_{n+1} - x_n).$$

On déduit d'après le théorème des valeurs intermédiaires appliqué à la fonction $f(x_n) + f'(x_n)(x - x_n)$ que

$$l < x_{n+1} < x_n.$$

Par récurrence, on en déduit que $(x_n)_{n \in \mathbb{N}}$ est une suite décroissante minorée, donc elle converge. Soit x^* sa limite. Alors de l'équation,

$$f(x_n) + f'(x_n)(x_{n+1} - x_n) = 0$$

on déduit par passage à la limite que

$$f(x^*) = 0.$$

■

Remarque 4.4.1 Il existe une autre version du théorème précédent dans le cas $f'' < 0$ sur $[a, b]$, avec $f(x_0) < 0$.

Remarque 4.4.2 Dans le théorème précédent, on a l'hypothèse que $f(x_0) > 0$. Si $f' > 0$ cela veut dire que $x_0 > l$. Si on choisit $x_0 < l$, alors du fait de la convexité ($f'' > 0$), on aura $x_1 > l$. Pour avoir la convergence, il suffit de vérifier que $x_1 < b$. Ce est bien assuré par la condition : $a - \frac{f(a)}{f'(a)} < b$. Cela est résumé dans le théorème suivant.

Théorème 4.4.2 Sous les conditions du théorème précédent, si on suppose de plus que

$$f' > 0 \text{ et } a - \frac{f(a)}{f'(a)} < b$$

alors pour tout $x_0 \in [a, b]$, la suite des itérés de Newton converge vers l .

Preuve :

- Pour $l < x_0 < b$, le résultat découle directement du théorème précédent.
- Pour $a < x_0 < l$, du fait de la convexité et de la croissance de f , x_1 est une fonction décroissante de x_0 . En effet, en posant :

$$x_1 = \phi(x_0) = x_0 - \frac{f(x_0)}{f'(x_0)}.$$

En dérivant par rapport à x_0 , on trouve :

$$\phi' = 1 - \frac{(f')^2 - ff''}{(f')^2} = \frac{ff''}{(f')^2} < 0.$$

■

4.4.2 Convergence de la méthode de Lagrange

Théorème 4.4.3 Soit f une fonction de classe C^2 sur l'intervalle $[a, b]$. On suppose que :

- $f(a)f(b) < 0$
- $f' \neq 0$ sur $[a, b]$
- $f'' > 0$ sur $[a, b]$.

Alors, si $f(x_0) > 0$ et $f(x_1) > 0$, la suite des itérés de Lagrange converge vers l'unique solution l de (4.1) sur cet intervalle.

Preuve : Supposons que $f' > 0$ sur $[a, b]$. (le cas $f' < 0$ se traite d'une manière analogue). D'après les hypothèses f réalise une bijection de $[a, b]$ sur $[f(a), f(b)]$. Donc, il existe un unique $l \in]a, b[$ tel que $f(l) = 0$. Puisque f est convexe, alors,

$$0 = f(l) > f(x_n) + \frac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}}(l - x_n).$$

Donc $x_n > x_{n+1} > l$. Alors, $(x_n)_n \in \mathbb{N}$ est une suite décroissante minorée, donc elle converge. Soit x^* sa limite. Alors de l'équation

$$f(x_n) + \frac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}}(x_{n+1} - x_n),$$

et par passage à la limite, on déduit que

$$f(x^*) = 0.$$

Théorème 4.4.4 Sous les hypothèses du théorème précédent, si on suppose de plus que

$$f' > 0 \text{ et } \frac{f(a)}{f'(a)} < b,$$

alors la suite des itérés de Lagrange converge vers l , pour tout $x_0, x_1 \in [a, b]$.

Preuve : On vérifie d'abord que la suite des itérés de Lagrange appartient à l'intervalle $[a, b]$. Puisque f est convexe et $f' > 0$, pour tout $a \leq x \leq b$,

$$\frac{f(y) - f(x)}{y - x} > f'(x).$$

On suppose que $x_{n-1} < x_n < l$. Alors

$$\begin{aligned} x_n - \frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})} f(x_n) &< x_n - \frac{f(x_n)}{f'(x_n)} \\ &< a - \frac{f(a)}{f'(a)} \\ &< b. \end{aligned}$$

On montre maintenant que la suite $(x_n)_{n \in \mathbb{N}}$ est convergente. supposons que $x_0 < x_1 < l$. Puisque f est convexe, alors

$$0 = f(l) > f(x_0) + \frac{f(x_1) - f(x_0)}{x_1 - x_0} (l - x_0).$$

Donc, $x_2 > l$. Puis de manière analogue, on trouve $x_3 < l$. Mais,

$$\begin{aligned} x_3 &= x_2 - \frac{x_2 - x_1}{f(x_2) - f(x_1)} f(x_2) \\ &= x_1 - \frac{x_2 - x_1}{f(x_2) - f(x_1)} f(x_1) \\ &> x_1. \end{aligned}$$

Puis de manière analogue, on trouve $x_4 < l$ et $x_4 > x_3$. Puis, $l < x_5 < x_2$, et $x_4 < x_6 < x_7 < l$. Alors on peut montrer par récurrence que la sous-suite $x_0, x_1, x_3, x_4, \dots, x_{3p}, x_{3p+1}, \dots$ est croissante majorée et la sous-suite $x_2, x_5, \dots, x_{3p+2}, \dots$, est décroissante minorée. Elles convergent donc vers deux limites l_1 et l_2 . Par passage à la limite dans l'équation,

$$f(x_{3p}) + \frac{f(x_{3p}) - f(x_{3p-1})}{x_{3p} - x_{3p-1}} (x_{3p+1} - x_{3p}) = 0,$$

on trouve,

$$f(l_1) = 0,$$

puis par passage à la limite dans l'équation

$$f(x_{3p+1}) + \frac{f(x_{3p+1}) - f(x_{3p})}{x_{3p+1} - x_{3p}} (x_{3p+2} - x_{3p+1}) = 0,$$

on obtient,

$$l_1 = l_2.$$

■

4.4.3 Convergence de la méthode de la corde

Théorème 4.4.5 *On suppose que f est de classe C^1 sur $[a, b]$ et*

- pour tout $x \in [a, b]$

$$\min(\lambda(x-a), \lambda(x-b)) < f(x) \max(\lambda(x-a), \lambda(x-b))$$

- pour tout $x \in [a, b]$

$$\min(0, 2\lambda) < f(x) \max(0, 2\lambda).$$

Alors la méthode de la corde converge vers l'unique solution l dans $[a, b]$.

Preuve : On a :

$$x_{n+1} = x_n - \frac{b-a}{f(b)-f(a)} f(x_n).$$

La suite des itérés de la méthode de la corde peut s'écrire comme une suite des itérés de la méthode du point fixe :

$$x_{n+1} = g(x_n)$$

où,

$$g(x) = x - \frac{f(x)}{\lambda} \text{ et } \lambda = \frac{f(b)-f(a)}{b-a}$$

D'après le théorème 4.3.1, pour que la méthode du point fixe converge, il suffit que :

1. $g([a, b]) \subset [a, b]$
2. $g'(x) < 1$ sur $[a, b]$. La première condition s'écrit :

$$g(x) \geq a \text{ et } g(x) \leq b.$$

Si $\lambda > 0$, on trouve :

$$\begin{aligned} a \leq g(x) \leq b &\iff a \leq x - \frac{f(x)}{\lambda} \leq b \\ &\iff \lambda(x-b) \leq f(x) \leq \lambda(x-a) \end{aligned}$$

Si $\lambda < 0$, on trouve :

$$a \leq g(x) \leq b \iff \lambda(x-a) \leq f(x) \leq \lambda(x-b).$$

Ou encore pour λ quelconque non nul :

$$\min(\lambda(x-a), \lambda(x-b)) < f(x) < \max(\lambda(x-a), \lambda(x-b)).$$

Pour la deuxième condition, on a :

$$\begin{aligned} |g'(x)| < 1 &\iff -1 < 1 - \frac{f'(x)}{\lambda} < 1 \\ &\iff 0 < \frac{f'(x)}{\lambda} < 2 \end{aligned}$$

Si $\lambda > 0$, on trouve :

$$0 < f'(x) < 2\lambda.$$

Si $\lambda < 0$, on trouve :

$$2\lambda < f'(x) < 0.$$

Ou encore pour λ quelconque non nul :

$$\min(0, 2\lambda) < f'(x) < \max(0, 2\lambda).$$

■

4.5 Systèmes d'équations non linéaires

On considère maintenant un système d'équations non linéaires donné par une fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$, $x = (x_1, \dots, x_n)^T \mapsto f(x) = (f_1(x_1, \dots, x_n), \dots, f_n(x_1, \dots, x_n))^T$. On cherche donc un vecteur $x = (x_1, \dots, x_n)^T \in \mathbb{R}^n$ tel que

$$f(x) = 0_{\mathbb{R}^n} \iff \begin{cases} f_1(x_1, \dots, x_n) = 0, \\ \vdots \\ f_n(x_1, \dots, x_n) = 0. \end{cases}$$

On définit l'itération

$$x^{(n+1)} = x^{(n)} + M^{-1}f(x^{(n)}) \quad (4.3)$$

où M est une matrice, et nous avons les mêmes résultats de convergence que dans le cas d'une seule équation.

Définition 4.5.1 La matrice jacobienne d'une fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ notée J_f est définie (lorsqu'elle existe) par :

$$\forall x = (x_1, \dots, x_n)^T \in \mathbb{R}^n, \quad J_f(x) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x) & \frac{\partial f_1}{\partial x_2}(x) & \cdots & \frac{\partial f_1}{\partial x_n}(x) \\ \frac{\partial f_2}{\partial x_1}(x) & \frac{\partial f_2}{\partial x_2}(x) & \cdots & \frac{\partial f_2}{\partial x_n}(x) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1}(x) & \frac{\partial f_n}{\partial x_2}(x) & \cdots & \frac{\partial f_n}{\partial x_n}(x) \end{pmatrix}$$

La méthode de Newton se généralise naturellement au cas des systèmes d'équations non linéaires de la manière suivante : on choisit $x^0 \in \mathbb{R}^n$ et on utilise la formule d'itération

$$x^{(n+1)} = x^{(n)} - J_f(x^{(n)})^{-1}f(x^{(n)}), \quad (4.4)$$

où $J_f(x^{(n)})^{-1}$ désigne l'inverse de la matrice jacobienne de f évaluée en $x^{(n)}$.

Théorème 4.5.1 Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ une fonction de classe C^2 sur une boule fermée B de \mathbb{R}^n . On suppose qu'il existe un zéro \tilde{x} de f dans B et que $J_f(\tilde{x})$ est inversible. Alors il existe $\epsilon > 0$ tel que pour tout $x^0 \in B$ tel que $\|x^0 - \tilde{x}\| \leq \epsilon$, la suite des itérés de la méthode de Newton définie par (4.4) est bien définie et converge vers \tilde{x} quand n tend vers l'infini.

Calculer l'itéré $n + 1$ à partir de l'itéré n en utilisant la formule (4.4) nécessite d'inverser la matrice $J_f(x^{(n)})$. Or, calculer l'inverse d'une matrice peut s'avérer coûteux. Par conséquent, nous ré-écrivons la formule d'itération (4.4) sous la forme :

$$J_f(x^{(n)})(x^{(n+1)} - x^{(n)}) = -f(x^{(n)}), \quad (4.5)$$

de sorte qu'à chaque itération, le calcul de l'inverse d'une matrice est remplacé par la résolution d'un système d'équations linéaires ce qui est asymptotiquement moins coûteux en nombre d'opérations.

Exemple 4.5.1 *Considérons le système d'équations non linéaires :*

$$(S) : \begin{cases} x_1^2 + 2x_1 - x_2^2 - 2 = 0, \\ x_1^3 + 3x_1x_2^2 - x_2^3 - 3 = 0. \end{cases}$$

Avec les notations précédentes, cela correspond à $n = 2$, $f_1(x_1, x_2) = x_1^2 + 2x_1 - x_2^2 - 2$, et $f_2(x_1, x_2) = x_1^3 + 3x_1x_2^2 - x_2^3 - 3$. La matrice jacobienne de f est alors :

$$J_f(x_1, x_2) = \begin{pmatrix} 2x_1 + 2 & -2x_2 \\ 3(x_1^2 + x_2^2) & 6x_1x_2 - 3x_2^2 \end{pmatrix}$$

Partant du point $x^{(0)} = (1, -1)^T$, calculons le premier itéré de la méthode de Newton pour résoudre le système (S). Pour $n = 1$, la formule d'itération (4.5) s'écrit :

$$J_f(x^{(0)})(x^{(1)} - x^{(0)}) = -f(x^{(0)}),$$

c'est-à-dire

$$\begin{pmatrix} 4 & 2 \\ 6 & -9 \end{pmatrix} \begin{pmatrix} x_1^{(1)} - 1 \\ x_2^{(1)} + 1 \end{pmatrix} = - \begin{pmatrix} 0 \\ 2 \end{pmatrix}.$$

En résolvant ce système linéaire, on trouve $x_1^{(1)} - 1 = -\frac{1}{12}$ et $x_2^{(1)} + 1 = \frac{1}{6}$ de sorte que :

$$x^{(1)} = \left(\frac{11}{12}, -\frac{5}{6}\right)^T.$$

Conclusion et perspectives

Les méthodes directes sont très coûteuses lorsqu'on se place dans le cas des systèmes linéaires de taille assez élevée, alors dans ce cas il serait très utile de passer aux méthodes itératives.

L'utilisation des méthodes itératives est généralement plus avantageuse lorsque celles-ci convergent. Les méthodes de Jaccobi, de Gauss-Seidel et de relaxation ne convergent le plus souvent que dans les cas que nous avons décrits dans le troisième chapitre. Il existe des méthodes itératives beaucoup plus élaborées et qui convergent dans des cas plus généraux que ceux décrits ici. Dans le cas de matrices creuses, i.e., les matrices contenant un nombre significatif de coefficients nuls, ces méthodes sont beaucoup plus économiques puisqu'il n'est pas nécessaire de stocker en mémoire les coefficients nuls ni d'effectuer les opérations utilisées.

La méthode de Gauss-Seidel converge généralement plus vite et plus souvent que la méthode de Jaccobi. La méthode de relaxation nécessite la connaissance d'une valeur optimale du paramètre de relaxation ω .

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 est assez petite :

$$\frac{\|x^{(k+1)} - x^{(k)}\|}{\|x^{(k)}\|} < \epsilon.$$

Une autre alternative est de tester le résidu :

$$\|Ax^{(k+1)} - b\| < \epsilon.$$

Bibliographie

- [1] A. Quarteroni, R. Sacco, F. Saleri, méthodes numériques : Algorithmes, analyse et applications, Springer-Verlag Italia, Milano (2007).
- [2] B. Flannery, W. Press, S. Teukolsky, W. Vetterling, Numerical recipes, Cambridge University Press (1988).
- [3] F. Jędrzejewski, Introduction aux méthodes numériques 2^{ème} édition, Springer-Verlag France, Paris (2005).
- [4] J. E. Rombaldi, Algorithmique numérique et Ada, Masson (1994).
- [5] J. E. Rombaldi, Analyse matricielle Cours et exercices résolus, Masson (1999).
- [6] J. F. Durand, Eléments de calcul matriciel et d'Analyse Factorielle de Données, Université Montpellier II ; Novembre (2002).
- [7] J. P. Grivet, méthodes numériques appliquées pour le scientifique et l'ingénieur, EDP Sciences (2009).
- [8] M. C .C. Moulin, J. Ezquerro, D. Fredon, A. Salinier, Algèbre linéaire, Hachette supérieur (1996).
- [9] M. Lakrib, Cours d'analyse numérique 4^{ème} édition, Office des publications Universitaires ; 06 – 2008.
- [10] P. G. Ciarlet, Introduction à l'analyse numérique matricielle et à l'optimisation, Masson (1982).

Résumé

Dans ce travail, nous avons exposé quelques méthodes numériques de résolution des systèmes linéaires et non linéaires. Pour les systèmes linéaires nous avons vu trois méthodes directes (éliminations de Gauss, décomposition LU et décomposition de Cholesky) et trois méthodes itératives (Jacobi, Gauss-Seidel et relaxation) qui sont très intéressantes dans le cas des systèmes de grandes tailles. Nous avons vu aussi quelques méthodes numériques de résolutions des équations non linéaires (Newton, corde, dichotomie et Lagrange) et enfin on a terminé notre travail par la méthode de Newton pour la résolution des systèmes non linéaires.

ملخص

في هذا العمل ، قدمنا بعض الطرق العددية لحل الجمل الخطية و غير الخطية. بالنسبة للطرق الخطية عرضنا ثلاث طرق مباشرة (الحذف لغوص، التحليل و شولسكي) و ثلاث طرق تكرارية (جاكوبي، غوص سيدل و الإسترخاء) ذات أهمية بالغة في حالة الجمل الكبيرة. رأينا كذلك بعض الطرق العددية لحل المعادلات غير الخطية (نيوتن، الحبل، التنصيف و لاغرانج) و في الأخير قدمنا طريقة نيوتن لحل الجمل غير الخطية.