

N° d'ordre :



Université Sultan Moulay Slimane
Faculté des Sciences et Techniques



Polycopié de Cours

Module : Mathématiques pour l'Intelligence Artificielle

MASTER : Artificial Intelligence & Digital Computation

Partie I

Rédigé par

Pr. ABDLEKBIR AFRAITES

Année Universitaire : 2023-2024

Table des matières

0	Révisions d'algèbre linéaire	4
0.1	Espace vectoriel	4
0.1.1	Définition de l'espace vectoriel	4
0.1.2	Sous-espace vectoriel	4
0.1.3	Famille libre, famille génératrice	5
0.1.4	Base d'un espace vectoriel	5
0.2	Applications Linéaires	6
0.2.1	Définition de l'application linéaire	6
0.2.2	Composition et réciproque des applications linéaires	6
0.3	Matrices	7
0.3.1	Définition des matrices	7
0.3.2	Somme et produit de matrices	8
0.3.3	Inverse et transposée d'une matrice	9
0.3.4	Rang d'une application linéaire, d'une matrice	10
0.4	Déterminants	11
0.4.1	Définition du déterminant par récurrence	11
0.4.2	Le déterminant et les colonnes	11
0.4.3	Propriétés essentielles du déterminant	12
0.4.4	Propriétés liées au rang	13
0.5	Systèmes linéaires	14
0.5.1	Systèmes linéaires à matrice carrée	14
0.6	Valeurs propres	14
0.6.1	Valeurs propres et matrices semblables	15
0.6.2	Diagonalisation des matrices	16
0.6.3	Valeurs propres et matrices définies positives	17
1	Résolution des systèmes linéaires : Méthodes directes	18
1.1	Motivations	18
1.1.1	Equation de la chaleur	18
1.2	Méthode d'élimination de Gauss	19
1.2.1	Résolution d'un système triangulaire	19
1.2.2	Principe de la méthode de Gauss	20
1.2.3	Méthode de Gauss : cas général	22
1.2.4	Algorithme d'élimination de Gauss	24
1.2.5	Écriture matricielle de l'élimination de Gauss	24
1.2.6	Unicité de la factorisation LU	26
1.3	Calcul direct de la factorisation $A = LU$	26

1.3.1	Motivations de la factorisation LU	26
1.3.2	Principe du calcul direct de la factorisation LU	27
1.3.3	Algorithme de Doolittle (factorisation LU)	28
1.4	Traitement des matrices symétriques	29
1.4.1	La factorisation LDL^T	29
1.4.2	La factorisation de Cholesky : existence	30
1.4.3	Algorithme de Cholesky	31
1.5	Normes matricielles, conditionnement des matrices	33
1.5.1	Normes vectorielles	33
1.5.2	Définition de la norme matricielle	33
1.5.3	Rayon spectral	34
1.5.4	Introduction au conditionnement d'une matrice	34
1.5.5	Lien entre le conditionnement et les erreurs	36
2	Résolution des systèmes linéaires : Méthodes itératives	38
2.1	Principes généraux	38
2.2	La méthode de Jacobi	39
2.3	La méthode de Gauss—Seidel	40
2.4	Etude de la convergence des méthodes itératives	41
2.4.1	Définitions et propositions	41
2.4.2	Etude de la convergence	42

Révisions d'algèbre linéaire

0.1 Espace vectoriel

0.1.1 Définition de l'espace vectoriel

Définition 0.1.1 *Un espace vectoriel sur K ($K = \mathbb{R}$ ou \mathbb{C}) est un ensemble E (dont les éléments sont appelés vecteurs) possédant les deux lois suivantes :*

1. *l'addition de vecteurs (loi interne qui donne à E une structure de groupe commutatif), c'est à dire que, pour tout $\vec{x}, \vec{y}, \vec{z}$ de E :*
 - *associativité : $(\vec{x} + \vec{y}) + \vec{z} = \vec{x} + (\vec{y} + \vec{z})$,*
 - *existence d'un élément neutre $\vec{e} \in E : \vec{x} + \vec{e} = \vec{e} + \vec{x} = \vec{x}$,*
 - *tout élément $\vec{x} \in E$ a un symétrique (opposé) $\vec{x}' \in E : \vec{x} + \vec{x}' = \vec{x}' + \vec{x} = \vec{e}$,*
 - *commutativité : $\vec{x} + \vec{y} = \vec{y} + \vec{x}$.*
2. *le produit d'un vecteur par un élément de K (loi externe dont le résultat est un vecteur de E) qui possède les propriétés suivantes : $\forall \lambda, \mu \in K ; \forall x, y \in E$:*
 - *$(\lambda\mu)\vec{x} = \lambda(\mu\vec{x})$,*
 - *$(\lambda + \mu)\vec{x} = \lambda\vec{x} + \mu\vec{x}$,*
 - *$\lambda(\vec{x} + \vec{y}) = \lambda\vec{x} + \lambda\vec{y}$,*
 - *$1\vec{x} = \vec{x}$ (1 élément unité de K).*

0.1.2 Sous-espace vectoriel

Soit E un espace vectoriel sur K (où K représente \mathbb{R} ou \mathbb{C}), on peut alors définir un sous-espace vectoriel de E

Définition 0.1.2 *On appelle sous espace vectoriel F de E un sous ensemble non vide de E qui reste stable pour les opérations sur E (que l'on notera s.e.v. en abrégé), c'est-à-dire :*

- $\forall \vec{x} \in F, \forall \vec{y} \in F, \vec{x} + \vec{y} \in F,$
- $\forall \lambda \in K, \forall \vec{x} \in F, \lambda\vec{x} \in F$

Cette définition est parfois donnée sous la forme synthétique suivante :

$$\forall \vec{x}, \vec{y} \in F, \forall \lambda, \mu \in K, \lambda\vec{x} + \mu\vec{y} \in F.$$

Par exemple, dans \mathbb{R}^2 , les droites de vecteur directeur \vec{x} sont des sous-espaces vectoriels ($F = \{\lambda\vec{x}, \lambda \in \mathbb{R}\}$).

Dans \mathbb{R}^3 , les plans engendrés par deux vecteurs \vec{x} et \vec{y} non proportionnels sont des sous-espaces vectoriels ($F = \{\lambda\vec{x} + \mu\vec{y}, \lambda, \mu \in \mathbb{R}\}$).

L'espace des polynômes P_k est un sous-espace vectoriel de P_n si $k \leq n$.

0.1.3 Famille libre, famille génératrice

Définition 0.1.3 On dit que la famille $\{\vec{x}_1, \dots, \vec{x}_p\}$ de vecteurs d'un espace vectoriel E est liée s'il existe $\lambda_1, \dots, \lambda_p \in K$ non tous nuls tels que :

$$\lambda_1 \vec{x}_1 + \lambda_2 \vec{x}_2 + \dots + \lambda_p \vec{x}_p = \vec{0}$$

Dans le cas où la famille possède plus de deux vecteurs, cette définition équivaut à dire qu'il existe un des vecteurs de la famille qui est combinaison linéaire des autres.

Définition 0.1.4 Une famille qui n'est pas liée est appelée famille libre, dans ce cas on a

$$\{\lambda_1 \vec{x}_1 + \lambda_2 \vec{x}_2 + \dots + \lambda_p \vec{x}_p = \vec{0}\} \iff \{\lambda_1 = \lambda_2 = \dots = \lambda_p = 0\}.$$

Ainsi la famille de vecteurs de \mathbb{R}^2 $\{\vec{x}_1, \vec{x}_2, \vec{x}_3\} = \{(1, 0), (0, 1), (1, 2)\}$ est liée puisque $\vec{x}_3 = \vec{x}_1 + 2\vec{x}_2$. Par contre la famille $\{\vec{x}_1, \vec{x}_2\} = \{(1, 0), (0, 1)\}$ est libre (le démontrer).

Définition 0.1.5 On dit que la famille $\{\vec{x}_1, \vec{x}_2, \dots, \vec{x}_p\}$ de vecteurs d'un espace vectoriel E est génératrice si tout vecteur \vec{x} de E est combinaison linéaire de $\vec{x}_1, \vec{x}_2, \dots, \vec{x}_p$, c'est-à-dire s'il existe $\lambda_1, \dots, \lambda_p \in K$ tels que :

$$\vec{x} = \lambda_1 \vec{x}_1 + \lambda_2 \vec{x}_2 + \dots + \lambda_p \vec{x}_p$$

On dit alors que E est engendré par la famille $\{\vec{x}_1, \dots, \vec{x}_p\}$ et on note :

$$E = Vect < \vec{x}_1, \dots, \vec{x}_p > .$$

0.1.4 Base d'un espace vectoriel

Nous considérons un espace vectoriel E qui admet une famille génératrice avec un nombre fini d'éléments. Par exemple $\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$ est une famille génératrice de \mathbb{R}^3 et $\{1, x, x^2, \dots, x^n\}$ est une famille génératrice de P_n . Certains espaces vectoriels, par exemple les fonctions continues réelles, n'ont pas de famille génératrice avec un nombre fini d'éléments, mais ce n'est pas l'objet de ce cours.

Définition 0.1.6 Une base $\mathbb{B} = \{\vec{e}_1, \dots, \vec{e}_n\}$ d'un espace vectoriel E est une famille libre et génératrice.

Un vecteur \vec{x} de E est une combinaison linéaire d'éléments de \mathbb{B} puisque \mathbb{B} est une famille génératrice et on peut montrer que cette décomposition est unique (à démontrer).

Définition 0.1.7 Un espace vectoriel de dimension finie E est un espace vectoriel qui admet une base ayant un nombre fini d'éléments : ce nombre est appelé la dimension de E .

Remarquons que, pour un espace vectoriel donné de dimension finie, on peut montrer que toutes les bases ont le même nombre d'éléments, ce qui justifie la définition de la dimension d'un espace vectoriel.

Par exemple, on définit le vecteur \vec{e}_i de \mathbb{R}^n dont toutes les composantes sont nulles sauf la i ème composante qui vaut 1, alors la famille $\{\vec{e}_1, \dots, \vec{e}_n\}$ est une base de \mathbb{R}^n , ce qui signifie que \mathbb{R}^n est un espace vectoriel de dimension n . De même la famille $\{1, x, x^2, \dots, x^n\}$ est une base de P_n , donc la dimension de l'espace vectoriel P_n est égale à $n + 1$.

Un résultat intéressant, car il simplifie les démonstrations est le suivant :

Proposition 0.1.1 Soit E un espace vectoriel de dimension n , alors

- toute famille libre de n vecteurs est une base,
- toute famille génératrice de n vecteurs est une base.

0.2 Applications Linéaires

0.2.1 Définition de l'application linéaire

On considère deux espaces vectoriels E et F définis tous les deux soit sur $K = \mathbb{R}$, soit sur $K = \mathbb{C}$. On définit alors l'application linéaire :

Définition 0.2.1 On appelle application linéaire $u : E \longrightarrow F$, une application possédant les propriétés suivantes :

$$\begin{aligned} u(\vec{x} + \vec{y}) &= u(\vec{x}) + u(\vec{y}), \quad \forall \vec{x}, \vec{y} \in E, \\ u(\lambda \vec{x}) &= \lambda u(\vec{x}), \quad \forall \vec{x} \in E, \quad \forall \lambda \in K \end{aligned}$$

L'application de \mathbb{R} dans \mathbb{R} qui à x associe e^x n'est pas linéaire, par contre celle qui à x associe $3x$ est linéaire.

Notons que l'image par u du vecteur nul de E est le vecteur nul de F (il suffit de prendre $\lambda = 0$ dans la définition).

Définition 0.2.2 On appelle noyau de u le sous espace vectoriel de E noté $\text{Ker } u$ tel que :

$$\vec{x} \in \text{Ker } u \iff u(\vec{x}) = 0$$

On appelle image de u le sous espace vectoriel de F noté $\text{Im } u$ formé des éléments $u(\vec{x})$, quand \vec{x} parcourt l'espace E , soit

$$\vec{y} \in \text{Im } u \iff \exists \vec{x} \in E, \text{ tel que } \vec{y} = u(\vec{x})$$

Par exemple, si u est l'application linéaire de \mathbb{R}^3 dans \mathbb{R} définie par $u(\vec{x}) = x_1 + 3x_2 + 5x_3$ où $\vec{x} = (x_1, x_2, x_3)$, alors $\text{Ker } u$ est le plan $x_1 + 3x_2 + 5x_3 = 0$ et $\text{Im } u$ est \mathbb{R} .

Il existe une relation entre les dimensions de ces sous-espaces vectoriels qui est donnée par la proposition suivante :

Proposition 0.2.1 Soit $u : E \longrightarrow F$ une application linéaire, alors :

$$\dim E = \dim (\text{Ker } u) + \dim (\text{Im } u).$$

0.2.2 Composition et réciproque des applications linéaires

Rappel : Injectivité, surjectivité

Définition 0.2.3 une application $f : E \longrightarrow F$ est surjective si $\text{Im } f = F$.

Définition 0.2.4 une application $f : E \longrightarrow F$ est injective si

$$\forall \vec{x} \in E, \quad \forall \vec{y} \in E, \quad f(\vec{x}) = f(\vec{y}) \implies \vec{x} = \vec{y}.$$

ce qui veut dire que deux éléments distincts ont des images distinctes.

Maintenant, On considère trois espaces vectoriels E , F et G définis sur le même corps K (\mathbb{R} ou \mathbb{C}).

Définition 0.2.5 Soient $f : E \longrightarrow F$ et $g : F \longrightarrow G$ deux applications linéaires, la composée (gof) de f et g est une application linéaire définie par :

$$\begin{aligned} gof : E &\longrightarrow G \\ \vec{x} &\longmapsto g[f(\vec{x})]. \end{aligned} \quad (1)$$

La composée est, en effet une application linéaire puisque :

$$\begin{aligned} (gof)(\vec{x} + \vec{y}) &= go[f(\vec{x} + \vec{y})] = go[f(\vec{x}) + f(\vec{y})] = go[f(\vec{x})] + go[f(\vec{y})] = (gof)(\vec{x}) + (gof)(\vec{y}) \\ (gof)(\lambda\vec{x}) &= go[f(\lambda\vec{x})] = go[\lambda f(\vec{x})] = \lambda go[f(\vec{x})] = \lambda (gof)(\vec{x}) \end{aligned}$$

Rappelons que l'inverse d'une application (non nécessairement linéaire) se définit par composition :

Définition 0.2.6 Soit f une application de E dans F , elle est inversible s'il existe une application de F dans E notée f^{-1} telle que :

$$f^{-1}of = i_E \quad \text{et} \quad fof^{-1} = i_F$$

i_E est l'identité de E dans E .

L'application f^{-1} est appelée application réciproque (ou inverse) de f .

Rappelons aussi qu'une application est inversible si et seulement si elle est injective et surjective, c'est-à-dire bijective et alors f^{-1} est aussi bijective. Dans le cas des applications linéaires, on a une caractérisation de l'injectivité par le noyau :

Proposition 0.2.2 Une application linéaire u est injective si et seulement si $\text{Ker } u = \{\vec{0}\}$.

Démonstration : Supposons que u est injective et soit $\vec{x} \in \text{Ker } u$, alors on a $u(\vec{x}) = 0$ et $u(\vec{0}) = 0$ d'où $\vec{x} = \vec{0}$ par injectivité et $\text{Ker } u = \{\vec{0}\}$.

Réciproquement, soient \vec{x} et \vec{x}' deux vecteurs tels que $u(\vec{x}) = u(\vec{x}')$, alors, par linéarité, on a $u(\vec{x} - \vec{x}') = 0$ et comme $\text{Ker } u = \{\vec{0}\}$, $\vec{x} - \vec{x}' = \vec{0}$, d'où u est injective.

Pour les applications linéaires u bijectives, on démontre que l'application réciproque u^{-1} est aussi linéaire.

0.3 Matrices

0.3.1 Définition des matrices

On considère deux espaces vectoriels E et F définis sur le même corps K (\mathbb{R} ou \mathbb{C}) et tels que $\dim E = n$ et $\dim F = m$.

Définition 0.3.1 Soit $u : E \longrightarrow F$ une application linéaire, on appelle matrice associée à u dans les bases $\mathcal{E} = \{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ et $\mathcal{F} = \{\vec{f}_1, \vec{f}_2, \dots, \vec{f}_m\}$ le tableau A de scalaires (c'est-à-dire d'éléments de K) à m lignes et n colonnes construit de la façon suivante : la j ème colonne de A est constituée par les composantes dans la base de F du vecteur $u(\vec{e}_j)$. Autrement dit si on note a_{ij} l'élément du tableau A situé à la i ème ligne et la j ème colonne, on a :

$$u(\vec{e}_j) = \sum_{i=1}^m a_{ij} \vec{f}_i, \quad 1 \leq j \leq n.$$

La matrice A est donc représentée par :

$$A = \begin{pmatrix} a_{11} & a_{12} & . & . & . & a_{1j} & . & . & . & a_{1n} \\ a_{21} & a_{22} & . & . & . & a_{2j} & . & . & . & a_{2n} \\ . & . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . & . \\ a_{i1} & a_{i2} & . & . & . & a_{ij} & . & . & . & a_{in} \\ . & . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . & . \\ a_{m1} & a_{m2} & . & . & . & a_{mj} & . & . & . & a_{mn} \end{pmatrix} \quad (2)$$

On dit que la matrice A est de format (m, n) , ou de type (m, n) si elle a m lignes et n colonnes et on appellera \mathcal{M}_{mn} l'ensemble de telles matrices. (On pourra préciser le corps K si c'est nécessaire.)

- Si $m = n$ la matrice est dite carrée.
- Si $m = 1$, la matrice est une matrice ligne ou encore un vecteur ligne.
- Si $n = 1$, la matrice est une matrice colonne ou encore un vecteur colonne.
- Si $m = n$ et $a_{ij} = 0$ pour $i \neq j$, la matrice est diagonale.
- Si $m = n$, si $a_{ij} = 0$ pour $i \neq j$ et si $a_{ii} = 1$, la matrice est la matrice identité.
- Si $m = n$ et $a_{ij} = 0$ pour $i < j$, la matrice est triangulaire inférieure.
- Si $m = n$ et $a_{ij} = 0$ pour $i > j$, la matrice est triangulaire supérieure.

Enfin, on notera dans la suite A_j , $1 \leq j \leq n$, la j ème colonne de A et \underline{A}_i , $1 \leq i \leq m$, la i ème ligne de A .

0.3.2 Somme et produit de matrices

Une matrice se présente donc comme un tableau de nombres (de \mathbb{R} ou \mathcal{C}), mais on ne peut raisonnablement rien démontrer sur les matrices sans faire référence aux applications linéaires qu'elles peuvent représenter. Nous ferons un usage constant de cette référence en considérant que toute matrice à m lignes et n colonnes peut être considérée comme la matrice de l'application linéaire $u : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ (ou $u : \mathcal{C}^n \longrightarrow \mathcal{C}^m$) définie par

$$u(\vec{e}_j) = \sum_{i=1}^m a_{ij} \vec{e}_i \quad 1 \leq j \leq n.$$

où $\{\vec{e}_1, \dots, \vec{e}_n\}$ et $\{\vec{f}_1, \dots, \vec{f}_m\}$ sont respectivement les bases de E et F .

On peut ainsi définir la somme de deux matrices et le produit d'une matrice par un scalaire.

Définition 0.3.2 Soient deux matrices A et B de même format (m, n) , on définit $A + B$, somme de A et B comme étant la matrice C , également de format (m, n) , dont le coefficient c_{ij} est donné par

$$c_{ij} = a_{ij} + b_{ij}, 1 \leq i \leq m, 1 \leq j \leq n$$

(C est la matrice de l'application linéaire somme des applications linéaires représentées par A et B). De même le produit du scalaire λ par la matrice A est la matrice, notée λA obtenue en multipliant chaque coefficient de A par λ .

Soient trois espaces vectoriels E , F et G tels que $\dim E = n$, $\dim F = m$ et $\dim G = q$.

Définition 0.3.3 Soient $u : F \longrightarrow G$ et $v : E \longrightarrow F$, on pose $w = u \circ v$ on a donc $w : E \longrightarrow G$. Soit A la matrice de u dans les bases \mathcal{F} et \mathcal{G} et B la matrice de v dans les bases \mathcal{E} et \mathcal{F} . Alors, par définition, la matrice de w dans les bases \mathcal{E} et \mathcal{G} est égale au produit de A par B noté AB .

On peut démontrer le résultat suivant :

Proposition 0.3.1 *Le produit $C = AB$ des deux matrices est donné par la formule suivante :*

$$c_{ij} = \sum_{k=1}^m a_{ik}b_{kj}, \quad i = 1 \dots q, \quad j = 1 \dots n$$

Proposition 0.3.2 *Soit u une application linéaire de E dans F , on note $y = u(\vec{x})$, soient \mathcal{E} et \mathcal{F} deux bases respectives de E et F , soient x_1, x_2, \dots, x_n (resp y_1, y_2, \dots, y_p) les composantes de \vec{x} (resp \vec{y}) dans la base \mathcal{E} (resp \mathcal{F}), on note*

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \quad \text{et} \quad Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_p \end{pmatrix} \quad (3)$$

Si A est la matrice associée à u dans les bases \mathcal{E} et \mathcal{F} , on a alors : $Y = AX$

0.3.3 Inverse et transposée d'une matrice

Soit u une application linéaire et bijective de E dans E , alors elle est inversible et la matrice A qui lui correspond est carrée. La matrice qui correspond à u^{-1} est notée A^{-1} et vérifie

$$AA^{-1} = A^{-1}A = I,$$

d'où la définition.

Définition 0.3.4 *Soit A une matrice carrée, la matrice inverse de A notée A^{-1} , si elle existe, est définie par*

$$AA^{-1} = A^{-1}A = I.$$

L'inverse d'une matrice, s'il existe est unique. D'autre part, puisque

$$ABB^{-1}A^{-1} = B^{-1}A^{-1}AB,$$

on en déduit que :

$$(AB)^{-1} = B^{-1}A^{-1}.$$

Définition 0.3.5 *Soit $A \in \mathcal{M}_{mn}$, on appelle **transposée** de A la matrice notée A^T appartenant à \mathcal{M}_{nm} obtenue en échangeant les lignes et les colonnes de A , on a donc*

$$(A^T)_{ij} = a_{ji}, \quad i = 1, \dots, n, \quad j = 1, \dots, m.$$

On peut démontrer très facilement les propriétés suivantes :

$$(A^T)^T = A, \quad (A + B)^T = A^T + B^T, \quad (\lambda A)^T = \lambda A^T.$$

Si X et Y sont deux vecteurs colonnes alors $X^TY = Y^TX$ est un scalaire de K (\mathbb{R} ou \mathbb{C}) et plus précisément on a : $X^TY = Y^TX = \sum_{i=1}^m x_i y_i$. Lorsque le produit AB existe, alors on a $(AB)^T = B^T A^T$. Cela se démontre aisément en comparant les éléments des deux matrices.

0.3.4 Rang d'une application linéaire, d'une matrice

Définition 0.3.6 Soit u une application linéaire de E dans F , on appelle **rang** de u et on note $\text{rang}(u)$ la dimension de $\text{Im } u$ (image de u).

Puisque $\text{Im } u$ est un sous-espace vectoriel de F , on a $\text{rang}(u) \leq \dim F$ avec égalité si u est surjective. D'autre part, on a vu que

$$\dim E = \dim (\text{Ker } u) + \dim (\text{Im } u)$$

donc on a aussi $\text{rang}(u) \leq \dim E$.

Définition 0.3.7 Soit $A \in \mathcal{M}_{mn}$, on appellera $\text{Im } A$, le sous espace vectoriel constitué des vecteurs colonnes qui s'écrivent $Y = AX$.

$\text{Im } A$ est le sous-espace vectoriel engendré par les colonnes de A , notées A_1, A_2, \dots, A_n , puisque

$$Y \in \text{Im } A \iff Y = AX \iff Y = \sum_{i=1}^n x_i A_i$$

On peut aussi définir le noyau $\text{Ker } A$ d'une matrice A de la façon suivante :

Définition 0.3.8 On appelle noyau d'une matrice A le sous espace vectoriel noté $\text{Ker } A$, défini par :

$$\text{Ker } A = \{X \in \mathcal{M}_{n1}, AX = 0\}$$

On a le résultat suivant (déjà énoncé dans le paragraphe référencé pour les applications linéaires) :

Théorème 0.3.1 Soit $A \in \mathcal{M}_{mn}$, alors :

$$\dim [\text{Ker } A] + \dim [\text{Im } A] = n.$$

Définition 0.3.9 Soit $A \in \mathcal{M}_{mn}$, on appelle rang de A et on note $\text{rang}(A)$, la dimension du sous-espace vectoriel $\text{Im } A$.

Proposition 0.3.3 Soit $u \in \mathcal{L}(E, F)$ et A une matrice représentant u (dans des bases arbitraires de E et F) alors :

- $\text{rang}(u) = \text{rang}(A)$
- A et A^T ont le même rang

Calcul pratique du rang d'une matrice : Il résulte de la définition et des résultats précédents que le calcul pratique du rang d'une matrice se ramène au calcul du nombre de colonnes linéairement indépendantes ou du nombre de lignes linéairement indépendantes suivant que l'un se calcule plus facilement que l'autre. On peut aussi déduire le rang d'une matrice à partir du rang de l'application linéaire u associée si celui-ci est connu.

On verra dans le paragraphe sur les déterminants une autre façon de calculer le rang d'une matrice.

0.4 Déterminants

0.4.1 Définition du déterminant par récurrence

Comme vous allez le comprendre dès la définition, la notion de déterminant ne peut être introduite que pour les matrices carrées.

Définition 0.4.1 Soit $A \in \mathcal{M}_{n,n}$, on définit par récurrence une application :

$$\begin{aligned} \det : \mathcal{M}_{n,n} &\longrightarrow K \\ A &\longmapsto \det A. \end{aligned} \quad (4)$$

de la manière suivante :

- si $n = 1$, $A = (a)$, on pose $\det A = a$,
- si $n > 1$, notons $A_{[i,j]}$ la matrice obtenue à partir de A en supprimant la i ème ligne et la j ème colonne, on pose alors :

$$\det A = a_{11}\det A_{[1,1]} + \dots + (-1)^{k+1}a_{1k}\det A_{[1,k]} + \dots + (-1)^{n+1}a_{1n}\det A_{[1,n]}$$

Le scalaire **det A** est dit **déterminant** de A et on le note habituellement :

$$\begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix}. \quad (5)$$

Ainsi,

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21} \quad (6)$$

et

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} \quad (7)$$

ce qui donne

$$\det A = a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + a_{12}a_{31}a_{23} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}$$

Proposition 0.4.1 La définition permet d'obtenir immédiatement les propriétés suivantes :

- Le déterminant d'une matrice triangulaire inférieure est égal au produit des termes diagonaux.
- Le déterminant d'une matrice diagonale est égal au produit des termes diagonaux. En particulier le déterminant de la matrice identité est égal à 1.
- Si \bar{A} est la matrice dont les termes sont les conjugués de ceux de A alors

$$\det \bar{A} = \overline{\det A}$$

0.4.2 Le déterminant et les colonnes

Théorème 0.4.1 Le déterminant est une fonction linéaire de chaque colonne, donc une application multi-linéaire de l'ensemble des colonnes, c'est à dire :

$$\det (A_1, \dots, A_{k-1}, \lambda A_k, A_{k+1}, \dots, A_n) = \lambda \det (A_1, \dots, A_{k-1}, A_k, A_{k+1}, \dots, A_n)$$

$$\det(A_1, \dots, A_{k-1}, B+C, A_{k+1}, \dots, A_n) = \det(A_1, \dots, A_{k-1}, B, A_{k+1}, \dots, A_n) + \det(A_1, \dots, A_{k-1}, C, A_{k+1}, \dots, A_n)$$

où $\lambda \in K$, B et C sont des vecteurs colonnes.

En conséquence, si $A \in \mathcal{M}_{n,n}$, on a :

$$\det(\lambda A) = \det(\lambda A_1, \lambda A_2, \dots, \lambda A_n) = \lambda^n \det(A)$$

puisque "on sort un λ par colonne" et

$$\det(A_1, \dots, A_{k-1}, 0, A_{k+1}, \dots, A_n) = 0$$

puisque une colonne nulle peut être considérée comme le produit du réel 0 par une colonne quelconque.

Théorème 0.4.2 Soit $A \in \mathcal{M}_{n,n}$

1. Si deux colonnes sont égales, le déterminant est nul.
2. Si on échange entre elles deux colonnes de la matrice, le déterminant change de signe.
3. Le déterminant d'une matrice ne change pas si à une colonne on ajoute une combinaison linéaire des autres colonnes.

0.4.3 Propriétés essentielles du déterminant

Les résultats de ce paragraphe sont à la base de la théorie des systèmes linéaires et ne seront pas démontrés.

Théorème 0.4.3 Soit $A \in \mathcal{M}_{n,n}$, alors $\det A = \det A^T$.

Ce résultat permet d'étendre de manière évidente aux lignes les propriétés du déterminant liées aux colonnes de la matrice. En particulier

1. le déterminant est une fonction multilinéaire de chacune des lignes,
2. si une matrice a deux lignes égales, le déterminant est nul,
3. si l'on échange deux lignes de A , le déterminant change de signe,
4. le déterminant d'une matrice ne change pas si à une ligne, on ajoute une combinaison linéaire des autres lignes.

Les propriétés précédentes permettent de calculer pratiquement le déterminant par récurrence à partir de n'importe quelle ligne ou n'importe quelle colonne :

Théorème 0.4.4 On a les formules suivantes :

1. Développement suivant la i ème ligne

$$\det A = a_{i1} \text{cof}(a_{i1}) + a_{i2} \text{cof}(a_{i2}) + \dots + a_{in} \text{cof}(a_{in})$$

2. Développement suivant la j ème colonne

$$\det A = a_{1j} \text{cof}(a_{1j}) + a_{2j} \text{cof}(a_{2j}) + \dots + a_{nj} \text{cof}(a_{nj})$$

où

$$\text{cof}(a_{ij}) = (-1)^{i+j} \det A_{[i,j]}$$

s'appelle le **cofacteur** de l'élément a_{ij} .

Le théorème suivant est l'un des plus utilisés de l'algèbre linéaire :

Théorème 0.4.5

$$A \text{ est inversible} \iff \det A \neq 0$$

Le déterminant d'un produit est donné par le théorème suivant :

Théorème 0.4.6 Soient $A, B \in \mathcal{M}_{n,n}$ alors

$$\det (AB) = \det A \det B.$$

De ce théorème on peut déduire que si une matrice est inversible, alors

$$\det (AA^{-1}) = \det A \det A^{-1} = \det I = 1.$$

0.4.4 Propriétés liées au rang

La notion de rang est essentielle dans la résolution des problèmes de lissage (dits aussi problèmes de moindres carrés) que nous traiterons par la suite.

Définition 0.4.2 Soit $A \in \mathcal{M}_{m,n}$, on appelle **matrice extraite** de A une matrice obtenue en sélectionnant des lignes et des colonnes de A . On peut donc se définir une matrice extraite par deux ensembles d'indices

$$I = \{i_1, i_2, \dots, i_p\} \subset \{1, 2, \dots, m\}, J = \{j_1, j_2, \dots, j_q\} \subset \{1, 2, \dots, n\}$$

les éléments de la matrice extraite $\hat{A} \in \mathcal{M}_{p,q}$ sont alors $\hat{a}_{kl} = a_{i_k, j_l}$.

Par exemple si $I = \{1, 3\}, J = \{2, 3, 5\}$

$$A = \begin{pmatrix} 1 & 7 & 5 & 6 & 8 \\ 3 & 4 & 8 & -1 & 6 \\ 7 & 5 & 4 & 3 & 2 \end{pmatrix} \text{ alors : } \hat{A} = \begin{pmatrix} 7 & 5 & 8 \\ 5 & 4 & 2 \end{pmatrix}$$

Notation : En relation avec cette notion de matrice extraite on peut définir une matrice par "blocs", par exemple : $A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$

où $A_{ij} \in \mathcal{M}_{m_i, n_j}$, on a donc $A \in \mathcal{M}_{m,n}$ avec $m = m_1 + m_2$ et $n = n_1 + n_2$.

Théorème 0.4.7 Soit $A \in \mathcal{M}_{m,n}$ alors le rang de A est le plus grand entier r tel qu'il existe une matrice carrée inversible $\hat{A} \in M_{r,r}$ extraite de A .

Ce théorème sert parfois à déterminer le rang d'une matrice mais plutôt, si on connaît le rang r de A , on sait qu'il est possible d'extraire de A une matrice $r \times r$ inversible. Ce théorème sert aussi à démontrer la proposition qui avait déjà été énoncée dans le paragraphe précédent.

Proposition 0.4.2 On a $\text{rang}(A) = \text{rang}(A^T)$.

Démonstration : Si \hat{A} est une matrice carrée inversible extraite de A alors \hat{A}^T est une matrice carrée inversible extraite de A^T et donc le résultat est une conséquence du théorème précédent.

On peut alors caractériser l'inversibilité d'une matrice par son rang.

Proposition 0.4.3 Soit $A \in \mathcal{M}_{n,n}$

$$A \text{ est inversible} \iff \det A \neq 0 \iff \text{Ker } A = \{0\} \iff \text{rang } A = n.$$

0.5 Systèmes linéaires

0.5.1 Systèmes linéaires à matrice carrée

Dans la résolution des systèmes linéaires on a l'habitude de noter x et non pas X le vecteur des inconnues dont les composantes sont (x_1, \dots, x_n) et b le second membre dont les composantes sont (b_1, \dots, b_n) .

De la définition de l'image de A , on déduit que, si $A \in \mathcal{M}_{nn}$, le système $Ax = b$ admet une solution si et seulement si $b \in \text{Im } A$. Mais ce résultat théorique est difficilement vérifiable ! Le théorème suivant, par contre, est essentiel.

Théorème 0.5.1 *Soit $A \in \mathcal{M}_{nn}$, le système $Ax = b$ admet une solution unique si et seulement si $\det A \neq 0$.*

Démonstration : Si $\det A \neq 0$, A est inversible et on constate alors que $x = A^{-1}b$ est l'unique solution de $Ax = b$.

Réciproquement, si $\det A = 0$, A n'est pas inversible et $\text{Ker } A$ n'est pas réduit à 0. Donc il existe un x^* non nul vérifiant $Ax^* = 0$ et le système $Ax = b$ ne peut admettre une solution unique puisque, si x est solution, $x + x^*$ est une autre solution.

Le théorème suivant donne la solution "théorique" d'un système à matrice carrée, encore dit **système de Cramer**.

Théorème 0.5.2 *Soit $A \in \mathcal{M}_{n,n}$ et soit le système $Ax = b$. Alors pour tout $i = 1, 2, \dots, n$, on a la formule de Cramer :*

$$\det(A_1, A_2, \dots, A_{i-1}, b, A_{i+1}, \dots, A_n) = x_i \det A$$

Ceci permet aussi de calculer l'inverse d'une matrice A , puisque chaque colonne X_i de A^{-1} est solution de $AX_i = I_i$ où I_i est la i ème colonne de I .

Théorème 0.5.3 *L'inverse de A est donnée par*

$$A^{-1} = \frac{1}{\det A} (\text{Co}(A))^T;$$

où

$$\text{Co}(A)_{ij} = \text{cof}(a_{ij}) = (-1)^{i+j} \det A_{[i,j]}$$

(La matrice $\text{co}(A)$ s'appelle co-matrice de A .)

Comme pour les systèmes linéaires, les formules de Cramer ne sont pas utilisées pour calculer numériquement l'inverse, on leur préfère des méthodes plus économiques par exemple des méthodes numériques. En revanche les formules de Cramer sont utilisées pour le calcul formel.

0.6 Valeurs propres

Définition 0.6.1 *On dit que $\lambda \in K$ est une **valeur propre** de $A \in \mathcal{M}_{n,n}$ s'il existe un vecteur $Y \in \mathcal{M}_{n,1}$ $Y \neq 0$ tel que :*

$$AY = \lambda Y$$

*Dans ce cas on dit que Y est un **vecteur propre** associé à la valeur propre λ et que (λ, Y) est un couple propre de A .*

Définition 0.6.2 On appelle **polynôme caractéristique** de A le polynôme

$$P_A(s) = \det (sI - A)$$

Proposition 0.6.1 Une condition nécessaire et suffisante pour que λ soit valeur propre de A est qu'elle soit racine du polynôme caractéristique, c'est à dire :

$$\det (\lambda I - A) = 0$$

Démonstration : Si l'on écrit la définition, on a :

$$\lambda \text{ est valeur propre de } A \iff \exists Y \neq 0 \text{ tel que } (\lambda I - A)Y = 0$$

$$\iff (\lambda I - A) \text{ n'est pas inversible } \iff \det (\lambda I - A) = 0$$

Pratiquement, les valeurs propres d'une matrice A sont les racines du polynôme $\det (A - sI)$ (qui a évidemment les mêmes racines que le polynôme caractéristique). La recherche des vecteurs propres se fait en résolvant alors le système $AY_i = \lambda_i Y_i$.

On a les propriétés suivantes :

- Les valeurs propres d'une matrice triangulaire sont ses termes diagonaux.
- A non inversible $\iff 0$ est valeur propre de A .
- A et A^T ont les mêmes valeurs propres mais pas les mêmes vecteurs propres.

Proposition 0.6.2 Une matrice $A \in \mathcal{M}_{nn}$ admet n valeurs propres complexes comptées avec leur ordre de multiplicité.

Si on appelle $\lambda_1, \lambda_2, \dots, \lambda_n$ ces valeurs propres (distinctes ou confondues), on a :

$$\text{Trace}(A) = \sum_{i=1}^n \lambda_i, \quad \det(A) = \prod_{i=1}^n \lambda_i.$$

0.6.1 Valeurs propres et matrices semblables

Définition 0.6.3 Deux matrices $A \in \mathcal{M}_{n,n}$ et $B \in \mathcal{M}_{nn}$ sont dites *semblables*, s'il existe une matrice inversible $P \in \mathcal{M}_{n,n}$ telle que $B = P^{-1}AP$.

Il est évident que deux matrices semblables ont le même déterminant. Par contre, la réciproque est fausse c'est-à-dire que deux matrices qui ont le même déterminant ne sont pas toujours semblables.

Proposition 0.6.3 Deux matrices semblables ont le même polynôme caractéristique et en particulier les mêmes valeurs propres.

Démonstration : Si A et B sont des matrices semblables, alors

$$sI - B = sI - P^{-1}AP = sP^{-1}P - P^{-1}AP = P^{-1}(sI - A)P$$

et donc

$$\det (sI - B) = \det (P^{-1}(sI - A)P) = \det P^{-1} \det (sI - A) \det (P) = \det (sI - A)$$

c'est-à-dire $P_A = P_B$.

Attention ! Si A et B sont semblables, elles ont bien les mêmes valeurs propres mais pas les mêmes vecteurs propres, cependant nous allons exhiber une relation entre les deux. En effet si (λ, Y) est un couple propre de A , on a

$$AY = \lambda Y \implies PBP^{-1}Y = \lambda Y \implies BP^{-1}Y = \lambda P^{-1}Y$$

donc $Z = P^{-1}Y$ est un vecteur propre de B associé à la valeur propre λ (ce qui, au passage, constitue une autre démonstration de la proposition précédente).

Remarquons que la méthode d'élimination de Gauss transforme un système $Ax = b$ en un système équivalent $Ux = c$ qui a la même solution et dont la matrice U est triangulaire supérieure. Malheureusement les matrices A et U ne sont pas semblables et l'élimination de Gauss ne permet pas de calculer les valeurs propres de A .

0.6.2 Diagonalisation des matrices

On a souvent l'habitude de ne considérer que des matrices à coefficients réels. Mais elles ne sont que des cas particuliers des matrices à coefficients complexes. Ce paragraphe a pour but de savoir dans quel cas on peut trouver une matrice diagonale D semblable à une matrice donnée A . Les éléments de la diagonale de D seront alors les valeurs propres de A qui sont intéressantes à connaître, en particulier dans de nombreuses applications de la physique, même si elles sont complexes.

Définition 0.6.4 On dit que la matrice $A \in \mathcal{M}_{n,n}$ est diagonalisable dans \mathbb{R} (resp. dans \mathbb{C}) si elle est semblable à une matrice $D \in \mathcal{M}_{n,n}(\mathbb{R})$ (resp. $D \in \mathcal{M}_{n,n}(\mathbb{C})$) diagonale.

Puisque A et D sont semblables, il existe une matrice P telle que

$$D = P^{-1}AP \iff AP = PD$$

ce qui est équivalent à $AY_i = \lambda_i Y_i$ (où Y_1, \dots, Y_n sont les colonnes de P et $\lambda_1, \dots, \lambda_n$ sont les termes diagonaux de D). Donc la diagonale de D est constituée des valeurs propres de A et les colonnes de P sont les vecteurs propres correspondants.

Les résultats suivants utiles mais délicats à démontrer (voir cours d'algèbre 2).

Théorème 0.6.1 Soit $A \in \mathcal{M}_{nn}(\mathbb{R})$ une matrice symétrique, alors

1. A a toutes ses valeurs propres réelles,
2. il existe une base orthonormée de \mathbb{R}^n constituée de vecteurs propres de A (on utilise le produit scalaire usuel de \mathbb{R}^n),
3. A est diagonalisable dans \mathbb{R} . Plus précisément, il existe une matrice P orthogonale telle que $D = P^T A P$ soit une matrice diagonale.

On rappelle que :

- le produit scalaire usuel de \mathbb{R}^n est celui qui à deux vecteurs x et y fait correspondre le réel $x^T y = \sum_{i=1}^n x_i y_i$. À ce produit scalaire on associe la norme euclidienne définie par :

$$(\|x\| = \sqrt{x^T x} = \sqrt{\sum_{i=1}^n x_i^2})$$

- Deux vecteurs sont orthogonaux si et seulement si leur produit scalaire est nul.
- Une base orthonormée de vecteur est telle que les vecteurs sont orthogonaux deux à deux et que la norme de chacun d'eux est égale à 1.
- Une matrice $Q \in \mathcal{M}_{nn}(\mathbb{R})$ est orthogonale si et seulement si $Q^T Q = I$.
- Si Q est orthogonale, alors ses colonnes Q_i sont des vecteurs orthonormés de \mathbb{R}^n (la notion de matrice orthogonale sera reprise dans le chapitre sur les moindres carrés).

0.6.3 Valeurs propres et matrices définies positives

Définition 0.6.5 On dit qu'une matrice $A \in \mathcal{M}_{nn}(\mathbb{R})$ symétrique est semi-définie positive si

$$x^T A x \geq 0, \quad \forall x \in \mathbb{R}^n$$

On dit qu'elle est définie positive si de plus

$$x^T A x = 0 \implies x = 0$$

Si $B \in \mathcal{M}_{mn}(\mathbb{R})$, alors les matrices $BB^T \in \mathcal{M}_{mm}(\mathbb{R})$ et $B^T B \in \mathcal{M}_{nn}(\mathbb{R})$ sont symétriques.

Proposition 0.6.4 Soit $B \in \mathcal{M}_{mn}(\mathbb{R})$, alors

- $B^T B$ est symétrique et semi-définie positive,
- supposons $n \leq m$ et B de rang n , alors $B^T B$ est définie positive.

Démonstration : $(B^T B)^T = B^T B$, ce qui prouve que $B^T B$ est symétrique.

$x^T B^T B x = (Bx)^T Bx = \|Bx\|^2 \geq 0$; ce qui montre que $B^T B$ est semi-définie positive.

$x^T B^T B x = 0 \implies \|Bx\|^2 = 0 \implies Bx = 0$, et puisque B est de rang n , son noyau est réduit au vecteur nul et donc $x = 0$. La matrice $B^T B$ est donc définie positive.

Proposition 0.6.5 Soit $A \in \mathcal{M}_{nn}(\mathbb{R})$ une matrice symétrique. Alors A est définie positive (resp. semi-définie positive) si et seulement si ses valeurs propres sont strictement positives (resp. positives ou nulles).

Démonstration : La démonstration se fait en deux parties :

- Supposons que A soit définie positive (resp. semi-définie positive) et soit (λ, Y) un couple propre de A , c'est-à-dire

$$AY = \lambda Y \implies Y^T AY = \lambda Y^T Y$$

Puisque Y est un vecteur propre, ce n'est pas un vecteur nul et donc on a $Y^T AY > 0$ (resp. $Y^T AY \geq 0$) et $Y^T Y > 0$, ce qui donne $\lambda > 0$ (resp. $\lambda \geq 0$).

- Réciproque : Supposons que les valeurs propres de A soient strictement positives (resp. positives ou nulles). Puisque A est symétrique, il existe une base orthonormée de vecteurs propres (voir le paragraphe référencé) que l'on note Y_1, \dots, Y_n . Considérons un vecteur x quelconque de \mathbb{R}^n , alors

$$x = \sum_{i=1}^n \alpha_i Y_i \implies x^T A x = \left(\sum_{i=1}^n \alpha_i Y_i^T \right) A \left(\sum_{i=1}^n \alpha_i Y_i \right) = \sum_{i=1}^n \alpha_i^2 \lambda_i$$

Et donc, $x^T A x \geq 0$ pour tout $x \in \mathbb{R}^n$ et, si les valeurs propres sont strictement positives,

$$x^T A x = 0 \implies \alpha_i = 0 \implies x = 0$$

Remarquons qu'une matrice symétrique définie positive n'a pas de valeur propre nulle, elle est donc inversible.

Résolution des systèmes linéaires : Méthodes directes

1.1 Motivations

1.1.1 Equation de la chaleur

Avant de commencer, voyons un exemple simple qui montre que l'on peut être amené à résoudre des systèmes linéaires de grande taille. Il s'agit de la résolution de l'équation de la chaleur stationnaire en dimension 1 :

$$\begin{cases} -\frac{d^2 u}{dx^2}(x) &= f(x), \quad x \in]0, 1[, \\ u(0) &= 0, \\ u(1) &= 0. \end{cases} \quad (1.1)$$

On peut imaginer $u(x)$ comme la distribution de température d'une barre de longueur 1 (dont toutes les constantes physiques ont été normalisées), chauffée de l'intérieur par un flux de chaleur $f(x)$. Les extrémités de la barre sont maintenues à une température nulle. Dans certaines situations particulières il est possible d'obtenir la solution exacte ; c'est le cas lorsque l'on connaît une double primitive de f . Dans le cas général on peut chercher une approximation de $u(x)$ en certains points (x_k) , avec $k = 0, \dots, N$ (où N le nombre de subdivisions). Pour simplifier, on choisit des points x_k équidistants, on a donc $h = \frac{1}{N}$ (le pas du maillage), et $x_k = kh$.

Aux points x_k , l'équation se réécrit :

$$\begin{cases} -\frac{d^2 u}{dx^2}(x_k) &= f(x_k), \quad k = 1, \dots, N-1, \\ u(x_0) &= 0, \\ u(x_N) &= 0. \end{cases} \quad (1.2)$$

On remarque par cette équation que la valeur de u est connue sur les bords de l'intervalle $[0, 1]$. Les dérivées apparaissant dans l'équation sont approchées par des formules n'utilisant que la valeur de u sur le réseau de points (x_k) . On peut approcher la dérivée seconde de u à l'aide de la formule classique :

$$\frac{d^2 u}{dx^2}(x_k) = \frac{u(x_{k+1}) - 2u(x_k) + u(x_{k-1}))}{h^2} + O(h^2). \quad (1.3)$$

Si l'on se donne le droit de négliger le reste $O(h^2)$ (on peut penser que ceci est raisonnable si le nombre N de subdivisions est assez grand), on peut obtenir une approximation $v_k \approx u(x_k)$

en écrivant les équations suivantes :

$$\begin{cases} -\frac{v_{k+1} - 2v_k + v_{k-1}}{h^2} = f(x_k), & k = 1, \dots, N-1, \\ v_0 = 0, \\ v_N = 0. \end{cases} \quad (1.4)$$

qui peuvent se mettre sous la forme matricielle suivante :

$$\frac{1}{h^2} \begin{pmatrix} 2 & -1 & & & & 0 \\ -1 & 2 & -1 & & & \\ & \cdot & \cdot & \cdot & & \\ & & \cdot & \cdot & \cdot & \\ & & & \cdot & \cdot & \\ & & & & -1 & 2 & -1 \\ 0 & & & & & -1 & 2 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \cdot \\ \cdot \\ \cdot \\ v_{N-2} \\ v_{N-1} \end{pmatrix} = \begin{pmatrix} f(x_1) \\ f(x_2) \\ \cdot \\ \cdot \\ \cdot \\ f(x_{N-2}) \\ f(x_{N-1}) \end{pmatrix} \quad (1.5)$$

et $v_0 = v_N = 0$.

Voilà donc un exemple d'un problème physique dont la résolution approchée conduit à la résolution d'un système linéaire (ici la matrice du système a une forme assez particulière : c'est une matrice tridiagonale). Nous allons maintenant voir quelles sont les différentes méthodes de résolution des systèmes linéaires "raisonnables", raisonnables dans le sens où ces méthodes peuvent être facilement implantées sur ordinateur.

1.2 Méthode d'élimination de Gauss

1.2.1 Résolution d'un système triangulaire

Dans tout ce qui suit on considérera une matrice $A \in \mathcal{M}_{nn}(\mathbb{R})$ qu'on supposera inversible. On cherche à résoudre le système linéaire

$$Ax = b,$$

où $b \in \mathbb{R}^n$ est donné.

Un cas facile à traiter est le cas où A est une matrice triangulaire inférieure (pour fixer les idées) : on a alors

$$a_{ij} = 0; \quad \text{pour } j > i.$$

De plus, comme A est inversible on a nécessairement $a_{ii} \neq 0 \quad \forall i$. La méthode de résolution est immédiate :

$$x_1 = \frac{b_1}{a_{11}},$$

puis

$$a_{21}x_1 + a_{22}x_2 = b_2 \quad \Rightarrow \quad x_2 = \frac{1}{a_{22}}(b_2 - a_{21}x_1),$$

On peut donc écrire la formule générale :

$$x_i = \frac{1}{a_{ii}} \left(b_i - \sum_{j=1}^{i-1} a_{ij}x_j \right) \quad \text{pour } i = 2, 3, \dots, n.$$

On obtient donc les valeurs des inconnues dans l'ordre naturel.

Maintenant, nous pouvons écrire l'algorithme qui permet de résoudre le système linéaire $Ax = b$ dans le cas où la matrice A est triangulaire inférieure, en n'oubliant pas de vérifier au départ que ce système a une solution (c'est à dire la matrice A est inversible).

Algorithm 1 Système linéaire : cas d'une matrice triangulaire inférieure

```

1: for  $i = 1$  to  $n$  do
2:   if  $|a_{ii}| < \epsilon$  then
3:     Arrêter l'algorithme et afficher un message d'erreur
4:   end if
5: end for
6:  $x_1 \leftarrow \frac{b_1}{a_{11}}$ 
7: for  $i = 2$  to  $n$  do
8:    $x_i \leftarrow \frac{b_i - \sum_{j=1}^{i-1} a_{ij}x_j}{a_{ii}}$ 
9: end for

```

Dans le cas où la matrice A est triangulaire supérieure, le calcul des inconnues est donné par

$$\begin{cases} x_n = \frac{b_n}{a_{nn}} \\ x_i = \frac{1}{a_{ii}} \left(b_i - \sum_{j=i+1}^n a_{ij}x_j \right) \quad \text{pour } i = n-1, n-2, \dots, 1. \end{cases} \quad (1.6)$$

On notera que l'on obtient dans ce dernier cas les valeurs des inconnues dans l'ordre des indices décroissants.

Algorithm 2 Système linéaire : cas d'une matrice triangulaire supérieure

```

1: for  $i = 1$  to  $n$  do
2:   if  $|a_{ii}| < \epsilon$  then
3:     Arrêter l'algorithme et afficher un message d'erreur
4:   end if
5: end for
6:  $x_n \leftarrow \frac{b_n}{a_{nn}}$ 
7: for  $i = n-1$  to  $1$  do
8:    $x_i \leftarrow \frac{b_i - \sum_{j=i+1}^n a_{ij}x_j}{a_{ii}}$ 
9: end for

```

1.2.2 Principe de la méthode de Gauss

Le principe de la méthode d'élimination de Gauss consiste alors à mettre le système linéaire $Ax = b$ sous la forme

$$\tilde{A}x = \tilde{b}, \quad (1.7)$$

où la matrice \tilde{A} est triangulaire supérieure, la résolution de ce nouveau système étant assez simple comme on vient de le voir précédemment.

Attention : La matrice \tilde{A} n'est pas semblable à A . Autrement dit, il ne faut pas croire que les matrices A et \tilde{A} ont les mêmes valeurs propres.

Exemple Considérons un exemple avec $N = 3$ pour montrer comment faire la transformation d'un système linéaire en un système triangulaire supérieur équivalent.

Nous choisissons :

$$A = \begin{pmatrix} 4 & 8 & 12 \\ 3 & 8 & 13 \\ 2 & 9 & 18 \end{pmatrix}, \quad b = \begin{pmatrix} 4 \\ 5 \\ 11 \end{pmatrix} \quad (1.8)$$

Le système $Ax = b$ devient dans ce cas :

$$\begin{cases} 4x_1 + 8x_2 + 12x_3 = 4 \\ 3x_1 + 8x_2 + 13x_3 = 5 \\ 2x_1 + 9x_2 + 18x_3 = 11 \end{cases} \quad (1.9)$$

Première étape La première opération consiste à diviser la première équation de (1.9) par a_{11} (appelé premier pivot) pour obtenir :

$$x_1 + 2x_2 + 3x_3 = 1 \quad (1.10)$$

ensuite, nous soustrayons 3 fois l'équation (1.10) à la deuxième équation de (1.9) et 2 fois à l'équation (1.10) à la troisième équation de (1.9) Nous obtenons un système équivalent à (1.9) qui est :

$$\begin{cases} x_1 + 2x_2 + 3x_3 = 1 \\ 2x_2 + 4x_3 = 2 \\ 5x_2 + 12x_3 = 9 \end{cases} \quad (1.11)$$

Nous observons que les deux dernières équations de (1.11) font apparaître seulement deux inconnues x_2 et x_3 (l'inconnue x_1 a été éliminée) et nous pouvons recommencer le procédé en laissant inchangée la première équation.

Deuxième étape Nous divisons la deuxième équation de (1.11) par 2 (le deuxième pivot). Nous obtenons :

$$x_2 + 2x_3 = 1. \quad (1.12)$$

Nous soustrayons 5 fois (1.12) à la troisième équation de (1.11) et nous obtenons le système suivant :

$$\begin{cases} x_1 + 2x_2 + 3x_3 = 1 \\ x_2 + 2x_3 = 1 \\ 2x_3 = 4 \end{cases} \quad (1.13)$$

qui est équivalent au système (1.11).

Dernière étape Finalement, il suffit de diviser la troisième équation de (1.13) par le troisième pivot, qui est encore ici 2, pour obtenir :

$$\begin{cases} x_1 + 2x_2 + 3x_3 = 1 \\ x_2 + 2x_3 = 1 \\ x_3 = 2 \end{cases} \quad (1.14)$$

De (1.14), il est facile de déduire successivement les inconnues x_3, x_2, x_1 comme nous l'avons décrit dans le paragraphe précédent (cas de matrice triangulaire supérieure). Nous obtenons :

$$x_3 = 2, x_2 = -3, x_1 = 1. \quad (1.15)$$

1.2.3 Méthode de Gauss : cas général

Le système (1.7) est simplement équivalent au système $Ax = b$ d'où nous sommes partis et la méthode consiste à éliminer successivement des inconnues entre les équations.

Première étape Éliminons dans un premier temps la première inconnue, à savoir x_1 : on part de

$$A^{(1)} = A = \begin{pmatrix} a_{11} & a_{12} & \cdot & \cdot & a_{1j} & \cdot & a_{1n} \\ a_{21} & a_{22} & \cdot & \cdot & a_{2j} & \cdot & a_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{i1} & a_{i2} & \cdot & \cdot & a_{ij} & \cdot & a_{in} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \cdot & \cdot & a_{nj} & \cdot & a_{nn} \end{pmatrix}, \quad b^{(1)} = b = \begin{pmatrix} b_1 \\ b_2 \\ \cdot \\ b_i \\ \cdot \\ \cdot \\ b_n \end{pmatrix} \quad (1.16)$$

On peut alors éliminer l'inconnue x_1 de la deuxième équation en retranchant de la deuxième ligne de A la première pré-multipliée par le coefficient $\frac{a_{21}}{a_{11}}$, à condition que a_{11} soit non nul, ce que l'on supposera pour le moment. Ceci conduit donc à remplacer

$$a_{2j}^{(1)} = a_{2j} \rightarrow a_{2j}^{(2)} = a_{2j} - \frac{a_{21}}{a_{11}} a_{1j} \quad \text{pour } j = 1, 2, \dots, n$$

et parallèlement

$$b_2^{(1)} = b_2 \rightarrow b_2^{(2)} = b_2 - \frac{a_{21}}{a_{11}} b_1.$$

La matrice et le vecteur second membre du système prennent les formes suivantes :

$$\begin{pmatrix} a_{11} & a_{12} & \cdot & \cdot & a_{1j} & \cdot & a_{1n} \\ 0 & a_{22}^{(2)} & \cdot & \cdot & a_{2j}^{(2)} & \cdot & a_{2n}^{(2)} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{i1} & a_{i2} & \cdot & \cdot & a_{ij} & \cdot & a_{in} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \cdot & \cdot & a_{nj} & \cdot & a_{nn} \end{pmatrix}, \quad \begin{pmatrix} b_1 \\ b_2^{(2)} \\ \cdot \\ \cdot \\ b_i \\ \cdot \\ b_n \end{pmatrix} \quad (1.17)$$

où on notera que le coefficient $a_{21}^{(2)}$ est nul (on a fait ce qu'il fallait pour cela), ce qui signifie que l'inconnue x_1 ne figure plus dans la deuxième équation du système.

On peut recommencer l'opération sur la troisième ligne en retranchant de celle-ci la première pré-multipliée par le coefficient $\frac{a_{31}}{a_{11}}$, et ainsi de suite. Pour la ligne i cela donne les formules :

$$\begin{cases} a_{ij}^{(1)} = a_{ij} \rightarrow a_{ij}^{(2)} = a_{ij} - \frac{a_{i1}}{a_{11}} a_{1j} & \text{pour } j = 1, 2, \dots, n \\ b_i^{(1)} = b_i \rightarrow b_i^{(2)} = b_i - \frac{a_{i1}}{a_{11}} b_1. \end{cases} \quad (1.18)$$

Lorsqu'on a fait l'élimination jusqu'à la dernière ligne on obtient un système linéaire sous la forme $A^{(2)}x = b^{(2)}$ avec

$$A^{(2)} = \begin{pmatrix} a_{11} & a_{12} & \cdot & \cdot & a_{1j} & \cdot & a_{1n} \\ 0 & a_{22}^{(2)} & \cdot & \cdot & a_{2j}^{(2)} & \cdot & a_{2n}^{(2)} \\ 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & a_{i2}^{(2)} & \cdot & \cdot & a_{ij}^{(2)} & \cdot & a_{in}^{(2)} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & a_{n2}^{(2)} & \cdot & \cdot & a_{nj}^{(2)} & \cdot & a_{nn}^{(2)} \end{pmatrix}, \quad b^{(2)} = \begin{pmatrix} b_1 \\ b_2^{(2)} \\ \cdot \\ \cdot \\ b_i^{(2)} \\ \cdot \\ b_n^{(2)} \end{pmatrix} \quad (1.19)$$

On dit qu'on vient de réaliser la première étape de l'élimination de Gauss, ce qui a consisté à faire apparaître des 0 sur la première colonne en dessous de la diagonale.

Deuxième étape La deuxième étape consiste à éliminer la deuxième inconnue x_2 des équations 3, 4, ..., n . Pour cela on commence par la troisième ligne, en supposant maintenant que $a_{22}^{(2)}$ est non nul. Ce qui conduit aux formules :

$$\begin{cases} a_{3j}^{(2)} & \rightarrow a_{3j}^{(3)} = a_{3j}^{(2)} - \frac{a_{32}^{(2)}}{a_{22}^{(2)}} a_{2j}^{(2)} \quad \text{pour } j = 2, 3, \dots, n \\ b_3^{(2)} & \rightarrow b_3^{(3)} = b_3^{(2)} - \frac{a_{32}^{(2)}}{a_{22}^{(2)}} b_2^{(2)}. \end{cases} \quad (1.20)$$

La formule pour la ligne i ($i = 3, \dots, n$) s'écrit alors

$$\begin{cases} a_{ij}^{(2)} & \rightarrow a_{ij}^{(3)} = a_{ij}^{(2)} - \frac{a_{i2}^{(2)}}{a_{22}^{(2)}} a_{2j}^{(2)} \quad \text{pour } j = 2, 3, \dots, n \\ b_i^{(2)} & \rightarrow b_i^{(3)} = b_i^{(2)} - \frac{a_{i2}^{(2)}}{a_{22}^{(2)}} b_2^{(2)}. \end{cases} \quad (1.21)$$

Ce calcul étant fait jusqu'à la ligne n , on obtient un système sous la forme $A^{(3)}x = b^{(3)}$ avec

$$A^{(3)} = \begin{pmatrix} a_{11} & a_{12} & \cdot & \cdot & a_{1j} & \cdot & a_{1n} \\ 0 & a_{22}^{(2)} & \cdot & \cdot & a_{2j}^{(2)} & \cdot & a_{2n}^{(2)} \\ 0 & 0 & a_{33}^{(3)} & \cdot & a_{3j}^{(3)} & \cdot & a_{3n}^{(3)} \\ 0 & 0 & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & a_{n3}^{(3)} & \cdot & a_{nj}^{(3)} & \cdot & a_{nn}^{(3)} \end{pmatrix}, \quad b^{(3)} = \begin{pmatrix} b_1 \\ b_2^{(2)} \\ b_3^{(3)} \\ \cdot \\ \cdot \\ b_n^{(3)} \end{pmatrix} \quad (1.22)$$

Étape générale On peut alors définir l'algorithme général en supposant qu'on a effectué $k - 1$ étapes de l'élimination de Gauss, et on a un système qui s'écrit $A^{(k)}x = b^{(k)}$, avec $A^{(k)}$ et $b^{(k)}$ donnés par

$$A^{(k)} = \begin{pmatrix} a_{11} & a_{12} & a_{13}^{(13)} & \cdot & a_{1k} & \cdot & a_{1n} \\ 0 & a_{22}^{(2)} & a_{23}^{(2)} & \cdot & a_{2k}^{(2)} & \cdot & a_{2n}^{(2)} \\ 0 & 0 & a_{33}^{(3)} & \cdot & a_{3k}^{(3)} & \cdot & a_{3n}^{(3)} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdot & a_{kk}^{(k)} & \cdot & a_{kn}^{(k)} \\ 0 & 0 & \cdot & \cdot & a_{nk}^{(k)} & \cdot & a_{nn}^{(k)} \end{pmatrix}, \quad b^{(k)} = \begin{pmatrix} b_1 \\ b_2^{(2)} \\ b_3^{(3)} \\ \cdot \\ b_k^{(k)} \\ \cdot \\ b_n^{(k)} \end{pmatrix} \quad (1.23)$$

L'étape k consiste à faire l'élimination de la variable x_k dans les équations $k + 1, k + 2, \dots, n$. Ceci conduit aux formules suivantes, définies pour $i = k + 1, k + 2, \dots, n$

$$\begin{cases} a_{ij}^{(k)} & \rightarrow a_{ij}^{(k+1)} = a_{ij}^{(k)} - \frac{a_{ik}^{(k)}}{a_{kk}^{(k)}} a_{kj}^{(k)} \quad \text{pour } j = k, k + 1, \dots, n \\ b_i^{(k)} & \rightarrow b_i^{(k+1)} = b_i^{(k)} - \frac{a_{ik}^{(k)}}{a_{kk}^{(k)}} b_k^{(k)}. \end{cases} \quad (1.24)$$

- On voit que la ligne i de la matrice $A^{(k)}$ et du vecteur $b^{(k)}$ n'est plus modifiée par l'algorithme dès que $i < k$.
- À l'étape k on pratique l'élimination sur une matrice carrée dont la taille est $n - k + 1$.
- Les modifications effectuées sur la matrice A au cours des étapes ne changent pas la valeur du déterminant puisqu'une ligne est toujours remplacée par elle-même plus une combinaison de la ligne "pivot". On a donc $\det(A) = \det(A^{(k)})$, pour $k = 1, \dots, n$.

L'algorithme se termine à l'étape $k = n$. La matrice $A^{(n)}$ ainsi obtenue est la matrice triangulaire supérieure \tilde{A} cherchée et $b^{(n)}$ est le vecteur \tilde{b} du système $\tilde{A}x = \tilde{b}$, qui est bien équivalent au système $Ax = b$.

1.2.4 Algorithme d'élimination de Gauss

En partant de $A^{(1)} = A$ et $b^{(1)} = b$ on construit par récurrence, pour $k = 1, 2, \dots, n$, des matrices $A^{(k)}$ et des vecteurs $b^{(k)}$ tels que le système $Ax = b$ soit équivalent à $A^{(k)}x = b^{(k)}$; de plus la matrice $A^{(n)}$ est triangulaire supérieure.

Définition 1.2.1 On appelle pivots les nombres $a_{kk}^{(k)}$.

Comme nous l'avons déjà noté, l'étape k de l'élimination ne modifie que les lignes et les colonnes $k + 1$ à n . Cela veut dire que de façon pratique, on n'a pas besoin de conserver les différentes versions de $A^{(k)}$ et $b^{(k)}$. On travaille donc avec la matrice A et le vecteur b originaux, dans lesquels on écrase au fur et à mesure les anciens termes.

Algorithm 3 Système linéaire : Méthode d'élimination de Gauss

```

1: for  $k = 1$  to  $n - 1$  do
2:   if  $|a_{kk}| < \epsilon$  then
3:     Arrêter l'algorithme et afficher un message d'erreur
4:   else
5:     for  $i = k + 1$  to  $n$  do
6:        $c \leftarrow \frac{a_{ik}}{a_{kk}}$ 
7:        $b_i \leftarrow b_i - cb_k$ 
8:        $a_{ik} \leftarrow 0$ 
9:       for  $j = k + 1$  to  $n$  do
10:         $a_{ij} \leftarrow a_{ij} - ca_{kj}$ 
11:      end for
12:    end for
13:  end if
14: end for
15: if  $|a_{nn}| < \epsilon$  then
16:   Arrêter l'algorithme et afficher un message d'erreur
17: end if

```

Attention :

- On ne peut pas a priori savoir si les pivots seront non nuls, en effet l'étape k de l'algorithme modifie les valeurs de termes diagonaux des lignes $k + 1, k + 2, \dots, n$.
- L'algorithme peut s'arrêter si à une étape quelconque k on a $a_{kk}^{(k)} = 0$. Cette situation peut malheureusement se produire même si la matrice est régulière, comme le montre l'exemple suivant

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \quad (1.25)$$

où $a_{11}^{(1)} = 0$! En examinant ce petit exemple, on voit qu'en permutant les deux lignes de A le système est triangulaire et l'élimination de Gauss est toute faite ! On peut donc espérer qu'en permutant éventuellement les lignes et/ou les colonnes on puisse toujours trouver un pivot non nul.

1.2.5 Écriture matricielle de l'élimination de Gauss

L'élimination de Gauss consiste à construire une suite de matrice $A^{(1)}, A^{(2)}, \dots, A^{(n)}$ où $A^{(n)}$ est une matrice triangulaire supérieure.

L'indice i étant donné, on va exprimer la i ème ligne de A à partir des lignes de la matrice $A^{(n)}$. Pour une notation lisible, nous noterons la i ème ligne d'une matrice C par \underline{C}_i .

On peut tout d'abord remarquer que, à partir de la matrice $A^{(i)}$, la i ème ligne n'est plus modifiée, c'est-à-dire :

$$\underline{A}_i^{(i)} = \underline{A}_i^{(i+1)} = \dots = \underline{A}_i^{(n)} \quad (1.26)$$

Explicitons successivement la i ème ligne des matrices $A^{(1)}, A^{(2)}, \dots$

$$\underline{A}_i^{(1)} = \underline{A}_i$$

$$\underline{A}_i^{(2)} = \underline{A}_i^{(1)} - m_{i1} \underline{A}_1^{(1)} \quad \text{avec} \quad m_{i1} = \frac{a_{i1}^{(1)}}{a_{11}^{(1)}}$$

$$\underline{A}_i^{(3)} = \underline{A}_i^{(2)} - m_{i2} \underline{A}_2^{(2)} \quad \text{avec} \quad m_{i2} = \frac{a_{i2}^{(2)}}{a_{22}^{(2)}}$$

.....

$$\underline{A}_i^{(i)} = \underline{A}_i^{(i-1)} - m_{i,i-1} \underline{A}_{i-1}^{(i-1)} \quad \text{avec} \quad m_{i,i-1} = \frac{a_{i,i-1}^{(i-1)}}{a_{i-1,i-1}^{(i-1)}}$$

En ajoutant les égalités précédentes, on obtient :

$$\underline{A}_i = m_{i1} \underline{A}_1^{(1)} + m_{i2} \underline{A}_2^{(2)} + \dots + m_{i,i-1} \underline{A}_{i-1}^{(i-1)} + \underline{A}_i^{(i)}$$

En utilisant l'équation (1.26), on a :

$$\underline{A}_i = m_{i1} \underline{A}_1^{(n)} + m_{i2} \underline{A}_2^{(n)} + \dots + m_{i,i-1} \underline{A}_{i-1}^{(n)} + \underline{A}_i^{(n)}$$

Si l'on pose $m_{ii} = 1$, $m_{ij} = 0$ pour $j > i$, on obtient :

$$\underline{A}_i = m_{i1} \underline{A}_1^{(n)} + m_{i2} \underline{A}_2^{(n)} + \dots + m_{i,i-1} \underline{A}_{i-1}^{(n)} + m_{ii} \underline{A}_i^{(n)} + m_{i,i+1} \underline{A}_{i+1}^{(n)} + \dots + m_{in} \underline{A}_n^{(n)}$$

En utilisant les propriétés du produit matriciel, cela s'écrit :

$$\underline{A}_i = \left(m_{i1} m_{i2} \dots m_{in} \right) A^{(n)} \quad \text{donc} \quad A = M A^{(n)}$$

où M est la matrice définie par :

$$\begin{pmatrix} 1 & 0 & . & . & . & . & 0 \\ m_{21} & 1 & 0 & . & . & . & 0 \\ m_{31} & m_{32} & 1 & 0 & . & . & 0 \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ m_{n1} & m_{n2} & . & . & . & m_{n,n-1} & 1 \end{pmatrix} \quad (1.27)$$

La matrice M est triangulaire inférieure avec des 1 sur la diagonale, cette matrice est souvent notée L , la matrice $A^{(n)}$ est triangulaire supérieure, elle est souvent notée U , on vient d'écrire une factorisation de A sous la forme $A = LU$.

1.2.6 Unicité de la factorisation LU

Dans le paragraphe précédent, nous avons montré qu'une matrice A pouvait se factoriser en un produit d'une matrice triangulaire inférieure L et d'une matrice triangulaire supérieure U à condition que l'on puisse appliquer l'élimination de Gauss à la matrice A , ce qui est possible si les pivots sont non nuls. Nous avons alors le théorème suivant :

Théorème 1.2.1 *Si au cours de l'élimination de Gauss les pivots sont non nuls, alors il existe une matrice triangulaire inférieure L et une matrice triangulaire supérieure U , telles que l'on ait :*

$$A = LU.$$

De plus si on impose à L d'avoir ses éléments diagonaux égaux à 1 alors la factorisation est unique.

Démonstration : Il reste à démontrer l'unicité. Soit donc une autre factorisation :

$$A = \bar{L}\bar{U},$$

\bar{L} ayant ses éléments diagonaux égaux à 1. On a donc :

$$LU = \bar{L}\bar{U},$$

et comme les matrices L et U sont inversibles, on peut écrire :

$$\bar{L}^{-1}L = \bar{U}U^{-1}$$

or $\bar{L}^{-1}L$ est une matrice triangulaire inférieure qui a ses éléments diagonaux égaux à 1, et $\bar{U}U^{-1}$ est une matrice triangulaire supérieure. Ces deux matrices ne peuvent être égales que si, d'une part, elles sont diagonales et que, d'autre part, il y a égalité des éléments diagonaux (égaux à 1 pour $\bar{L}^{-1}L$), d'où

$$\bar{L}^{-1}L = \bar{U}U^{-1} = I,$$

et donc $L = \bar{L}$ et $U = \bar{U}$.

Remarque : Dans cette démonstration on a utilisé des résultats sur les matrices triangulaires. En particulier, le produit de deux matrices triangulaires inférieures (resp. supérieures) est triangulaire inférieure (resp. supérieure) et l'inverse d'une matrice triangulaire inférieure (resp. supérieure) est triangulaire inférieure (resp. supérieure) avec ses éléments diagonaux qui sont les inverses des éléments diagonaux de la matrice originale.

1.3 Calcul direct de la factorisation $A = LU$

1.3.1 Motivations de la factorisation LU

Dans certaines situations on a besoin de résoudre plusieurs fois le système linéaire $Ax = b$ avec la même matrice A mais pour plusieurs seconds membres différents. Il semble alors inutile de recommencer les opérations d'élimination sur A qui est inchangée et l'on n'effectue celles-ci que sur le second membre b . Ces opérations étant linéaires on doit pouvoir les mettre sous forme matricielle.

Reprenons le système :

$$Ax = b$$

Si l'on dispose d'une factorisation $A = LU$, telle que celle obtenue dans le paragraphe précédent, mais que nous allons calculer directement dans les paragraphes suivants, alors on peut récrire $Ax = b$ sous la forme :

$$LUx = b$$

La solution de ce système s'obtient en résolvant successivement les deux systèmes :

$$\begin{cases} Ly = b \\ Ux = y \end{cases} \quad (1.28)$$

ce qui correspond à la résolution de deux systèmes triangulaires.

Calcul du déterminant Comme nous l'avons déjà souligné la matrice U n'est pas semblable à A , cependant on a :

$$\det(A) = \det(LU) = \det(L) * \det(U) = \det(U) = \prod_{k=1}^n a_{kk}^{(k)}$$

ce qui fournit une méthode économique (par rapport au développement suivant une ligne ou une colonne) de calcul du déterminant de A . A titre d'exemple dans le cas $n = 10$ il faut de l'ordre de 10^3 opérations en utilisant la méthode de *Gauss* et de l'ordre de 10^7 opérations par un calcul direct. Dans ce cas on n'a pas besoin d'explicitier L puisque son déterminant vaut a priori 1.

1.3.2 Principe du calcul direct de la factorisation LU

Rappelons la présentation classique de l'algorithme de *Doolittle*. On oublie l'algorithme d'élimination de *Gauss*, pour chercher directement une décomposition de A de la forme $A = LU$, où L est triangulaire inférieure et U triangulaire supérieure. Pour assurer l'unicité de la décomposition, nous demandons que la diagonale de L soit unitaire ($l_{ii} = 1, i = 1, \dots, n$).

Exemple Traitons un exemple, soit la matrice

$$A = \begin{pmatrix} 2 & 1 & -2 \\ 4 & 5 & -3 \\ -2 & 5 & 3 \end{pmatrix} \quad (1.29)$$

On cherche $L = \begin{pmatrix} 1 & 0 & 0 \\ \times & 1 & 0 \\ \times & \times & 1 \end{pmatrix}$, $U = \begin{pmatrix} \times & \times & \times \\ 0 & \times & \times \\ 0 & 0 & \times \end{pmatrix}$ telles que $A = LU$.

- On identifie la première ligne de A et la première ligne de LU , cela permet d'obtenir la première ligne de U :

$$LU = \begin{pmatrix} 1 & 0 & 0 \\ \times & 1 & 0 \\ \times & \times & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 & -2 \\ 0 & \times & \times \\ 0 & 0 & \times \end{pmatrix} = \begin{pmatrix} 2 & 1 & -2 \\ 4 & 5 & -3 \\ -2 & 5 & 3 \end{pmatrix} \quad (1.30)$$

- On identifie la première colonne de A avec la première colonne de LU , cela permet d'obtenir la première colonne de L :

$$LU = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & \times & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 & -2 \\ 0 & \times & \times \\ 0 & 0 & \times \end{pmatrix} = \begin{pmatrix} 2 & 1 & -2 \\ 4 & 5 & -3 \\ -2 & 5 & 3 \end{pmatrix} \quad (1.31)$$

- On identifie la deuxième ligne de A avec la deuxième ligne de LU , cela permet d'obtenir la deuxième ligne de U :

$$LU = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & \times & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 & -2 \\ 0 & 3 & 1 \\ 0 & 0 & \times \end{pmatrix} = \begin{pmatrix} 2 & 1 & -2 \\ 4 & 5 & -3 \\ -2 & 5 & 3 \end{pmatrix} \quad (1.32)$$

- On identifie la deuxième colonne de A avec la deuxième colonne de LU , cela permet d'obtenir la deuxième colonne de L :

$$LU = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & 2 & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 & -2 \\ 0 & 3 & 1 \\ 0 & 0 & \times \end{pmatrix} = \begin{pmatrix} 2 & 1 & -2 \\ 4 & 5 & -3 \\ -2 & 5 & 3 \end{pmatrix} \quad (1.33)$$

- On identifie la troisième ligne de A avec la troisième ligne de LU , cela permet d'obtenir la troisième ligne de U :

$$LU = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & 2 & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 & -2 \\ 0 & 3 & 1 \\ 0 & 0 & -1 \end{pmatrix} = \begin{pmatrix} 2 & 1 & -2 \\ 4 & 5 & -3 \\ -2 & 5 & 3 \end{pmatrix} \quad (1.34)$$

Une autre méthode consiste à calculer toujours explicitement une décomposition LU de la matrice A , mais cette fois-ci c'est la diagonale de U qui est formée de 1 mais non plus celle de L . Cette méthode conduit à un algorithme qui s'appelle algorithme de Crout.

1.3.3 Algorithme de Doolittle (factorisation LU)

En écrivant $A = LU$ et en se souvenant que les matrices L et U sont triangulaires et que les termes diagonaux de L valent 1, on obtient :

$$\underline{A}_i = \underline{L}_i U = \sum_{k=1}^{i-1} l_{ik} \underline{U}_k + \underline{U}_i \Leftrightarrow \underline{U}_i = \underline{A}_i - \sum_{k=1}^{i-1} l_{ik} \underline{U}_k$$

Nous voyons que pour calculer les éléments de la i ème ligne de U, \underline{U}_i , il nous faut connaître préalablement les éléments des lignes 1 à $i-1$ de U ainsi que les éléments des colonnes 1 à $i-1$ de L . On peut remarquer de plus que, pratiquement, on ne calculera que les termes u_{ij} pour j compris entre i et n , puisque les autres termes de la ligne sont connus car nuls.

De manière similaire on a :

$$A_i = LU_i = \sum_{k=1}^{i-1} u_{ki} L_k + u_{ii} L_i \Leftrightarrow L_i = \frac{1}{u_{ii}} \left(A_i - \sum_{k=1}^{i-1} u_{ki} L_k \right).$$

Nous voyons que, pour calculer les éléments de la i ème colonne de L, L_i , il nous faut connaître préalablement les éléments des lignes 1 à i de U ainsi que les éléments des colonnes 1 à $i-1$ de L . On peut remarquer de plus que, pratiquement, on ne calculera que les termes l_{ji} pour j compris entre $i+1$ et n , puisque on sait déjà que $l_{ii} = 1$ et que les autres termes de la colonne sont nuls.

Nous allons donc calculer :

- la première ligne de U , puis la première colonne de L ,
- la deuxième ligne de U , puis la deuxième colonne de L ,... et ainsi de suite.

Dans l'algorithme décrit ci-dessous, il faut compléter les matrices L et U , pour U en mettant des zéros pour la partie triangulaire inférieure, pour L en mettant des zéros pour la partie triangulaire supérieure et des uns sur la diagonale.

Algorithm 4 Système linéaire : Algorithme de DOOLITTLE

```

1: for  $i = 1$  to  $n - 1$  do
2:   for  $j = i$  to  $n$  do
3:      $u_{ij} \leftarrow a_{ij} - \sum_{k=1}^{i-1} l_{ik} u_{kj}$ 
4:   end for
5:   for  $j = i + 1$  to  $n$  do
6:      $l_{ji} \leftarrow \frac{1}{u_{ii}} \left( a_{ji} - \sum_{k=1}^{i-1} l_{jk} u_{ki} \right)$ 
7:   end for
8: end for
9:  $u_{nn} \leftarrow a_{nn} - \sum_{k=1}^{n-1} l_{nk} u_{kn}$ 

```

1.4 Traitement des matrices symétriques

1.4.1 La factorisation LDL^T

Lorsqu'une matrice A est symétrique, alors la factorisation $A = LU$, lorsque celle-ci est possible, peut prendre une forme particulière tenant compte de cette symétrie.

Proposition 1.4.1 *Si A est symétrique avec toutes ses sous-matrices principales régulières, alors elle admet une unique factorisation sous la forme :*

$$A = LDL^T$$

où L est une matrice triangulaire inférieure à diagonale unité et D une matrice diagonale régulière.

Démonstration : D'après l'hypothèse, A vérifie les conditions pour avoir une unique factorisation LU , avec L triangulaire inférieure à diagonale unité. Par ailleurs si on note D la matrice diagonale dont la diagonale est égale à celle de U on peut écrire

$$A = LDV$$

avec $V = D^{-1}U$ qui est une matrice triangulaire supérieure à diagonale unité (car U est une matrice régulière et donc ses éléments diagonaux sont non nuls). Comme A est symétrique, on a

$$A = LDV = V^T DL^T = A^T.$$

Or, on a démontré l'unicité de la factorisation LU (dans le cas où L est à diagonale unité), d'où $L = V^T$; puisque V^T est une matrice triangulaire inférieure à diagonale unité et que DL^T est une matrice triangulaire supérieure. L'unicité de la factorisation LDL^T découle de l'unicité de la factorisation LU .

Le résultat précédent est uniquement un résultat d'existence : en pratique on calcule les matrices

L et D en procédant différemment. L'algorithme consiste à calculer les termes de D et L colonne par colonne (ou ligne par ligne) en identifiant les termes de A dans l'égalité

$$A = LDL^T$$

Nous allons maintenant voir une méthode de factorisation adaptée aux matrices symétriques définies positives, qui est directement issue de la factorisation LDL^T pour les matrices symétriques.

1.4.2 La factorisation de Cholesky : existence

Dans tout ce paragraphe A désigne une matrice symétrique définie positive. Nous aurons besoin du lemme suivant pour pouvoir faire la démonstration du théorème principal concernant l'existence et l'unicité de la factorisation.

Lemme 1.4.1 *Une matrice A symétrique définie positive a toutes ses sous-matrices principales $[A]_k$ régulières.*

Démonstration : Soit k fixé et $\bar{z} \in R^k$ un vecteur quelconque non nul. On définit

$$z = \left(\bar{z}^T, 0, \dots, 0 \right)^T = \left(z_1, z_2, z_k, 0, \dots, 0 \right)^T.$$

Comme A est définie positive

$$0 < z^T A z = \bar{z}^T [A]_k \bar{z}$$

ce qui montre que $[A]_k$ est définie positive donc inversible.

Théorème 1.4.1 *Si A est une matrice symétrique définie positive elle admet une unique factorisation sous la forme*

$$A = BB^T;$$

où B est une matrice triangulaire inférieure dont tous les éléments diagonaux sont positifs.

Démonstration : Tout d'abord, comme A est définie positive toutes ses sous-matrices principales sont régulières (voir le Lemme précédent) et il résulte de la proposition qu'elle admet une factorisation sous la forme

$$A = LDL^T \tag{1.35}$$

La matrice D a tous ses éléments diagonaux positifs.

En effet, pour i fixé ($1 \leq i \leq n$), soit z tel que $L^T z = e_i$. Alors on a

$$0 < z^T A z = (L^T z)^T D L^T z = (e_i)^T D e_i = d_{ii};$$

puisque A est définie positive. En introduisant alors la matrice diagonale Λ telle que

$$\lambda_{ii} = \sqrt{d_{ii}}$$

on peut récrire (1.35) sous la forme

$$A = L \Lambda L^T = (L \Lambda)(L \Lambda)^T = B B^T,$$

obtenue en posant

$$B = L \Lambda.$$

Démontrons l'unicité. Soit donc une autre factorisation

$$A = \bar{B}\bar{B}^T,$$

où \bar{B} est une matrice triangulaire à éléments diagonaux positifs. On peut donc écrire

$$BB^T = \bar{B}\bar{B}^T \Rightarrow B^{-1}BB^T[\bar{B}^T]^{-1} = B^{-1}\bar{B}\bar{B}^T[\bar{B}^T]^{-1} \Rightarrow B^T[\bar{B}^T]^{-1} = B^{-1}\bar{B}$$

or, dans cette dernière égalité les matrices sont, à gauche, triangulaires supérieures et, à droite, triangulaires inférieures, ceci n'est possible que si elles sont diagonales et donc

$$B^T[\bar{B}^T]^{-1} = B^{-1}\bar{B} = D,$$

ce qui implique que

$$\forall i = 1, 2, \dots, n \quad \frac{b_{ii}}{\bar{b}_{ii}} = \frac{\bar{b}_{ii}}{b_{ii}} = d_{ii} \Rightarrow (b_{ii})^2 = (\bar{b}_{ii})^2,$$

soit finalement $b_{ii} = \bar{b}_{ii}$, puisque ces nombres sont positifs, et donc

$$d_{ii} = 1.$$

D est donc la matrice identité, donc $B = \bar{B}$ d'où l'unicité.

Comme dans le cas de la factorisation LDL^T le résultat précédent est uniquement un résultat d'existence et en pratique on calcule la matrice B colonne par colonne (ou ligne par ligne) en identifiant les termes de A dans l'égalité

$$A = BB^T.$$

1.4.3 Algorithme de Cholesky

Il s'agit d'une adaptation de la factorisation LU , à un système dont la matrice A est symétrique définie positive. L'algorithme se décompose en trois étapes :

- calcul de la décomposition de la matrice du système,
- résolution du premier système triangulaire,
- résolution du deuxième système triangulaire.

Plus précisément, on cherche une décomposition de la matrice A du système, de la forme $A = BB^T$ (voir le paragraphe précédent), où B est une matrice triangulaire inférieure dont les termes diagonaux sont positifs (B^T désigne sa transposée). On obtient cette décomposition par identification :

$$a_{11} = \underline{B}_1(B^T)_1 = b_{11}^2.$$

Puisque b_{11} doit être positif, on a

$$b_{11} = \sqrt{a_{11}}$$

$$A_1 = B(B^T)_1 = b_{11}B_1 \iff B_1 = \frac{1}{b_{11}}A_1,$$

ce qui permet de déterminer la première colonne de B .

De façon similaire on définit la deuxième colonne de B .

$$a_{22} = \underline{B}_2(B^T)_2 = b_{21}^2 + b_{22}^2.$$

Puisque b_{22} doit être positif, on a

$$b_{22} = \sqrt{a_{22} - b_{21}^2}.$$

$$A_2 = B(B^T)_2 = b_{21}B_1 + b_{22}B_2 \iff B_2 = \frac{1}{b_{22}}(A_2 - b_{21}B_1),$$

ce qui termine de déterminer la deuxième colonne de B .

De façon générale, lorsque l'on a déterminé les $j - 1$ premières colonnes de B , on écrit :

$$a_{jj} = \underline{B}_j(B^T)_j = \sum_{k=1}^j b_{jk}^2 \iff b_{jj}^2 = a_{jj} - \sum_{k=1}^{j-1} b_{jk}^2.$$

Puisque b_{jj} doit être positif, on a

$$b_{jj} = \sqrt{a_{jj} - \sum_{k=1}^{j-1} b_{jk}^2}.$$

$$A_j = B(B^T)_j = \sum_{k=1}^j b_{jk}B_k \iff B_j = \frac{1}{b_{jj}}\left(A_j - \sum_{k=1}^{j-1} b_{jk}B_k\right),$$

ce qui termine de déterminer la j ème colonne de B . On déterminera successivement les colonnes $1, 2, \dots, n - 1$, puis on terminera par le calcul de

$$b_{nn} = \sqrt{a_{nn} - \sum_{k=1}^{n-1} b_{nk}^2}.$$

Comme dans la factorisation de *Doolittle*, dans chaque colonne j on ne calcule que les termes b_{ij} pour i supérieur ou égal à j , puisque les autres termes de cette colonne sont connus car nuls.

On a donc l'algorithme :

Algorithm 5 Système linéaire : Algorithme de Cholesky

```

1: for  $j = 1$  to  $n - 1$  do
2:    $b_{jj} \leftarrow \sqrt{a_{jj} - \sum_{k=1}^{j-1} b_{jk}^2};$ 
3:   for  $i = j + 1$  to  $n$  do
4:      $b_{ij} \leftarrow \frac{1}{b_{jj}}\left(a_{ij} - \sum_{k=1}^{j-1} b_{ik}b_{jk}\right)$ 
5:   end for
6: end for
7:  $b_{nn} \leftarrow \sqrt{a_{nn} - \sum_{k=1}^{n-1} b_{nk}^2}.$ 

```

Si la matrice est définie positive, les nombres dont on prend la racine carrée sont positifs, les nombres par lesquels on divise sont non nuls (puisque l'on a montré que la décomposition existait). Cependant, il est prudent d'inclure dans l'algorithme des tests qui s'assurent que ces propriétés sont vérifiées.

1.5 Normes matricielles, conditionnement des matrices

1.5.1 Normes vectorielles

Pour pouvoir écrire que deux vecteurs sont proches ou que deux matrices sont proches, il faut pouvoir mesurer la "distance" entre ces deux objets. Ceci se fait généralement en utilisant une norme qu'elle soit vectorielle ou matricielle. Les normes matricielles sont aussi un outil indispensable pour le calcul du conditionnement d'une matrice, indicateur de la "bonne résolution" d'un système linéaire.

Définition 1.5.1 Soit E un espace vectoriel sur K ($K = \mathbb{R}$ ou \mathbb{C}). On appelle norme sur E une application de E dans \mathbb{R}^+ notée

$$x \mapsto \|x\|$$

possédant les propriétés suivantes : quel que soit $x \in E, y \in E, \lambda \in K$:

1. $\|x\| = 0 \Leftrightarrow x = 0$,
2. $\|\lambda x\| = |\lambda| \|x\|$ (la norme est positivement homogène)
3. $\|x + y\| \leq \|x\| + \|y\|$ (inégalité triangulaire)

La notation $|\lambda|$ représente la valeur absolue ou le module de λ suivant que λ est réel ou complexe.

Exemples :

$$\begin{aligned} \text{--- } E = \mathbb{R}^n \quad \|x\|_1 &= \sum_{i=1}^n |x_i|, \\ \text{--- } E = \mathbb{R}^n \quad \|x\|_2 &= \sqrt{\sum_{i=1}^n x_i^2}, \text{ (norme euclidienne),} \\ \text{--- } E = \mathbb{R}^n \quad \|x\|_\infty &= \max_{i=1 \dots n} |x_i|, \\ \text{--- } E = \mathbb{C}^n \quad \|x\|_2 &= \sqrt{\sum_{i=1}^n |x_i|^2} \text{ avec } |x_i|^2 = x_i \bar{x}_i \end{aligned}$$

Pour montrer l'inégalité triangulaire de la norme euclidienne, on utilise l'inégalité de Cauchy-Schwarz :

$$\left| \sum_{i=1}^n x_i y_i \right| \leq \sqrt{\sum_{i=1}^n x_i^2} \sqrt{\sum_{i=1}^n y_i^2}$$

1.5.2 Définition de la norme matricielle

Définition 1.5.2 On appelle norme matricielle une application qui à une matrice quelconque, associe un nombre réel positif. On note :

$$A \mapsto \|A\|$$

possédant les propriétés suivantes : quel que soit $A \in \mathcal{M}_{nn}, B \in \mathcal{M}_{nn}, \lambda \in K$:

1. $\|A\| = 0 \Leftrightarrow A = 0$,
2. $\|\lambda A\| = |\lambda| \|A\|, \forall A \in \mathcal{M}_{mn}, \forall \lambda \in \mathbb{R}$
3. $\|A + B\| \leq \|A\| + \|B\|, \forall A \in \mathcal{M}_{mn}, \forall B \in \mathcal{M}_{mn},$
4. $\|AB\| \leq \|A\| \|B\|, \forall A \in \mathcal{M}_{mn}, \forall B \in \mathcal{M}_{np}$

Les normes matricielles vérifient donc, en plus des propriétés des normes vectorielles, une relation sur le produit des matrices.

Définition 1.5.3 Soit $A \in \mathcal{M}_{mn}(\mathbb{R})$. on appelle **norme matricielle subordonnée**, une norme matricielle définie par :

$$\|A\| = \max_{x \in \mathbb{R}^n, x \neq 0} \frac{\|Ax\|}{\|x\|}$$

Quelques exemples des normes matricielles :

- On peut montrer que la norme matricielle subordonnée à la norme vectorielle notée $\|\cdot\|_1$ vaut

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

- la norme matricielle subordonnée à la norme vectorielle notée $\|\cdot\|_\infty$ vaut

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

- Il existe des normes matricielles qui ne sont pas subordonnées à une norme vectorielle, c'est le cas de la norme de Frobenius :

$$\|A\|_F = \sqrt{\sum_{i=1}^n \sum_{j=1}^n |a_{ij}|^2} = \sqrt{\text{trace}(A^T A)}$$

1.5.3 Rayon spectral

Définition 1.5.4 Soit $A \in \mathcal{M}_{nn}(C)$, on appelle **rayon spectral** de A et on note $\rho(A)$ le nombre réel :

$$\rho(A) = \max_{1 \leq i \leq n} |\lambda_i|$$

où $(\lambda_i)_{1 \leq i \leq n}$ sont les valeurs propres (complexes) de A .

En effet si λ_k est la plus grande valeur propre (en module) de A , si on pose $r = |\lambda_k|$, alors toutes les valeurs propres sont dans le disque (du plan complexe) de rayon r , cette quantité s'appelle le rayon spectral (l'ensemble des valeurs propres d'une matrice s'appelle aussi le spectre de la matrice).

Théorème 1.5.1 — Soit $A \in \mathcal{M}_{nn}(\mathbb{R})$, alors :

$$\|A\|_2^2 = \rho(AA^T) = \rho(A^T A)$$

- Si la matrice C est symétrique, alors :

$$\|C\|_2 = \rho(C)$$

- Soit $A \in \mathcal{M}_{nm}(\mathbb{R})$, alors

$$\|A\|_2^2 = \|AA^T\|_2 = \|A^T A\|_2.$$

1.5.4 Introduction au conditionnement d'une matrice

Lorsque l'on a résolu numériquement un système linéaire il est possible que la solution trouvée soit très différente de la solution exacte. En effet dans certains cas une petite modification

du second membre ou de la matrice du système entraîne une grande modification de la solution. Par exemple la résolution de $Ax = b$ avec le choix de A et b suivant

$$A = \begin{pmatrix} 10 & 7 & 8 & 7 \\ 7 & 5 & 6 & 5 \\ 8 & 6 & 10 & 9 \\ 7 & 5 & 9 & 10 \end{pmatrix}, \quad b = \begin{pmatrix} 32 \\ 23 \\ 33 \\ 31 \end{pmatrix} \quad \text{donne pour solution} \quad x = \begin{pmatrix} 32 \\ 23 \\ 33 \\ 31 \end{pmatrix} \quad (1.36)$$

Le choix d'un nouveau second membre $b + \delta b$

$$b + \delta b = \begin{pmatrix} 32.1 \\ 22.9 \\ 33.1 \\ 30.9 \end{pmatrix} \quad \text{donne pour solution} \quad x + x\delta = \begin{pmatrix} 9.2 \\ -12.6 \\ 4.5 \\ 1.1 \end{pmatrix} \quad (1.37)$$

Une faible erreur relative sur la norme de b induit une erreur relative sur la norme de x beaucoup plus importante, par exemple si on prend la norme $\|\cdot\|_\infty$ on a :

$$\frac{\|\delta b\|}{\|b\|} = 3.10^{-3}, \quad \frac{\|\delta x\|}{\|x\|} = 13.6$$

Une telle situation est due au fait que la matrice A a un conditionnement grand (on dit également que A est mal conditionnée), au sens de la définition suivante :

Définition 1.5.5 On appelle **conditionnement** de A relatif à une norme subordonnée le nombre :

$$\chi(A) = \|A\| \|A^{-1}\|.$$

Il résulte de la définition d'une norme subordonnée que

$$1 = \|I\| = \|AA^{-1}\| \leq \|A\| \|A^{-1}\|$$

Donc, on a toujours $\chi(A) \geq 1$.

Dans l'exemple précédent on a

$$A^{-1} = \begin{pmatrix} 25 & -41 & 10 & -6 \\ -41 & 68 & -17 & 10 \\ 10 & -17 & 5 & -3 \\ -6 & 10 & -3 & 2 \end{pmatrix} \quad (1.38)$$

et $\chi_\infty(A) = \|A\|_\infty \|A^{-1}\|_\infty = 4488$.

Proposition 1.5.1 Si la matrice A est symétrique et si l'on prend la norme subordonnée à la norme euclidienne, on a :

$$\chi_2(A) = \frac{\max_{1 \leq i \leq n} |\lambda_i|}{\min_{1 \leq i \leq n} |\lambda_i|}$$

pour les matrices symétriques. Le résultat montre donc que si le spectre d'une matrice, c'est-à-dire l'ensemble des valeurs propres, est très étalée la matrice est mal conditionnée. Par contre si ce spectre (en module) est bien regroupé, la matrice sera bien conditionnée.

1.5.5 Lien entre le conditionnement et les erreurs

On va démontrer maintenant des inégalités liant le conditionnement de A et les erreurs relatives sur b et x :

Théorème 1.5.2 *On suppose que l'on a $Ax = b$ et $A(x + \delta x) = b + \delta b$, (on perturbe le second membre), Alors :*

$$\frac{\|\delta x\|}{\|x\|} \leq \chi(A) \frac{\|\delta b\|}{\|b\|} \quad (1.39)$$

On suppose que l'on a $Ax = b$ et $(A + \delta A)(x + \delta x) = b$, (on perturbe la matrice), alors :

$$\frac{\|\delta x\|}{\|x + \delta x\|} \leq \chi(A) \frac{\|\delta A\|}{\|A\|} \quad (1.40)$$

Il découle immédiatement de ce théorème que si le conditionnement de A est faible (très proche de 1), si la variation relative de b ou de A est faible alors la variation de x sera faible elle aussi.

Remarquons que les inégalités démontrées dans le théorème précédent sont optimales en ce sens que la matrice A étant donnée, il existe b et δb tels qu'il y ait égalité dans la relation (1.39). On peut remarquer, à ce sujet, que l'exemple du paragraphe référencé est révélateur puisque avec un conditionnement de 4488 et une erreur relative sur b de 3.10^{-3} la borne supérieure de l'erreur relative sur x est de l'ordre de 13.6 ce qui est précisément ce que l'on a trouvé ! De même il existe δA et b tels qu'il y ait égalité dans la relation (1.40).

Résolution des systèmes linéaires : Méthodes itératives

2.1 Principes généraux

Dans ce chapitre nous considérons, comme dans le chapitre 1, un système d'équations linéaires d'ordre n de la forme :

$$Ax = b, \quad (2.1)$$

où $A \in \mathcal{M}_{n,n}(\mathbb{R})$ est une matrice régulière donnée et $b \in \mathbb{R}^n$ un vecteur donné également.

Nous avons vu que pour résoudre numériquement le système (2.1), il suffit de procéder à une élimination de Gauss avec ou sans changement de pivot, ou éventuellement, de procéder à une décomposition LU de A (ou BB^T si A est symétrique définie positive), puis résoudre deux systèmes triangulaires.

Dans tous les cas, si la matrice A est pleine, le nombre d'opérations nécessaires à la mise en oeuvre de ces algorithmes est de l'ordre de n^3 , ce qui peut être énorme lorsque n est grand. Dans ce cas, nous allons étudier d'autres méthodes dites **itératives** pour résoudre (2.1).

Les méthodes itératives sont utilisées soit pour la résolution de systèmes linéaires de très grande taille, soit lorsque l'on dispose d'une estimation de la solution que l'on veut améliorer. Une méthode itérative consiste à construire une suite de vecteurs $x^{(0)}, x^{(1)}, \dots, x^{(k)}, \dots$ qui, on l'espère, ou mieux on le démontre, convergera vers la solution du système linéaire (2.1), c'est à dire

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

On se propose de résoudre le problème suivant

$$Ax = b$$

par une méthode itérative de la forme suivante :

$$\begin{cases} x^{(0)} & \text{donné} \\ Mx^{(k+1)} &= Nx^{(k)} + b, \end{cases}$$

où les matrices $M, N \in \mathcal{M}_{nn}(\mathbb{R})$, M inversible, sont convenablement choisies.

Si cette méthode itérative converge, c'est-à-dire si la suite $x^{(k)}$ converge vers \bar{x} alors on a, à la limite,

$$(M - N)\bar{x} = b,$$

ce qui correspondra à la résolution de $Ax = b$ si

$$A = M - N.$$

Il y a une infinité de choix possibles pour M et N vérifiant $A = M - N$, nous allons en donner deux à titre d'exemples. L'idée est bien sûr de choisir une matrice M particulièrement facile à inverser, par exemple diagonale, ou bien triangulaire inférieure. Ces deux choix correspondent respectivement à la méthode de *Jacobi* et à la méthode de *Gauss – Seidel*

2.2 La méthode de Jacobi

Le système linéaire $Ax = b$ s'écrit

$$\begin{cases} \sum_{j=1}^n a_{1j}x_j = b_1 \\ \dots \\ \sum_{j=1}^n a_{ij}x_j = b_i, \text{ pour } 2 \leq i \leq n-1, \\ \dots \\ \sum_{j=1}^n a_{nj}x_j = b_n. \end{cases}$$

La méthode de *Jacobi* consiste, à chaque itération k , à résoudre chaque équation par rapport à l'une des variables, les autres étant fixées à leurs valeurs obtenues à l'itération précédente. Soit donc le vecteur $x^{(k)}$ donné, alors on détermine successivement les composantes de $x^{(k+1)}$ par les formules :

$$\begin{cases} a_{11}x_1^{(k+1)} = b_1 - \sum_{j=2}^n a_{1j}x_j^{(k)} \\ \dots \\ a_{ii}x_i^{(k+1)} = b_i - \sum_{j=1, j \neq i}^n a_{ij}x_j^{(k)}, \text{ pour } 2 \leq i \leq n-1, \\ \dots \\ a_{nn}x_n^{(k+1)} = b_n - \sum_{j=1}^{n-1} a_{nj}x_j^{(k)}. \end{cases}$$

Les formules précédentes ne définissent effectivement $x^{(k+1)}$ que si les coefficients diagonaux de A sont tous non nuls. Ceci n'est pas une restriction, car on peut démontrer que si une matrice est inversible, il existe une permutation de ses lignes telle que tous les éléments de la diagonale de la matrice ainsi obtenue soient non nuls.

On peut écrire le système précédant composante par composante de la manière suivante :

$$x_i^{(k+1)} = \frac{1}{a_{ii}} \left(b_i - \sum_{j=1, j \neq i}^n a_{ij}x_j^{(k)} \right), \quad 1 \leq i \leq n. \quad (2.2)$$

les relations (2.2) permettent de calculer explicitement les composantes $x_i^{(k+1)}$, $1 \leq i \leq n$, du vecteur $x^{(k+1)}$ à partir des composantes $x_i^{(k)}$, $1 \leq i \leq n$ du vecteur $x^{(k)}$.

On peut écrire les relations précédentes sous forme matricielle. Pour cela introduisons la décomposition suivante de A :

$$A = D - E - F$$

avec

- D matrice diagonale contenant la diagonale de A ,
- E matrice triangulaire inférieure (triangle inférieur de $-A$),
- F matrice triangulaire supérieure (triangle supérieur de $-A$).

Avec ces notations on peut écrire le système $Ax = b$ sous la forme

$$Dx = (E + F)x + b,$$

La méthode de *Jacobi* s'écrit

$$\begin{cases} x^{(0)} & \text{donné} \\ Dx^{(k+1)} &= (E + F)x^{(k)} + b, \end{cases} \quad (2.3)$$

La matrice D est diagonale, donc, $D^{-1} = \text{diag}(1/a_{11}, 1/a_{22}, \dots, 1/a_{nn})$.

2.3 La méthode de Gauss–Seidel

Il s'agit d'une modification de la méthode de *Jacobi* qui consiste à utiliser pour chaque équation, les composantes de $x^{(k+1)}$ déjà calculées, ceci conduit aux formules

$$\begin{cases} a_{11}x_1^{(k+1)} &= b_1 - \sum_{j=2}^n a_{1j}x_j^{(k)} \\ \dots & \\ a_{ii}x_i^{(k+1)} &= b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{(k+1)} - \sum_{j=i+1}^n a_{ij}x_j^{(k)}, \text{ pour } 2 \leq i \leq n-1, \\ \dots & \\ a_{nn}x_n^{(k+1)} &= b_n - \sum_{j=1}^{n-1} a_{nj}x_j^{(k+1)}. \end{cases}$$

Nous avons alors :

$$x_i^{(k+1)} = \frac{1}{a_{ii}} \left(b_i - \sum_{j < i} a_{ij}x_j^{(k+1)} - \sum_{j > i} a_{ij}x_j^{(k)} \right), \quad 1 \leq i \leq n. \quad (2.4)$$

Sous forme matricielle cela revient à écrire le système $Ax = b$ sous la forme

$$(D - E)x = Fx + b,$$

où les matrices D , E et F ont été données dans le paragraphe précédent sur la méthode de *Jacobi*.

La méthode de *Gauss – Seidel* s'écrit donc

$$\begin{cases} x^{(0)} & \text{donné} \\ (D - E)x^{(k+1)} &= Fx^{(k)} + b, \end{cases} \quad (2.5)$$

chaque itération, la matrice du système à résoudre est triangulaire inférieure.

Remarques : On observe que les méthodes de *Jacobi* et *Gauss – Seidel* que nous venons de voir peuvent se mettre sous la forme

$$Mx^{(k+1)} = Nx^{(k)} + b$$

où

- $M = D$, $N = E + F$, pour la méthode de *Jacobi*,
- $M = D - E$, $N = F$, pour la méthode de *Gauss – Seidel*
- La matrice J définie par : $J = M^{-1}N = D^{-1}(E + F)$ est appelée la matrice de *Jacobi*.
- La matrice G définie par : $G = M^{-1}N = (D - E)^{-1}F$ est appelée la matrice de *Gauss – Seidel*.

A priori, la méthode de Gauss-Seidel devrait être plus performante que la méthode de Jacobi puisqu'on tient compte, au fur et à mesure, des valeurs $x_i^{(k+1)}$ déjà calculées.

2.4 Etude de la convergence des méthodes itératives

Avant, d'énoncer quelques résultats de convergence, nous donnons les définitions suivantes :

2.4.1 Définitions et propositions

Définition 2.4.1 Soit $\|\cdot\|$ une certaine norme vectorielle. On définit une norme matricielle, appelée norme matricielle subordonnée à la norme $\|\cdot\|$, par

$$\|A\| = \max_{x \in \mathbb{R}^n - \{0\}} \frac{\|Ax\|}{\|x\|}$$

Définition 2.4.2 On dit que la matrice A est à diagonale strictement dominante si

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

Définition 2.4.3 On dit que λ est une valeur propre de $A \in M_{n,n}$ s'il existe un vecteur $Y \in M_{n,1}$, $Y \neq 0$ tel que

$$AY = \lambda Y.$$

Dans ce cas on dit que Y est un vecteur propre associé à la valeur propre λ et que (λ, Y) est un couple propre de A .

Définition 2.4.4 On appelle polynôme caractéristique de A le polynôme

$$P_A(s) = \det(sI - A)$$

Proposition 2.4.1 Une condition nécessaire et suffisante pour que λ soit valeur propre de A est qu'elle soit racine du polynôme caractéristique, c'est à dire :

$$P_A(\lambda) = 0$$

Définition 2.4.5 Si $A \in M_{n,n}(R)$ est une matrice de valeurs propres $\lambda_1, \lambda_2, \dots, \lambda_n$, nous appelons rayon spectral de A la quantité :

$$\rho(A) = \max_{1 \leq j \leq n} |\lambda_j|,$$

où $|\lambda_j|$ est le module de λ_j , $1 \leq j \leq n$.

2.4.2 Etude de la convergence

Commençons par étudier la convergence de la méthode itérative suivante :

$$\begin{cases} x^{(0)} & \text{donné} \\ x^{(k+1)} & = Cx^{(k)} + d, \end{cases} \quad (2.6)$$

Proposition 2.4.2 *S'il existe une norme matricielle subordonnée telle que*

$$\|C\| < 1$$

la méthode (2.6) est convergente quel que soit le choix de $x^{(0)}$ et elle converge vers la solution de

$$(I - C)\bar{x} = d.$$

Démonstration : Tout d'abord la solution \bar{x} du système d'équations $(I - C)\bar{x} = d$ existe et elle est unique, car la matrice $I - C$ est inversible. En effet :

$$x \in \text{Ker}(I - C) \Leftrightarrow x = Cx \Rightarrow \|x\| = \|Cx\| \leq \|C\|\|x\|$$

Si $\|x\| \neq 0$, on aurait, après simplification $1 \leq \|C\|$, ce qui est contraire à l'hypothèse, donc $\|x\| = 0$, donc $x = 0$. On vient de montrer que $\text{ker}(I - C) = \{0\}$, ce qui est équivalent à $I - C$ inversible puisque la matrice $I - C$ est carrée.

Montrons maintenant que $x^{(k)}$ converge vers \bar{x} solution de $(I - C)\bar{x} = d$. On a

$$x^{(k)} - \bar{x} = C(x^{(k-1)} - \bar{x}) = C^2(x^{(k-2)} - \bar{x}) = \dots = C^k(x^{(0)} - \bar{x})$$

d'où

$$\|x^{(k)} - \bar{x}\| \leq \|C^k\|\|x^{(0)} - \bar{x}\| \leq \|C\|^k\|x^{(0)} - \bar{x}\|$$

Passons à la limite quand k tend vers l'infini. Puisque $\|C\| < 1$, $\|C\|^k$ tend vers 0 et donc $x^{(k)}$ tend vers \bar{x} .

Un résultat intéressant et admis sur la convergence des méthodes itératives donne une caractérisation à partir du rayon spectral de la matrice C .

Théorème 2.4.1 *Une condition nécessaire et suffisante pour que la méthode itérative (2.6) soit convergente est que*

$$\rho(C) < 1$$

Proposition 2.4.3 *Si la matrice A est à diagonale strictement dominante alors les méthodes de Jacobi et Gauss–Seidel sont convergentes.*

Démonstration : Démontrons cette proposition pour la méthode de Jacobi. Pour la méthode de Gauss–Seidel, la démonstration sera faite en travaux dirigés.

La méthode de Jacobi s'écrit :

$$x^{(k+1)} = D^{-1} \left[b + (E + F)x^{(k)} \right],$$

et donc la matrice d'itération est

$$C = D^{-1}(E + F).$$

Puisque A est à diagonale dominante, on a

$$\|D^{-1}(E + F)\|_{\infty} = \max\left(\frac{1}{a_{ii}} \sum_{j \neq i} |a_{ij}|\right) < 1$$

et donc la méthode est convergente.

Un autre résultat intéressant pour les applications est le suivant (nous l'admettrons sans démonstration).

Proposition 2.4.4 *Si la matrice A est symétrique définie positive, alors la méthode de Gauss–Seidel est convergente.*

