Francis Filbet

# Analyse numérique

## Algorithme et étude mathématique

**2e édition**

- **Cours**
- **exercices corrigés**

Licence 3
Master
École d'ingénieurs
Agrégation

DUNOD

Francis Filbet

# Analyse numérique

## Algorithme et étude mathématique



2e édition

DUNOD

Tout le catalogue sur
**www.dunod.com**

DUNOD
**ÉDITEUR DE SAVOIRS**

Illustration de couverture : © vege - Fotolia.com


ISBN 978-2-10-059910-3

# Analyse numérique

# AVANT-PROPOS

L'objet de cet ouvrage est de fournir au lecteur une présentation et une analyse des méthodes numériques les plus couramment utilisées pour la résolution de problèmes de mathématiques appliquées. Chaque chapitre fait l'objet d'une introduction et d'exemples puisés dans des applications à divers domaines des mathématiques. Il s'agit ici d'illustrer l'utilisation d'outils mathématiques pour des problèmes actuels (modèles économiques, physiques, algorithme de Google, etc).

Cet ouvrage s'adresse tout particulièrement aux étudiants de Licence 3, Master ou préparation à l'agrégation ainsi qu'aux élèves des écoles d'Ingénieurs intéressés par des problèmes issus de la physique, de la biologie ou encore de l'économie, et qui souhaitent avoir une idée des méthodes de bases utilisées aujourd'hui et être en mesure de les mettre en œuvre sur ordinateur.

Nous avons volontiers glissé plusieurs rappels théoriques qui devraient aider le lecteur à mieux appréhender les difficultés mathématiques de certaines démonstrations. Notons toutefois que certains résultats sont présentés dans un cadre simplifié de manière à ne pas noyer le lecteur dans des difficultés techniques.

L'ensemble des sujets abordés ici à déjà largement été traité dans la littérature et ce cours doit beaucoup à d'autres ouvrages et cours dispensés en France, comme les cours d'analyse numérique de Jean-Marie Thomas, Michel Pierre et Abderrahmane Bendali.

Lors de la rédaction de cet ouvrage, de nombreux collègues et étudiants ont bien voulu me faire part de leurs avis, commentaires, suggestions. À ce propos, je tiens tout particulièrement à exprimer ma reconnaissance à Jean-Pierre Bourgade, Nicolas Crouseilles, Thomas Lepoutre, Céline Parzani, Miguel Rodrigues, qui m'ont fait part de leurs conseils éclairés et pour certains ont bien voulu effectuer une relecture de ce document.

Je dédie ce livre à Raphaël et Flavien.

# Table des matières

# LES SYSTÈMES LINÉAIRES

# 1

La résolution d'un système linéaire algébrique est au cœur de la plupart des calculs en analyse numérique. Il paraît donc naturel de débuter un cours de calcul scientifique par là. Ici, nous décrivons les algorithmes de résolution les plus populaires qui sont appliqués à des systèmes généraux. Nous considérons le problème suivant : trouver le vecteur $x$ solution de

$$Ax = b,$$

où $A$ est une matrice carrée et $b$ un vecteur donné à coefficients réels ou complexes. La discrétisation d'équations différentielles ordinaires ou d'équations aux dérivées partielles, la modélisation de problèmes en physique, chimie ou économie conduit souvent à la résolution de systèmes linéaires de grande taille avec plusieurs milliers d'inconnues et il devient pratiquement impossible de résoudre ces systèmes d'équations sans l'aide d'un ordinateur. Il s'agit alors de trouver des algorithmes de résolution efficaces où le nombre d'opérations et donc le temps de calcul, n'est pas prohibitif. C'est un problème classique mais difficile en analyse numérique.

L'objectif de ce chapitre est de proposer différentes méthodes numériques de résolution des systèmes linéaires et de sensibiliser le lecteur à l'importance du choix de la méthode en fonction des propriétés du système. Nous distinguerons deux types de méthodes : *les méthodes directes* où nous calculons exactement la solution et *les méthodes itératives* où nous calculons une solution approchée.

## 1.1 EXEMPLE D'UN SYSTÈME LINÉAIRE

Commençons par présenter un problème provenant de l'économie, où nous sommes amenés à la résolution d'un système linéaire.

### Exemple d'une analyse de l'offre et de la demande

Imaginons que plusieurs artisans décident de coopérer pour fabriquer différents produits utiles à chacun d'eux. Afin d'éviter le coût du stockage des produits fabriqués, nous recherchons une situation d'équilibre entre l'offre et la demande. Pour cela, considérons $n$ artisans, sachant que chacun fabrique un produit spécifique. Ces artisans ont choisi de coopérer, c'est-à-dire qu'ils souhaitent adapter leur production, d'une part, à leurs propres besoins déterminés par la quantité de produit nécessaire aux autres artisans pour qu'ils puissent fabriquer leur propre produit et, d'autre part, aux besoins du marché.

Puisque chaque artisan fabrique un produit différent, nous notons par $i \in \{1, \dots, n\}$ le produit fabriqué par un artisan, $x_i$ désigne le nombre total de produits $i$ fabriqués par un artisan et $b_i$ la demande du marché en ce produit. D'autre part $(c_{i,j})_{1 \leqslant i,j \leqslant n}$ exprime la quantité de produit $i$ nécessaire à la confection d'une unité de produit $j$. En supposant que la relation qui lie les différents produits est linéaire, nous recherchons l'équilibre entre les besoins et la production, c'est-à-dire que nous imposons que la quantité de produit $i$ fabriquée soit égale à la somme des besoins des autres artisans en produit $i$ pour fabriquer leur propre produit et des besoins du marché en $i$

$$x_i = \sum_{j=1}^{n} C_{i,j}\, x_j + b_i, \quad i = 1, \dots, n$$

ou encore

$$x = C\,x + b,$$

où la matrice $C$ est formée par les coefficients $(c_{i,j})_{1 \leqslant i,j \leqslant n}$ tandis que le vecteur $b$ correspond à la demande du marché $b := (b_1, \dots, b_n)^T$. Par conséquent, la production totale $x = (x_1, \dots, x_n)^T$ est la solution du système linéaire $A\,x = b$, où la matrice $A$ est donnée par $A = I_n - C$ et $I_n$ est la matrice identité composée de 1 sur sa diagonale et de 0 ailleurs.

Cette modélisation pose plusieurs difficultés mathématiques. Tout d'abord nous nous interrogeons sur les propriétés de la matrice $A$ permettant d'assurer que ce problème a bien une solution. Il vient ensuite la question du calcul de cette solution. En pratique, lorsque le nombre d'artisans considérés devient grand, il n'est plus possible de calculer l'inverse de la matrice $A$, il faudra donc fournir des algorithmes qui ne nécessitent pas l'inversion de cette matrice. C'est l'objectif de ce premier chapitre.

## 1.2 Rappels sur les matrices

Avant de s'intéresser à la résolution numérique de systèmes linéaires, présentons quelques rappels d'algèbre linéaire.

Par la suite, nous noterons $\mathscr{M}_{m,n}(\mathbb{K})$ l'ensemble des matrices à $m$ lignes, $n$ colonnes et à coefficient dans le corps des réels $\mathbb{K} = \mathbb{R}$ ou des complexes $\mathbb{K} = \mathbb{C}$. Pour un vecteur colonne $u \in \mathbb{K}^n$, la valeur $u_i$ avec $1 \leqslant i \leqslant n$, désigne la $i$-ème composante dans la base canonique de $\mathbb{K}^n$ du vecteur $u$. Nous appelons *adjoint* du vecteur colonne $u$, le vecteur ligne $u^\star$ de $\mathbb{K}^n$ tel que $u^\star = (\overline{u}_1, \dots, \overline{u}_n)$ où $\overline{u}_i$ désigne le conjugué de $u_i$ et transposé de $u$ est le vecteur ligne $u^T = (u_1, \dots, u_n)$.

Rappelons également le produit matrice-vecteur. Soient $A$ une matrice à $m$ lignes et $n$ colonnes et $u$ un vecteur de $\mathbb{K}^n$, nous définissons $v = A\,u \in \mathbb{K}^m$ le vecteur dont les composantes sont données par

$$v_i = \sum_{j=1}^{n} a_{i,j}\, u_j, \quad i = 1, \dots, m$$

et pour $B$ une matrice à $n$ lignes et $p$ colonnes, le produit $A\,B$ fournit une matrice $C \in \mathcal{M}_{m,p}$ donnée par

$$c_{i,j} = \sum_{k=1}^{n} a_{i,k}\, b_{k,j}, \quad i = 1,\ldots,m, \quad j = 1,\ldots,p.$$

Introduisons également la notion de matrice adjointe et transposée.

**Définition 1.1**

Soit $A \in \mathcal{M}_{m,n}(\mathbb{K})$, la matrice adjointe de $A$, notée $A^\star$, à $n$ lignes et $m$ colonnes, est donnée par $A^\star = (a^\star_{i,j})_{\substack{1\leqslant i\leqslant n\\ 1\leqslant j\leqslant m}}$ et $a^\star_{i,j} = \overline{a}_{j,i}$, pour $1 \leqslant i \leqslant n$ et $1 \leqslant j \leqslant m$.
La matrice transposée de $A$, notée $A^T$, à $n$ lignes et $m$ colonnes, est donnée par $A^T = (a^T_{i,j})_{\substack{1\leqslant i\leqslant n\\ 1\leqslant j\leqslant m}}$ et $a^T_{i,j} = a_{j,i}$, pour $1 \leqslant i \leqslant n$ et $1 \leqslant j \leqslant m$.

Dans la base canonique de $\mathbb{K}^n$, définissons le produit scalaire ou produit hermitien entre deux vecteurs $u$ et $v \in \mathbb{K}^n$, le scalaire de $\mathbb{K}$ donné par

$$\langle u, v\rangle = u^\star\, v = \sum_{i=1}^{n} u^\star_i\, v_i.$$

### 1.2.1 Cas des matrices carrées

Considérons maintenant le cas particulier des matrices carrées, pour lesquelles le nombre de lignes est égal au nombre de colonnes. Nous rappelons les définitions suivantes :

**Définition 1.2**

Une matrice carrée $A \in \mathcal{M}_{n,n}(\mathbb{K})$ est dite inversible, ou régulière, ou encore non singulière, s'il existe une matrice $B \in \mathcal{M}_{n,n}(\mathbb{K})$ telle que $A\,B = B\,A = I_n$. Dans ce cas, la matrice $B$ est unique et s'appelle la matrice inverse de $A$, elle est notée $A^{-1}$.

**Définition 1.3**

Soit $A \in \mathcal{M}_{n,n}(\mathbb{K})$, alors

- $A$ est une matrice normale si $A^\star A = A\,A^\star$.
- $A$ est une matrice unitaire si $A^\star A = A\,A^\star = I_n$, où $I_n$ désigne la matrice identité. Dans le cas où $\mathbb{K} = \mathbb{R}$, nous parlons de matrice orthogonale et $A^T A = A\,A^T = I_n$, c'est-à-dire $A^T = A^{-1}$.
- $A$ est une matrice hermitienne si $A^\star = A$. Dans le cas où $\mathbb{K} = \mathbb{R}$, nous parlons de matrice symétrique et $A^T = A$.

Il est alors facile de faire le lien entre les matrices hermitiennes et normales.

**Proposition 1.4**

Toute matrice hermitienne est une matrice normale.

Comme nous l'avons vu en début de chapitre, l'objectif de cette partie est de mettre au point des algorithmes de résolution numérique pour un système de la forme

$$A\,x = b, \tag{1.1}$$

où $A \in \mathscr{M}_{n,n}(\mathbb{K})$ et $b \in \mathbb{K}^n$ sont donnés et $x \in \mathbb{K}^n$ est l'inconnue. Nous donnons d'abord une condition nécessaire et suffisante sur la matrice $A$ pour que ce système admette une solution unique. Le système linéaire (1.1) a une solution unique dès que la matrice $A$ est inversible et cette solution est donnée par $x = A^{-1}\,b$. Il est alors important de connaître quelques propriétés des matrices inversibles [16].

**Proposition 1.5**

Soient $A$, $B \in \mathscr{M}_{n,n}(\mathbb{K})$ inversibles, nous avons alors

- pour tout $\alpha \in \mathbb{K}^*$, $(\alpha\,A)^{-1} = \dfrac{1}{\alpha}\,A^{-1}$.
- $(A\,B)^{-1} = B^{-1}\,A^{-1}$.
- $(A^\star)^{-1} = (A^{-1})^\star$.

Énonçons quelques propriétés des matrices inversibles, qui sont autant d'outils permettant de s'assurer que le problème (1.1) admet bien une solution [16].

**Théorème 1.6** ***Théorème des matrices inversibles***

Soit $A \in \mathscr{M}_{n,n}(\mathbb{K})$, les propositions suivantes sont équivalentes :

- $A$ est inversible,
- le déterminant de $A$ est non nul,
- le rang de $A$ est égal à $n$, c'est-à-dire que les $n$ vecteurs colonnes forment une famille libre,
- le système homogène $A\,x = 0$ a pour unique solution $x = 0$,
- pour tout $b$ dans $\mathbb{K}^n$, le système linéaire $A\,x = b$ a exactement une solution.

Pour conclure cette partie, rappelons qu'une valeur propre de $A$ est donnée par $\lambda \in \mathbb{K}$ telle que $\det(A - \lambda\,I_n) = 0$. Ainsi, il existe au moins un vecteur $v$ non nul, dit *vecteur propre*, vérifiant $A\,v = \lambda\,v$. Définissons alors le spectre d'une matrice.

**Définition 1.7**

Soit $A \in \mathcal{M}_{n,n}(\mathbb{K})$. Nous appelons spectre de $A$ l'ensemble des valeurs propres de $A$ :

$$\mathrm{Sp}(A) = \{\lambda \in \mathbb{K};\quad \exists v \in \mathbb{K}^n \quad v \neq 0 \quad Av = \lambda v\}$$

Nous appelons rayon spectral de $A$ le nombre réel positif $\rho(A)$ tel que

$$\rho(A) = \max\{|\lambda|, \lambda \in \mathbb{C}, v \in \mathbb{C} \setminus \{0\} Av = \lambda v\}$$

À partir de la définition d'une valeur propre, nous vérifions facilement qu'une matrice est inversible si et seulement si 0 n'est pas valeur propre.

La conséquence de ces propriétés est que l'ensemble des matrices carrées inversibles forme un groupe, appelé le groupe linéaire et noté habituellement $Gl_n(\mathbb{K})$. En général, « presque toutes » les matrices sont inversibles. Sur le corps $\mathbb{K}$, cela peut être formulé de façon plus précise : l'ensemble des matrices non inversibles, considéré comme sous-ensemble de $\mathcal{M}_{n,n}(\mathbb{K})$, est un ensemble négligeable, c'est-à-dire de mesure de Lebesgue nulle. Intuitivement, cela signifie que si vous choisissez au hasard une matrice carrée à coefficients réels, la probabilité pour qu'elle soit non inversible est égale à zéro. La raison est que des matrices non inversibles peuvent être considérées comme racines d'une fonction polynôme donnée par le déterminant [16].

Par la suite, nous nous intéresserons au calcul proprement dit de la solution de (1.1). Proposons d'abord un premier algorithme naïf qui permet de calculer la solution d'un système linéaire ; cet algorithme est dû à Cramer[1] [23].

**Théorème 1.8** ***Résolution de Cramer***

Soit $A \in \mathcal{M}_{n,n}(\mathbb{K})$ une matrice inversible. Alors la solution du système $A\,x = b$ est donnée par $x_i = \det(A_i)/\det(A)$, pour tout $i = 1, \ldots, n$, où $A_i$ est la matrice $A$ pour laquelle la $i$-ème colonne est remplacée par le vecteur $b$.

Cette méthode bien que très élégante est très coûteuse puisqu'elle nécessite plus de $n!$ opérations où $n!$ est la factorielle de $n$, donnée par $n! = n \times (n-1) \times \ldots \times 2 \times 1$. Elle n'est donc jamais utilisée en pratique sauf en dimension $n = 2$. Cet algorithme revient à calculer explicitement la matrice $A^{-1}$. Hélas, ce calcul est souvent long et fastidieux, même pour un ordinateur, c'est pourquoi nous avons recours à des algorithmes de résolution exacte d'une complexité moindre (nous parlons alors d'une méthode directe), ou des méthodes itératives qui consistent à construire une suite de solutions approchées qui converge vers la solution exacte. Avant de présenter de tels algorithmes, nous introduirons des matrices dont la structure particulière est bien adaptée à la résolution du système linéaire (1.1). Puis nous verrons comment se ramener à ces cas particuliers.

1. En référence au mathématicien suisse Gabriel Cramer (1704-1752).

### 1.2.2 Quelques matrices particulières

Cette partie constitue la base théorique permettant de mettre au point des méthodes pour la résolution exacte de (1.1).

- *Matrices triangulaires*

Nous introduisons la notion de matrice triangulaire et montrons que pour les matrices de ce type, la résolution du système linéaire (1.1) devient très facile.

**Définition 1.9**

Soit $A \in \mathscr{M}_{n,n}(\mathbb{K})$,

- la matrice $A$ est une matrice triangulaire inférieure si $A = (a_{i,j})_{1 \leqslant i,j \leqslant n}$ avec $a_{i,j} = 0, \quad 1 \leqslant i < j \leqslant n$.
- $A$ est une matrice triangulaire supérieure lorsque $a_{i,j} = 0, 1 \leqslant j < i \leqslant n$.
- $A$ est une matrice diagonale dès lors que $a_{i,j} = 0$ pour $i \neq j$.

Nous vérifions d'abord que l'ensemble des matrices triangulaires supérieures (resp. inférieures) est stable par la somme et le produit. En outre, nous avons le résultat suivant.

**Proposition 1.10**

Soit $L \in \mathscr{M}_{n,n}(\mathbb{K})$ une matrice triangulaire inférieure inversible, ce qui signifie que tous les éléments diagonaux sont non nuls. Alors $L^{-1}$ est aussi une matrice triangulaire inférieure.
Soit $U$ une matrice triangulaire supérieure inversible. Alors $U^{-1}$ est aussi une matrice triangulaire supérieure.

Les matrices triangulaires jouent un rôle important en analyse numérique car elles sont facilement inversibles ou du moins, nous pouvons facilement trouver la solution $x \in \mathbb{K}^n$ du système linéaire $A\,x = b$. En effet, considérons le cas d'une matrice triangulaire supérieure, alors la solution $x$ se calcule par un algorithme dit de remontée. Nous observons d'abord qu'une matrice triangulaire inversible a tous ses éléments diagonaux non nuls, c'est pourquoi nous pouvons écrire $x_n = b_n / a_{n,n}$ puis pour tout $i = n-1, n-2, \ldots, 1$, nous pouvons calculer $x_i$ de la manière suivante :

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

***Exemple***

Soit $A \in \mathscr{M}_{3,3}(\mathbb{R})$ donnée par

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 2 \\ 0 & 0 & 3 \end{pmatrix},$$