

Analyse numérique élémentaire

Équipe de Mathématiques Appliquées

UTSEUS

Printemps 2020



Sommaire

1	Algèbre linéaire - révisions	5
1.1	Espace vectoriel	6
1.2	Applications linéaires	14
1.3	Matrices	19
1.4	Déterminants	32
1.5	Systèmes linéaires	41
1.6	Valeurs propres	46
	Documents du chapitre 1	55
	Exercices du chapitre 1	59
2	Résolution des systèmes linéaires	104
2.1	Motivations	105
2.2	L'élimination de Gauss	109
2.3	Calcul direct de la factorisation $A = LU$	127
2.4	Factorisation LU avec recherche de pivots	134
2.5	Traitement des matrices symétriques	152

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)



	Exemples du chapitre 2	162
	Exercices du chapitre 2	167
3	Résolution des problèmes de moindres carrés	185
3.1	Formulation générale des problèmes de moindres carrés	186
	Documents du chapitre 3	194
	Exercices du chapitre 3	197
4	Méthodes itératives	201
4.1	Méthodes itératives de résolution des systèmes linéaires	202
4.2	Résolution d'une équation non-linéaire à une inconnue	212
4.3	Résolution d'un système d'équations non-linéaires	228
	Documents du chapitre 4	237
	Exercices du chapitre 4	242
5	Interpolation	258
5.1	Motivations	259
5.2	Interpolation polynomiale	264
5.3	Interpolation avec les splines cubiques	282
	Exemples du chapitre 5	286
	Documents du chapitre 5	294
	Exercices du chapitre 5	297
6	Intégration	311
6.1	Motivations et principe des méthodes numériques d'intégration	312
6.1	Méthodes utilisant le polynôme d'interpolation	314

Sommaire
Concepts

Exemples
Exercices
Documents

	Exemples du chapitre 6	325
	Documents du chapitre 6	331
	Exercices du chapitre 6	337
7	Équations différentielles	345
7.1	Principes généraux de résolution	346
7.2	Les schémas à un pas	358
7.3	Les schémas multi-pas	376
	Exemples du chapitre 7	387
	Documents du chapitre 7	394
	Exercices du chapitre 7	405
8	Détermination des valeurs propres et des vecteurs propres	413
8.1	Rappels et notations	414
8.2	Méthode de la puissance itérée - énoncé et hypothèses	416
8.3	Méthode de la puissance itérée - convergence et remarques	418
8.4	Méthode de la puissance itérée inverse - principe	421
8.5	Méthode de la puissance itérée inverse	423
8.6	Les méthodes de déflation	425
	Exercices du chapitre 8	428

Sommaire
Concepts

Exemples
Exercices
Documents

Chapitre 1

Algèbre linéaire - révisions

1.1	Espace vectoriel	6
1.2	Applications linéaires	14
1.3	Matrices	19
1.4	Déterminants	32
1.5	Systèmes linéaires	41
1.6	Valeurs propres	46
	Documents du chapitre 1	55
	Exercices du chapitre 1	59

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

1.1 Espace vectoriel

1.1.1	Généralités	7
1.1.2	Sous-espace vectoriel	9
1.1.3	Famille libre, famille génératrice	10
1.1.4	Base d'un espace vectoriel	12

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

1.1.1 Généralités

Exercices :[Exercice 1.1](#)[Exercice 1.2](#)[Exercice 1.3](#)**Documents :**[Document 1.1](#)

La définition précise d'un espace vectoriel est donnée dans le document référencé. Puisque les principales applications que vous rencontrerez ne feront appel qu'à \mathbb{R}^n (ou \mathbb{C}^n) et P_n (les polynômes de degré inférieur ou égal à n), nous allons illustrer la définition dans ces deux cas particuliers.

- Commençons par étudier l'espace \mathbb{R}^3 des vecteurs \vec{x} que vous pouvez identifier à \overrightarrow{OM} dans un repère orthonormé (voir figure 8.0.1). Vous connaissez la somme des vecteurs et vous vérifierez en exercice qu'elle est associative, commutative, qu'il existe un élément neutre ($\vec{0}$) et que tout vecteur (\vec{x}) admet un symétrique ($-\vec{x}$). C'est la loi de composition interne de l'espace vectoriel \mathbb{R}^3 .

Vous savez multiplier un vecteur par un nombre réel λ (homothétie de rapport λ), c'est la deuxième loi de l'espace vectoriel dont il est facile de montrer qu'elle possède les bonnes propriétés.

- L'ensemble P_n (polynômes de degré inférieur ou égal à n à coefficients réels) est muni aussi d'une addition associative, commutative, ayant un élément neutre (le polynôme nul) et telle que tout polynôme p a un symétrique $-p$. Vous savez aussi multiplier un polynôme par un

[Sommaire](#)
[Concepts](#)[Exemples](#)
[Exercices](#)
[Documents](#)

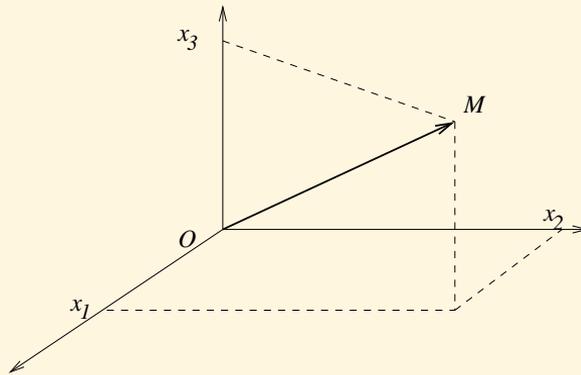


FIGURE 1.1.1 – Espace

nombre réel λ (il suffit de multiplier tous ses coefficients par λ) et ce produit possède aussi les bonnes propriétés pour faire de P_n un espace vectoriel.

Dans les deux exemples précédents la deuxième loi est constituée du produit d'un élément de l'ensemble par un nombre réel. On dit donc que les deux ensembles \mathbb{R}^3 et P_n sont des espaces vectoriels sur \mathbb{R} .

Remarque - Considérons maintenant l'ensemble \mathbb{C} des nombres complexes. La loi de composition interne est l'addition des nombres complexes. Pour la deuxième loi, il y a deux choix possibles. Si l'on considère le produit d'un nombre complexe par un nombre réel on obtient alors un espace vectoriel sur \mathbb{R} , mais, si l'on considère le produit d'un nombre complexe par un nombre complexe, on obtient alors un espace vectoriel sur \mathbb{C} .

Généralités

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

1.1.2 Sous-espace vectoriel

Exercices :

[Exercice 1.4](#)

[Exercice 1.5](#)

[Exercice 1.6](#)

Cours :

[Espace vectoriel - Généralités](#)

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 1.1.1. *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

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

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 ($F = \{\lambda \vec{x} + \mu \vec{y}, \lambda \in \mathbb{R}, \mu \in \mathbb{R}\}$).

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

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

1.1.3 Famille libre, famille génératrice

Exercices :[Exercice 1.7](#)[Exercice 1.8](#)

Définition 1.1.2. 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 1.1.3. 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 1.1.4. On dit que la famille $\{\vec{x}_1, \dots, \vec{x}_p\}$ 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

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

[Sommaire](#)
[Concepts](#)[Exemples](#)
[Exercices](#)
[Documents](#)

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

$$E = \text{vect} \langle \vec{x}_1, \dots, \vec{x}_p \rangle .$$

**Famille libre,
famille
génératrice**

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

1.1.4 Base d'un espace vectoriel

Exercices :[Exercice 1.9](#)[Exercice 1.10](#)[Exercice 1.11](#)**Cours :**[Famille libre, famille génératrice](#)

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 \mathbf{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 1.1.5. Une base $B = \{\vec{e}_1, \dots, \vec{e}_n\}$ d'un espace vectoriel E est une famille libre et génératrice.

Un vecteur \vec{x} de E est une combinaison linéaire d'éléments de B puisque B est une famille génératrice et on peut montrer que cette décomposition est unique (voir exercice 1.10).

Définition 1.1.6. 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 .

[Sommaire](#)
[Concepts](#)[Exemples](#)
[Exercices](#)
[Documents](#)

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 (non démontré), car il simplifie les démonstrations est le suivant :

Proposition 1.1.1. *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.*

Base d'un espace vectoriel

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

1.2 Applications linéaires

1.2.1	Définition de l'application linéaire	15
1.2.2	Composition et réciproque des applications linéaires	17

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

1.2.1 Définition de l'application linéaire

Exercices :

[Exercice 1.12](#)

[Exercice 1.13](#)

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 1.2.1. On appelle application linéaire $u : E \rightarrow F$, une application possédant les propriétés suivantes :

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

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

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

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

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

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

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 (non démontrée) suivante :

Proposition 1.2.1. *Soit $u : E \rightarrow F$ une application linéaire, alors*

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

Définition de l'application linéaire

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

1.2.2 Composition et réciproque des applications linéaires

Exercices :
[Exercice 1.14](#)

Documents :
[Document 1.2](#)

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

Définition 1.2.3. Soient $f : E \rightarrow F$ et $g : F \rightarrow G$ deux applications linéaires, la composée $(g \circ f)$ de f et g est une application linéaire définie par

$$\begin{aligned} g \circ f : E &\rightarrow G \\ \vec{x} &\mapsto g[f(\vec{x})]. \end{aligned}$$

La composée est, en effet une application linéaire puisque

$$(g \circ f)(\vec{x} + \vec{y}) = g[f(\vec{x} + \vec{y})] = g[f(\vec{x}) + f(\vec{y})] = g[f(\vec{x})] + g[f(\vec{y})] = (g \circ f)(\vec{x}) + (g \circ f)(\vec{y}),$$

$$(g \circ f)(\lambda \vec{x}) = g[f(\lambda \vec{x})] = g[\lambda f(\vec{x})] = \lambda g[f(\vec{x})] = \lambda (g \circ f)(\vec{x}).$$

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

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Définition 1.2.4. 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

$$f^{-1} \circ f = i_E \text{ et } f \circ f^{-1} = i_F, \quad (i_E \text{ est l'identité de } E \text{ dans } E).$$

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 (voir le document référencé pour les définitions) et alors f^{-1} est aussi bijective. Dans le cas des applications linéaires, on a une caractérisation de l'injectivité par le noyau (la démonstration est donnée dans le document référencé) :

Proposition 1.2.2. L'application linéaire u est injective si et seulement si

$$\text{Ker } u = \{\vec{0}\}.$$

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

Composition et réciproque des applications linéaires

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

1.3 Matrices

1.3.1	Définition des matrices	20
1.3.2	Somme et produit de matrices	23
1.3.3	Inverse et transposée d'une matrice	26
1.3.4	Définition d'un changement de base	28
1.3.5	Rang d'une application linéaire, d'une matrice	30

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

1.3.1 Définition des matrices

Exercices :

[Exercice 1.15](#)

[Exercice 1.16](#)

Cours :

[Application linéaire, noyau, image](#)

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 1.3.1. Soit $u : E \rightarrow F$ 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

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

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

La matrice A est donc représentée par

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2j} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3j} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{i1} & a_{i2} & a_{i3} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mj} & \cdots & a_{mn} \end{pmatrix}$$

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.

Par exemple, dans le plan E on peut écrire la matrice d'une rotation θ en prenant un repère orthonormé (on aura $F = E$, $\mathcal{F} = \mathcal{E} = \{\vec{e}_1, \vec{e}_2\}$). On a alors :

$$u(\vec{e}_1) = \cos \theta \vec{e}_1 + \sin \theta \vec{e}_2, \quad u(\vec{e}_2) = -\sin \theta \vec{e}_1 + \cos \theta \vec{e}_2,$$

Définition des matrices

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

ce qui donne la matrice

$$A = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}.$$

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 .

Définition des matrices

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

1.3.2 Somme et produit de matrices

Exercices :[Exercice 1.17](#)[Exercice 1.18](#)[Exercice 1.19](#)[Exercice 1.20](#)**Cours :**[Matrices - définition](#)

Une matrice se présente donc comme un tableau de nombres (de \mathbb{R} ou \mathbb{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 $u : \mathbb{R}^n \rightarrow \mathbb{R}^m$ (ou $u : \mathbb{C}^n \rightarrow \mathbb{C}^m$) définie par

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

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 1.3.2. 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

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

[Sommaire](#)
[Concepts](#)[Exemples](#)
[Exercices](#)
[Documents](#)

(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 1.3.3. Soient $u : F \rightarrow G$ et $v : E \rightarrow F$, on pose $w = u \circ v$ on a donc $w : E \rightarrow G$. 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 .

$$\begin{array}{ccc} & v & u \\ E & \longrightarrow & F & \longrightarrow & G \\ \mathcal{E} & B & \mathcal{F} & A & \mathcal{G} \end{array} \qquad \begin{array}{ccc} & w = u \circ v & \\ E & \longrightarrow & G \\ \mathcal{E} & C = AB & \mathcal{G} \end{array}$$

On peut démontrer le résultat suivant :

Proposition 1.3.1. 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 1.3.2. Soit u une application linéaire de E dans F , on note $\vec{y} = u(\vec{x})$, 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

**Somme et
produit de
matrices**

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

\vec{y}) dans la base \mathcal{E} (resp \mathcal{F}) on note

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \text{ et } Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_p \end{pmatrix}.$$

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

La démonstration est proposée en exercice.

**Somme et
produit de
matrices**

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

1.3.3 Inverse et transposée d'une matrice

Exercices :[Exercice 1.21](#)[Exercice 1.22](#)[Exercice 1.23](#)**Cours :**[Matrices - somme et produit](#)

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 1.3.4. 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 (voir exercice 1.22). D'autre part, puisque

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

on en déduit que

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

[Sommaire](#)
[Concepts](#)[Exemples](#)
[Exercices](#)
[Documents](#)

Définition 1.3.5. 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 $X^T Y = Y^T X$ est un scalaire de K (\mathbb{R} ou \mathbb{C}) et plus précisément on a : $X^T Y = Y^T X = \sum_{i=1}^m x_i y_i$.

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.

Inverse et transposée d'une matrice

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

1.3.4 Définition d'un changement de base

Exercices :[Exercice 1.24](#)[Exercice 1.25](#)

Définition 1.3.6. Soit $\mathcal{E} = \{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ et $\mathcal{E}' = \{\vec{e}'_1, \vec{e}'_2, \dots, \vec{e}'_n\}$ deux bases de E , on appelle **matrice de passage** de la base \mathcal{E} à la base \mathcal{E}' la matrice P définie par :

$$P = \begin{pmatrix} & \vec{e}'_1 & \vec{e}'_2 & \dots & \vec{e}'_j & \dots & \vec{e}'_n \\ p_{11} & p_{12} & \dots & p_{1j} & \dots & p_{1n} \\ p_{21} & p_{22} & \dots & p_{2j} & \dots & p_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ p_{i1} & p_{i2} & \dots & p_{ij} & \dots & p_{in} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ p_{n1} & p_{n2} & \dots & p_{nj} & \dots & p_{nn} \end{pmatrix} \begin{matrix} \vec{e}_1 \\ \vec{e}_2 \\ \dots \\ \vec{e}_i \\ \dots \\ \vec{e}_n \end{matrix}$$

qui est donc obtenue en mettant en colonnes les composantes, dans "l'ancienne base" \mathcal{E} des vecteurs de la "nouvelle base" \mathcal{E}' .

La matrice P est inversible et son inverse P^{-1} est la matrice de passage de \mathcal{E}' à \mathcal{E} . Les résultats suivants (voir le chapitre 2 de MT23) permettent de calculer les composantes d'un vecteur (resp. la matrice d'une application linéaire u) dans la nouvelle base à partir des composantes (resp. la matrice de u) dans l'ancienne base.

[Sommaire](#)
[Concepts](#)[Exemples](#)
[Exercices](#)
[Documents](#)

Proposition 1.3.3. - Transformation des composantes par changement de base - Soit $\vec{x} \in E$ et soit (x_1, x_2, \dots, x_n) et $(x'_1, x'_2, \dots, x'_n)$ les composantes de \vec{x} dans les bases \mathcal{E} et \mathcal{E}' , alors on a

$$X = PX' \text{ ou encore } X' = P^{-1}X, \text{ avec } X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \text{ et } X' = \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}.$$

Théorème 1.3.1. - Transformation des matrices par changement de base - Soit $u \in \mathcal{L}(E; E)$, A et A' les matrices représentant u dans les bases \mathcal{E} et \mathcal{E}' , alors si P est la matrice de passage de \mathcal{E} à \mathcal{E}' , on a la relation

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

(A et A' sont dites semblables.)

Définition d'un changement de base

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

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

Exercices :
[Exercice 1.26](#)

Cours :
[Application linéaire, noyau, image](#)

Définition 1.3.7. 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 1.3.8. 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 :

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Définition 1.3.9. 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 1.3.2. Soit $A \in \mathcal{M}_{mn}$, alors $\dim [\text{Ker } A] + \dim [\text{Im } A] = n$.

Définition 1.3.10. Soit $A \in \mathcal{M}_{mn}$, on appelle rang de A et on note $\text{rang}(A)$, la dimension du sous-espace vectoriel $\text{Im } A$.

Les résultats suivants sont démontrés dans le chapitre 2 de MT23 :

- 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.

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

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

1.4 Déterminants

1.4.1	Définition du déterminant par récurrence	33
1.4.2	Le déterminant et les colonnes	35
1.4.3	Propriétés essentielles du déterminant	37
1.4.4	Propriétés liées au rang	39

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

1.4.1 Définition du déterminant par récurrence

Exercices :

[Exercice 1.27](#)

[Exercice 1.28](#)

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

Définition 1.4.1. Soit $A \in \mathcal{M}_{n,n}$, on définit par récurrence une application :

$$\begin{array}{lcl} \det : \mathcal{M}_{n,n} & \longrightarrow & K \\ A & \longrightarrow & \det A \end{array}$$

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^e ligne et la j^e colonne, on pose alors

$$\det A = a_{11} \det A_{|1,1|} + \dots + (-1)^{k+1} a_{1k} \det A_{|1,k|} + \dots + (-1)^{n+1} a_{1n} \det A_{|1,n|}. \quad (1.4.1)$$

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

$$\begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix}.$$

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

$$\text{Ainsi, } \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21} \text{ et}$$

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

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 1.4.1. *La définition permet d'obtenir immédiatement les propriétés suivantes :*

1. *Le déterminant d'une matrice triangulaire inférieure est égal au produit des termes diagonaux.*
2. *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.*
3. *Si \bar{A} est la matrice dont les termes sont les conjugués de ceux de A alors*

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

Définition du déterminant par récurrence

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

1.4.2 Le déterminant et les colonnes

Exercices :

[Exercice 1.29](#)

[Exercice 1.30](#)

[Exercice 1.31](#)

Une illustration du théorème suivant a été donnée dans l'exercice [1.27](#) (pour la démonstration voir le chapitre 3 de MT23).

Théorème 1.4.1. *Le déterminant est une fonction linéaire de chaque colonne, donc une **application multi-linéaire** de l'ensemble des colonnes, c'est à dire*

$$\det(A_1, \dots, A_{k-1}, \lambda A_k, A_{k+1}, \dots, A_n) = \lambda \det(A_1, \dots, A_{k-1}, A_k, A_{k+1}, \dots, A_n) \quad (1.4.2)$$

$$\det(A_1, \dots, A_{k-1}, B + C, A_{k+1}, \dots, A_n) = \det(A_1, \dots, A_{k-1}, B, A_{k+1}, \dots, A_n) + \det(A_1, \dots, A_{k-1}, C, A_{k+1}, \dots, A_n). \quad (1.4.3)$$

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

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

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

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

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. Le résultat important suivant a aussi été démontré dans le chapitre 3 de MT23 :

Théorème 1.4.2. *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.*

Le déterminant et les colonnes

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

1.4.3 Propriétés essentielles du déterminant

Exercices :[Exercice 1.32](#)[Exercice 1.33](#)[Exercice 1.34](#)**Cours :**[Déterminant et colonnes](#)

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

Théorème 1.4.3. *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 :

[Sommaire](#)
[Concepts](#)[Exemples](#)
[Exercices](#)
[Documents](#)

Théorème 1.4.4. *On a les formules suivantes :*

(i) *Développement suivant la i^e ligne*

$$\det A = a_{i1} \operatorname{cof}(a_{i1}) + a_{i2} \operatorname{cof}(a_{i2}) + \dots + a_{in} \operatorname{cof}(a_{in}). \quad (1.4.4)$$

(ii) *Développement suivant la j^e colonne*

$$\det A = a_{1j} \operatorname{cof}(a_{1j}) + a_{2j} \operatorname{cof}(a_{2j}) + \dots + a_{nj} \operatorname{cof}(a_{nj}). \quad (1.4.5)$$

où

$$\operatorname{cof}(a_{ij}) = (-1)^{i+j} \det A_{|i,j|}$$

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

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

Théorème 1.4.5.

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

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

Théorème 1.4.6. *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

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

**Propriétés
essentielles du
déterminant**

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

1.4.4 Propriétés liées au rang

Exercices :
[Exercice 1.35](#)

Cours :
[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 1.4.2. 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

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

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 1.4.7. *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 $\hat{A} \in \mathcal{M}_{r, r}$ 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 référencé.

Proposition 1.4.2. *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 1.4.3. *Soit $A \in \mathcal{M}_{n, n}$,*

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

**Propriétés liées
au rang**

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

1.5 Systèmes linéaires

1.5.1	Systemes linéaires à matrice carrée	42
1.5.2	Systemes linéaires de dimension quelconque	44

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

1.5.1 Systèmes linéaires à matrice carrée

Exercices :

[Exercice 1.36](#)

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 1.5.1. *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** et la démonstration de ce système est donnée dans le chapitre 3 du cours de MT23.

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Théorème 1.5.2. Soit $A \in \mathcal{M}_{n,n}$ et soit le système $Ax = b$. Alors pour tout $j = 1, 2, \dots, n$, on a la formule de Cramer :

$$\det(A_1, A_2, \dots, A_{i-1}, b, A_{i+1}, \dots, A_n) = x_i \det A. \quad (1.5.1)$$

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 1.5.3. L'inverse de A est donnée par

$$A^{-1} = \frac{1}{\det A} (\text{co}(A))^T, \quad (1.5.2)$$

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.

**Systèmes
linéaires à
matrice carrée**

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

1.5.2 Systèmes linéaires de dimension quelconque

Exercices :

[Exercice 1.37](#)

[Exercice 1.38](#)

Cours :

[rang](#)

On veut résoudre un système linéaire de n équations à p inconnues ($A \in \mathcal{M}_{n,p}$) :

$$Ax = b \quad x \in \mathbb{R}^p \quad b \in \mathbb{R}^n.$$

- $b = 0$ - On dit que le système est homogène. On est donc amené à chercher les vecteurs x tels que $Ax = 0$ ce qui correspond au noyau de A . Ce système a donc au moins toujours la solution nulle. La résolution pratique utilise la méthode d'élimination de Gauss pour résoudre $Ax = 0$.

On va traiter maintenant en détail l'exemple suivant :

$$\begin{cases} x_1 & +x_2 & +x_3 & = 0 \\ x_1 & +2x_2 & -x_3 & = 0 \\ x_1 & & +3x_3 & = 0 \\ 2x_1 & +x_2 & +4x_3 & = 0 \end{cases} \iff \begin{cases} x_1 & +x_2 & +x_3 & = 0 \\ & +x_2 & -2x_3 & = 0 \\ & -x_2 & +2x_3 & = 0 \\ & -x_2 & +2x_3 & = 0 \end{cases}$$

$$\iff \begin{cases} x_1 & +x_2 & +x_3 & = 0 \\ & +x_2 & -2x_3 & = 0 \end{cases} \iff \begin{cases} x_2 & = 2x_3 \\ x_1 & = -3x_3 \end{cases} \iff x = x_3 \begin{pmatrix} -3 \\ 2 \\ 1 \end{pmatrix}$$

On a donc obtenu l'ensemble des vecteurs de $\text{Ker } A$, qui sont solutions de $Ax = 0$. On voit

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

que la dimension de $\text{Ker } A$ est 1, donc on obtient le rang de A à savoir $(3 - 1)$. En fait c'est de cette façon que très souvent on détermine le rang de A .

- $b \neq 0$ - On dit que le système est inhomogène. Cette fois il n'y a plus de solution évidente et il est possible qu'il n'existe aucune solution. La méthode consiste à utiliser la méthode d'élimination de Gauss pour résoudre $Ax = b$ et à voir s'il n'existe pas d'équations incompatibles (c'est-à-dire avec le même premier membre mais des seconds membres différents).

Regardons sur un exemple :

$$\begin{cases} x_1 + x_2 + x_3 = 1 \\ x_1 + 2x_2 - x_3 = -1 \\ x_1 + 3x_3 = 3 \\ 2x_1 + x_2 + 4x_3 = 4 \end{cases} \iff \begin{cases} x_1 + x_2 + x_3 = 1 \\ +x_2 - 2x_3 = -2 \\ -x_2 + 2x_3 = 2 \\ -x_2 + 2x_3 = 2 \end{cases}$$

$$\iff \begin{cases} x_1 + x_2 + x_3 = 1 \\ +x_2 - 2x_3 = -2 \end{cases} \iff \begin{cases} x_2 = 2x_3 - 2 \\ x_1 = -3x_3 + 3 \end{cases}$$

$$\iff x = \begin{pmatrix} -3x_3 + 3 \\ 2x_3 - 2 \\ x_3 \end{pmatrix} = x_3 \begin{pmatrix} -3 \\ 2 \\ 1 \end{pmatrix} + \begin{pmatrix} 3 \\ -2 \\ 0 \end{pmatrix}.$$

On trouve donc les solutions du système $Ax = b$ en sommant les solutions du système homogène et une solution particulière du système non homogène. Par contre le système suivant n'a pas de solution (voir l'exercice 1.38) :

$$\begin{cases} x_1 + x_2 + x_3 = 1 \\ x_1 + 2x_2 - x_3 = -1 \\ x_1 + 3x_3 = 3 \\ 2x_1 + x_2 + 4x_3 = 2 \end{cases}$$

Systèmes linéaires de dimension quelconque

Sommaire
Concepts

Exemples
Exercices
Documents

1.6 Valeurs propres

1.6.1	Valeurs propres	47
1.6.2	Valeurs propres et matrices semblables	49
1.6.3	Diagonalisation des matrices	51
1.6.4	Valeurs propres et matrices définies positives	53

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

1.6.1 Valeurs propres

Exercices :

[Exercice 1.39](#)

[Exercice 1.40](#)

Définition 1.6.1. 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

$$AY = \lambda Y. \quad (1.6.1)$$

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 1.6.2. On appelle **polynôme caractéristique** de A le polynôme

$$\Pi_A(s) = \det(sI - A).$$

Proposition 1.6.1. 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 :

$$\det(\lambda I - A) = 0. \quad (1.6.2)$$

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

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

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

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

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

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

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

Valeurs propres

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

1.6.2 Valeurs propres et matrices semblables

Exercices :

[Exercice 1.41](#)

Définition 1.6.3. Deux matrices $A \in \mathcal{M}_{n,n}$ et $B \in \mathcal{M}_{n,n}$ 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 fautive c'est-à-dire que deux matrices qui ont le même déterminant ne sont pas toujours semblables.

Proposition 1.6.3. 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 $\Pi_A = \Pi_B$.

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

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 .

Valeurs propres et matrices semblables

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

1.6.3 Diagonalisation des matrices

Exercices :

[Exercice 1.42](#)

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 1.6.4. *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, se trouvent dans le chapitre 5 de MT23 :

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Théorème 1.6.1. 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 à deux vecteurs x et y fait correspondre le réel $x^T y = \sum_{i=1}^n x_i y_i$. A 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).

Diagonalisation des matrices

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

1.6.4 Valeurs propres et matrices définies positives

Exercices :

[Exercice 1.43](#)

[Exercice 1.44](#)

Cours :

[Diagonalisation des matrices](#)

Définition 1.6.5. 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

$$x^T A x = 0 \Rightarrow x = 0.$$

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

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

La démonstration est laissée en exercice.

Proposition 1.6.5. 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).

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

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

$$AY = \lambda Y \Rightarrow Y^T AY = \lambda Y^T Y.$$

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

$$x = \sum_{i=1}^n \alpha_i Y_i \Rightarrow x^T Ax = \left(\sum_{i=1}^n \alpha_i Y_i^T \right) A \left(\sum_{j=1}^n \alpha_j Y_j \right) = \sum_{i=1}^n \alpha_i^2 \lambda_i.$$

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

$$x^T Ax = 0 \Rightarrow \alpha_i^2 = 0 \Rightarrow x = 0.$$

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

**Valeurs propres
et matrices
définies
positives**

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Documents du chapitre 1

1.1	Définition de l'espace vectoriel	56
1.2	Injectivité, surjectivité	57

Sommaire
Concepts

Exemples
Exercices
Documents

Document 1.1 Définition de l'espace vectoriel

Définition 1.6.6. *Un espace vectoriel sur K ($K = \mathbb{R}$ ou \mathbb{C}) est un ensemble E (dont on appelle vecteurs les éléments) possédant les deux lois suivantes :*

- *l'addition de vecteurs (loi interne qui donne à E une structure de groupe commutatif), c'est à dire que, pour tout x, y, z de E , :*
 - *associativité : $(x + y) + z = x + (y + z)$,*
 - *existence d'un élément neutre $e \in E : x + e = e + x = x$*
 - *tout élément $x \in E$ a un symétrique $x' \in E : x + x' = x' + x = e$,*
 - *commutativité : $x + y = y + x$.*
- *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 :*
 - $\forall \lambda, \mu \in K, \forall x, y \in E :$
 - $(\lambda\mu)x = \lambda(\mu x)$
 - $(\lambda + \mu)x = \lambda x + \mu x$
 - $\lambda(x + y) = \lambda x + \lambda y$
 - $1x = x$ (1 élément unité de K).

[retour au cours](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Document 1.2 Injectivité, surjectivité

Par définition une application $f : E \rightarrow F$ est surjective si $\text{Im } f = F$. Elle est injective si

$$\forall x \in E, \forall y \in E \ f(x) = f(y) \Leftrightarrow x = y,$$

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

Pour les applications linéaires, on a la caractérisation suivante de l'injectivité :

Proposition 1.6.6.

Une application linéaire u est injective si et seulement si $\ker u = \{\vec{0}\}$.

Démonstration - - Supposons que u est injective et soit $\vec{x} \in \ker u$, alors on a

$$u(\vec{x}) = \vec{0} \text{ et } u(\vec{0}) = \vec{0}$$

d'où $\vec{x} = \vec{0}$ par injectivité et $\ker u = \{\vec{0}\}$.

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

[retour au cours](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercices du chapitre 1

1.1	60
1.2	61
1.3	62
1.4	63
1.5	64
1.6	65
1.7	66
1.8	67
1.9	68
1.10	69
1.11	70
1.12	71
1.13	72
1.14	73
1.15	74
1.16	75
1.17	76
1.18	77
1.19	78
1.20	79
1.21	80
1.22	81

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

1.23	82
1.24	83
1.25	84
1.26	85
1.27	86
1.28	87
1.29	88
1.30	89
1.31	90
1.32	91
1.33	92
1.34	93
1.35	94
1.36	95
1.37	96
1.38	97
1.39	98
1.40	99
1.41	100
1.42	101
1.43	102
1.44	103

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 1.1

Montrer que la somme de vecteurs et le produit d'un vecteur par un nombre réel donnent à \mathbb{R}^3 une structure d'espace vectoriel sur \mathbb{R} .

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 1.2

Montrer que la somme de polynômes et le produit d'un polynôme par un nombre réel donnent à P_n (polynômes de degré inférieur ou égal à n à coefficients réels) une structure d'espace vectoriel sur \mathbb{R} .

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 1.3

Montrer que l'ensemble des polynômes de degré exactement égal à n n'est pas un espace vectoriel.

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 1.4

- Montrer que si \vec{x} est un vecteur de \mathbb{R}^2 , alors $F = \{\lambda\vec{x}, \lambda \in \mathbb{R}\}$ est un sous-espace vectoriel de \mathbb{R}^2 .
- Montrer que si \vec{x} et \vec{y} sont deux vecteurs de \mathbb{R}^3 , alors $F = \{\lambda\vec{x} + \mu\vec{y}, \lambda, \mu \in \mathbb{R}\}$ est un sous-espace vectoriel de \mathbb{R}^3 .
- Montrer que l'espace des polynômes P_k est un sous-espace vectoriel de P_n si $k \leq n$.

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 1.5

Montrer que si F et G sont deux sous-espaces vectoriels de E alors $F \cap G$ est un sous-espace vectoriel de E alors que $F \cup G$ ne l'est pas (en général).

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 1.6

Soient F et G deux sous-espaces vectoriels de E et soit

$$H = \{\vec{x} = \vec{y} + \vec{z}, \vec{y} \in F, \vec{z} \in G\}.$$

On note alors $H = F + G$ et on appelle H la somme de F et G .

- Représenter graphiquement un élément de H lorsque $E = \mathbb{R}^3$, F est un plan vectoriel et G une droite vectorielle (non contenue dans le plan).
- Montrer que H est un sous-espace vectoriel de E .
- Que vaut H dans le cas particulier de la première question ? Que vaut $F \cup G$ dans ce même cas ?
- Montrer que si $F \cap G = \{\vec{0}\}$, \vec{y} et \vec{z} sont uniques pour un \vec{x} donné. On dit alors que $F + G$ est une somme directe et on note $F \oplus G$.

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 1.7

- Donner une famille liée de vecteurs de \mathbb{R}^3 (donner les vecteurs par leurs composantes).
- Montrer que la famille $\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$ de vecteurs de \mathbb{R}^3 donnés par leurs composantes est libre. Montrer qu'elle est aussi génératrice.
- En vous inspirant de la question précédente, donner une famille libre et génératrice de \mathbb{R}^n .

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 1.8

La famille $S = \{(1, 1, -1), (1, 1, 1), (\sqrt{2}, \sqrt{2}, 3)\}$ est-elle liée ?

[retour au cours](#)

Solution

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 1.9

Montrer que la famille $\{1, x, x^2, \dots, x^n\}$ de polynômes de P_n est une base de P_n .

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 1.10

Soit $B = \{\vec{e}_1, \dots, \vec{e}_n\}$ une base de l'espace vectoriel E et soit $\vec{x} \in E$. En raisonnant par l'absurde, montrer que la décomposition de \vec{x} sur B est unique, c'est-à-dire qu'il existe $\lambda_1, \lambda_2, \dots, \lambda_n$ uniques tels que

$$\vec{x} = \lambda_1 \vec{e}_1 + \dots + \lambda_n \vec{e}_n.$$

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 1.11

À partir des vecteurs $\vec{x} = (1, 1, -1)$ et $\vec{y} = (1, 1, 1)$ de \mathbb{R}^3 , trouver un vecteur \vec{z} tel que la famille $\{\vec{x}, \vec{y}, \vec{z}\}$ soit une base de \mathbb{R}^3 .

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 1.12

Vérifier rapidement que les applications suivantes sont linéaires. Calculer leur noyau et leur image.

1. $u_1 : \mathbb{R}^n \rightarrow \mathbb{R}$, définie par

$$u_1(x) = a_1x_1 + a_2x_2 + \dots + a_nx_n$$

où a_1, a_2, \dots, a_n sont des réels donnés.

2. $u_2 : P_k \rightarrow P_{k-1}$, définie par

$$u_2(p) = p',$$

où p' est la dérivée du polynôme p .

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 1.13

Soit F et G deux sous-espaces vectoriels de E tels que $E = F + G$ et $F \cap G = \{\vec{0}\}$ (voir l'exercice 1.6). On a montré que si $\vec{x} \in E$, alors il existe deux vecteurs uniques $\vec{y} \in F$ et $\vec{z} \in G$ tels que $\vec{x} = \vec{y} + \vec{z}$. Soit l'application

$$u : E \rightarrow F, \text{ telle que } u(\vec{x}) = \vec{y}.$$

1. Montrer que u est une application linéaire.
2. Calculer le noyau et l'image de u .
3. Donner leur dimension et vérifier le résultat

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

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 1.14

On note u la rotation d'un angle θ dans \mathbb{R}^2 , utiliser les propriétés géométriques pour traiter l'exercice :

1. Montrer que l'application u est linéaire.
2. Quel est son noyau, quelle est son image ?
3. Montrer qu'elle est bijective.
4. Donner l'application inverse et en déduire qu'elle est linéaire et bijective.

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 1.15

Montrer que la matrice de l'application $i_E : E \rightarrow E$ est la matrice identité I lorsque l'on munit E de la même base.

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 1.16

On suppose que $E = F = P_2$, on munit P_2 de la base canonique $\{1, x, x^2\}$ et on définit u telle que $u(p) = p'$. Déterminer alors la matrice de u .

[retour au cours](#)

Solution

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 1.17

Calculer le produit AB (et BA lorsque cela est possible) dans les cas suivants :

1. $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$

2. $A = \begin{pmatrix} 1 & 2 & -1 \\ 0 & 1 & 1 \\ 1 & -1 & 0 \end{pmatrix}, B = \begin{pmatrix} 1 & 0 \\ -1 & 1 \\ 1 & -1 \end{pmatrix}.$

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 1.18

Montrer, en utilisant le produit de matrices, que la composée de deux rotations planes d'angles θ_1 et θ_2 est une rotation plane d'angle $\theta_1 + \theta_2$.

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 1.19

Montrer la proposition 1.3.2. Pour cela on calculera la décomposition de $u(\vec{x})$ dans la base de F et on la comparera à la décomposition de \vec{y} dans la même base.

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 1.20

Soit la matrice $A = \begin{pmatrix} 3 & 4 & 5 \\ 1 & 2 & 6 \end{pmatrix}$, u est l'application linéaire de \mathbb{R}^3 dans \mathbb{R}^2 dont la matrice est A lorsque l'on munit \mathbb{R}^3 et \mathbb{R}^2 de leur base canonique. Calculer $u(\vec{x})$ pour $\vec{x} = (1, -1, 2)$.

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 1.21

On a montré, dans l'exercice 1.14 que la rotation plane d'angle θ est bijective. Donner la matrice inverse de la matrice de cette rotation.

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 1.22

Montrer que la matrice inverse d'une matrice, si elle existe, est unique.

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 1.23

Soient X et Y deux vecteurs colonnes, les produits $X^T Y$ et XY^T existent-ils ? et, si oui, sont-ils égaux ?

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 1.24

Soit E un espace vectoriel muni d'une base $\mathcal{E} = \{\vec{e}_1, \vec{e}_2\}$. On définit les vecteurs $\vec{e}'_1 = \vec{e}_1 + \vec{e}_2$, $\vec{e}'_2 = 2\vec{e}_1 - \vec{e}_2$ d'une nouvelle base $\mathcal{E}' = \{\vec{e}'_1, \vec{e}'_2\}$. Donner P , matrice de passage de \mathcal{E} dans \mathcal{E}' .

[retour au cours](#)

Solution

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 1.25

On reprend les données de l'exercice 1.24. On définit $u \in \mathcal{L}(E; E)$ par $u(\vec{e}_1) = \vec{e}_1 + 3\vec{e}_2, u(\vec{e}_2) = 2\vec{e}_1 - \vec{e}_2$

- Quelle est la matrice A de u dans la base \mathcal{E} ?
- Exprimer $u(\vec{e}'_1), u(\vec{e}'_2)$ en fonction de \vec{e}_1 et \vec{e}_2 .
- En déduire $u(\vec{e}'_1), u(\vec{e}'_2)$ en fonction de \vec{e}'_1 et \vec{e}'_2 .
- En déduire A' .
- Calculer $P^{-1}AP$.

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 1.26

Déterminer le rang des matrices A suivantes :

– A est la matrice de la rotation dans le plan.

$$- A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 3 & 1 & 1 \end{pmatrix}.$$

$$- A = \begin{pmatrix} 1 & 3 & 5 & 1 \\ 2 & 2 & 6 & -2 \\ 1 & 2 & 4 & 0 \end{pmatrix}.$$

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 1.27

Calculer les déterminants suivants :

$$\begin{vmatrix} a & c \\ b & d \end{vmatrix}, \begin{vmatrix} 3a & c \\ 3b & d \end{vmatrix}, \begin{vmatrix} 4 & 2 & 3 \\ 0 & 3 & 4 \\ 0 & 0 & 5 \end{vmatrix}, \begin{vmatrix} 4 & 2 & 3 \\ 0 & 1 & 2 \\ 4 & 1 & 2 \end{vmatrix}, \begin{vmatrix} 4 & 3 & 2 \\ 0 & 2 & 1 \\ 4 & 2 & 1 \end{vmatrix}, \begin{vmatrix} 1 & 2 & 2 \\ 3 & 1 & 1 \\ 4 & 1 & 1 \end{vmatrix}, \begin{vmatrix} 4 & 2 & 3\lambda \\ 0 & 1 & 2\lambda \\ 4 & 1 & 2\lambda \end{vmatrix}.$$

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 1.28

Calculer les déterminants suivants :

$$\begin{vmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{vmatrix}, \begin{vmatrix} \lambda \cos\theta & -\sin\theta \\ \lambda \sin\theta & \cos\theta \end{vmatrix}, \begin{vmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{vmatrix}.$$

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 1.29

1. Considérons une matrice $A \in \mathcal{M}_{n,n}$ triangulaire inférieure ($a_{ij} = 0$ pour $i < j$), en utilisant la définition du déterminant montrer que $\det A = \prod_{i=1}^n a_{ii}$.
2. En déduire que :
 - pour les matrices diagonales ($a_{ij} = 0$ pour $i \neq j$) on a aussi $\det A = \prod_{i=1}^n a_{ii}$,
 - la matrice identité a pour déterminant $\det I = 1$.

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 1.30

Montrer par un exemple que, en général, $\det(A + B) \neq \det(A) + \det(B)$.

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 1.31

$$\text{Calculer } \begin{vmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 2 & 2 \\ 2 & 1 & 1 & 1 \end{vmatrix}.$$

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 1.32

Montrer que le déterminant d'une matrice triangulaire supérieure est égal au produit de ses termes diagonaux.

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 1.33

Montrer que deux matrices semblables ont même déterminant.

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 1.34

Calculer à nouveau le déterminant de la matrice de l'exercice 1.31, en développant par rapport à une ligne ou une colonne de votre choix.

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 1.35

Quel est le rang de la matrice

$$A = \begin{pmatrix} 1 & 3 & 5 & 1 \\ 2 & 2 & 6 & -2 \\ 1 & 2 & 4 & 0 \end{pmatrix}?$$

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 1.36

Calculer l'inverse de $A = \begin{pmatrix} 2 & 1 & 2 \\ 1 & 2 & 3 \\ 3 & 1 & -2 \end{pmatrix}$ en utilisant les co-facteurs.

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 1.37

Résoudre le système linéaire $Ax = 0$ où

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

En déduire le rang de la matrice A .

Un système $Ax = 0$, dont la matrice $A \in \mathcal{M}_{n,p}$ est telle que $n > p$, a-t-il toujours une solution ? si oui est-elle unique ?

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 1.38

Résoudre le système linéaire

$$\begin{cases} x_1 + x_2 + x_3 = 1 \\ x_1 + 2x_2 - x_3 = -1 \\ x_1 + 3x_3 = 3 \\ 2x_1 + x_2 + 4x_3 = 2 \end{cases}$$

Ce système a-t-il une solution ?

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 1.39

Calculer les valeurs propres et les vecteurs propres de la matrice $A = \begin{pmatrix} 5 & 2 \\ 2 & 2 \end{pmatrix}$.

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 1.40

Montrer que :

- 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.

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 1.41

Montrer que deux matrices semblables ont le même déterminant et que la réciproque est fausse.

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 1.42

Montrer que si $A = PDP^{-1}$ où la matrice D est diagonale, alors les colonnes de P sont vecteurs propres de A , les valeurs propres étant les éléments de la diagonale de D .

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 1.43

Soit $B \in \mathcal{M}_{mn}(\mathbb{R})$, montrer que $B^T B$ est symétrique et semi-définie positive.

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 1.44

Soit $B \in \mathcal{M}_{mn}(\mathbb{R})$, avec $n \leq m$ et $\text{rang } B = n$, montrer que $B^T B$ est symétrique et définie positive.

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Chapitre 2

Résolution des systèmes linéaires

2.1	Motivations	105
2.2	L'élimination de Gauss	109
2.3	Calcul direct de la factorisation $A = LU$	127
2.4	Factorisation LU avec recherche de pivots	134
2.5	Traitement des matrices symétriques	152
	Exemples du chapitre 2	162
	Exercices du chapitre 2	167

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

2.1 Motivations

2.1.1 Equation de la chaleur 106

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

2.1.1 Equation de la chaleur

Exercices :

[Exercice 2.1](#)

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 :

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

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 $k = 0 \dots N$. Pour simplifier, on choisit des points x_k équidistants, on a donc $h = \frac{1}{N}$, et $x_k = kh$.

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Aux points x_k , l'équation se récrit :

$$\begin{cases} -\frac{d^2 u}{dx^2}(x_k) = f(x_k), & k = 1, \dots, N-1, \\ u(x_0) = 0, \\ u(x_N) = 0. \end{cases} \quad (2.1.1)$$

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 :

$$\frac{d^2 u}{dx^2}(x_k) = \frac{u(x_{k-1}) - 2u(x_k) + u(x_{k+1}))}{h^2} + \mathcal{O}(h^2) \quad (2.1.2)$$

Si l'on se donne le droit de négliger le reste $\mathcal{O}(h^2)$ (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 :

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

qui peuvent se mettre sous la forme matricielle suivante :

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

Equation de la chaleur

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

et $v_0 = v_N = 0$.

Voilà donc un exemple de 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 "raisonnables" de résolution des systèmes linéaires, raisonnables dans le sens où ces méthodes peuvent être facilement implantées sur ordinateur.

Equation de la chaleur

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

2.2 L'élimination de Gauss

2.2.1	Résolution d'un système triangulaire	110
2.2.2	Principe de la méthode de Gauss et première étape	112
2.2.3	Deuxième étape	115
2.2.4	Étape générale	117
2.2.5	Algorithme d'élimination de Gauss	119
2.2.6	Écriture matricielle de l'élimination de Gauss	122
2.2.7	Unicité de la factorisation LU	125

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

2.2.1 Résolution d'un système triangulaire

Exercices :

[Exercice 2.2](#)

[Exercice 2.3](#)

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

$$Ax = b, \quad (2.2.1)$$

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

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

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

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

puis

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

On peut donc écrire la formule générale :

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

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

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

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

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

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

Résolution d'un système triangulaire

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

2.2.2 Principe de la méthode de Gauss et première étape

Exercices :

[Exercice 2.4](#)

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

$$\hat{A}x = \hat{b} \quad (2.2.2)$$

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

Attention! La matrice \hat{A} n'est pas semblable à A . Autrement dit, il ne faut pas croire que les matrices A et \hat{A} ont les mêmes valeurs propres. C'est dommage, puisque les valeurs propres de \hat{A} sont ses termes diagonaux. À ce sujet, il est à noter qu'il n'est pas possible de calculer en un nombre fini d'étapes les valeurs propres d'une matrice de rang supérieur à 5.

Le système (2.2.2) est simplement équivalent au système $Ax = b$ d'où nous sommes partis et la méthode consiste à éliminer successivement des inconnues entre les équations.

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

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

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

On peut alors éliminer l'inconnue x_1 entre les deux premières équations 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

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

et parallèlement

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

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

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

**Principe de la
méthode de
Gauss et
première étape**

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

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

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

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

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

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.

Principe de la méthode de Gauss et première étape

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

2.2.3 Deuxième étape

Exercices :

[Exercice 2.5](#)

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 :

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

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

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

Ce calcul étant fait jusqu'à la ligne n , on obtient un système sous la forme

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

$A^{(3)}x = b^{(3)}$ avec

$$A^{(3)} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1j} & \dots & a_{1n} \\ 0 & a_{22}^{(2)} & a_{23}^{(2)} & \dots & a_{2j}^{(2)} & \dots & a_{2n}^{(2)} \\ 0 & 0 & a_{33}^{(3)} & \dots & a_{3j}^{(3)} & \dots & a_{3n}^{(3)} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & a_{n3}^{(3)} & \dots & a_{nj}^{(3)} & \dots & a_{nn}^{(3)} \end{pmatrix}, \quad b^{(3)} = \begin{pmatrix} b_1 \\ b_2^{(2)} \\ b_3^{(3)} \\ \vdots \\ b_n^{(3)} \end{pmatrix}$$

Deuxième étape

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

2.2.4 Étape générale

Exercices :

[Exercice 2.6](#)

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

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

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 $i = k + 1, k + 2, \dots, n$,

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

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)

[Documents](#)

- 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 $i \leq k$.
- 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

$$\det A = \det A^{(k)}, \text{ pour } k = 1, \dots, n.$$

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

Étape générale

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

2.2.5 Algorithme d'élimination de Gauss

Exercices :

[Exercice 2.7](#)[Exercice 2.8](#)

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 2.2.1. 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.

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 $k+1, k+2, \dots, n$.

[Sommaire](#)
[Concepts](#)[Exemples](#)
[Exercices](#)
[Documents](#)

- 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 :

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \quad (2.2.3)$$

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.

Algorithme d'élimination de GAUSS

```

1: pour  $k = 1$  jusqu'à  $n - 1$  faire
2:   si  $|a_{kk}| < \varepsilon$  alors
3:     Arrêter l'algorithme et donner un message d'erreur
4:   sinon
5:     pour  $i = k + 1$  jusqu'à  $n$  faire
6:        $c \leftarrow a_{ik} / a_{kk}$ 
7:        $b_i \leftarrow b_i - cb_k$ 
8:        $a_{ik} \leftarrow 0$ 
9:     pour  $j = k + 1$  jusqu'à  $n$  faire
10:       $a_{ij} \leftarrow a_{ij} - ca_{kj}$ 
11:     fin pour
12:   fin pour
13: fin si

```

Algorithme d'élimination de Gauss

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

14: **fin pour**
15: **si** $|a_{nn}| < \varepsilon$ **alors**
16: Arrêter l'algorithme et donner un message d'erreur
17: **fin si**

Algorithme d'élimination de Gauss

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

2.2.6 Écriture matricielle de l'élimination de Gauss

Exercices :

[Exercice 2.9](#)

L'élimination de Gauss consiste à construire une suite de matrices $A^{(1)}, A^{(2)}, \dots, A^{(n)}$. La matrice du système que l'on résout est la matrice triangulaire $A^{(n)}$. 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 : \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 :

$$\underline{A}_i^{(i)} = \underline{A}_i^{(i+1)} = \dots = \underline{A}_i^{(n)} \quad (2.2.4)$$

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

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

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

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

....

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

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

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 2.2.4, 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 :

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

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

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

où M est la matrice définie par :

$$\begin{pmatrix} 1 & 0 & \dots & \dots & \dots & 0 \\ m_{21} & 1 & 0 & \dots & \dots & 0 \\ m_{31} & m_{32} & 1 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ m_{n1} & m_{n2} & \dots & \dots & m_{nn-1} & 1 \end{pmatrix}$$

Écriture matricielle de l'élimination de Gauss

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

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 $A = LU$.

On pourrait montrer de manière similaire que $A = L^{(k-1)} A^{(k)}$, avec

$$L^{(k-1)} = \begin{pmatrix} 1 & 0 & \dots & \dots & \dots & \dots & 0 \\ m_{21} & 1 & 0 & \dots & \dots & \dots & 0 \\ \dots & \dots & 1 & 0 & \dots & \dots & 0 \\ m_{k1} & \dots & m_{kk-1} & 1 & 0 & \dots & 0 \\ m_{k+11} & \dots & m_{k+1k-1} & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ m_{n1} & \dots & m_{k+1k-1} & 0 & \dots & 0 & 1 \end{pmatrix}$$

Écriture matricielle de l'élimination de Gauss

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

2.2.7 Unicité de la factorisation LU

Exercices :
[Exercice 2.10](#)

Cours :
[Élimination de Gauss / factorisation](#)
 $A = LU$

Dans le paragraphe référencé, 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 2.2.2. *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*

$$A = LU. \quad (2.2.5)$$

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 :

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

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

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

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

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

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

or $\tilde{L}^{-1}L$ est une matrice triangulaire inférieure qui a ses éléments diagonaux égaux à 1, et $\tilde{U}U^{-1}$ 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 $\tilde{L}^{-1}L$), d'où

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

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

Dans cette démonstration on a utilisé des résultats sur les matrices triangulaires qui ont été démontrés en exercice dans le chapitre de révision d'algèbre linéaire. 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.

Unicité de la factorisation LU

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

2.3 Calcul direct de la factorisation $A = LU$

2.3.1	Motivations de la factorisation LU	128
2.3.2	Principe du calcul direct de la factorisation LU	130
2.3.3	Algorithme de Doolittle	132

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

2.3.1 Motivations de la factorisation LU

Exercices :
[Exercice 2.11](#)

Cours :
[Élimination de Gauss / factorisation](#)
 $A = LU$

– *Résolution des systèmes linéaires*

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

$$Ax = b. \quad (2.3.1)$$

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

$$LUX = b. \quad (2.3.2)$$

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

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

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

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 le 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.

Motivations de la factorisation LU

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

2.3.2 Principe du calcul direct de la factorisation LU

Exercices :

[Exercice 2.12](#)

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 ($l_{ii} = 1, i = 1, \dots, n$).

Traitons un exemple, soit la matrice $A = \begin{pmatrix} 2 & 1 & -2 \\ 4 & 5 & -3 \\ -2 & 5 & 3 \end{pmatrix}$,

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 :

$$LU = \begin{pmatrix} 1 & 0 & 0 \\ \times & 1 & 0 \\ \times & \times & 1 \end{pmatrix} \begin{pmatrix} \mathbf{2} & \mathbf{1} & \mathbf{-2} \\ 0 & \times & \times \\ 0 & 0 & \times \end{pmatrix} = \begin{pmatrix} \mathbf{2} & \mathbf{1} & \mathbf{-2} \\ 4 & 5 & -3 \\ -2 & 5 & 3 \end{pmatrix}.$$

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

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

$$LU = \begin{pmatrix} 1 & 0 & 0 \\ \mathbf{2} & 1 & 0 \\ -1 & \times & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 & -2 \\ 0 & \times & \times \\ 0 & 0 & \times \end{pmatrix} = \begin{pmatrix} 2 & 1 & -2 \\ \mathbf{4} & 5 & -3 \\ -2 & 5 & 3 \end{pmatrix}.$$

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

$$LU = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & \times & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 & -2 \\ 0 & \mathbf{3} & \mathbf{1} \\ 0 & 0 & \times \end{pmatrix} = \begin{pmatrix} 2 & 1 & -2 \\ 4 & \mathbf{5} & -\mathbf{3} \\ -2 & 5 & 3 \end{pmatrix}.$$

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

$$LU = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & \mathbf{2} & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 & -2 \\ 0 & 3 & 1 \\ 0 & 0 & \times \end{pmatrix} = \begin{pmatrix} 2 & 1 & -2 \\ 4 & 5 & -3 \\ -2 & \mathbf{5} & 3 \end{pmatrix}.$$

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

$$LU = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & 2 & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 & -2 \\ 0 & 3 & 1 \\ 0 & 0 & -\mathbf{1} \end{pmatrix} = \begin{pmatrix} 2 & 1 & -2 \\ 4 & 5 & -3 \\ -2 & 5 & \mathbf{3} \end{pmatrix}.$$

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, vue en exercice, conduit à l'algorithme de Crout.

Principe du calcul direct de la factorisation LU

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

2.3.3 Algorithme de Doolittle

Exercices :

[Exercice 2.13](#)

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

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

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 :

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

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 .

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

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. Cet algorithme est décrit ci-dessous, si besoin est, 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.

Algorithme de DOOLITTLE

```
1: pour  $i = 1$  jusqu'à  $n - 1$  faire
2:   pour  $j = i$  jusqu'à  $n$  faire
3:      $u_{ij} \leftarrow a_{ij} - \sum_{k=1}^{i-1} l_{ik}u_{kj}$ 
4:   fin pour
5:   pour  $j = i + 1$  jusqu'à  $n$  faire
6:      $l_{ji} \leftarrow \frac{1}{u_{ii}} (a_{ji} - \sum_{k=1}^{i-1} l_{jk}u_{ki})$ 
7:   fin pour
8: fin pour
9:  $u_{nn} \leftarrow a_{nn} - \sum_{k=1}^{n-1} l_{nk}u_{kn}$ 
```

Algorithme de Doolittle

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

2.4 Factorisation LU avec recherche de pivots

2.4.1	Faisabilité de la factorisation $A = LU$	135
2.4.2	Pivot nul. Permutation de lignes	138
2.4.3	Choix du pivot. Pivot maximal	140
2.4.4	La factorisation $PA = LU$	142
2.4.5	Algorithme de la factorisation $PA = LU$	146
2.4.6	Application à la résolution d'un système linéaire	148
2.4.7	La méthode du pivot total	150

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

2.4.1 Faisabilité de la factorisation $A = LU$

Exercices :

[Exercice 2.14](#)

Tout ce que nous avons fait jusqu'ici suppose qu'à aucun moment on ne rencontre de pivot nul. Nous allons voir ici qu'il existe une condition simple pour garantir la non-nullité de ces pivots. Voici tout d'abord une définition qui va nous servir par la suite :

Définition 2.4.1. On appelle sous-matrice principale d'ordre k de la matrice A et on note $[A]_k$ la matrice

$$[A]_k = (a_{ij})_{1 \leq i \leq k, 1 \leq j \leq k}.$$

On a le théorème suivant :

Théorème 2.4.2. Les trois propositions suivantes sont équivalentes :

– Proposition 1 :

Tous les pivots $a_{kk}^{(k)}$ sont définis et $a_{kk}^{(k)} \neq 0$, $k = 1, \dots, n$.

– Proposition 2 :

A admet une factorisation $A = LU$ avec U inversible.

– Proposition 3 :

$[A]_1, [A]_2, \dots, [A]_n$ inversibles.

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Démonstration : Remarquons tout d'abord que l'équivalence entre les deux premières propositions est évidente puisque la factorisation est possible lorsque les pivots sont non nuls et que d'autre part les pivots $a_{kk}^{(k)}$ sont les termes diagonaux de U .

Montrons que la proposition 3 implique la proposition 1.

$a_{11}^{(1)} = a_{11}$ est toujours défini, de plus $[A]_1$ inversible, donc

$$\det [A]_1 = a_{11} = a_{11}^{(1)} \neq 0.$$

Supposons maintenant que les pivots $a_{ii}^{(i)}$ ont été définis pour $i = 1, \dots, k-1$ et qu'ils sont non nuls : $a_{ii}^{(i)} \neq 0$ pour $i = 1, \dots, k-1$. Il est donc possible de définir $a_{kk}^{(k)}$, montrons que $a_{kk}^{(k)} \neq 0$.

Nous avons remarqué que le déterminant de A n'est pas modifié par chaque étape de l'élimination de Gauss et, plus précisément, les déterminants des sous-matrices principales puisque l'on ne fait aucune permutation de lignes ou de colonnes. On a donc

$$\det ([A]_k) = \det ([A^{(k)}]_k).$$

Or la sous matrice principale $[A^{(k)}]_k$ est triangulaire supérieure donc

$$\det ([A]_k) = \det ([A^{(k)}]_k) = \left(\prod_{i=1}^{k-1} a_{ii}^{(i)} \right) a_{kk}^{(k)},$$

et, puisque $[A]_k$ est inversible, son déterminant est non nul, on en déduit que $a_{kk}^{(k)} \neq 0$. On a donc montré par récurrence que la proposition 3 impliquait la proposition 1.

On peut montrer que la proposition 2 implique la proposition 3 :

Faisabilité de la factorisation

$$A = LU$$

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)

[Documents](#)

$A = LU \Rightarrow [A]_k = [L]_k [U]_k$: démontrez ce résultat. Donc $\det [A]_k = 1 \times \det [U]_k$. Or U est triangulaire supérieure inversible, donc les termes diagonaux de U sont non nuls, donc $[U]_k$ est elle aussi une matrice triangulaire supérieure inversible, donc $\det [A]_k \neq 0$, donc les sous matrices principales $[A]_k$ sont inversibles. \square

Faisabilité de la factorisation

$$A = LU$$

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

2.4.2 Pivot nul. Permutation de lignes

On suppose toujours que la matrice A est inversible, mais cette fois la condition énoncée dans le théorème 2.4.2 n'est plus vérifiée, supposons que l'on a :

$$a_{11}^{(1)} \neq 0, a_{22}^{(2)} \neq 0, \dots, a_{k-1, k-1}^{(k-1)} \neq 0, a_{kk}^{(k)} = 0.$$

On a donc pu construire la matrice $A^{(k)}$ comme précédemment :

$$A^{(k)} = \begin{pmatrix} a_{11}^{(1)} & \times & \dots & \times & \dots & \times \\ 0 & a_{22}^{(2)} & \dots & \times & \dots & \times \\ \vdots & \ddots & \ddots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & a_{kk}^{(k)} & \dots & a_{kn}^{(k)} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & a_{nk}^{(k)} & \dots & a_{nn}^{(k)} \end{pmatrix}$$

Par contre puisque $a_{kk}^{(k)} = 0$, il n'est pas possible de calculer $A^{(k+1)}$ comme précédemment. Existe-t-il un coefficient $a_{ik}^{(k)}$, $i = k+1, \dots, n$ non nul ? Si c'est le cas, comme l'ordre des équations composant un système d'équations linéaires ne joue aucun rôle, nous pouvons permuter la $k^{\text{ème}}$ ligne avec l'une des lignes suivantes, il faut bien sûr permuter les seconds membres en parallèle. On obtiendra un nouveau coefficient $a_{kk}^{(k)}$ non nul et on pourra poursuivre.

Existe-t-il un coefficient $a_{ik}^{(k)}$, $i > k$ non nul ? La réponse est oui, montrons-le.

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Si l'on pose

$$S_k = \begin{pmatrix} a_{k,k}^{(k)} & \cdots & a_{k,n}^{(k)} \\ \vdots & \vdots & \vdots \\ a_{n,k}^{(k)} & \cdots & a_{nn}^{(k)} \end{pmatrix},$$

on montre aisément que

$$\det A^{(k)} = a_{11}^{(1)} a_{22}^{(2)} \cdots a_{k-1,k-1}^{(k-1)} \det S_k.$$

Or $\det A^{(k)} = \det A \neq 0$, donc $\det S_k \neq 0$, donc la première colonne de S_k n'est pas nulle, donc il existe un coefficient $a_{ik}^{(k)}$, $i > k$ non nul. Il est donc possible de continuer l'élimination de Gauss à condition d'échanger l'ordre des équations.

Proposition 2.4.1. *À toute étape k de la méthode d'élimination de Gauss il existe au moins un élément non nul dans la colonne k et situé sur la ligne i avec $i \geq k$.*

**Pivot nul.
Permutation de
lignes**

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

2.4.3 Choix du pivot. Pivot maximal

Exemples :

[Exemple 2.1](#)

Exercices :

[Exercice 2.15](#)

En théorie, ce n'est que lorsqu'un pivot est nul que l'on est obligé de permuter les équations donc les lignes de A . Pratiquement, il y a des cas où le pivot n'est pas nul, mais où l'échange des lignes est souhaitable à cause des erreurs numériques, voir l'exemple référencé. On peut démontrer que l'on a intérêt, au moment de l'élimination de chacune des inconnues x_k , à rechercher puis choisir, le coefficient $a_{ik}^{(k)}$ le plus grand en module pour $k \leq i \leq n$, puis l'ayant trouvé de permuter les lignes correspondantes. En outre, l'expérience numérique confirme ce résultat mathématique.

En pratique donc, on fera toujours (même si $a_{kk}^{(k)}$ est non nul) des échanges de lignes de façon à avoir le pivot le plus grand en valeur absolue : à l'étape k de l'élimination de Gauss, on recherchera dans la colonne k , pour les indices de ligne i variant de k à n , l'élément $a_{ik}^{(k)}$ tel que

$$\forall i \in \{k, k+1, \dots, n\}, |a_{ik}^{(k)}| \leq |a_{lk}^{(k)}|$$

et on échangera les lignes l et k . Une étape de l'algorithme d'élimination de Gauss avec permutation de lignes peut donc se représenter par

$$A^{(k+1)} = M^{(k)} P^{(k)} A^k,$$

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

où $P^{(k)}$ est la matrice de permutation permettant de permuter les lignes l et k où $l \geq k$. Comme on peut le montrer dans l'exercice référencé, cette matrice a tous ses éléments nuls sauf

$$\begin{cases} p_{ii} = 1, & i \neq k, i \neq l \\ p_{kl} = 1, \\ p_{lk} = 1. \end{cases}$$

Cette matrice étant une matrice de permutation symétrique elle possède la propriété

$$P^{(k)} = P^{(k)T} = P^{(k)-1}.$$

Choix du pivot.
Pivot maximal

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

2.4.4 La factorisation $PA = LU$

Exemples :

[Exemple 2.2](#)

Vous pouvez commencer par étudier l'exemple référencé. Dans le cas général :

- – On commence par une recherche de pivot maximal dans la première colonne de A , si ce pivot se trouve en ligne numéro ℓ , on échange les lignes 1 et ℓ de A , c'est-à-dire si $P^{(1)}$ est la matrice de permutation qui échange la ligne 1 et la ligne ℓ , on calcule

$$\hat{A}^{(1)} = P^{(1)} A.$$

- On effectue la première étape de la méthode de Gauss sur la matrice $\hat{A}^{(1)}$, c'est-à-dire on détermine $L^{(1)}$ et $A^{(2)}$ telles que

$$P^{(1)} A = L^{(1)} A^{(2)}.$$

- – On effectue une recherche de pivot maximal dans la deuxième colonne de $A^{(2)}$, si ce pivot se trouve en ligne numéro ℓ , on échange les lignes 2 et ℓ de $A^{(2)}$, c'est-à-dire on calcule $P^{(2)} A^{(2)}$ où $P^{(2)}$ est la matrice de permutation qui échange les lignes 2 et ℓ .

On peut alors remarquer que $P^{(2)} P^{(1)} A = P^{(2)} L^{(1)} P^{(2)} P^{(1)} A^{(2)}$, en appliquant la proposition 2.4.2 $P^{(2)} L^{(1)} P^{(2)} = \hat{L}^{(1)}$, on peut poser $\hat{A}^{(2)} = P^{(2)} A^{(2)}$, on a donc

$$P^{(2)} P^{(1)} A = \hat{L}^{(1)} \hat{A}^{(2)}.$$

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)

[Documents](#)

- On effectue la deuxième étape de la méthode de Gauss sur la matrice $\hat{A}^{(2)}$, c'est-à-dire on détermine $L^{(2)}$ et $A^{(3)}$ telles que

$$P^{(2)} P^{(1)} A = L^{(2)} A^{(3)}.$$

- De façon générale lorsque l'on a obtenu

$$P^{(k-1)} \dots P^{(1)} A = L^{(k-1)} A^{(k)},$$

- on pivote sur la colonne k de $A^{(k)}$, on obtient

$$P^{(k)} P^{(k-1)} \dots P^{(1)} A = \hat{L}^{(k-1)} \hat{A}^{(k)}$$

- on effectue la k ème étape de la méthode de Gauss, on obtient

$$P^{(k)} P^{(k-1)} \dots P^{(1)} A = L^{(k)} A^{(k+1)}.$$

- On obtient après $n - 1$ étapes :

$$P^{(n-1)} P^{(n-2)} \dots P^{(1)} A = L^{(n-1)} A^{(n)},$$

si l'on pose $P = P^{(n-1)} P^{(n-2)} \dots P^{(1)}$, $L = L^{(n-1)}$, $U = A^{(n)}$, on a obtenu la factorisation recherchée $PA = LU$.

Théorème 2.4.3. *Soit A une matrice inversible. Alors il existe une matrice triangulaire inférieure L , une matrice triangulaire supérieure U , et une matrice de permutation P telles que*

$$PA = LU. \tag{2.4.1}$$

La factorisation

$$PA = LU$$

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)

[Documents](#)

Remarquons qu'il n'y a pas unicité de la décomposition $PA = LU$.

Nous avons utilisé la proposition suivante :

Proposition 2.4.2. Soit $L^{(k-1)}$ une matrice triangulaire inférieure à diagonale unité vérifiant de plus

$$[L^{(k-1)}]_{ij} = 0, \quad i > j > k - 1,$$

et soit $P^{(k)}$ la matrice de permutation de l'étape k de l'algorithme de Gauss avec recherche de pivot, permutant les lignes k et l où $l \geq k$. Alors on a

$$P^{(k)} L^{(k-1)} P^{(k)} = \hat{L}^{(k-1)},$$

où $\hat{L}^{(k-1)}$ est une matrice triangulaire inférieure obtenue à partir de $L^{(k-1)}$ en permutant les termes des lignes k et l dans les colonnes 1 à $k - 1$.

Démonstration : On peut écrire $L^{(k-1)}$ sous la forme

$$L^{(k-1)} = I + \sum_{p=1}^{k-1} z^{(p)} (e^p)^T,$$

où $z_i^{(p)} = 0$ pour $i \leq p$. On a ensuite

$$\begin{aligned} P^{(k)} L^{(k-1)} P^{(k)} &= I + \sum_{p=1}^{k-1} P^{(k)} z^{(p)} (e^p)^T P^{(k)}, \\ &= I + \sum_{p=1}^{k-1} P^{(k)} z^{(p)} (e^p)^T, \end{aligned}$$

La factorisation

$$PA = LU$$

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

car $P^{(k)}P^{(k)} = I$ et $(e^p)^T P^{(k)} = (e^p)^T$ puisque $p < k \leq l$ (les coefficients permutés de e^p sont nuls). \square

Remarquons que, dans la pratique, il est inutile de stocker toutes les matrices $P^{(k)}$, il suffit à chaque échange de mettre à jour un vecteur p , reprenez l'exemple.

La factorisation

$$PA = LU$$

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

2.4.5 Algorithme de la factorisation $PA = LU$

Comme dans le cas de la factorisation $A = LU$, il n'y a pas besoin de conserver les matrices successives $A^{(k)}$ et $L^{(k)}$. En ce qui concerne les permutations, on en garde l'historique en mettant à jour un vecteur p valant initialement

$$p = [1, 2, \dots, n]^T,$$

sur lequel on effectuera les permutations successives représentées par les matrices $P^{(k)}$, $k = 1, \dots, n-1$.

En ce qui concerne la recherche de l'élément pivot, on recherchera le pivot le plus grand en valeur absolue pour les raisons déjà évoquées précédemment.

Algorithme de factorisation $PA = LU$

- 1: $L \leftarrow I$
- 2: $p \leftarrow [1, 2, \dots, n]^T$
- 3: **pour** $k = 1$ jusqu'à $n - 1$ **faire**
- 4: Trouver l tel que $|a_{lk}| \geq |a_{ik}|$, $i = k \dots n$
- 5: **si** $|a_{lk}| < \varepsilon$ **alors**
- 6: Arrêter l'algorithme et donner un message d'erreur
- 7: **sinon**

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

```
8:   Appliquer la permutation  $(k, l)$  au vecteur  $p$ , aux lignes de  $A$  et aux lignes de  $L$  dans les
    colonnes 1 à  $k - 1$ 
9:   pour  $i = k + 1$  jusqu'à  $n$  faire
10:      $l_{ik} \leftarrow a_{ik} / a_{kk}$ 
11:     pour  $j = k$  jusqu'à  $n$  faire
12:        $a_{ij} \leftarrow a_{ij} - l_{ik} a_{kj}$ 
13:     fin pour
14:   fin pour
15: fin si
16: fin pour
17: si  $|a_{nn}| < \varepsilon$  alors
18:   Arrêter l'algorithme et donner un message d'erreur
19: fin si
```

Le dernier test est important car la valeur de a_{nn} , qui n'est pas un pivot, n'est pas testée dans la boucle principale. Le message d'erreur à afficher dans les deux cas pourrait être du type «la matrice est singulière à la précision machine».

Algorithme de la factorisation

$$PA = LU$$

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

2.4.6 Application à la résolution d'un système linéaire

Comment utilise-t-on cette factorisation pour résoudre un système linéaire ? C'est aussi simple que pour la factorisation LU "classique". On a $Ax = b$ que l'on multiplie à gauche par la matrice de permutation P . On a donc

$$PAx = Pb,$$

soit $LUx = Pb$. La solution de ce système s'obtient en résolvant successivement les deux systèmes

$$\begin{cases} Ly = Pb, \\ Ux = y, \end{cases}$$

ce qui correspond toujours à la résolution de deux systèmes triangulaires. La seule différence est qu'ici on doit permuter les composantes du vecteur b avant résolution.

De façon pratique, on n'explicite jamais la matrice de permutation P mais on dispose (voir algorithme précédent) du vecteur p défini par

$$p = P \begin{pmatrix} 1 \\ 2 \\ \vdots \\ n \end{pmatrix}.$$

Le vecteur $c = Pb$ est donc le vecteur

$$c = \begin{pmatrix} b_{p_1} \\ b_{p_2} \\ \vdots \\ b_{p_n} \end{pmatrix}.$$

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Pour le calcul du déterminant de A , on a

$$\det PA = \det P \det A = \prod_{k=1}^n u_{kk},$$

avec $\det P = (-1)^q$ où q est le nombre de permutations différentes de l'identité qui ont été effectuées sur A .

**Application à la
résolution d'un
système linéaire**

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

2.4.7 La méthode du pivot total

Dans certaines situations la permutation des lignes est insuffisante. En effet il est possible que dans une colonne on ne puisse pas trouver d'élément suffisamment grand en valeur absolue (par rapport à un seuil que l'on se fixe). Ceci conduit à rechercher à chaque étape k de la méthode de Gauss, l'élément de la sous-matrice

$$\begin{pmatrix} a_{k,k} & \dots & a_{k,n} \\ \vdots & \ddots & \vdots \\ a_{n,k} & \dots & a_{nn} \end{pmatrix},$$

dont la valeur absolue est maximum. Ceci nécessite alors une permutation des lignes *et* des colonnes. Cela correspond à échanger des équations du système, ainsi que des inconnues. En termes de factorisation de type $A = LU$, cette approche permet d'obtenir L et U respectivement triangulaire inférieure et triangulaire supérieure, telles que

$$LU = PAQ,$$

où P et Q sont des matrices de permutation. Cette méthode est rarement utilisée, car la recherche du pivot peut très vite devenir coûteuse (quand la matrice est grande) par rapport au gain de précision éventuellement obtenu.

Ce problème du choix de pivot maximal : ***pivot partiel*** (recherché uniquement sur la $k^{\text{ème}}$ colonne : $k \leq i \leq n$) ou ***pivot total*** (recherché sur toute la sous-matrice : $k \leq i, j \leq n$) est un exemple typique de situation classique en analyse numérique. Plus précisément, on sait obtenir mathématiquement des majorations d'erreur pour la méthode de pivot total bien meilleures que pour

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

la méthode de pivot partiel. On sait exhiber des matrices pour lesquelles la majoration d'erreur relative au pivot partiel est atteinte. Pourtant, l'expérience numérique montre que dans toutes les situations pratiques, la méthode du pivot partiel suffit.

La méthode du pivot total

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

2.5 Traitement des matrices symétriques

2.5.1	La factorisation LDL^T	153
2.5.2	La factorisation de Cholesky : existence	155
2.5.3	Algorithme de Cholesky	158

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

2.5.1 La factorisation LDL^T

Exercices :

[Exercice 2.16](#)

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 2.5.1. *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

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

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Or, on a démontré l'unicité de la factorisation LU (dans le cas où L est à diagonale unité), d'où $L = V$, 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.

La factorisation LDL^T

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

2.5.2 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 2.5.1. *Une matrice A symétrique définie positive a toutes ses sous-matrices principales $[A]_k$ régulières.*

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

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

Comme A est définie positive

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

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

Théorème 2.5.2. *Si A est une matrice symétrique définie positive elle admet une unique factorisation sous la forme*

$$A = B B^T,$$

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

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Démonstration - Tout d'abord, comme A est définie positive toutes ses sous-matrices principales sont régulières (voir le Lemme 2.5.1 précédent) et il résulte de la proposition 2.5.1 qu'elle admet une factorisation sous la forme

$$A = LDL^T. \quad (2.5.1)$$

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 (2.5.1) 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

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

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

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

La factorisation de Cholesky : existence

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)

[Documents](#)

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

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

ce qui implique que

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

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

$$d_{ii} = 1.$$

D est donc la matrice identité, donc $B = \tilde{B}$ 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

$$A = BB^T.$$

La factorisation de Cholesky : existence

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

2.5.3 Algorithme de Cholesky

Exercices :
[Exercice 2.17](#)

Cours :
[Factorisation de Cholesky / existence](#)

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 référencé), 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

Algorithme de Cholesky

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

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

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

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)

[Documents](#)

Algorithme de Cholesky

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

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

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 , puis les autres termes de cette colonne sont connus car nuls.

On a donc l'algorithme :

1: **pour** $j = 1$ jusqu'à $n-1$ **faire**

$$2: \quad b_{jj} \leftarrow \sqrt{a_{jj} - \sum_{k=1}^{j-1} b_{jk}^2};$$

3: **pour** $i = j+1$ jusqu'à n **faire**

$$4: \quad b_{ij} \leftarrow \frac{1}{b_{jj}} \left(a_{ij} - \sum_{k=1}^{j-1} b_{ik}b_{jk} \right);$$

5: **fin pour**

6: **fin pour**

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

$$7: b_{nn} \leftarrow \sqrt{a_{nn} - \sum_{k=1}^{n-1} b_{nk}^2};$$

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.

Algorithme de Cholesky

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exemples du chapitre 2

2.1	Influence des arrondis	163
2.2	Exemple de factorisation $PA = LU$	165

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exemple 2.1 Influence des arrondis

Prenons le système linéaire $Ax = b$ avec

$$A = \begin{pmatrix} 10^{-4} & 1 \\ 1 & 1 \end{pmatrix}, b = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

Supposons que l'on travaille sur une machine où les nombres sont représentés avec une mantisse de 4 chiffres significatifs. Cela signifie que chaque nombre est représenté sous la forme

$$\pm 0.d_1 d_2 d_3 d_4 10^e,$$

où d_i représente un nombre entier positif et e l'exposant (le nombre de chiffres significatifs pour l'exposant n'intervient pas dans notre exemple). Lorsque l'on effectue la première (et dernière ici) étape de l'élimination de Gauss, on obtient le système équivalent

$$\hat{A} = \begin{pmatrix} 10^{-4} & 1 \\ 0 & 1 - 10^4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 - 10^4 \end{pmatrix}.$$

On a l'addition suivante à effectuer :

$$1 - 10^4 = 0.1 \cdot 10^1 - 0.1 \cdot 10^5 = 0.00001 \cdot 10^5 - 0.1 \cdot 10^5,$$

mais, comme on ne dispose que de 4 chiffres significatifs, le nombre $0.00001 \cdot 10^5$ est représenté en machine par $0.0000 \cdot 10^5$, soit 0. On a donc $1 - 10^4 = -10^4$. Comme le raisonnement est le même pour $2 - 10^4$, on a donc, en machine, le système suivant :

$$\begin{pmatrix} 10^{-4} & 1 \\ 0 & -10^4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ -10^4 \end{pmatrix}.$$

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Lorsque l'on effectue la remontée de ce système, on obtient $x_2 = 1$, et la première équation donne $10^{-4}x_1 = 0$. On a donc $x = [0, 1]^T$. Voyons maintenant ce qu'il se passe si on échange les lignes 1 et 2 dans la matrice A et le vecteur b . On obtient donc le système équivalent

$$\begin{pmatrix} 1 & 1 \\ 10^{-4} & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

L'élimination de Gauss donne, en respectant les problèmes d'arrondi dûs à la mantisse de 4 chiffres, le système équivalent

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix},$$

qui donne $x = [1, 1]^T$. Si l'on regarde la solution *théorique*, du système, on constate que l'on obtient

$$x_1 = \frac{10000}{9999} \approx 1, \quad x_2 = \frac{9998}{9999} \approx 1.$$

Donc la première méthode sans échange de ligne donne une solution complètement erronée, alors que la deuxième méthode donne une approximation tout à fait raisonnable.

[retour au cours](#)

Exemple 2.1

Influence des arrondis

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exemple 2.2 Exemple de factorisation $PA = LU$

On considère la matrice de départ (on omet les numéros des étapes pour L)

$$A = \begin{pmatrix} 0 & 1 & -1 \\ 1 & 1 & -1 \\ -1 & 1 & 1 \end{pmatrix}.$$

On a initialement $L = I$ et $p = \{1, 2, 3\}$.

- $k = 1$: on effectue la permutation (1, 2) pour avoir un pivot non nul et on effectue l'élimination, ce qui donne

$$A^{(2)} = \begin{pmatrix} 1 & 1 & -1 \\ 0 & 1 & -1 \\ 0 & 2 & 0 \end{pmatrix}, \quad L = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}, \quad p = \{2, 1, 3\}.$$

- $k = 2$: on effectue la permutation (2, 3) pour avoir un pivot maximum. Attention : avant de remplir la 2^{ème} colonne de L on permute ses coefficients des lignes (2, 3) dans la colonne 1, ce qui donne

$$A^{(3)} = \begin{pmatrix} 1 & 1 & -1 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \quad L = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & \frac{1}{2} & 1 \end{pmatrix}, \quad p = \{2, 3, 1\}.$$

Si on effectue LU on a bien

$$LU = \begin{pmatrix} 1 & 1 & -1 \\ -1 & 1 & 1 \\ 0 & 1 & -1 \end{pmatrix} = \begin{pmatrix} A_2 \\ A_3 \\ A_1 \end{pmatrix}.$$

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

[retour au cours](#)

Exemple 2.2

Exemple de
factorisation

$$PA = LU$$

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercices du chapitre 2

2.1	168
2.2	169
2.3	170
2.4	171
2.5	172
2.6	173
2.7	174
2.8	175
2.9	176
2.10	177
2.11	178
2.12	179
2.13	180
2.14	181
2.15	182
2.16	183
2.17	184

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 2.1

On définit la matrice A , à n lignes et n colonnes par

$$A = \begin{pmatrix} 2 & -1 & & & 0 \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ 0 & & & -1 & 2 \end{pmatrix}$$

On veut résoudre $Ax = 0$.

1. Montrer en résolvant les $n - 1$ premières équations que $x_i = ix_1$, $i = 1, \dots, n$.
2. Résoudre la dernière équation et en déduire que $x = 0$.
3. En déduire que A est inversible.

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 2.2

Soit A une matrice triangulaire inférieure. Écrire l'algorithme permettant de résoudre le système linéaire $Ax = b$ (b vecteur donné) en n'oubliant pas de vérifier au départ que ce système a une solution.

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 2.3

Soit A une matrice triangulaire supérieure, montrer que le calcul du vecteur inconnu est donné par :

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

Écrire alors l'algorithme correspondant.

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 2.4

Soit le système $Ax = b$. On considère la première étape de l'élimination de Gauss. Montrer que la i ème équation (pour $i \geq 2$) est modifiée de la manière suivante :

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

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 2.5

Soit le système $Ax = b$. On considère la deuxième étape de l'élimination de Gauss. Montrer que la i ème équation (pour $i \geq 3$) est modifiée de la manière suivante :

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

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 2.6

Soit le système $Ax = b$. On considère la k ème étape de l'élimination de Gauss. Montrer que la i ème équation (pour $i \geq k + 1$) est modifiée de la manière suivante :

$$\left\{ \begin{array}{l} a_{ij}^{(k)} \longrightarrow a_{ij}^{(k+1)} = a_{ij}^{(k)} - \frac{a_{ik}^{(k)}}{a_{kk}^{(k)}} a_{kj}^{(k)} \text{ pour } j = k, k+1, \dots, n \\ b_i^{(k)} \longrightarrow b_i^{(k+1)} = b_i^{(k)} - \frac{a_{ik}^{(k)}}{a_{kk}^{(k)}} b_k^{(k)}. \end{array} \right.$$

[retour au cours](#)

Solution

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 2.7

Soit la matrice $A = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 2 & 1 \\ 1 & 1 & 1 \end{pmatrix}$ et le vecteur $b = \begin{pmatrix} 4 \\ 5 \\ 3 \end{pmatrix}$, appliquez l'algorithme de Gauss "à la main" pour calculer la solution de $Ax = b$.

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 2.8

Calculer le nombre d'opérations effectuées pour réaliser l'élimination de Gauss en fonction de n en séparant multiplications/divisions et additions/ soustractions. Pour cela on pourra utiliser les deux formules

$$\sum_{k=1}^n k = \frac{n(n+1)}{2},$$
$$\sum_{k=1}^n k^2 = \frac{n(n+1)(2n+1)}{6}.$$

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 2.9

Soient L une matrice triangulaire inférieure et U une matrice triangulaire supérieure et on pose $A = LU$. Montrer que, pour la colonne j de A , on a

$$a_{ij} = \sum_{k=1}^i l_{ik} u_{kj}, \text{ pour } i \leq j,$$

et

$$a_{ij} = \sum_{k=1}^j l_{ik} u_{kj}, \text{ pour } i > j.$$

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 2.10

Soit A une matrice inversible qui admet une factorisation $A = LU$ où L est triangulaire inférieure, U est triangulaire supérieure et la diagonale de U ne comporte que des 1, alors cette factorisation est unique.

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 2.11

Résoudre le système $Ax = b$ dont la factorisation LU de A est donnée :

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & 2 & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 & -2 \\ 0 & 3 & 1 \\ 0 & 0 & -1 \end{pmatrix} \quad b = \begin{pmatrix} 1 \\ 6 \\ 6 \end{pmatrix}$$

[retour au cours](#)

Solution

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 2.12

Soit $A = \begin{pmatrix} 2 & 1 & -2 \\ 4 & 5 & -3 \\ -2 & 5 & 3 \end{pmatrix}$, en vous inspirant de ce qui a été fait pour l'algorithme de Doolittle

dans le paragraphe "[Factorisation \$A = LU\$ / calcul direct](#)", effectuez la factorisation de Crout de la matrice A , c'est-à-dire déterminez L et U telles que $A = LU$ avec les termes diagonaux de U égaux à 1 (ceux de L sont quelconques).

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 2.13

Dans le calcul direct de la factorisation LU , on suppose maintenant que c'est la matrice U dont tous les éléments de la diagonale sont égaux à 1 et non pas la matrice L . Calculer les éléments des matrices U et L à partir d'éléments de A et d'éléments de U et L de colonnes ou de lignes précédentes. Comment modifier l'algorithme de Doolittle pour le calcul des éléments l_{ij} et u_{ij} des matrices L et U . Cet algorithme s'appelle l'algorithme de Crout.

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 2.14

Montrez que si la factorisation $A = LU$ existe (L triangulaire inférieure avec une diagonale unitaire et U triangulaire supérieure inversible), alors les sous-matrices principales de A sont inversibles.

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 2.15

Soit σ une permutation de $\{1, 2, \dots, n\}$ et soit g l'application linéaire telle que $g(\vec{e}_j) = \vec{e}_{\sigma(j)}$ où $\{\vec{e}_1, \dots, \vec{e}_n\}$ est la base canonique de \mathbb{R}^n . Montrer que la matrice P de l'application g est telle que

$$p_{ij} = \delta_{i, \sigma(j)}$$

et que $P^{-1} = P^T$.

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 2.16

1. Soit A une matrice symétrique admettant une factorisation LDL^T . Montrer que pour $i \geq j$ on a

$$a_{ij} = \sum_{k=1}^j d_k l_{ik} l_{jk},$$

où on a noté d_k le $k^{\text{ème}}$ élément de la diagonale de D .

2. Dédurre de la question précédente que les coefficients de L et ceux de D peuvent être obtenus par les formules (on considère que les sommes ne sont pas effectuées quand $j = 1$)

$$d_j = a_{jj} - \sum_{k=1}^{j-1} d_k l_{jk}^2,$$

et pour $i > j$

$$l_{ij} = \frac{a_{ij} - \sum_{k=1}^{j-1} d_k l_{ik} l_{jk}}{d_j}.$$

Indication : ne pas oublier que $l_{jj} = 1$ par définition.

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 2.17

Soit A une matrice symétrique définie positive. On considère sa factorisation de Cholesky $A = BB^T$. Montrer que tous les éléments de la diagonale de B sont non nuls.

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Chapitre 3

Résolution des problèmes de moindres carrés

3.1	Formulation générale des problèmes de moindres carrés	186
	Documents du chapitre 3	194
	Exercices du chapitre 3	197

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

3.1 Formulation générale des problèmes de moindres carrés

3.1.1	Un exemple : un problème de lissage.	187
3.1.2	Formulation matricielle	189
3.1.3	Les équations normales	191

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

3.1.1 Un exemple : un problème de lissage.

Exercices :

[Exercice 3.1](#)

Si on se donne une famille de points du plan $(t_i, b_i)_{1 \leq i \leq m}$ (les t_i étant distincts) alors il existe un unique polynôme $P(t)$ de degré inférieur ou égal à $m - 1$ tel que

$$P(t_i) = b_i, \quad i = 1, \dots, m,$$

qui est le polynôme d'interpolation des points (t_i, b_i) . Si le nombre de points m est trop grand, ou si les ordonnées sont bruitées, on préfère en général chercher une fonction $f(t)$ qui, dans une classe donnée (polynômes, fractions rationnelles, polynômes trigonométriques, exponentielles ...), approche "au mieux" les points (t_i, b_i) , on parle alors d'approximation, de lissage ou bien de *régression* (voir la figure 3.1.1).

Soit donc une famille de fonctions

$$f_1(t), f_2(t), \dots, f_n(t), \quad n \leq m,$$

linéairement indépendantes. Etant donné n nombres réels x_1, x_2, \dots, x_n on peut introduire le nombre $E(x)$

$$E(x) = \sum_{i=1}^m [f(t_i) - b_i]^2,$$

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

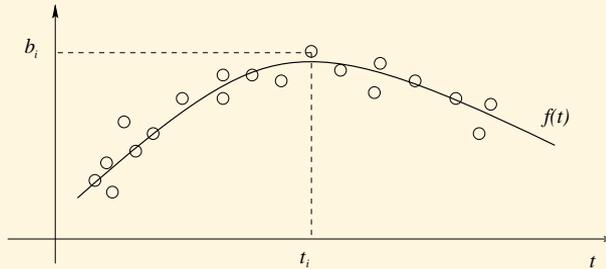


FIGURE 3.1.1 – un problème de lissage

avec $f(t) = \sum_{k=1}^n x_k f_k(t)$. La quantité $E(x)$ représente la somme des erreurs quadratiques entre les valeurs données et celles prises par f aux points t_i . Le problème d'approximation se formule alors de la façon suivante :

Trouver $\hat{x} \in \mathbb{R}^n$, tel que $E(\hat{x}) \leq E(x)$, $\forall x \in \mathbb{R}^n$.

Par exemple si l'on désire faire de la régression polynômiale, c'est-à-dire prendre pour $f(t)$ un polynôme de degré $\leq n - 1$, on a

$$f_k(t) = t^{k-1}, \quad k = 1, \dots, n, \quad f(t) = x_1 + x_2 t + \dots + x_n t^{n-1}.$$

L'écriture inhabituelle du polynôme avec x_k comme coefficients permet de mettre en évidence que les inconnues du problème de moindres carrés sont ces coefficients.

Un exemple : un problème de lissage.

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

3.1.2 Formulation matricielle

Exercices :
[Exercice 3.2](#)

Cours :
[Lissage](#)

On suppose que l'on a à résoudre un système linéaire $Ax = b$ ($A \in \mathcal{M}_{mn}$), avec un second membre b non nul et on suppose que le nombre d'équations est supérieur strictement au nombre d'inconnues ($m > n$). Dans la plupart des cas, ce système n'a pas de solution. On cherche alors une approximation de la solution qui réduise la différence $Ax - b$. Un des choix possible est de minimiser la norme euclidienne de cette différence. Dans tout ce chapitre, nous n'utiliserons que la norme euclidienne.

Définition 3.1.1. Soit $A \in M_{mn}(\mathbb{R})$ et $b \in \mathbb{R}^m$ donnés. On appelle problème de moindres carrés le problème

$$\min_{x \in \mathbb{R}^n} \|Ax - b\|_2^2. \quad (3.1.1)$$

On notera \hat{x} la solution de ce problème. On montrera dans la suite qu'elle existe, et qu'elle est unique sous l'hypothèse fondamentale suivante :

$$\text{Les colonnes de } A \text{ sont linéairement indépendantes.} \quad (3.1.2)$$

ou, de façon équivalente

$$\text{rang } A = n.$$

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Formulation matricielle

Un cas particulier est le cas $m = n$ et A inversible, alors \hat{x} est la solution unique de $Ax = b$, mais ce n'est pas ce cas particulier qui nous intéresse dans ce chapitre.

Un autre exemple, pour le problème de lissage, est introduit dans le paragraphe référencé. Il n'est pas possible en général de faire passer une fonction $f(t)$ dépendant de n inconnues (par exemple un polynôme de degré $n - 1$) par m points ($m > n$). Il n'existe donc sans doute pas de $x \in \mathbb{R}^n$ qui annule les quantités

$$\sum_{k=1}^n x_k f_k(t_i) - b_i, \quad 1 \leq i \leq m.$$

Cette expression s'écrit bien $Ax - b$ où

$$a_{ij} = f_j(t_i), \quad 1 \leq i \leq m, \quad 1 \leq j \leq n,$$

et l'on a bien à minimiser

$$E(x) = \sum_{i=1}^m \left[\sum_{k=1}^n x_k f_k(t_i) - b_i \right]^2.$$

Cette fois-ci le minimum ne sera pas nul.

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

3.1.3 Les équations normales

Exercices :

[Exercice 3.3](#)

Documents :

[Document 3.1](#)

Nous ferons plus tard une résolution purement algébrique du problème des moindres carrés. Cependant nous allons en donner une approche analytique qui permet de voir ce problème sous un angle différent. Tout d'abord, explicitons la fonction à minimiser. On a

$$\begin{aligned} E(x) &= \|Ax - b\|_2^2 = (Ax - b)^T (Ax - b) = x^T A^T Ax - b^T Ax - x^T A^T b + \|b\|_2^2 \\ &= x^T A^T Ax - 2(A^T b)^T x + \|b\|_2^2. \end{aligned}$$

Le problème de moindres carrés peut donc se reformuler en

$$\min_{x \in \mathbb{R}^n} J(x) = x^T G x - 2h^T x,$$

où $G = A^T A$ est symétrique et $h = A^T b$ est un vecteur donné.

Définition 3.1.2. On appelle fonction quadratique une fonction $J : \mathbb{R}^n \rightarrow \mathbb{R}$, de la forme

$$J(x) = x^T G x - 2h^T x,$$

où G est une matrice $n \times n$ symétrique et h est un vecteur donné de \mathbb{R}^n .

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Rappelons que si $J : \mathbb{R} \rightarrow \mathbb{R}$, continûment dérivable, admet un minimum $\hat{x} \in \mathbb{R}$, alors $J'(\hat{x}) = 0$. De même soit $J : \mathbb{R}^n \rightarrow \mathbb{R}$ continûment dérivable, alors

$$J(\hat{x}) \leq J(x) \quad \forall x \in \mathbb{R}^n \Rightarrow \nabla J(\hat{x}) = 0, \quad \nabla \text{ opérateur gradient .}$$

Ici le calcul du gradient de J donne (voir le document référencé)

$$\nabla J(x) = 2(Gx - h)$$

et la solution \hat{x} du problème de moindres carrés vérifie donc nécessairement

$G\hat{x} = h$, soit

$$A^T A \hat{x} = A^T b.$$

Ces relations sont appelées *équations normales* du problème, ce sont des conditions nécessaires pour que \hat{x} soit minimum, on montre de plus que ces conditions sont suffisantes d'où le théorème.

Théorème 3.1.1. Soit $A \in M_{mn}(\mathbb{R})$ et $b \in \mathbb{R}^m$ donnés, si l'on suppose que la matrice A est de rang n , le problème de moindres carrés $\min_{x \in \mathbb{R}^n} \|Ax - b\|_2^2$, admet une solution unique \hat{x} donnée par

$$A^T A \hat{x} = A^T b. \tag{3.1.3}$$

Démonstration : On remarque tout d'abord que si A est de rang n , alors $G = A^T A$ est définie positive en effet :

$$x^T G x = \|Ax\|_2^2 \geq 0 \quad \forall x \in \mathbb{R}^n,$$

Les équations normales

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

d'autre part puisque le rang de A vaut n , alors la dimension de $\text{Ker } A$ est nulle donc

$$x^T Gx = 0 \Leftrightarrow \|Ax\|_2^2 = 0 \Leftrightarrow Ax = 0 \Leftrightarrow x = 0.$$

La matrice G étant définie positive, elle est inversible, donc il existe une unique solution \hat{x} vérifiant $A^T A\hat{x} = A^T b \Leftrightarrow G\hat{x} = h$.

Montrons maintenant que

$$\forall x \in \mathbb{R}^n, x \neq \hat{x} \Rightarrow J(x) > J(\hat{x}).$$

Posons $y = x - \hat{x}$, on a donc $y \neq 0$.

$$\begin{aligned} J(\hat{x} + y) &= (\hat{x} + y)^T G(\hat{x} + y) - 2h^T (\hat{x} + y), \\ &= \hat{x}^T G\hat{x} + y^T Gy + 2\hat{x}^T Gy - 2h^T y - 2h^T \hat{x}, \\ &= \hat{x}^T G\hat{x} + y^T Gy + 2(G\hat{x} - h)^T y - 2h^T \hat{x}, \\ &= y^T Gy + \hat{x}^T G\hat{x} - 2h^T \hat{x}, \end{aligned}$$

puisque $G\hat{x} = h$. d'où

$$J(\hat{x} + y) = J(\hat{x}) + y^T Gy.$$

Puisque G est une matrice définie positive, et que $y \neq 0$, on a $y^T Gy > 0$, et donc

$$J(\hat{x} + y) > J(\hat{x}), \forall y \in \mathbb{R}^n, y \neq 0,$$

ce qui montre que \hat{x} réalise le minimum de J . □

Les équations normales

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Documents du chapitre 3

3.1 [Le gradient d'une forme quadratique](#) 195

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Document 3.1 Le gradient d'une forme quadratique

Soit la fonction quadratique

$$J(x) = x^T Gx - 2h^T x,$$

où G est une matrice symétrique. On veut calculer son gradient, soit

$$\nabla J(x) = \left(\frac{\partial J}{\partial x_1}, \frac{\partial J}{\partial x_2}, \dots, \frac{\partial J}{\partial x_n} \right)$$

Développons la fonction J

$$J(x) = \sum_{i=1}^n x_i (Gx)_i - 2 \sum_{i=1}^n h_i x_i.$$

Alors

$$\frac{\partial J}{\partial x_k} = (Gx)_k + \sum_{i=1}^n x_i \frac{\partial}{\partial x_k} (Gx)_i - 2h_k,$$

et

$$\frac{\partial}{\partial x_k} (Gx)_i = \frac{\partial}{\partial x_k} \left(\sum_{j=1}^n g_{ij} x_j \right) = g_{ik} = g_{ki}.$$

Donc

$$\frac{\partial J}{\partial x_k} = (Gx)_k + \sum_{i=1}^n x_i g_{ki} - 2h_k = (Gx)_k + (Gx)_k - 2h_k = 2(Gx)_k - 2h_k,$$

d'où le résultat

$$\nabla J(x) = 2(Gx - h).$$

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

[retour au cours](#)

Document 3.1
Le gradient d'une
forme
quadratique

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercices du chapitre 3

3.1	198
3.2	199
3.3	200

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 3.1

1. Ecrire le problème de la régression linéaire comme un problème de moindres carrés : plus précisément on se donne une famille de points $(t_i, b_i)_{1 \leq i \leq m}$ (les t_i étant distincts) et on cherche à faire passer une droite le plus près possible de ces points.
2. La question précédente conduit à une fonction de deux variables à minimiser. On admet que ce minimum est donné en annulant les deux dérivées partielles. Donner le système linéaire de deux équations à deux inconnues ainsi obtenu.

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 3.2

On cherche à approcher les données $(t_i, b_i)_{1 \leq i \leq m}$ (les t_i étant distincts) à l'aide d'un polynôme de degré inférieur ou égal à $n - 1$, on suppose $m \geq n$, on note

$$p(t) = \alpha_1 + \alpha_2 t + \dots + \alpha_n t^{n-1},$$

et on cherche les coefficients $\alpha_1, \alpha_2, \dots, \alpha_n$ qui minimisent

$$E(\alpha_1, \alpha_2, \dots, \alpha_n) = \sum_{i=1}^m (p(t_i) - b_i)^2.$$

Ecrire le problème de moindres carrés sous forme matricielle. Montrer alors que la matrice A est bien de rang n . (On rappelle que la matrice carrée de Van der Monde $V, v_{ij} = t_i^{j-1}$, est inversible si tous les t_i sont distincts).

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 3.3

Donner les équations normales du problème de moindres carrés associé à la régression linéaire. Montrer que l'on retrouve les équations de l'exercice 3.1.

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Chapitre 4

Méthodes itératives

4.1	Méthodes itératives de résolution des systèmes linéaires	202
4.2	Résolution d'une équation non-linéaire à une inconnue	212
4.3	Résolution d'un système d'équations non-linéaires	228
	Documents du chapitre 4	237
	Exercices du chapitre 4	242

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

4.1 Méthodes itératives de résolution des systèmes linéaires

4.1.1	Principes généraux	203
4.1.2	La méthode de Jacobi	205
4.1.3	La méthode de Gauss-Seidel	207
4.1.4	Etude de la convergence des méthodes itératives	209

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

4.1.1 Principes généraux

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 à résoudre.

Soit $A \in \mathcal{M}_{nn}(\mathbb{R})$ une matrice régulière et $b \in \mathbb{R}^n$ donnés. On se propose de résoudre le problème

$$f(x) = Ax - b = 0$$

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

$$A = M - N.$$

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

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*.

Principes généraux

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

4.1.2 La méthode de Jacobi

Exercices :

[Exercice 4.1](#)[Exercice 4.2](#)

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 (voir exercice).

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

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

La méthode de Jacobi

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

4.1.3 La méthode de Gauss-Seidel

Exercices :[Exercice 4.3](#)**Cours :**[Jacobi](#)

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

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

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 référencé sur la méthode de Jacobi.

La méthode de Gauss-Seidel s'écrit donc

$$\left\{ \begin{array}{l} x^{(0)} \text{ donné,} \\ (D - E)x^{(k+1)} = (Fx^{(k)} + b), \end{array} \right.$$

[Sommaire](#)
[Concepts](#)[Exemples](#)
[Exercices](#)
[Documents](#)

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

On observe que les méthodes de Jacobi et Gauss-Seidel que nous venons de voir peuvent se mettre sous la forme $Mx^{(k+1)} = Nx^{(k)} + b$:

- $M = D$, $N = E + F$, pour la méthode de Jacobi,
- $M = D - E$, $N = F$, pour la méthode de Gauss-Seidel.

La méthode de Gauss-Seidel

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

4.1.4 Etude de la convergence des méthodes itératives

Exercices :

[Exercice 4.4](#)

Commençons par étudier la convergence de la méthode itérative

$$\begin{cases} x^{(0)} \text{ donné,} \\ x^{(k+1)} = Cx^{(k)} + d. \end{cases} \quad (4.1.1)$$

Proposition 4.1.1. *S'il existe une norme matricielle subordonnée telle que*

$$\|C\| < 1$$

la méthode (4.1.1) 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

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

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

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 puisque la matrice $I - C$ est carrée à $I - C$ inversible

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} .

Définition 4.1.1. On dit que la matrice A est à diagonale strictement dominante si

$$|a_{ii}| > \sum_{j \neq i} |a_{ij}|, \quad \forall i, 1 \leq i \leq n. \quad (4.1.2)$$

Proposition 4.1.2. 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. La démonstration sera faite en travaux dirigés pour la méthode de Gauss-Seidel.

La méthode de Jacobi s'écrit

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

**Etude de la
convergence des
méthodes
itératives**

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

et donc la matrice d'itération est

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

Puisque A est à diagonale dominante, on a

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

et donc la méthode est convergente.

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

Proposition 4.1.3. *Si la matrice A est symétrique définie positive, alors la méthode de Gauss-Seidel est convergente.*

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

Théorème 4.1.2. *Une condition nécessaire et suffisante pour que la méthode itérative (4.1.1) soit convergente est que*

$$\rho(C) < 1.$$

Etude de la convergence des méthodes itératives

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

4.2 Résolution d'une équation non-linéaire à une inconnue

4.2.1	Méthode de la dichotomie	213
4.2.2	Le principe des méthodes de point fixe	215
4.2.3	Convergence des méthodes de point fixe	217
4.2.4	La méthode de Newton pour résoudre une équation	221
4.2.5	Convergence de la méthode de Newton pour résoudre une équation	223
4.2.6	La méthode de la sécante	225

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

4.2.1 Méthode de la dichotomie

Exercices :

[Exercice 4.5](#)

On veut résoudre $f(x) = 0$, où f est une fonction de \mathbb{R} dans \mathbb{R} non linéaire (sinon c'est évident !). On recherche donc un nombre réel x^* tel que $f(x^*) = 0$. Comme pour toute méthode itérative, on va donc construire une suite de nombres réels $x^{(0)}, x^{(1)}, \dots, x^{(k)}, \dots$ qui converge vers x^* . Puisqu'ici les termes de la suite sont des nombres réels, on les notera plus simplement $x_0, x_1, \dots, x_k, \dots$

Le principe de la dichotomie est très simple. On suppose que $f : [a, b] \rightarrow \mathbb{R}$ est continue et que $f(a)f(b) < 0$, c'est-à-dire qu'il y a au moins une racine réelle de f dans $[a, b]$. On prend alors le milieu de l'intervalle $c = \frac{a+b}{2}$. Si $f(c) = 0$ (ou numériquement très proche de 0) on considère que l'on a résolu le problème. Autrement deux cas peuvent se présenter. Si $f(a)f(c) < 0$, alors f a une racine réelle dans $[a, c]$ (au moins une, en tous cas, un nombre impair de racines), autrement $f(a)f(c) > 0$ et donc $f(a)f(b)f(a)f(c) < 0$, ce qui donne $f(b)f(c) < 0$ et f a une racine réelle dans $[c, b]$. On itère alors le processus avec l'intervalle qui contient au moins une racine.

La méthode de dichotomie s'écrit :

$$a_0 = a, \quad b_0 = b, \quad x_0 = \frac{a_0 + b_0}{2},$$

si $f(x_k)f(a_k) < 0$ alors $a_{k+1} = a_k$ et $b_{k+1} = x_k$ sinon $a_{k+1} = x_k$ et $b_{k+1} = b_k$.

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

$$x_{k+1} = \frac{a_{k+1} + b_{k+1}}{2},$$

Il est clair qu'à chaque itération la longueur l_k de l'intervalle de localisation de la racine x^* est divisée par 2. On sait de plus que $|x_k - x^*| \leq l_k$, d'où le résultat de convergence de la suite (x_k) vers x^* si on montre que l_k tend vers 0

Proposition 4.2.1. *La suite de segments $([a_k, b_k])_{k \geq 0}$ est telle que*

$$l_k = b_k - a_k = \frac{b - a}{2^k}, \quad k \geq 1.$$

Cette proposition est utile sur le plan numérique; en effet elle renseigne sur le nombre minimal d'itérations nécessaires pour calculer la racine à ϵ près, $\epsilon > 0$ étant une précision fixée à l'avance par l'utilisateur (test d'arrêt). En effet, on peut montrer (en exercice) que ce nombre n doit vérifier

$$n \geq \log_2 \left(\frac{b - a}{\epsilon} \right).$$

Méthode de la dichotomie

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

4.2.2 Le principe des méthodes de point fixe

Exercices :

[Exercice 4.6](#)

Pour résoudre $f(x) = 0$, on peut se ramener à un problème équivalent de la forme $g(x) = x$ (par exemple en posant $g(x) = f(x) + x$). Mais il y a beaucoup de choix possibles pour définir g ! La résolution de $g(x) = x$ s'appelle recherche des points fixes de g .

Par exemple pour calculer la racine carrée d'un nombre réel $a > 0$, on résout $x^2 = a$. Il existe alors différentes manières de se ramener à un problème de point fixe :

–

$$x = \frac{a}{x}, \text{ et } g_1(x) = \frac{a}{x},$$

–

$$2x = x + \frac{a}{x} \text{ soit } x = 2x - \frac{a}{x} \text{ et } g_2(x) = 2x - \frac{a}{x},$$

–

$$x = \frac{x}{2} + \frac{a}{2x}, \text{ et } g_3(x) = \frac{x}{2} + \frac{a}{2x}.$$

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Pour résoudre $g(x) = x$, problème équivalent à $f(x) = 0$, les méthodes dites de point fixe, consistent à construire la suite $x_0, x_1, \dots, x_k, \dots$ de la façon suivante :

$$\begin{cases} x_0 \text{ donné,} \\ x_{k+1} = g(x_k), k > 0 \end{cases}$$

Si la suite (x_k) converge, la continuité de g implique qu'elle converge vers un point fixe x^* de g . En effet

$$x^* = \lim_{k \rightarrow \infty} x_{k+1} = g(\lim_{k \rightarrow \infty} x_k) = g(x^*).$$

Vous verrez en exercice comment se comportent graphiquement les différentes suites dans le cas des exemples cités plus haut. Vous vous apercevrez qu'elles ne convergent pas toujours et qu'il faut donc choisir soigneusement la fonction g . Le paragraphe suivant donne quelques conditions sur la fonction g pour que la suite converge.

Le principe des méthodes de point fixe

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

4.2.3 Convergence des méthodes de point fixe

Exercices :[Exercice 4.7](#)[Exercice 4.8](#)[Exercice 4.9](#)[Exercice 4.10](#)**Cours :**[Point fixe/ Principe](#)

Proposition 4.2.2. *Soit une fonction $g : [a, b] \rightarrow [a, b]$ continûment dérivable sur $[a, b]$ et supposons que*

$$|g'(x)| \leq k < 1, \forall x \in [a, b].$$

Alors g possède un point fixe unique $x^ \in [a, b]$ et la suite*

$$\begin{cases} x_0 \text{ donné dans } [a, b], \\ x_{n+1} = g(x_n), n > 0 \end{cases}$$

converge vers x^ .*

Démonstration - La première partie est une conséquence immédiate du théorème des valeurs intermédiaires. En effet, si l'on n'est pas dans les cas évidents $g(a) = a$ ou $g(b) = b$, la fonction $H(x) = g(x) - x$ est telle que $H(a) (= g(a) - a) > 0$ (car $g(a) \in [a, b]$) et $H(b) (= g(b) - b) < 0$ (car $g(b) \in [a, b]$). Il existe donc un point x^* de $]a, b[$ tel que $H(x^*) = 0$, c'est-à-dire un point fixe de g . On montre aisément en exercice que ce point fixe est unique.

[Sommaire](#)
[Concepts](#)[Exemples](#)
[Exercices](#)
[Documents](#)

Pour montrer la convergence de la suite vers x^* , on calcule

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

En itérant le processus, on trouve

$$|x_n - x^*| \leq k^n |x_0 - x^*|.$$

Puisque $k < 1$, la suite k^n tend vers 0 quand $n \rightarrow \infty$, ce qui démontre la convergence de la suite.

Remarquons que si l'on a une application $g : [a, b] \rightarrow \mathbb{R}$ telle que $|g'(x)| \geq k \geq 1$, $\forall x \in [a, b]$ et qui possède un point fixe dans $[a, b]$, alors la suite (x_n) précédemment définie est divergente au sens suivant. Soit x_n sort du domaine de définition de g , soit

$$|x_n - x^*| \geq k^n |x_0 - x^*|$$

et donc la suite (x_n) s'éloigne de x^* (au mieux elle reste à une distance constante si $k = 1$).

On dit que les méthodes de point fixe ont une convergence linéaire car

$$|x_{n+1} - x^*| = |g(x_n) - g(x^*)| \leq k|x_n - x^*|.$$

De même que pour la dichotomie, on peut estimer le nombre d'itérations nécessaires pour obtenir une précision donnée. En effet, vous pourrez montrer en exercice, que

$$|x_n - x^*| \leq \frac{k^n}{1-k} |x_1 - x_0|.$$

Convergence des méthodes de point fixe

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Donc, pour une précision ϵ donnée, le nombre d'itérations n est donné par

$$\frac{k^n}{1-k} |x_1 - x_0| \leq \epsilon \Rightarrow n \leq \frac{\ln \frac{(1-k)\epsilon}{|x_1 - x_0|}}{\ln k}.$$

Dans la proposition précédente, on a supposé que la valeur absolue de la dérivée était majorée par 1 sur l'intervalle $[a, b]$ tout entier. Si l'on considère l'exemple de la racine carrée d'un nombre réel a traité dans le paragraphe référencé, on a décrit trois fonctions g pour lesquelles le point fixe x^* correspond à cette racine carrée. Calculons les dérivées correspondantes

- $g_1(x) = \frac{a}{x}$, $g'_1(x) = -\frac{a}{x^2}$, $g'_1(x^*) = -1$,
- $g_2(x) = 2x - \frac{a}{x}$, $g'_2(x) = 2 + \frac{a}{x^2}$, $g'_2(x^*) = 3$,
- $g_3(x) = \frac{x}{2} + \frac{a}{x}$, $g'_3(x) = \frac{1}{2} - \frac{a}{2x^2}$, $g'_3(x^*) = 0$.

Proposition 4.2.3. Soit x^* un point fixe de g , fonction continûment dérivable, tel que $|g'(x^*)| < 1$, alors la suite (x_n) définie par $x_{n+1} = g(x_n)$ converge vers x^* si x_0 est suffisamment proche de x^* .

Démonstration - Puisque g' est continue, il existe un intervalle contenant x^* tel que $|g'(x)| < 1$. Plus précisément, il existe $\alpha > 0$ tel que

$$|g'(x)| < 1, \quad \forall x \in [x^* - \alpha, x^* + \alpha].$$

Il est alors facile de montrer que g est une application de $I = [x^* - \alpha, x^* + \alpha]$ dans lui-même. En effet si $x \in I$,

$$|g(x) - x^*| = |g(x) - g(x^*)| \leq \max_{x \in I} |g'(x)| |x - x^*| \leq \alpha.$$

Convergence des méthodes de point fixe

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

D'autre part, puisque la fonction $|g'|$ est continue, elle atteint son maximum sur l'intervalle I et donc $|g'(x)| \leq k < 1$, pour tout x de I . Ceci permet d'appliquer la proposition du paragraphe référencé et de montrer ainsi que la suite $x_{n+1} = g(x_n)$ converge vers x^* si x_0 est dans I , c'est-à-dire suffisamment proche de x^* .

Remarquons que si $|g'(x^*)| > 1$, il existe un intervalle $[x^* - \alpha, x^* + \alpha]$ sur lequel $|g'(x)| \geq k > 1$. La suite (x_n) définie précédemment a peu de chance de converger vers x^* . En effet si on part proche de x^* , on a

$$|x_1 - x^*| \geq k|x_0 - x^*| > |x_0 - x^*|$$

et x_1 s'éloigne de x^* . Il en sera ainsi jusqu'à ce que l'on sorte de l'intervalle $[x^* - \alpha, x^* + \alpha]$. Alors, comme on ne connaît pas le comportement de g , il se peut qu'un itéré postérieur "re-rentre" dans l'intervalle. La seule chance que l'on a de converger, c'est alors de "tomber" sur x^* . Cette méthode n'est donc pas à utiliser dans ce cas.

Convergence des méthodes de point fixe

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

4.2.4 La méthode de Newton pour résoudre une équation

Exercices :

[Exercice 4.11](#)

La méthode de Newton s'utilise pour calculer une racine de $f(x) = 0$, où f est une fonction deux fois continûment dérivable de $[a, b]$ dans \mathbb{R} . En voici le principe : soit x_0 donné, on écrit le développement de Taylor en x_0 :

$$f(x) = f(x_0) + (x - x_0)f'(x_0) + (x - x_0)\epsilon(x - x_0).$$

La méthode de Newton consiste à négliger les termes d'ordre supérieur à 1 (ce qui revient à assimiler localement la courbe à sa tangente) et à considérer alors que x_1 est le point qui annule l'approximation de f ainsi obtenue, c'est-à-dire

$$0 = f(x_0) + (x_1 - x_0)f'(x_0),$$

ce qui donne

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

L'étape ainsi décrite qui permet de passer de x_0 à x_1 permet aussi de passer de x_n à x_{n+1} .

L'algorithme de Newton pour résoudre une équation $f(x) = 0$ est le suivant

$$\begin{cases} x_0 \text{ donné,} \\ x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \quad n \geq 0 \end{cases}$$

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Il s'agit donc d'une méthode de point fixe $x_{n+1} = g(x_n)$ pour laquelle g est définie par

$$g(x) = x - \frac{f(x)}{f'(x)}.$$

La méthode de Newton est une méthode de point fixe telle que $g'(x^*) = 0$ c'est-à-dire que la dérivée s'annule au point fixe de la fonction g .

Une autre manière de présenter la méthode de Newton, plus graphique, est d'obtenir x_{n+1} comme l'intersection de la tangente à la courbe $y = f(x)$ au point x_n avec l'axe Ox (voir l'exercice).

Si la méthode de point fixe converge vers x^* , alors on a en passant à la limite quand n tend vers $+\infty$:

$$x^* = x^* - \frac{f(x^*)}{f'(x^*)}.$$

On voit alors que l'on obtient bien $f(x^*) = 0$ si $f'(x^*) \neq 0$. On suppose dorénavant que cette condition est remplie c'est-à-dire que la courbe $y = f(x)$ n'est pas tangente à l'axe Ox lorsqu'elle coupe cet axe.

La méthode de Newton pour résoudre une équation

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

4.2.5 Convergence de la méthode de Newton pour résoudre une équation

Exercices :

[Exercice 4.12](#)

Proposition 4.2.4. Soit f une fonction deux fois continûment dérivable sur $[a, b]$ et $x^* \in [a, b]$ tel que $f(x^*) = 0$ et $f'(x^*) \neq 0$. Alors il existe $\alpha > 0$ tel que la suite générée par la méthode de Newton

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

converge vers x^* pour tout $x_0 \in [x^* - \alpha, x^* + \alpha]$.

Démonstration - Comme il s'agit d'une méthode de point fixe de la fonction $g(x) = x - \frac{f(x)}{f'(x)}$, on calcule la dérivée de cette fonction au point fixe

$$g'(x) = 1 - \frac{(f'(x))^2 - f''(x)f(x)}{f'(x)^2} = \frac{f''(x)f(x)}{(f'(x))^2}$$

et puisque $f(x^*) = 0$ on a $g'(x^*) = 0$. On peut donc appliquer la proposition 4.2.3 qui montre l'existence de α tel que la suite $x_{n+1} = g(x_n)$ converge vers le point fixe si $x_0 \in [x^* - \alpha, x^* + \alpha]$.

Cette proposition donne une condition suffisante de convergence mais heureusement la méthode de Newton converge souvent même si l'on part loin de la solution. Montrons maintenant que la vitesse de convergence est quadratique, ce qui est mieux que linéaire.

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Proposition 4.2.5. Soit g une fonction deux fois continûment dérivable de $[a, b]$ dans $[a, b]$, soit $x^* \in [a, b]$ tel que $g(x^*) = x^*$ et $g'(x^*) = 0$, soit $x_0 \in [a, b]$, on définit $x_{n+1} = g(x_n)$, alors :

$$|x_{n+1} - x^*| \leq \frac{M}{2} |x_n - x^*|^2$$

où $M = \max_{x \in [a, b]} |g''(x)|$.

Démonstration - Le développement de Taylor au point x^* est donné par

$$g(x_n) = g(x^*) + (x_n - x^*)g'(x^*) + \frac{(x_n - x^*)^2}{2} g''(\xi)$$

où ξ est strictement compris entre x_n et x^* . Etant données les hypothèses sur g , on obtient

$$|x_{n+1} - x^*| = \frac{(x_n - x^*)^2}{2} |g''(\xi)| \leq \frac{M}{2} |x_n - x^*|^2.$$

On dit alors que la vitesse de convergence est quadratique.

**Convergence de
la méthode de
Newton pour
résoudre une
équation**

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

4.2.6 La méthode de la sécante

Exercices :

[Exercice 4.13](#)

[Exercice 4.14](#)

Cours :

[Newton/ Convergence](#)

Cette méthode peut être vue comme une approximation de la méthode de Newton, présentée dans le paragraphe référencé. En effet la méthode de Newton nécessite le calcul de la dérivée de la fonction en un point à chaque itération. Si le calcul de la dérivée est coûteux en temps ou peu précis (par exemple c'est le résultat d'une approximation numérique), on peut approcher la dérivée de la manière suivante

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

On obtient ainsi la méthode de la sécante :

$$\begin{cases} x_0, x_1 \text{ donnés tels que } f(x_0)f(x_1) < 0, \\ x_{n+1} = x_n - \frac{f(x_n)(x_n - x_{n-1})}{f(x_n) - f(x_{n-1})}. \end{cases}$$

Ceci signifie, graphiquement, que l'on remplace la tangente par la "corde". Il s'agit en effet de l'intersection de la droite passant par les points $(x_n, f(x_n))$ et $(x_{n-1}, f(x_{n-1}))$ avec l'axe Ox (exercice).

Cette méthode n'est pas une méthode de point fixe puisque chaque point de la suite est construit à partir des deux précédents. La convergence de la suite ne peut être assurée que pour

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

certaines valeurs de x_0 et x_1 suffisamment proches de x^* , solution de $f(x^*) = 0$. L'étude de l'erreur est compliquée mais elle montre que

$$|x_{n+1} - x^*| \leq K|x_n - x^*|^{\frac{1+\sqrt{5}}{2}},$$

où K est une constante qui dépend de la fonction f et de la racine x^* . La vitesse de convergence, égale au nombre d'or, est plus grande que pour une convergence linéaire mais moins importante que pour une convergence quadratique puisque

$$1 < \frac{1 + \sqrt{5}}{2} < 2.$$

Cette méthode devrait donc converger moins vite que la méthode de Newton. Mais une itération de la sécante ne demande qu'une évaluation de f au point x_n (sauf pour la première itération) tandis que la méthode de Newton demande l'évaluation de f et f' . Donc, même si on suppose que l'évaluation de f et de f' a le même coût, une itération de Newton représente deux itérations de la sécante. Donc pour la comparaison des vitesses de convergence c'est $\sqrt{2} = 1,414$ qu'il faut comparer à $\frac{1+\sqrt{5}}{2} = 1,618$ et ainsi la comparaison tourne à l'avantage de la sécante.

Toutefois un inconvénient de la méthode de la sécante, telle qu'on l'a décrite, est que si deux points successifs sont sur une droite "presque parallèle" à l'axe Ox , l'intersection avec l'axe Ox part à l'"infini". Une variante de la méthode consiste à s'assurer que deux points successifs encadrent toujours la racine de f . On part donc de x_0 et x_1 qui encadrent x^* et pour calculer x_{n+1} on n'utilise x_n et x_{n-1} que s'ils entourent x^* . Sinon, on utilise x_n et x_{n-2} . On peut montrer alors que pour une fonction f dont la concavité a un signe constant autour de x^* , un des points servant à calculer x_{n+1} se fige (exercice) de sorte que, la méthode se comporte comme

$$x_{n+1} = x_n - \frac{f(x_n)(x_n - x_0)}{f(x_n) - f(x_0)}.$$

La méthode de la sécante

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

On retombe alors sur une méthode de point fixe qui a une vitesse de convergence linéaire. On obtient ainsi une convergence globale au détriment de la vitesse de convergence.

La méthode de la sécante

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

4.3 Résolution d'un système d'équations non-linéaires

4.3.1	La méthode de Newton pour deux équations non-linéaires	229
4.3.2	Les méthodes de point fixe	232
4.3.3	La méthode de Newton pour un système d'équations non-linéaires .	235

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

4.3.1 La méthode de Newton pour deux équations non-linéaires

Exercices :
[Exercice 4.15](#)

Cours :
[Newton/ Convergence](#)

On veut maintenant résoudre deux équations à deux inconnues :

$$f_1(x_1, x_2) = 0, \quad f_2(x_1, x_2) = 0.$$

On manipule maintenant des vecteurs de \mathbb{R}^2 que l'on peut noter

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad f(x) = \begin{pmatrix} f_1(x) \\ f_2(x) \end{pmatrix} = \begin{pmatrix} f_1(x_1, x_2) \\ f_2(x_1, x_2) \end{pmatrix}$$

Comme pour toutes les méthodes itératives que l'on a vues dans ce chapitre, on va construire une suite de vecteurs censée converger vers une solution de $f(x) = 0$, on notera les vecteurs de la suite $x^{(0)}, x^{(1)}, \dots, x^{(k)}, \dots$

L'approximation par un développement de Taylor au premier ordre que nous avons utilisée pour introduire la méthode de Newton pour une équation peut être étendue à deux équations voire à n équations. Explicitons-la pour 2 équations :

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Soit $x^{(0)}$ un couple donné. Ecrivons le développement de Taylor en ce point

$$\begin{cases} f_1(x) = f_1(x^{(0)}) + \frac{\partial f_1}{\partial x_1}(x^{(0)})(x_1 - x_1^{(0)}) + \frac{\partial f_1}{\partial x_2}(x^{(0)})(x_2 - x_2^{(0)}) + \text{reste}, \\ f_2(x) = f_2(x^{(0)}) + \frac{\partial f_2}{\partial x_1}(x^{(0)})(x_1 - x_1^{(0)}) + \frac{\partial f_2}{\partial x_2}(x^{(0)})(x_2 - x_2^{(0)}) + \text{reste}. \end{cases}$$

Si l'on note $Df(x^{(0)})$ la matrice jacobienne de f au point $x^{(0)}$ définie par :

$$Df(x^{(0)}) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x^{(0)}) & \frac{\partial f_1}{\partial x_2}(x^{(0)}) \\ \frac{\partial f_2}{\partial x_1}(x^{(0)}) & \frac{\partial f_2}{\partial x_2}(x^{(0)}) \end{pmatrix} = \begin{pmatrix} (\nabla f_1(x^{(0)}))^T \\ (\nabla f_2(x^{(0)}))^T \end{pmatrix},$$

avec

$$\nabla f_1(x^{(0)}) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x^{(0)}) \\ \frac{\partial f_1}{\partial x_2}(x^{(0)}) \end{pmatrix} \text{ et } \nabla f_2(x^{(0)}) = \begin{pmatrix} \frac{\partial f_2}{\partial x_1}(x^{(0)}) \\ \frac{\partial f_2}{\partial x_2}(x^{(0)}) \end{pmatrix},$$

les relations précédentes peuvent s'écrire :

$$f(x) = f(x^{(0)}) + Df(x^{(0)})(x - x^{(0)}) + \text{Reste},$$

soit encore :

$$f(x) = f(x^{(0)}) + \begin{pmatrix} (\nabla f_1(x^{(0)}))^T (x - x^{(0)}) \\ (\nabla f_2(x^{(0)}))^T (x - x^{(0)}) \end{pmatrix} + \text{Reste}.$$

Attention : $f(x)$, $f(x^{(0)})$, $x - x^{(0)}$, " Reste ", $\nabla f_1(x^{(0)})$, $\nabla f_2(x^{(0)})$ sont des vecteurs colonnes. $Df(x^{(0)})$ est une matrice. $Df(x^{(0)})(x - x^{(0)})$ est un vecteur colonne. Tous les produits et sommes sont donc bien cohérents.

**La méthode de
Newton pour
deux équations
non-linéaires**

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

On néglige le reste, on calcule l'itéré suivant $x^{(1)}$ en annulant l'approximation de f ainsi obtenue.

$$0 = f(x^{(0)}) + Df(x^{(0)})(x^{(1)} - x^{(0)})$$

Le vecteur $x^{(1)} - x^{(0)}$ se calcule en résolvant un système linéaire :

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

Quand on a obtenu le vecteur $h = x^{(1)} - x^{(0)}$, on calcule $x^{(1)} = h + x^{(0)}$.

On recommence les mêmes étapes pour obtenir le vecteur $x^{(2)}$ à partir du vecteur $x^{(1)}$. On s'arrête lorsque la différence de deux itérés successifs, c'est-à-dire le vecteur h est inférieur en norme à un epsilon fixé à l'avance.

La méthode de Newton pour résoudre le système de deux équations non-linéaires $f(x) = 0$ est donnée par :

On se donne $x^{(0)}$ dans \mathbb{R}^2 , puis pour $k \geq 0$

1. On résout le système linéaire $Df(x^{(k)})h = -f(x^{(k)})$.
2. On pose $x^{(k+1)} = x^{(k)} + h$.

Lorsqu'il s'agit de deux équations non-linéaires, il est facile d'inverser la matrice pour résoudre le système. Par contre lorsque l'on a n équations non-linéaires à résoudre, on utilise une méthode de Gauss pour résoudre le système linéaire qui permet de calculer la différence entre deux itérés successifs.

La méthode de Newton pour deux équations non-linéaires

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

4.3.2 Les méthodes de point fixe

Cours :[Point fixe/ Principe](#)[Point fixe/ Convergence](#)**Documents :**[Document 4.1](#)

Les résultats présentés dans ce paragraphe sont une généralisation des résultats en dimension n présentés dans les paragraphes référencés. g est une fonction de \mathbb{R}^n dans \mathbb{R}^n , x et $g(x)$ sont des vecteurs. Comme précédemment on construit une suite définie par :

$$\begin{cases} x^{(0)} \text{ donné,} \\ x^{(k+1)} = g(x^{(k)}). \end{cases}$$

Quand “tout se passe bien”, cette suite converge vers \hat{x} vérifiant $\hat{x} = g(\hat{x})$.

Nous allons donner deux théorèmes principaux dont vous trouverez les démonstrations en document.

Théorème 4.3.1. *Soit g une application de \mathbb{R}^n dans \mathbb{R}^n , et un réel λ , $0 \leq \lambda < 1$, tels que pour tout z et tout y dans \mathbb{R}^n on ait*

$$\|g(z) - g(y)\| \leq \lambda \|z - y\|. \quad (4.3.1)$$

Alors il existe un \hat{x} unique appartenant à \mathbb{R}^n tel que $\hat{x} = g(\hat{x})$ et la suite $\{x^{(k)}\}_{k \geq 0}$ définie par

$$\begin{cases} x^{(0)} \text{ donné,} \\ x^{(k+1)} = g(x^{(k)}) \end{cases} ,$$

[Sommaire](#)
[Concepts](#)[Exemples](#)
[Exercices](#)
[Documents](#)

converge vers \hat{x} .

Définition 4.3.2. Soit g une fonction de \mathbb{R}^n dans \mathbb{R}^n . On appelle matrice jacobienne de g la matrice $Dg(x) \in \mathcal{M}_{nn}(\mathbb{R})$ définie par

$$Dg(x) = \begin{pmatrix} \frac{\partial g_1}{\partial x_1}(x) & \dots & \frac{\partial g_1}{\partial x_n}(x) \\ \vdots & \ddots & \vdots \\ \frac{\partial g_n}{\partial x_1}(x) & \dots & \frac{\partial g_n}{\partial x_n}(x) \end{pmatrix},$$

où g_1, \dots, g_n désignent les composantes de g .

Définition 4.3.3. La fonction g est différentiable en x si g admet des dérivées partielles en x et si la matrice jacobienne $Dg(x)$ vérifie

$$g(y) = g(x) + Dg(x)(y - x) + \|y - x\|\varepsilon(y - x)$$

où $\varepsilon(y - x)$ est un vecteur de \mathbb{R}^n qui tend vers 0 quand y tend vers x .

Le théorème suivant dont la démonstration est délicate relie la convergence de la méthode de point fixe au rayon spectral de la matrice jacobienne :

Théorème 4.3.4. Soit g une application de \mathbb{R}^n dans \mathbb{R}^n qui admet un point fixe \hat{x} , on suppose que g est différentiable en \hat{x} et que $\rho[Dg(\hat{x})]$, le rayon spectral de la matrice jacobienne de g , vérifie $\rho[Dg(\hat{x})] < 1$. Alors il existe η tel que si x^0 vérifie $\|x^{(0)} - \hat{x}\| < \eta$, la suite définie par

$$x^{(k+1)} = g(x^{(k)})$$

converge vers \hat{x} .

Les méthodes de point fixe

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

On pourrait montrer que la convergence obtenue est *linéaire*, c'est-à-dire qu'il existe $C > 0$ tel que

$$\|x^{(k+1)} - \hat{x}\| \leq C \|x^{(k)} - \hat{x}\|.$$

On parle de convergence d'*ordre* n quand on a la majoration

$$\|x^{(k+1)} - \hat{x}\| \leq C \|x^{(k)} - \hat{x}\|^n.$$

Les méthodes de point fixe

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

4.3.3 La méthode de Newton pour un système d'équations non-linéaires

Cours :

[Newton/ Convergence](#)

[Newton/ Système 2 équations non-linéaires](#)

On cherche à résoudre le système d'équations non-linéaires $f(x) = 0$. Comme pour le cas $n = 1$, l'idée de la méthode de Newton consiste à considérer l'approximation affine de f en $x^{(k)}$, soit

$$f(x^{(k)} + h) = f(x^{(k)}) + Df(x^{(k)})h + \|h\|\varepsilon(h),$$

si f est différentiable, où on a noté $Df(x)$ la matrice jacobienne de f . On détermine alors h tel que $f(x^{(k)}) + Df(x^{(k)})h = 0$, puis $x^{(k+1)} = x^{(k)} + h$.

La méthode de Newton pour résoudre le système d'équations non-linéaires $f(x) = 0$ est donnée par :

On se donne $x^{(0)}$ dans \mathbb{R}^n , puis pour $k \geq 0$

1. On résout le système linéaire $Df(x^{(k)})h = -f(x^{(k)})$.
2. On pose $x^{(k+1)} = x^{(k)} + h$.

C'est ce que l'on a vu dans le paragraphe référencé qui traite du cas particulier de deux équations non-linéaires.

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Théorème 4.3.5. Soit f une application de \mathbb{R}^n dans \mathbb{R}^n . on cherche à résoudre $f(x) = 0$ par la méthode de Newton. S'il existe \hat{x} tel que

- $f(\hat{x})$ est nul,
- f est différentiable dans un voisinage de \hat{x} et $\|Df(x) - Df(\hat{x})\| \leq \alpha \|x - \hat{x}\|$,
- $Df(\hat{x})$ est inversible,

alors il existe un réel strictement positif η tel que si $x^{(0)}$ vérifie $\|x^{(0)} - \hat{x}\| < \eta$ la suite construite par la méthode de Newton converge vers \hat{x} . De plus la convergence obtenue est d'ordre 2, ou quadratique, c'est-à-dire qu'il existe une constante $C > 0$ telle que

$$\|x^{(k+1)} - \hat{x}\| \leq C \|x^{(k)} - \hat{x}\|^2.$$

Nous ne démontrerons pas ce résultat, qui est la "traduction" des hypothèses correspondantes dans le cas $n = 1$. On notera aussi que la convergence est très rapide (quadratique).

La méthode de Newton pour un système d'équations non-linéaires

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Documents du chapitre 4

4.1 Les théorèmes de point fixe pour les systèmes non-linéaires 238

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Document 4.1 Les théorèmes de point fixe pour les systèmes non-linéaires

On sait définir la convergence d'une suite à l'aide de sa limite. Si l'on ne connaît pas cette limite, on peut montrer qu'une suite est convergente si et seulement si c'est une suite de Cauchy, dont la définition suit.

Définition 4.3.6. On dit que la suite $\{x^{(k)}\}$ est une suite de Cauchy si

$$\forall \varepsilon > 0, \exists N(\varepsilon) \in \mathbb{N} \text{ tel que } \forall k \geq N(\varepsilon), \forall p \geq 0, \|x^{(k+p)} - x^{(k)}\| \leq \varepsilon,$$

Ceci nous permet d'énoncer quelques résultats sur les méthodes de point fixe.

Théorème 4.3.7. Soit g une application de \mathbb{R}^n dans \mathbb{R}^n , et un réel $\lambda, 0 \leq \lambda < 1$, tels que pour tout x et tout y dans \mathbb{R}^n on ait

$$\|g(x) - g(y)\| \leq \lambda \|x - y\|. \tag{4.3.2}$$

Alors il existe un \hat{x} unique appartenant à \mathbb{R}^n tel que $\hat{x} = g(\hat{x})$ et la suite $\{x^{(k)}\}_{k \geq 0}$ définie par

$$\begin{cases} x^{(0)} & \text{donné} \\ x^{(k+1)} & = g(x^{(k)}) \end{cases}$$

converge vers \hat{x} .

Démonstration - On a

$$\|x^{(k+1)} - x^{(k)}\| \leq \lambda \|x^{(k)} - x^{(k-1)}\| \leq \dots \leq \lambda^k \|x^{(1)} - x^{(0)}\|.$$

On a d'autre part

$$x^{(k+p)} - x^{(k)} = x^{(k+p)} - x^{(k+p-1)} + x^{(k+p-1)} - x^{(k+p-2)} + \dots + x^{(k+2)} - x^{(k+1)} + x^{(k+1)} - x^{(k)}.$$

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)



On peut donc écrire que

$$\begin{aligned} \|x^{(k+p)} - x^{(k)}\| &\leq \|x^{(k+p)} - x^{(k+p-1)}\| + \dots + \|x^{(k+2)} - x^{(k+1)}\| + \|x^{(k+1)} - x^{(k)}\|, \\ &\leq (\lambda^{k+p-1} + \dots + \lambda^{k+1} + \lambda^k) \|x^{(1)} - x^{(0)}\|, \\ &\leq \lambda^k (1 + \lambda + \dots + \lambda^{p-1}) \|x^{(1)} - x^{(0)}\| = \lambda^k \frac{1 - \lambda^p}{1 - \lambda} \|x^{(1)} - x^{(0)}\|, \\ &\leq \frac{\lambda^k}{1 - \lambda} \|x^{(1)} - x^{(0)}\|. \end{aligned}$$

La quantité $\|x^{(k+p)} - x^{(k)}\|$ tend donc vers zéro quand k tend vers l'infini ; ainsi la suite est une suite de Cauchy, de sorte qu'elle est convergente. On note \hat{x} la limite de cette suite. On a

$$\|g(x^{(k)}) - g(\hat{x})\| \leq \|x^{(k)} - \hat{x}\|,$$

donc la limite de $g(x^{(k)})$ est $g(\hat{x})$. On a bien $g(\hat{x}) = \hat{x}$ car

$$\hat{x} = \lim_{k \rightarrow \infty} x^{(k+1)} = \lim_{k \rightarrow \infty} g(x^{(k)}) = g(\hat{x}).$$

La suite ainsi construite converge donc vers une solution de $g(x) = x$. Montrons que cette équation admet une solution unique : on suppose qu'il existe 2 solutions u et v , on a donc $u = g(u)$, $v = g(v)$, donc $g(u) - g(v) = u - v$. Or $\|g(u) - g(v)\| \leq \lambda \|u - v\|$, on a donc

$$\|u - v\| \leq \lambda \|u - v\|, \quad \text{soit} \quad \|u - v\| (1 - \lambda) \leq 0$$

avec $\lambda < 1$. Ceci n'est possible que si $\|u - v\| = 0$ donc $u = v$. Ceci termine la démonstration du théorème .

Le théorème suivant dont la démonstration est délicate relie la convergence de la méthode de point fixe au rayon spectral de la matrice jacobienne :

Document 4.1

Les théorèmes de point fixe pour les systèmes non-linéaires

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Théorème 4.3.8. Soit g une application de \mathbb{R}^n dans \mathbb{R}^n qui admet un point fixe \hat{x} , on suppose que g est différentiable en \hat{x} et que $\rho[Dg(\hat{x})]$, le rayon spectral de la matrice jacobienne de g , vérifie $\rho[Dg(\hat{x})] < 1$. Alors il existe η tel que si x^0 vérifie $\|x^{(0)} - \hat{x}\| < \eta$, la suite définie par

$$x^{(k+1)} = g(x^{(k)})$$

converge vers \hat{x} .

Démonstration - Grâce à des résultats sur les normes matricielles, on peut montrer que, si l'on note

$$\hat{\varepsilon} = \frac{1 - \rho[Dg(\hat{x})]}{3} > 0,$$

il est possible de trouver une norme matricielle subordonnée telle que

$$0 \leq \|Dg(\hat{x})\| \leq \rho[Dg(\hat{x})] + \hat{\varepsilon},$$

on utilisera donc dans la suite cette norme matricielle et la norme vectorielle correspondante.

En utilisant la définition de la différentiabilité on obtient donc :

$$\begin{aligned} \|g(x) - g(\hat{x})\| &= \|Dg(\hat{x})(x - \hat{x}) + \|x - \hat{x}\|\varepsilon(x - \hat{x})\| \leq \|Dg(\hat{x})(x - \hat{x})\| + \|x - \hat{x}\|\|\varepsilon(x - \hat{x})\| \\ &\leq \|Dg(\hat{x})\|\|x - \hat{x}\| + \|x - \hat{x}\|\|\varepsilon(x - \hat{x})\| = \|x - \hat{x}\|(\|Dg(\hat{x})\| + \|\varepsilon(x - \hat{x})\|). \end{aligned}$$

Or $\|\varepsilon(x - \hat{x})\|$ tend vers 0 quand x tend vers \hat{x} , il existe donc η tel que

$$\|x - \hat{x}\| < \eta \implies \|\varepsilon(x - \hat{x})\| \leq \hat{\varepsilon}.$$

On obtient alors :

$$(\|Dg(\hat{x})\| + \|\varepsilon(x - \hat{x})\|) \leq \rho[Dg(\hat{x})] + \hat{\varepsilon} + \hat{\varepsilon} = \lambda.$$

Document 4.1

Les théorèmes de point fixe pour les systèmes non-linéaires

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Or $\lambda = \rho[Dg(\hat{x})] + 2\hat{\varepsilon} = 1 - \hat{\varepsilon} < 1$, on a donc démontré que :

$$\|x - \hat{x}\| < \eta \implies \|g(x) - g(\hat{x})\| \leq \lambda \|x - \hat{x}\|,$$

avec $\lambda < 1$. La suite de la démonstration est analogue à ce qui se passe pour $n = 1$.

[retour au cours](#)

Document 4.1

Les théorèmes de point fixe pour les systèmes non-linéaires

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercices du chapitre 4

4.1	243
4.2	244
4.3	245
4.4	246
4.5	247
4.6	248
4.7	249
4.8	250
4.9	251
4.10	252
4.11	253
4.12	254
4.13	255
4.14	256
4.15	257

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 4.1

On rappelle que la définition du déterminant à partir des permutations est la suivante

$$\det A = \sum_{\sigma} a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n}$$

où la somme est faite sur toutes les permutations de l'ensemble $\{1, 2, \dots, n\}$. En utilisant cette définition, montrer 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.

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 4.2

Donner les matrices D , E et F correspondant à la décomposition $A = D - E - F$ de la méthode de Jacobi dans le cas

$$A = \begin{pmatrix} 1 & 3 & 0 \\ 6 & 7 & -1 \\ 3 & 4 & 9 \end{pmatrix}.$$

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 4.3

Quelle est, pour la méthode de Gauss-Seidel, la forme de la matrice du système permettant de calculer $x^{(k+1)}$ à partir de $x^{(k)}$. Que pensez-vous de la résolution de ce système ?

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 4.4

Soit la décomposition $A = M - N$ avec M inversible et la méthode itérative

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

Donner une condition suffisante sur les matrices M et N pour que la suite $x^{(k)}$ converge vers la solution de $Ax = b$.

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 4.5

Montrer que pour la dichotomie, le nombre d'itérations nécessaires pour obtenir une précision inférieure à ϵ est supérieur ou égal à $\log_2 \frac{b-a}{\epsilon}$.

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 4.6

Pour calculer la racine carrée d'un nombre réel $a > 0$, on résout $x^2 = a$. Il existe alors différentes manières de se ramener à une méthode de point fixe :

-

$$x = \frac{a}{x} \text{ donc } g_1(x) = \frac{a}{x},$$

-

$$2x = x + \frac{a}{x}, \text{ soit } x = 2x - \frac{a}{x} \text{ donc } g_2(x) = 2x - \frac{a}{x},$$

-

$$x = \frac{x}{2} + \frac{a}{2x} \text{ donc } g_3(x) = \frac{x}{2} + \frac{a}{2x}.$$

Tracer les fonctions $g_i(x)$ et la bissectrice pour $x > 0$. Prendre u_0 quelconque et construire graphiquement la suite $u_{k+1} = g_i(u_k)$ pour $k > 0$. Conclusion ?

[retour au cours](#)

Solution

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 4.7

Soit $g : [a, b] \rightarrow [a, b]$, continument dérivable, telle que $|g'(x)| \leq k < 1, \forall x \in [a, b]$. On suppose que g admet deux points fixes dans $[a, b]$. Montrer, en utilisant le théorème des accroissements finis, que l'on arrive à une contradiction.

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 4.8

Pour calculer la racine x^* de $x^2 - 3x - 1 = 0$ sur $[-1, +1]$, on pose $g(x) = \frac{x^2 - 1}{3}$ pour appliquer une méthode de point fixe $x = g(x)$. Montrer que la suite $x_{n+1} = g(x_n)$, avec $x_0 \in [-1, +1]$, converge vers un unique point fixe $x^* \in [-1, 1]$.

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 4.9

Soit une suite (x_n) convergeant vers x^* . On suppose que

$$\exists k < 1, |x_{i+1} - x_i| \leq k|x_i - x_{i-1}| \forall i \in \mathbf{N}$$

1. Montrer que

$$|x_{n+1} - x_n| \leq k^n |x_1 - x_0|.$$

2. En déduire que, pour $p > n$,

$$|x_p - x_n| \leq (k^{p-1} + \dots + k^n) |x_1 - x_0|.$$

3. Après avoir calculé la somme du membre de droite, faites tendre p vers l'infini pour obtenir

$$|x^* - x_n| \leq \frac{k^n}{1 - k} |x_1 - x_0|.$$

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 4.10

On reprend la fonction $g(x) = \frac{x^2-1}{3}$ dont l'unique point fixe x^* de $[-1, 1]$ est racine du trinôme $x^2 - 3x - 1 = 0$. Soit $x_0 = 0$, calculer le nombre d'itérations de point fixe n nécessaires pour obtenir une précision inférieure à $\epsilon = 10^{-3}$ pour le calcul de x^* . Vérifier votre résultat en calculant les n premiers itérés de la méthode de point fixe.

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 4.11

Soit la courbe $y = f(x)$. Ecrire l'équation de la tangente à cette courbe au point d'abscisse $x = x_n$. Donner alors l'abscisse du point d'intersection de cette tangente avec l'axe Ox . Montrer que l'on retrouve ainsi l'itéré de la méthode de Newton. Pour illustrer graphiquement et numériquement cette méthode, tracer et calculer deux itérations avec $f(x) = x^2 - 2$ et $x_0 = 2$.

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 4.12

Soit la fonction $f(x) = \cos x - x$. Montrer qu'elle admet une racine sur $[0, \pi/2]$. Calculer trois itérations de la méthode de Newton en partant de $x_0 = \pi/4$ puis avec $x_0 = 0$.

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 4.13

Écrire l'équation de la droite passant par les points $(x_n, f(x_n))$ et $(x_{n-1}, f(x_{n-1}))$. Donner l'abscisse x_{n+1} de l'intersection de cette droite avec l'axe Ox. Montrer que l'on retrouve l'équation de la sécante.

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 4.14

Tracer une courbe convexe qui coupe l'axe Ox en un point x^* . Considérer deux points x_0 et x_1 qui entourent ce point. Effectuer alors graphiquement quelques itérations de la méthode de la sécante (modifiée) de telle sorte que la sécante passe toujours par deux points qui entourent x^* . Remarquez alors que soit x_0 , soit x_1 est pris en compte dans toutes les itérations.

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 4.15

On considère les deux équations non-linéaires

$$x_1^2 + (x_2 - 1)^2 = 4, \quad x_1^2 - x_2 + 1 = 0.$$

Ecrire une étape de la méthode de Newton. Prenez un vecteur $x^{(0)}$ et calculez le vecteur $x^{(1)}$.

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Chapitre 5

Interpolation

5.1	Motivations	259
5.2	Interpolation polynomiale	264
5.3	Interpolation avec les splines cubiques	282
	Exemples du chapitre 5	286
	Documents du chapitre 5	294
	Exercices du chapitre 5	297

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

5.1 Motivations

5.1.1	Approximation d'une fonction	260
5.1.2	Interpolation	262

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

5.1.1 Approximation d'une fonction

Exercices :

[Exercice 5.1](#)

Dans la résolution de certains problèmes numériques on rencontre la situation suivante : on veut calculer les valeurs d'une fonction $f(t)$ pour un très grand nombre de valeurs de t , mais on ne connaît pas f "explicitement". Ceci se produit lorsque :

- f n'est connue qu'en certains points expérimentaux t_0, t_1, \dots, t_n ,
- ou lorsque la fonction f est évaluée *numériquement* par un code de calcul dont l'exécution est coûteuse.

On veut alors "représenter" f par une fonction simple dont l'évaluation est aisée.

Autrement dit, il s'agit d'approcher la fonction f par une autre fonction, plus facile à calculer. Cette fonction approchée \tilde{f} est choisie dans une classe \tilde{F} de fonctions dont le calcul n'est pas trop coûteux. Citons l'ensemble des polynômes, l'ensemble des fractions rationnelles, l'ensemble des polynômes trigonométriques,...

La question est alors de savoir en quel sens on désire approcher f . Il y a un grand nombre de possibilités. Citons les plus courantes.

- On peut chercher une fonction $\tilde{f} \in \tilde{F}$ telle que :

$$\forall i = 0, 1, \dots, n, \quad \tilde{f}(t_i) = f(t_i).$$

C'est la technique de l'**interpolation**.

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

- On peut chercher une fonction $\tilde{f} \in \tilde{F}$ qui minimise l'écart entre les deux courbes aux points d'abscisse t_i ($0 \leq i \leq n$) :

$$\sum_{i=0}^n (\tilde{f}(t_i) - f(t_i))^2 = \min_{g \in \tilde{F}} \sum_{i=0}^n (g(t_i) - f(t_i))^2.$$

C'est la technique de la minimisation au sens des **moindres carrés** qui est étudiée dans un autre chapitre.

- On peut chercher une fonction $\tilde{f} \in \tilde{F}$ telle que

$$\int_a^b |f(x) - \tilde{f}(x)|^2 dx = \min_{g \in \tilde{F}} \int_a^b |f(x) - g(x)|^2 dx.$$

C'est la technique de l'**approximation quadratique**.

- Enfin, on peut chercher une fonction $\tilde{f} \in \tilde{F}$ telle que

$$\max_{x \in [a,b]} |f(x) - \tilde{f}(x)| = \min_{g \in \tilde{F}} \left(\max_{x \in [a,b]} |f(x) - g(x)| \right).$$

C'est la technique de l'**approximation uniforme**.

Dans la suite de ce chapitre, nous ne traitons que de l'interpolation.

Approximation d'une fonction

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

5.1.2 Interpolation

Exercices :

[Exercice 5.2](#)

Dans tout ce qui suit on suppose données deux familles de $n + 1$ réels :

- les abscisses : $\{t_0, t_1, \dots, t_n\}$, supposées **toutes distinctes**,
- les ordonnées correspondantes : $\{y_0, y_1, \dots, y_n\}$. Ces valeurs peuvent, par exemple, être les $\{f(t_i)\}$, où f est la fonction que l'on cherche à représenter.

On cherche alors une fonction h qui prend pour valeur y_i en chaque point t_i , ceci pour $i = 0, \dots, n$. Évidemment ce problème admet une infinité de solutions. Il faut donc préciser la classe \tilde{F} de fonctions dans laquelle on va chercher h .

Dans ce qui suit nous allons considérer deux classes particulières de fonctions :

- l'ensemble \mathcal{P}_n des polynômes de degré au plus n ,
- l'ensemble des splines cubiques qui sont des fonctions définies comme des “polynômes par morceaux”, chacun des polynômes étant de degré 3.

Il existe également d'autres méthodes d'interpolation. Citons celles qui utilisent des informations sur les dérivées de f aux points t_0, t_1, \dots, t_n . C'est l'**interpolation d'Hermite**, que nous ne verrons pas ici. Contentons-nous d'en donner un exemple très simple. Supposons connues les valeurs de f et de ses dérivées jusqu'à l'ordre n en un point t_0 . Alors, l'unique polynôme P_n de degré n vérifiant

$$\forall i = 0, 1, \dots, n, \quad P_n^{(i)}(t_0) = f^{(i)}(t_0),$$

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

est donné par

$$P_n(t) = f(t_0) + (t - t_0)f'(t_0) + \cdots + \frac{(t - t_0)^n}{n!} f^{(n)}(t_0).$$

Ce polynôme n'est rien d'autre en fait que la partie polynomiale du développement de Taylor de f à l'ordre n au point $t = t_0$.

Interpolation

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

5.2 Interpolation polynomiale

5.2.1	Polynôme d'interpolation dans la base de <i>Lagrange</i>	265
5.2.2	Calcul de l'interpolation	267
5.2.3	Polynôme d'interpolation dans la base de <i>Newton</i>	270
5.2.4	Différences divisées	272
5.2.5	Calcul pratique des différences divisées	276
5.2.6	Algorithme de Horner	278
5.2.7	Convergence du polynôme d'interpolation	280

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

5.2.1 Polynôme d'interpolation dans la base de *Lagrange*

Exercices :
[Exercice 5.3](#)

Exemples :
[Exemple 5.1](#)

Théorème 5.2.1. *Étant donnés $n + 1$ nombres réels (t_0, t_1, \dots, t_n) distincts et $n + 1$ nombres réels (y_0, y_1, \dots, y_n) , il existe un polynôme $p_n \in \mathcal{P}_n$ (ensemble des polynômes de degré $\leq n$) et un seul tel que*

$$\forall i = 0, 1, \dots, n, \quad p_n(t_i) = y_i. \quad (5.2.1)$$

Démonstration -

1. L'existence est démontrée directement en exhibant le polynôme cherché, soit

$$p_n(t) = \sum_{i=0}^n y_i \mathcal{L}_i(t) \quad (5.2.2)$$

où les polynômes \mathcal{L}_i sont les polynômes de base de **Lagrange** qui prennent la forme

$$\mathcal{L}_i(t) = \prod_{k=0, k \neq i}^n \frac{t - t_k}{t_i - t_k}. \quad (5.2.3)$$

Ces polynômes sont appelés polynômes de base, parce qu'ils engendrent l'espace \mathcal{P}_n des polynômes de degré inférieur ou égal à n . En outre, on les introduit plutôt que d'utiliser

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

la base canonique $\{1, t, t^2, \dots, t^n\}$, parce qu'ils permettent d'exprimer immédiatement tout polynôme de degré inférieur ou égal à n en fonction de ses valeurs aux points t_i . On vérifie bien en effet que, par construction même de $\mathcal{L}_i(t)$:

$$\mathcal{L}_i(t_j) = \begin{cases} 1, & \text{pour } j = i, \\ 0, & \text{pour } j \neq i. \end{cases}$$

2. L'unicité se montre de la façon suivante : supposons qu'il existe deux polynômes p_n et q_n appartenant à \mathcal{P}_n tels que $p_n(t_i) = q_n(t_i) = y_i$. Alors si on pose $d_n = p_n - q_n$, d_n appartient à \mathcal{P}_n et $d_n(t_i) = 0$, $i = 0, \dots, n$. Autrement dit, le polynôme d_n a au moins $n + 1$ racines distinctes. Il est de degré inférieur ou égal à n . Il est donc identiquement nul et l'on a ainsi $p_n = q_n$.

**Polynôme
d'interpolation
dans la base de
Lagrange**

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

5.2.2 Calcul de l'interpolation

Exercices :

[Exercice 5.4](#)

Soit $f : [a, b] \rightarrow \mathbb{R}$ une fonction donnée. On construit le polynôme $p_n(t)$ qui interpole les valeurs de f aux points t_0, t_1, \dots, t_n ($t_i \in [a, b]$), ce qui conduit à poser $p_n(t_i) = f(t_i)$ pour $i = 0, \dots, n$. On a ainsi approché la fonction f par le polynôme p_n . La question est alors la suivante : quelle erreur commet-on quand on approche f par p_n ?

Hypothèses et notations. Dans ce qui suit nous supposons que f est $(n + 1)$ fois continûment dérivable sur $[a, b]$ et nous noterons $e_n(t)$ l'erreur d'interpolation définie par :

$$e_n(t) = f(t) - p_n(t).$$

Soient alors $\text{Int}(t, t_0, \dots, t_n)$ le plus petit intervalle fermé contenant les points t, t_0, \dots, t_n et $\pi_n(t)$ la fonction définie par :

$$\pi_n(t) = (t - t_0)(t - t_1)(t - t_2) \dots (t - t_n).$$

Théorème 5.2.2. *Quel que soit $t \in [a, b]$, il existe $\xi_t \in \text{Int}(t, t_0, \dots, t_n)$, dépendant de t , tel que*

$$e_n(t) = \frac{1}{(n + 1)!} f^{(n+1)}(\xi_t) \pi_n(t). \quad (5.2.4)$$

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Démonstration : 1) S'il existe un j tel que $t = t_j$, l'erreur est nulle et la formule (5.2.4) est immédiate.

2) Si, quel que soit j , $t \neq t_j$, alors $\pi_n(t) \neq 0$, et nous pouvons définir, une fonction g_n de la variable z par :

$$g_n(z) = f(z) - p_n(z) - e_n(t) \frac{\pi_n(z)}{\pi_n(t)}.$$

Il est clair que $g_n(t) = 0$ et $g_n(t_j) = 0$ pour $j = 0, \dots, n$. Nous voyons ainsi que $g_n(z)$ s'annule $n+2$ fois dans l'intervalle $\text{Int}(t, t_0, \dots, t_n)$. En appliquant plusieurs fois le Théorème de Rolle¹ nous voyons de même que :

- g_n' s'annule au moins $n+1$ fois dans cet intervalle,
- g_n s'annule au moins n fois dans cet intervalle, etc
- $g_n^{(n+1)}$ s'annule au moins 1 fois dans l'intervalle $\text{Int}(t, t_0, \dots, t_n)$.

Notons alors $\xi_t \in \text{Int}(t, t_0, \dots, t_n)$ la (ou l'une des) valeur(s) de z telle que $g_n^{(n+1)}(\xi_t) = 0$. Comme la fonction g_n dépend de t , le point ξ_t aussi. Par ailleurs, comme p_n est un polynôme de degré n ,

$$p_n^{(n+1)}(z) = 0.$$

De même, comme le terme de plus haut degré de $\pi_n(t)$ est t^{n+1} nous avons :

$$\pi_n^{(n+1)}(z) = (n+1)!$$

de sorte que

$$g_n^{(n+1)}(\xi_t) = f^{(n+1)}(\xi_t) - e_n(t) \frac{(n+1)!}{\pi_n(t)} = 0,$$

ce qui démontre le théorème. □

1. Si f est continûment dérivable sur $[a, b]$ et si $f(a) = f(b)$ alors il existe au moins un réel $c \in]a, b[$ tel que $f'(c) = 0$.

Remarquons tout d'abord que cette expression de l'erreur n'a de sens que si la fonction f que l'on interpole est $n+1$ fois continûment dérivable. Si ce n'est pas le cas, l'expression précédente de l'erreur n'est plus définie. On peut d'ailleurs vérifier expérimentalement dans ce cas qu'il ne sert à rien d'interpoler f par un polynôme de degré n . C'est une règle générale en analyse numérique : il est inutile d'utiliser des méthodes d'ordre élevé, construites pour des fonctions régulières, lorsque la fonction à approcher n'a pas la régularité requise.

Remarquons aussi que l'expression de l'erreur obtenue dans le théorème ci-dessus est ce qu'on appelle une **estimation d'erreur**. Elle est précieuse en ce qu'elle donne la forme de l'erreur et montre comment varie (au ξ près) cette erreur quand on augmente le nombre de points d'interpolation. Par contre, elle ne permet pas de calculer la valeur de l'erreur $e(t)$ en un point t donné. En effet, le point ξ est inconnu (on sait seulement qu'il existe) et pire encore, on ne sait pas en général calculer la dérivée $(n+1)$ -ème de f . En effet, si l'on approche f c'est que l'on ne sait pas la calculer ou au moins que son calcul coûte très cher. Ceci exclut en général toute possibilité de calcul de $f^{(n+1)}$ à un coût abordable.

Calcul de l'interpolation

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

5.2.3 Polynôme d'interpolation dans la base de *Newton*

Exercices :[Exercice 5.5](#)[Exercice 5.6](#)

Nous avons introduit une nouvelle base de l'espace \mathcal{P}_n des polynômes de degré inférieur ou égal à n : la base de Lagrange. Cela nous a permis de démontrer l'existence d'un polynôme d'interpolation en donnant ses composantes dans cette base. Par contre, la base de Lagrange est très malcommode pour calculer les valeurs du polynôme d'interpolation en un ou plusieurs points. La base canonique n'étant pas plus adaptée, nous allons introduire une troisième base. Étant donnés $n + 1$ points t_0, t_1, \dots, t_n , la **base de Newton** est définie par :

$$\{ 1, (t - t_0), (t - t_0)(t - t_1), \dots, (t - t_0)(t - t_1) \dots (t - t_{n-1}) \}, \quad (5.2.5)$$

Nous allons montrer comment calculer les coefficients du polynôme de degré inférieur ou égal à n , p_n , interpolant une fonction f en les points t_0, t_1, \dots, t_n . Ce polynôme s'écrit :

$$p_n(t) = c_0 + c_1(t - t_0) + \dots + c_n(t - t_0)(t - t_1) \dots (t - t_{n-1}).$$

Notons que le coefficient c_n est le coefficient de t^n dans p_n . C'est ce qu'on appelle le *coefficient directeur* de p_n .

[Sommaire](#)
[Concepts](#)[Exemples](#)
[Exercices](#)
[Documents](#)

Un des avantages de l'expression du polynôme p_n exprimé dans la base de Newton est le suivant : si on note p_k les polynômes tronqués, c'est-à-dire :

$$p_k(t) = c_0 + c_1(t - t_0) + \cdots + c_k(t - t_0)(t - t_1) \cdots (t - t_{k-1}), k = 0, \dots, n,$$

alors p_k est le polynôme de degré inférieur ou égal à k qui interpole f aux points t_0, \dots, t_k , en effet on a

$$p_n(t) = p_k(t) + c_{k+1}(t - t_0)(t - t_1) \cdots (t - t_k) + \cdots + c_n(t - t_0)(t - t_1) \cdots (t - t_{n-1}).$$

donc $p_k(t_0) = p_n(t_0) = f(t_0)$, $p_k(t_1) = p_n(t_1) = f(t_1), \dots$, $p_k(t_k) = p_n(t_k) = f(t_k)$.

On a donc en particulier :

$$p_n(t) = p_{n-1}(t) + c_n(t - t_0)(t - t_1) \cdots (t - t_{n-1}),$$

donc si on connaît le polynôme p_{n-1} et que l'on rajoute un point d'interpolation t_n , les coefficients c_0, \dots, c_{n-1} sont les mêmes, il suffit de calculer le coefficient c_n et de rajouter au polynôme p_{n-1} le terme $c_n(t - t_0)(t - t_1) \cdots (t - t_{n-1})$.

Polynôme d'interpolation dans la base de Newton

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

5.2.4 Différences divisées

Exercices :

[Exercice 5.7](#)

[Exercice 5.8](#)

Documents :

[Document 5.1](#)

Définition 5.2.3. Soit une fonction f dont on connaît les valeurs en des points distincts t_0, t_1, t_2, \dots . On appelle **différence divisée d'ordre 0, 1, 2, ..., n** les expressions suivantes définies itérativement par :

$$\begin{aligned}f[t_0] &= f(t_0), f[t_1] = f(t_1), f[t_2] = f(t_2) \\f[t_0, t_1] &= \frac{f[t_0] - f[t_1]}{t_0 - t_1}, f[t_1, t_2] = \frac{f[t_1] - f[t_2]}{t_1 - t_2} \\f[t_0, t_1, t_2] &= \frac{f[t_0, t_1] - f[t_1, t_2]}{t_0 - t_2}.\end{aligned}$$

D'une manière générale, si X est un ensemble fini de points distincts si $a \notin X$ et $b \notin X$, si $a \neq b$ alors

$$f[a, X, b] = \frac{f[a, X] - f[X, b]}{a - b}.$$

Théorème 5.2.4. Soit une fonction f dont on connaît les valeurs en $n + 1$ points distincts t_i ($0 \leq i \leq n$) et soit p_n le polynôme d'interpolation :

$$p_n(t) = c_0 + c_1(t - t_0) + c_2(t - t_0)(t - t_1) + \dots + c_n(t - t_0)(t - t_1) \dots (t - t_{n-1}).$$

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

$$\forall k = 0, \dots, n, \quad f[t_0, t_1, \dots, t_k] = c_k \quad (5.2.6)$$

et

$$e_n(t) = f[t_0, t_1, \dots, t_n, t](t - t_0)(t - t_1)(t - t_2)\dots(t - t_n). \quad (5.2.7)$$

Démonstration - Elle se fait par récurrence sur le nombre de points d'interpolation.

- Si $n = 0$, p_0 est le polynôme constant qui interpole f en t_0 , donc

$$p_0(t) = f(t_0) = f[t_0].$$

- Si $n = 1$, p_1 est le polynôme de degré inférieur ou égal à 1 qui interpole f en t_0, t_1 , on écrit $p_1(t) = c_0 + c_1(t - t_0)$, $p_1(t_0) = f(t_0)$, $p_1(t_1) = f(t_1)$ et on obtient

$$c_0 = f(t_0) = f[t_0], c_1 = \frac{f(t_0) - f(t_1)}{t_0 - t_1} = f[t_0, t_1].$$

- Supposons que la formule soit exacte pour les polynômes qui interpolent en $k + 1$ points, on a par exemple p_k le polynôme de degré inférieur ou égal à k qui interpole f en t_0, \dots, t_k qui s'écrit :

$$p_k(t) = f[t_0] + f[t_0, t_1](t - t_0) + \dots + f[t_0, \dots, t_k](t - t_0)(t - t_1)\dots(t - t_{k-1}).$$

De même q_k le polynôme de degré inférieur ou égal à k qui interpole f en t_1, \dots, t_{k+1} s'écrit :

$$q_k(t) = f[t_1] + f[t_1, t_2](t - t_1) + \dots + f[t_1, \dots, t_{k+1}](t - t_1)(t - t_2)\dots(t - t_k).$$

Si on définit le polynôme p par :

$$p(t) = \frac{(t - t_{k+1})p_k(t) - (t - t_0)q_k(t)}{t_0 - t_{k+1}},$$

**Différences
divisées**

[Sommaire
Concepts](#)

[Exemples
Exercices
Documents](#)

alors $p = p_{k+1}$, en effet p est un polynôme de degré inférieur ou égal à $k + 1$ et de plus il interpole f aux points t_0, t_1, t_{k+1} , en effet :

$$p(t_0) = p_k(t_0) = f(t_0), p(t_{k+1}) = q_k(t_{k+1}) = f(t_{k+1}),$$

pour $i = 1, \dots, k$,

$$p(t_i) = \frac{(t_i - t_{k+1})p_k(t_i) - (t_i - t_0)q_k(t_i)}{t_0 - t_{k+1}} = \frac{(t_i - t_{k+1}) - (t_i - t_0)}{t_0 - t_{k+1}} f(t_i) = f(t_i).$$

Le coefficient directeur de p_k , c'est-à-dire le coefficient de t^k est $f[t_0, \dots, t_k]$, le coefficient directeur de q_k , c'est-à-dire le coefficient de t^k est $f[t_1, \dots, t_{k+1}]$, donc le coefficient directeur de p_{k+1} , c'est-à-dire le coefficient de t^{k+1} est

$$c_{k+1} = \frac{f[t_0, \dots, t_k] - f[t_1, \dots, t_{k+1}]}{t_0 - t_{k+1}} = f[t_0, \dots, t_{k+1}]$$

On sait d'autre part que

$$p_{k+1}(t) = p_k(t) + c_{k+1}(t - t_0)(t - t_1)\dots(t - t_k)$$

On a donc

$$p_{k+1}(t) = f[t_0] + f[t_0, t_1](t - t_0) + \dots + f[t_0, \dots, t_{k+1}](t - t_0)(t - t_1)\dots(t - t_k).$$

Ce qui termine la démonstration.

Quant au calcul de l'erreur, il est détaillé dans l'exercice (5.7). ◻

On pourra voir dans le document le calcul des différences divisées lorsque les points sont équadistants.

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Proposition 5.2.1. *Les différences divisées $f[t_0, t_1, \dots, t_k]$ sont des fonctions symétriques de leurs arguments.*

Démonstration - Le nombre $f[t_0, t_1, \dots, t_k] = c_k$ est le coefficient du terme en t^k dans le polynôme de degré inférieur ou égal à k qui interpole f aux points $\{t_0, t_1, \dots, t_k\}$. Or ce polynôme est unique et ne dépend pas de l'ordre dans lequel on a numéroté les points, de sorte que :

$$f[t_{\sigma(0)}, t_{\sigma(1)}, \dots, t_{\sigma(k)}] = f[t_0, t_1, \dots, t_k],$$

pour toute permutation σ des entiers $\{0, 1, 2, \dots, k\}$.

Différences divisées

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

5.2.5 Calcul pratique des différences divisées

Exercices :

[Exercice 5.9](#)

[Exercice 5.10](#)

Rappelons l'expression des différences divisées données dans la définition [5.2.3](#) :

$$f[t_i] = f(t_i), \quad f[t_0, t_1, \dots, t_k] = \frac{f[t_0, t_1, \dots, t_{k-1}] - f[t_1, t_2, \dots, t_k]}{t_0 - t_k}.$$

L'utilisation de cette formule permet de calculer les $f[t_0, t_1, \dots, t_k]$ de proche en proche à l'aide du tableau suivant (chaque colonne se déduit de la colonne précédente) :

$k = 0$	$k = 1$	$k = 2$...	$k = n$
$f[t_0]$				
$f[t_1]$	$f[t_0, t_1]$			
$f[t_2]$	$f[t_1, t_2]$	$f[t_0, t_1, t_2]$		
\vdots	\vdots	\vdots	\ddots	
$f[t_n]$	$f[t_{n-1}, t_n]$	$f[t_{n-2}, t_{n-1}, t_n]$...	$f[t_0, t_1, \dots, t_n]$

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

- Nous voyons que la détermination des coefficients c_k nécessite $\frac{n(n+1)}{2}$ divisions.
- Nous voyons aussi que nous obtenons sur la diagonale du tableau les coefficients $f[t_0]$, $f[t_0, t_1]$, $f[t_0, t_1, t_2]$, et finalement $f[t_0, t_1, \dots, t_n]$, c'est-à-dire les coefficients c_0, c_1, c_2 et finalement c_n de p_n dans la base de Newton.

Voici un algorithme succinct pour calculer les différences divisées d_i ($0 \leq i \leq n$) :

1: **pour** $i = 0$ jusqu'à n **faire**

2: $d_i = f(t_i)$

3: **fin pour**

4: **pour** $k = 1$ jusqu'à n **faire**

5: **pour** $i = n$ jusqu'à k par pas de -1 **faire**

6: $d_i = \frac{d_i - d_{i-1}}{t_i - t_{i-k}}$

7: **fin pour**

8: **fin pour**

Remarquer que les éléments d_i en fin d'algorithme correspondent bien aux coefficients c_i du polynôme d'interpolation écrit dans la base de Newton.

Calcul pratique des différences divisées

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

5.2.6 Algorithme de Horner

Exercices :

[Exercice 5.11](#)

Il faut remarquer tout d'abord que la forme

$$p_n(t) = c_0 + c_1(t - t_0) + c_2(t - t_0)(t - t_1) + \cdots + c_n(t - t_0)(t - t_1) \dots (t - t_{n-1}).$$

se prête bien à l'évaluation numérique en utilisant le **schéma de Horner** :

- 1: Les données sont : $c_0, \dots, c_n, t_0, \dots, t_{n-1}, t$
- 2: $p = c_n$
- 3: **pour** $k = n - 1$ jusqu'à 0 par pas de -1 **faire**
- 4: $p = c_k + (t - t_k) * p$
- 5: **fin pour**

Par exemple, pour $n=3$, on évalue successivement chacun des crochets en partant du plus interne.

$$p_3(t) = c_0 + (t - t_0) [c_1 + (t - t_1) [c_2 + (t - t_2) [c_3]]]. \quad (5.2.8)$$

L'écriture classique donnerait

$$p_3(t) = c_0 + c_1(t - t_0) + c_2(t - t_0)(t - t_1) + c_3(t - t_0)(t - t_1)(t - t_2)$$

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

ce qui correspond à 5 multiplications.

En exercice, vous montrerez que, de manière générale, pour le calcul de p_n , l'algorithme de Horner comporte n multiplications alors que le calcul "classique" en comporte $2n - 1$.

Algorithme de Horner

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

5.2.7 Convergence du polynôme d'interpolation

Exercices :

[Exercice 5.12](#)

L'expression de l'erreur :

$$e_n(t) = \frac{1}{(n+1)!} \pi_n(t) f^{(n+1)}(\xi),$$

obtenue précédemment nous permet déjà, dans quelques cas favorables, d'obtenir des résultats intéressants.

Soit donc $[a, b]$ un intervalle contenant tous les points d'interpolation t_i . Il vient :

$$E_n = \sup_{t \in [a, b]} |e_n(t)| \leq \frac{1}{(n+1)!} \sup_{t \in [a, b]} |\pi_n(t)| \sup_{t \in [a, b]} |f^{(n+1)}(t)|,$$

d'où découle la majoration :

$$E_n \leq \frac{(b-a)^{n+1}}{(n+1)!} \sup_{t \in [a, b]} |f^{(n+1)}(t)|$$

Comme

$$\lim_{n \rightarrow \infty} \frac{(b-a)^{n+1}}{(n+1)!} = 0,$$

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

il en résulte que $\lim_{n \rightarrow \infty} E_n = 0$, pour toutes les fonctions $f \in C^\infty([a, b])$, dont les dérivées sont bornées uniformément en n . Cette classe de fonctions est assez restreinte, mais contient des fonctions comme e^t , $\sin t$, $\cos t$.

Par contre, l'exercice (5.12) montre que pour une fonction aussi raisonnable que

$$f(t) = \frac{1}{1+t^2}$$

interpolée sur l'intervalle $[-5, 5]$, le polynôme d'interpolation ne converge pas uniformément vers f .

Convergence du polynôme d'interpolation

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

5.3 Interpolation avec les splines cubiques

5.3.1	Position du problème	283
5.3.2	Définition de la spline cubique d'interpolation	285

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

5.3.1 Position du problème

Exemples :

[Exemple 5.2](#)

L'interpolation polynômiale présente plusieurs inconvénients :

- elle ne prend pas très bien en compte les fonctions assez "raides" ,
- elle converge mal en ce sens qu'augmenter le nombre de points n'améliore guère l'erreur au delà d'un certain degré.

Qualitativement, ces difficultés proviennent de la trop grande rigidité des polynômes. Du fait de cette rigidité, le saut brusque de 0 à 1 engendre des oscillations sur tout son intervalle de définition. Il faut donc avoir recours à une approximation utilisant des fonctions approchées (les éléments de \tilde{F}) moins rigides. En d'autres termes, on demandera moins de régularité **globale** aux fonctions de \tilde{F} . Ceci, outre les propriétés sympathiques des polynômes, va nous conduire à utiliser dorénavant des **polynômes par morceaux**.

Concrètement, dans ce qui suit on considère une partition de l'intervalle $[a, b]$

$$\Delta = \{ a = t_0 < t_1 < t_2 < \dots < t_{n-1} < t_n = b \},$$

dont on appellera les points t_i de subdivision les **nœuds**. On suppose donnée une famille de nombres réels $\{y_0, \dots, y_n\}$ qui correspondent, comme précédemment, aux valeurs d'interpolation. Une idée assez naturelle consiste à considérer individuellement chacun des sous-intervalles

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

$[t_i, t_{i+1}]$ et à construire une approximation **par morceaux** : on obtient ainsi une fonction g vérifiant les conditions d'interpolation $g(t_i) = y_i$ et qui dans chaque intervalle $[t_i, t_{i+1}]$ coïncide avec un polynôme $g_i(t)$ de degré plus ou moins élevé. Dans le cas où l'on choisit $g_i(t)$ de degré 1, cette technique consiste à relier les points (t_i, y_i) par des segments de droite : il s'agit de la façon la plus élémentaire d'utiliser ce qu'on appelle les fonctions **splines**. Dans ce cas la courbe d'équation $y = g(t)$ est une ligne brisée, on a donc des points anguleux, la fonction g n'est pas dérivable aux abscisses t_i . On sent bien que si l'on prend des polynômes de degré plus élevé, le choix de $g_i(t)$ n'est plus unique. L'idée consiste à se servir de ces degrés de liberté supplémentaires pour assurer à g une certaine **régularité** aux points t_i . On pourra par exemple imposer à g d'être une ou plusieurs fois continûment dérivable. Dans le cadre de ce cours nous nous limiterons au cas où ces polynômes sont de degré 3.

Position du problème

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

5.3.2 Définition de la spline cubique d'interpolation

Exercices :
[Exercice 5.13](#)

Exemples :
[Exemple 5.3](#)

Définition 5.3.1. *Etant donné une partition de l'intervalle $[a, b]$*

$$\Delta = \{ a = t_0 < t_1 < t_2 < \cdots < t_{n-1} < t_n = b \},$$

*On appelle **spline cubique** d'interpolation une fonction notée g , qui vérifie les propriétés suivantes :*

- $g \in C^2[a, b]$ (g est deux fois continûment dérivable),
- g correspond, sur chaque intervalle $[t_i, t_{i+1}]$, à un polynôme de degré inférieur ou égal à 3,
- $g(t_i) = y_i$ pour $i = 0, 1, \dots, n$.

Remarquons que la définition qui précède ne suffit pas à déterminer de façon unique la fonction spline d'interpolation car il manque des conditions. Comme on va le voir la détermination de la spline nécessite la résolution d'un système linéaire et il manque encore deux équations pour déterminer cette fonction. Ces deux équations proviendront de conditions supplémentaires comme par exemple

$$g''(a) = g''(b) = 0. \tag{5.3.1}$$

Ces deux conditions, rajoutées à celles de la définition (5.3.1), définissent ce qu'on appelle une spline **naturelle**.

Vous verrez dans l'exercice comment construire les splines cubiques naturelles.

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exemples du chapitre 5

5.1	287
5.2	290
5.3	292

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exemple 5.1

Voici un exemple pour $n = 3$. On a choisi $t_i = i$, et on a

$$\mathcal{L}_0(t) = \frac{(t-1)(t-2)(t-3)}{-6}, \quad \mathcal{L}_1(t) = \frac{t(t-2)(t-3)}{2},$$

$$\mathcal{L}_2(t) = \frac{t(t-1)(t-3)}{-2}, \quad \mathcal{L}_3(t) = \frac{t(t-1)(t-2)}{6}.$$

La figure 5.3.1 montre les deux premiers de ces polynômes.

Prenons les données suivantes : $y_0 = \frac{1}{2}$, $y_1 = 1$, $y_2 = 2$, $y_3 = -\frac{1}{2}$. On obtient le polynôme d'interpolation

$$p(t) = -\frac{(t-1)(t-2)(t-3)}{12} + \frac{t(t-2)(t-3)}{2} \quad (5.3.2)$$

$$- t(t-1)(t-3) - \frac{t(t-1)(t-2)}{12}, \quad (5.3.3)$$

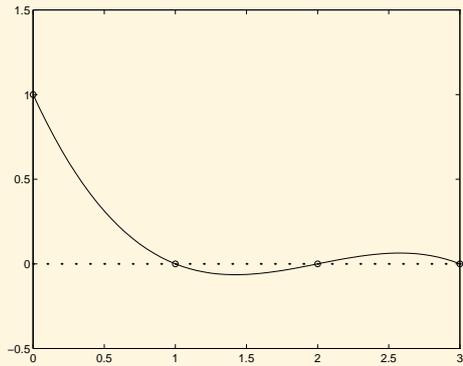
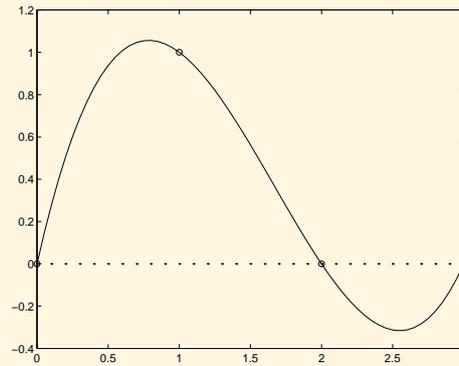
représenté sur la figure 5.3.2.

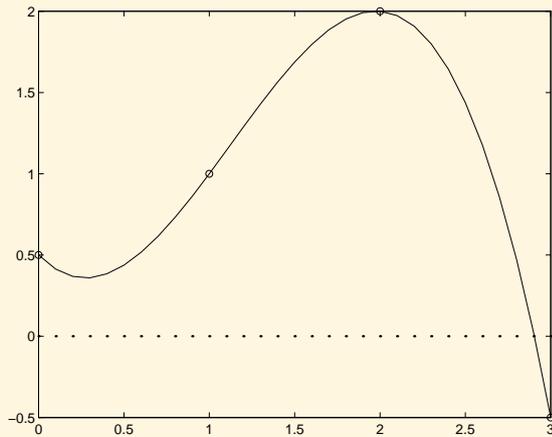
[retour au cours](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exemple 5.1

 $\mathcal{L}_0(t)$  $\mathcal{L}_1(t)$ FIGURE 5.3.1 – les polynômes de Lagrange \mathcal{L}_0 et \mathcal{L}_1 .[Sommaire](#)
[Concepts](#)[Exemples](#)
[Exercices](#)
[Documents](#)

$p(t)$ **Exemple 5.1**[Sommaire](#)
[Concepts](#)[Exemples](#)
[Exercices](#)
[Documents](#)

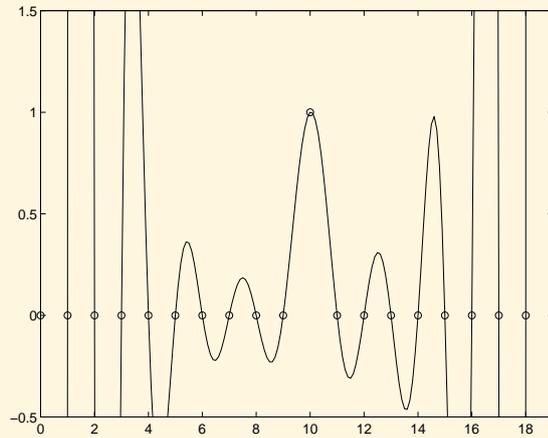
Exemple 5.2

On peut se faire une idée des problèmes que l'on peut rencontrer en regardant sur la figure 5.3.3 le polynôme de Lagrange \mathcal{L}_{10} construit sur 20 points équidistants $t_i = i$, $i = 0 \dots 19$: ce polynôme présente bien les propriétés requises, à savoir $\mathcal{L}_{10}(t_{10}) = 1$ et $\mathcal{L}_{10}(t_i) = 0$, $i \neq 10$, mais son comportement *entre les points d'interpolation* présente des oscillations d'amplitude très importante, laissant présager des oscillations pour le polynôme d'interpolation lui-même.

[retour au cours](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exemple 5.2[Sommaire](#)
[Concepts](#)[Exemples](#)
[Exercices](#)
[Documents](#)

Exemple 5.3

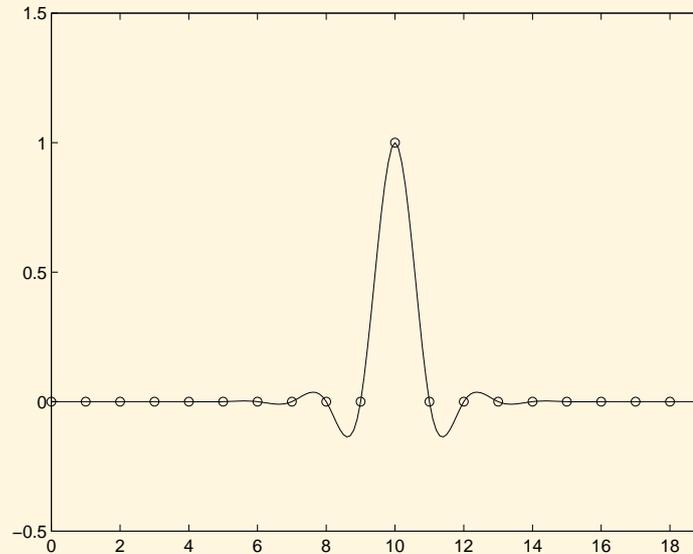


FIGURE 5.3.4 – La spline cubique $g(t)$.

On peut regarder ce que donne une spline cubique sur l'exemple suivant : $n = 19$, $t_i = i$, $i = 0 \dots 19$ et $y_{10} = 1$, $y_i = 0$ pour $i \neq 10$. On essaye en fait de construire une fonction $g(t)$ similaire au

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

polynôme de Lagrange $\mathcal{L}_{10}(t)$. La figure 5.3.4, à comparer avec la figure 5.3.3, montre l'efficacité des splines par rapport à l'interpolation polynomiale.

[retour au cours](#)

Exemple 5.3

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Documents du chapitre 5

5.1 [Analogie des expressions des polynômes d'interpolation et de Taylor](#) 295

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Document 5.1 Analogie des expressions des polynômes d'interpolation et de Taylor

L'expression de la partie polynomiale du développement de Taylor d'une fonction f suffisamment dérivable est classique. Elle s'écrit :

$$P(t) = P(t_0) + (t - t_0)f'(t_0) + \dots + \frac{(t - t_0)^n}{n!} f^{(n)}(t_0).$$

Nous avons vu que le polynôme d'interpolation admet l'expression :

$$P(t) = P(t_0) + (t - t_0)f[t, t_0] + \dots + (t - t_0)(t - t_1) \cdots (t - t_{n-1})f[t_0, t_1, \dots, t_n].$$

Nous allons voir que, dans le cas de points d'interpolation équidistants, il est possible d'obtenir une expression de ce polynôme ressemblant encore plus à celle du polynôme de Taylor. En effet, on montre que, en posant

$$h = t_1 - t_0 = t_2 - t_1 = \dots = t_n - t_{n-1}, \quad s = \frac{(t - t_0)}{h}$$

$$\Delta_1 f_i = f(t_{i+1}) - f(t_i), \quad \Delta_k f_i = \Delta_{k-1} f(t_{i+k}) - \Delta_{k-1} f(t_i),$$

il vient (à démontrer)

$$(t - t_0)(t - t_1) \cdots (t - t_{k-1})f[t_0, t_1, \dots, t_k] = \frac{s(s-1) \cdots (s-k+1)}{k!} \Delta_k f_0$$

d'où l'on déduit (à démontrer) l'expression du polynôme d'interpolation :

$$P(t) = f_0 + s\Delta f_0 + \frac{s(s-1)}{2!} \Delta_2 f_0 + \dots + \frac{s(s-1) \cdots (s-n+1)}{n!} \Delta_n f_0.$$

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)



Si du fait de contacts suffisants avec les sciences physiques ou mécaniques, on a été sensibilisé au problème des dimensions des diverses grandeurs rencontrées dans ses équations, on sera très attentif ici.

Les différences $(t - t_i)$ ont la dimension d'une longueur par exemple (ou d'un temps ou toute autre peu importe). Ainsi, dans la formule de Taylor, la dimension de $(t - t_0)^k$ est compensée par les k quotients par $t - t_0$ introduits dans la dérivée k -ème de f .

Par contre, les grandeurs $s, s - 1, \dots$ sont adimensionnelles. C'est pourquoi apparaissent des différences $\Delta_k f_0$ qui ne font intervenir aucune division.

[retour au cours](#)

Document 5.1

Analogie des expressions des polynômes d'interpolation et de Taylor

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercices du chapitre 5

5.1	298
5.2	299
5.3	300
5.4	301
5.5	302
5.6	303
5.7	304
5.8	305
5.9	306
5.10	307
5.11	308
5.12	309
5.13	310

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 5.1

Soit f une fonction connue aux points d'abscisse t_i ($0 \leq i \leq n$), supposées toutes distinctes. Soit l'ensemble \mathcal{P}_n des polynômes de degré inférieur ou égal à n . Tout polynôme p de \mathcal{P}_n peut s'écrire

$$p(t) = a_0 + a_1 t + \dots + a_n t^n.$$

On cherche $p \in \mathcal{P}_n$ tel que $p(t_i) = f(t_i), i = 0, \dots, n$.

Ecrire les conditions d'interpolation, montrer que le système linéaire obtenu admet une solution unique.

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 5.2

On suppose la fonction f connue aux points $\{-1, 0, 1\}$ où elle prend les valeurs $\{0, 1, 0\}$ et soit \mathcal{P}_m l'ensemble des polynômes de degré au plus m . Quelle est la valeur minimale de m qui conduit à une technique d'interpolation? Pour quelle valeur de m le polynôme d'interpolation est unique?

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 5.3

On considère les points du plan $\{(t_i, z_i), 0 \leq i \leq n\}$ avec $t_0 < t_1 < \dots < t_n$.

1. Écrire l'équation de la droite passant par les points (t_i, z_i) et (t_{i+1}, z_{i+1}) en utilisant la base de Lagrange.
2. Écrire l'équation de la ligne brisée qui interpole tous les points.

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 5.4

Soit n un entier naturel.

1. Calculer l'erreur commise en interpolant la fonction $f(t) = t^n$, définie sur l'intervalle $[0, 1]$, en les points $t_i = i/n$, $i = 0, 1, \dots, n$, à l'aide du polynôme d'interpolation de Lagrange de degré n . Expliquer le résultat.
2. Même question pour la fonction $g(t) = t^{n+1}$.

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 5.5

Montrer que les polynômes

$$1, (t - t_0), (t - t_0)(t - t_1), \dots, (t - t_0)(t - t_1) \cdots (t - t_{n-1}),$$

forment une base de \mathcal{P}_n , ensemble des polynômes de degré inférieur ou égal à n .

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 5.6

On a $t_0 = 1, t_1 = 2, t_2 = 3, f(t_0) = 1, f(t_1) = 3, f(t_2) = 4$.

Ecrire le polynome p_0 de degré 0 qui interpole f en t_0 .

Ecrire le polynome p_1 de degré 1 qui interpole f en t_0, t_1 .

Ecrire le polynome p_2 de degré 2 qui interpole f en t_0, t_1, t_2 .

Ecrire chacun des polynômes dans la base de Newton.

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 5.7

Soit $\theta \in \mathbb{R}$ donné, soit p_n le polynôme de degré inférieur ou égal à n qui interpole f en t_0, t_1, \dots, t_n , on veut évaluer l'erreur en θ , c'est-à-dire $e_n(\theta) = f(\theta) - p_n(\theta)$. Si θ est égal à l'un des t_i , l'erreur est nulle.

Supposons maintenant que $\theta \neq t_i, \forall i = 0, \dots, n$.

On définit alors le polynôme p par

$$p(t) = p_n(t) + \Pi_n(t) \frac{f(\theta) - p_n(\theta)}{\Pi_n(\theta)}$$

où $\Pi_n(t) = \prod_{i=0}^n (t - t_i)$.

1. Montrez que p interpole f aux points $\{t_0, t_1, \dots, t_n, \theta\}$. Quel est le degré de p ?
2. En déduire $p(t) - p_n(t)$ en fonction de $f[t_0, t_1, \dots, t_n, t]$.
3. En déduire le calcul de l'erreur $e_n(\theta) = f(\theta) - p_n(\theta)$.

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 5.8

Soit la fonction f connue aux trois points d'abscisse t_0 , t_1 et t_2 . On considère le polynôme d'interpolation dans la base de Newton avec les notations du cours. Montrer, par le calcul, que

$$c_2 = f[t_0, t_1, t_2]$$

en utilisant la définition et la symétrie des différences divisées.

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 5.9

Calculer les coefficients c_k du polynôme d'interpolation p_3 de l'exemple 5.1 , dans la base de Newton.

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 5.10

On a calculé à l'aide des différences divisées le polynôme p_n d'interpolation de f aux points $\{t_0, \dots, t_n\}$. On désire rajouter un point d'interpolation t_{n+1} . Doit-on refaire tout le tableau des différences divisées ? Et, si on utilisait la base des polynômes de Lagrange, devrait-on refaire tous les calculs ?

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 5.11

Calculer le nombre d'opérations arithmétiques nécessaires pour évaluer, en un point t , la valeur de

$$p_n(t) = c_0 + c_1(t - t_0) + c_2(t - t_0)(t - t_1) + \cdots + c_n(t - t_0)(t - t_1) \dots (t - t_{n-1}).$$

1. Par la méthode 'naturelle', en écrivant l'algorithme
2. Par le schéma de Horner.

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 5.12

Mettre en évidence expérimentalement, en utilisant un logiciel de calcul (Matlab, Scilab,...) les difficultés de l'interpolation polynomiale de la fonction $1/(1+t^2)$ sur l'intervalle $[-5, +5]$.

[retour au cours](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 5.13

1. Compter le nombre de degrés de liberté à déterminer pour définir complètement une spline cubique.
2. Compter le nombre d'équations disponibles pour ce faire. Comparer.

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Chapitre 6

Intégration

6.1	Motivations et principe des méthodes numériques d'intégration	312
6.1	Méthodes utilisant le polynôme d'interpolation	314
	Exemples du chapitre 6	325
	Documents du chapitre 6	331
	Exercices du chapitre 6	337

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

6.1 Motivations et principe des méthodes numériques d'intégration

Documents :

[Document 6.1](#)

Il existe deux situations où l'on a besoin de formules pour approcher l'intégrale d'une fonction f :

$$I(f) = \int_a^b f(t) dt$$

- On ne connaît la valeur de f qu'en certains points t_0, t_1, \dots, t_n , et il n'est pas possible d'avoir d'autres valeurs que celles-ci (c'est le cas quand la fonction f est tabulée).
- Il est possible de calculer $f(t)$ pour un t quelconque, mais la primitive de f n'est pas connue, ou bien l'expression analytique de f est trop compliquée pour être explicitée ($f(t)$ est par exemple le résultat d'un code de calcul trop complexe).

Dans ces deux situations, on doit donc chercher à approcher $I(f)$ à l'aide d'un nombre fini de valeurs de f en certains points t_0, t_1, \dots, t_n qui sont soit imposés, soit à choisir de façon optimale pour que l'approximation soit la meilleure possible. Plus précisément on aura recours à des combinaisons linéaires de valeurs de la fonction à intégrer en des points de l'intervalle d'intégration, soit

$$\int_a^b f(t) dt \simeq \sum_{i=0}^n w_i f(t_i). \quad (6.1)$$

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Les points t_i sont appelés les **nœuds** de la formule, les w_i sont les **poids** (weights, en anglais), parfois aussi dénommés **coefficients** de la formule. L'application

$$f \mapsto \sum_{i=0}^n w_i f(t_i)$$

qui à f fait correspondre une valeur 'approchée' de son intégrale sur l'intervalle considéré définit une **formule** (ou **méthode**) **d'intégration numérique**. On dit aussi formule ou méthode de **quadrature**. Quelques commentaires sont donnés en document.

**Motivations et
principe des
méthodes
numériques
d'intégration**

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

6.1 Méthodes utilisant le polynôme d'interpolation

6.1.1	Principe de la méthode	315
6.1.2	Intervalle de référence	317
6.1.3	Majoration de l'erreur de quadrature - ordre	319
6.1.4	Formules de Newton-Cotes	321
6.1.5	Intégration numérique composée	323

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

6.1.1 Principe de la méthode

Exercices :
[Exercice 6.1](#)

Exemples :
[Exemple 6.1](#)

Pour approcher l'intégrale $I(f)$ par une combinaison linéaire des valeurs de f en des points t_0, t_1, \dots, t_n , une première méthode consiste à approcher $I(f)$ par

$$I(f) = \int_a^b p_f(t) dt,$$

où $p_f(t)$ est le polynôme qui interpole f aux points distincts t_0, t_1, \dots, t_n . Ces points sont choisis le plus souvent dans l'intervalle $[a, b]$. Nous rencontrerons cependant, dans le chapitre sur les équations différentielles, des formules de quadrature utilisant des nœuds extérieurs à l'intervalle d'intégration. La justification en est très simple : cela permet d'utiliser les valeurs de la fonction à intégrer en des points où ces valeurs ont déjà été calculées.

Nous avons vu au chapitre précédent que le polynôme d'interpolation s'exprime dans la base de Lagrange par

$$p_f(t) = \sum_{i=0}^n f(t_i) \mathcal{L}_i(t).$$

Nous obtenons alors directement

$$I(f) = \sum_{i=0}^n w_i f(t_i), \text{ avec } w_i = \int_a^b \mathcal{L}_i(t) dt. \quad (6.1.1)$$

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

La formule (6.1.1) est appelée **formule de quadrature**.

De manière équivalente, nous pouvons aussi obtenir les coefficients w_i en écrivant que la formule doit être exacte pour tous les polynômes de degré inférieur ou égal à n , soit de façon équivalente pour chacun des monômes de la base canonique de \mathcal{P}_n . Cette équivalence se démontre facilement, voir l'exercice référencé.

On peut donc calculer les coefficients w_0, \dots, w_n , en écrivant que $I(p) = J(p)$ pour $p \in \{p_0, p_1, \dots, p_n\}$ où $\{p_0, p_1, \dots, p_n\}$ est la base canonique de \mathcal{P}_n . On obtient le système linéaire

$$\begin{bmatrix} 1 & 1 & \dots & 1 \\ t_0 & t_1 & \dots & t_n \\ t_0^2 & t_1^2 & \dots & t_n^2 \\ \vdots & \vdots & \vdots & \vdots \\ t_0^n & t_1^n & \dots & t_n^n \end{bmatrix} \begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ w_n \end{bmatrix} = \begin{bmatrix} v_0 \\ v_1 \\ \vdots \\ v_n \end{bmatrix}, \quad (6.1.2)$$

où

$$v_k = \int_a^b t^k dt.$$

La matrice de ce système est une matrice de **Vandermonde**. Son déterminant est non nul si et seulement si les points t_i sont deux à deux distincts. Le système (6.1.2) admet donc effectivement une solution et une seule.

L'exemple référencé vous permet de voir sur un cas particulier les deux méthodes pour calculer les w_i .

Principe de la méthode

[Sommaire Concepts](#)

[Exemples Exercices Documents](#)

6.1.2 Intervalle de référence

Exercices :
[Exercice 6.2](#)

Exemples :
[Exemple 6.2](#)

Une fois les nœuds t_i choisis, les poids w_i ne dépendent pas de la fonction f à intégrer. Par contre, ils dépendent de a et b . Nous allons voir maintenant, qu'il est possible de se ramener sans restriction aucune à approcher des intégrales sur un unique intervalle dit **intervalle de référence**. Nous choisirons l'intervalle $[-1, 1]$, car sa symétrie par rapport à 0 simplifie les calculs. Par contre, on aurait pu choisir tout autre intervalle, comme par exemple $[0, 1]$.

Nous allons donc ramener le calcul d'une intégrale sur un intervalle $[a, b]$ à celui d'une intégrale sur $[-1, 1]$, en utilisant un changement de variable affine. En effet si on pose

$$t = \frac{a + b + (b - a)u}{2},$$

on peut écrire que

$$\int_a^b f(t) dt = \frac{b - a}{2} \int_{-1}^1 g(u) du,$$

avec

$$g(u) = f\left(\frac{a + b + (b - a)u}{2}\right).$$

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Nous pouvons ainsi définir des formules de quadrature sur $[-1, 1]$, où ni les nœuds, ni les poids ne dépendent de a ou b . Ce sont ces formules que l'on retrouve dans les tables des livres consacrés à l'intégration numérique.

On ne retrouve pas une situation aussi simple quand on doit calculer des intégrales doubles ou triples. En effet, il n'existe pas de transformation de coordonnées *simple* permettant de passer d'un domaine bi ou tri-dimensionnel quelconque à un domaine de référence. On a alors des familles de formes de domaines (triangles, quadrangles, tétraèdres,...) et pour chacune de ces familles, un domaine de référence. Le passage de la dimension 1 à la dimension 2 ou 3 change le niveau de complexité du problème de la quadrature.

Intervalle de référence

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

6.1.3 Majoration de l'erreur de quadrature - ordre

Exercices :

[Exercice 6.3](#)

[Exercice 6.4](#)

[Exercice 6.5](#)

Exemples :

[Exemple 6.3](#)

L'erreur de quadrature est donnée par

$$E(f) = I(f) - J(f) = \int_a^b (f(t) - p_f(t)) dt.$$

Or si f est une fonction $n+1$ fois continûment dérivable et si p_f est son polynôme d'interpolation de degré n , nous avons vu dans le chapitre consacré à l'interpolation, que :

$$f(t) - p_f(t) = (t - t_0)(t - t_1) \dots (t - t_n) \frac{f^{(n+1)}(\xi(t))}{(n+1)!} = \pi_n(t) \frac{f^{(n+1)}(\xi(t))}{(n+1)!},$$

où $\xi(t) \in \text{Int}(t, t_0, \dots, t_n)$. On a donc

$$|E(f)| \leq \int_a^b |\pi_n(t)| \frac{|f^{(n+1)}(\xi(t))|}{(n+1)!} dt = \frac{|f^{(n+1)}(\eta)|}{(n+1)!} \int_a^b |\pi_n(t)| dt,$$

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

où $\eta \in \text{Int}(a, b, t_0, \dots, t_n)$ (on a appliqué le deuxième théorème de la moyenne¹). On peut encore majorer :

$$|E(f)| \leq \frac{M}{(n+1)!} \int_a^b |\pi_n(t)| dt,$$

où $M > 0$ est un majorant de $|f^{(n+1)}|$ sur $\text{Int}(a, b, t_0, \dots, t_n)$.

On retrouve bien que si f est un polynôme de degré inférieur ou égal à n , alors $E(f) = 0$, c'est à dire que l'approximation est exacte. Nous verrons plus loin, en particulier dans l'exemple et un exercice référencés ci-dessus, que l'on peut obtenir des estimations plus fines de l'erreur.

Définition 6.1.1. Une formule de quadrature exacte pour tous les polynômes de degré inférieur ou égal à n (n maximum) est dite **d'ordre n** ou de degré d'exactitude n .

1. Si f et g sont deux fonctions continues sur $[a, b]$ et que g ne change pas de signe sur $[a, b]$, alors il existe $c \in]a, b[$ tel que $\int_a^b f(x)g(x) dx = f(c) \int_a^b g(x) dx$.

Majoration de l'erreur de quadrature - ordre

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

6.1.4 Formules de Newton-Cotes

Exercices :

[Exercice 6.6](#)

Nous allons énumérer maintenant quelques formules classiques construites avec des t_i fixés équidistants. On suppose que $t_0 = a$, $t_n = b$ et que les points t_i sont équidistants :

$$t_{j+1} - t_j = h = \frac{b-a}{n}, \forall j = 0, 1, \dots, n-1. \text{ On a alors } t_i = a + ih, \text{ pour } i = 0, \dots, n.$$

On généralise à un nombre quelconque de points l'idée utilisée pour la méthode des trapèzes : on approche $I(f)$ par $J(f) = \int_a^b p_f(t) dt = \sum_{i=0}^n w_i f(t_i)$, où p_f interpole f aux points t_0, \dots, t_n .

Donnons maintenant les formules obtenues pour diverses valeurs de n . Nous donnerons en même temps, pour chacune d'elles, une estimation de l'erreur E correspondante. Formules de Newton-Cotes pour $n = 1, 2, 3$:

– $n = 1$: formule des *trapèzes* ($h = b - a$) :

$$J(f) = \frac{h}{2} (f(t_0) + f(t_1)); \exists \eta \in [a, b], E(f) = -\frac{h^3}{12} f''(\eta).$$

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

– $n = 2$: formule de *Simpson* ($h = (b - a)/2$) :

$$J(f) = \frac{h}{3} (f(t_0) + 4f(t_1) + f(t_2)); \exists \eta \in [a, b], E(f) = -\frac{h^5}{90} f^{(4)}(\eta).$$

– $n = 3$: formule de *Simpson 3/8* ($h = (b - a)/3$) :

$$J(f) = \frac{3h}{8} (f(t_0) + 3f(t_1) + 3f(t_2) + f(t_3)); \exists \eta \in [a, b], E(f) = -\frac{3h^5}{80} f^{(4)}(\eta).$$

On a le résultat général suivant :

Théorème 6.1.1 (Formules de Newton-Cotes avec reste).

Si n est pair et si f est $n + 2$ fois continûment dérivable, il existe $\eta \in [a, b]$ tel que

$$\int_a^b f(t) dt = \sum_{i=0}^n w_i f(t_i) + \frac{h^{n+3} f^{(n+2)}(\eta)}{(n+2)!} \int_0^n t^2(t-1)\dots(t-n) dt,$$

Si n est impair et si f est $n + 1$ fois continûment dérivable, il existe $\eta \in [a, b]$ tel que

$$\int_a^b f(t) dt = \sum_{i=0}^n w_i f(t_i) + \frac{h^{n+2} f^{(n+1)}(\eta)}{(n+1)!} \int_0^n t(t-1)\dots(t-n) dt,.$$

Les coefficients w_i sont obtenus par résolution du système (6.1.2) ou par le calcul de $\int_a^b p_f(t) dt$, où p_f est le polynôme d'interpolation.

On peut remarquer que pour n pair (donc un nombre impair de points) les formules de Newton-Cotes sont exactes pour les polynômes de degré inférieur ou égal à $n + 1$, alors que pour n impair, ces formules sont exactes pour les polynômes de degré inférieur ou égal à n . Le cas n pair donne donc des formules d'ordre plus élevé que prévu. Ceci a une conséquence directe : lorsque n est pair, ajouter un seul point n'améliore pas l'approximation de l'intégrale, il faut ajouter 2 points.

Formules de Newton-Cotes

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

6.1.5 Intégration numérique composée

Exercices :
[Exercice 6.7](#)

Documents :
[Document 6.2](#)

Dans la pratique on dépasse rarement $n = 2$, pour les mêmes raisons qui font préférer les splines au polynôme d'interpolation qui a tendance à osciller fortement quand n devient grand. Pour traiter de grands intervalles on utilise une approche locale : on découpe d'abord l'intervalle $[a, b]$ et on applique des formules d'ordre faible sur les sous-intervalles. Par exemple pour la méthode des trapèzes :

1. On subdivise l'intervalle $[a, b]$ en N sous-intervalles à l'aide de points de subdivision $t_i = a + ih$, pour $i = 0, \dots, N$, avec $h = (b - a)/N$,
2. puis sur chaque sous-intervalle $[t_i, t_{i+1}]$ on applique la méthode des trapèzes.

On obtient ainsi la formule composée suivante :

$$\begin{aligned} J(f) &= h \left(\frac{f(t_0) + f(t_1)}{2} + \frac{f(t_1) + f(t_2)}{2} + \dots + \frac{f(t_{N-1}) + f(t_N)}{2} \right) \\ &= h \left(\frac{1}{2} f(t_0) + \sum_{i=1}^{N-1} f(t_i) + \frac{1}{2} f(t_N) \right). \end{aligned}$$

Pour le calcul de l'erreur, on somme les erreurs commises sur chacun des sous-intervalles, soit

$$E(f) = \sum_{i=0}^{N-1} E_i(f), \text{ avec } E_i(f) = -\frac{h^3}{12} f''(\eta_i),$$

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

où $\eta_i \in [t_i, t_{i+1}]$. On applique alors le théorème de la valeur intermédiaire² pour invoquer l'existence d'un nombre $\eta \in [a, b]$ tel que

$$f''(\eta) = \frac{1}{N} \sum_{i=0}^{N-1} f''(\eta_i),$$

l'erreur globale s'écrit

$$E(f) = -\frac{b-a}{12} h^2 f''(\eta).$$

De même pour la méthode de Simpson : on suppose que N est pair, soit $N = 2M$. La formule obtenue est la suivante (à montrer en exercice) :

$$J(f) = \frac{h}{3} \left(f(t_0) + 4 \sum_{i=0}^{M-1} f(t_{2i+1}) + 2 \sum_{i=1}^{M-1} f(t_{2i}) + f(t_{2M}) \right),$$

où on a toujours posé $h = (b-a)/N$, et par la même technique que précédemment, on obtient l'estimation d'erreur :

$$E(f) = -\frac{b-a}{180} h^4 f^{(4)}(\eta), \quad \eta \in [a, b].$$

Dans le document référencé, on trouve un calcul d'estimation a posteriori de l'erreur et une technique d'utilisation de ces estimations pour construire des méthodes de quadrature **adaptatives**.

2. Si f est continue sur $[a, b]$ alors il existe $c \in]a, b[$ tel que $f(c) = [f(a) + f(b)]/2$.

Exemples du chapitre 6

6.1	Méthode des trapèzes	326
6.2	Méthode des trapèzes sur un intervalle quelconque	328
6.3	Calcul de l'erreur de la méthode des trapèzes	330

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exemple 6.1 Méthode des trapèzes

Soit l'intervalle de référence $[-1, +1]$ et soit φ une fonction définie sur cet intervalle. On cherche à approcher $I(\varphi) = \int_{-1}^1 \varphi(t) dt$ par une formule de quadrature à deux points, soit

$$J(\varphi) = w_0\varphi(-1) + w_1\varphi(1).$$

Notons p le polynôme (de degré inférieur ou égal à 1) tel que $p(-1) = \varphi(-1)$ et $p(1) = \varphi(1)$. Il s'écrit :

$$p(t) = \frac{1-t}{2}\varphi(-1) + \frac{t+1}{2}\varphi(1)$$

d'où, après calcul des deux intégrales,

$$J(\varphi) = \int_{-1}^1 p(t) dt = \varphi(-1) + \varphi(1).$$

Nous avons ainsi obtenu $w_0 = w_1 = 1$.

Nous aurions aussi pu obtenir les deux coefficients par la deuxième technique. Elle consiste à écrire que l'approximation doit être exacte pour les monômes 1 et t , soit

$$\int_{-1}^1 p(t) dt = w_0p(-1) + w_1p(1),$$

pour $p(t) = 1$ et $p(t) = t$. Ceci donne deux équations linéaires :

$$\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} w_0 \\ w_1 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$$

qui redonnent bien $w_0 = w_1 = 1$.

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

[retour au cours](#)

Exemple 6.1
Méthode des
trapèzes

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exemple 6.2 Méthode des trapèzes sur un intervalle quelconque

On a vu dans l'exemple (6.1) que la formule des trapèzes pour approcher $I(\varphi) = \int_{-1}^1 \varphi(\xi) d\xi$ est $J(\varphi) = \varphi(-1) + \varphi(1)$

Si maintenant nous nous plaçons sur un intervalle quelconque $[a, b]$, nous arrivons par un changement de variable à :

$$\int_a^b f(t) dt = \frac{b-a}{2} \int_{-1}^1 \varphi(\xi) d\xi,$$

avec

$$\varphi(\xi) = f\left(\frac{b-a}{2}\xi + \frac{b+a}{2}\right).$$

Nous pouvons ainsi approcher cette dernière intégrale

$$\frac{b-a}{2} \int_{-1}^1 \varphi(\xi) d\xi \approx \frac{b-a}{2} (\varphi(-1) + \varphi(1)) = \frac{b-a}{2} (f(a) + f(b)).$$

La raison pour laquelle cette formule est dite des trapèzes est que l'on approche l'intégrale $I(f)$ par l'aire du trapèze représenté sur la figure 6.1.1.

[retour au cours](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

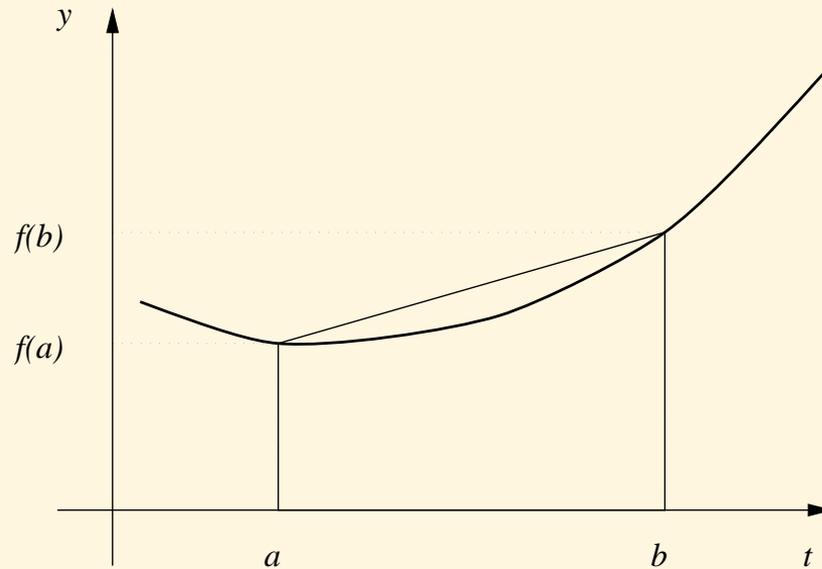


FIGURE 6.1.1 – Méthode des trapèzes

Exemple 6.2
Méthode des trapèzes sur un intervalle quelconque

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exemple 6.3 Calcul de l'erreur de la méthode des trapèzes

Dans certains cas particuliers il est possible d'obtenir, à peu de frais, une estimation plus fine de l'erreur. Pour la méthode des trapèzes, on a

$$E(f) = \int_a^b (t-a)(t-b) \frac{f''(\xi(t))}{2} dt,$$

or sur $[a, b]$ le produit $(t-a)(t-b)$ ne change pas de signe d'où il existe $\eta \in [a, b]$ tel que

$$E(f) = \frac{f''(\eta)}{2} \int_a^b (t-a)(t-b) dt = -\frac{h^3}{12} f''(\eta),$$

où on a noté $h = b - a$.

[retour au cours](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Documents du chapitre 6

6.1	Approchée en quel sens ?	332
6.2	Estimations a priori et a posteriori de l'erreur de quadrature	335

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Document 6.1 Approchée en quel sens ?

Le choix de la forme

$$\int_a^b f(t) dt \approx \sum_{i=0}^n w_i f(t_i). \quad (6.1.3)$$

ne doit pas surprendre. En effet, rappelons comment est définie l'intégrale d'une fonction intégrable. On se donne une subdivision

$$t_0 = a < t_1 < \dots < t_n = b$$

de l'intervalle $[a, b]$ et pour chaque sous-intervalle $[t_i, t_{i+1}]$, un point $\xi_i \in [t_i, t_{i+1}]$. On introduit alors la somme

$$\sum_{i=0}^{n-1} f(\xi_i)(t_{i+1} - t_i),$$

qui représente, une valeur approchée de l'aire sous-tendue par la courbe $y = f(t)$ entre les points a et b . Posons $h = \max_{0 \leq i < n-1} |t_{i+1} - t_i|$. On dit que la fonction f est **Riemann-intégrable** sur l'intervalle $[a, b]$, si la limite

$$\lim_{h \rightarrow 0} \sum_{i=0}^{n-1} f(\xi_i)(t_{i+1} - t_i),$$

existe et est la même,

1. quel que soit le choix des points de subdivision t_i (pourvu que ce choix vérifie simplement la condition $h \rightarrow 0$),
2. et, quel que soit le choix des points ξ_i dans chaque sous-intervalle $[t_i, t_{i+1}]$.

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

La valeur de cette limite est l'intégrale de f sur l'intervalle $[a, b]$.

Nous voyons ainsi, que chaque somme

$$\sum_{i=0}^{n-1} f(\xi_i)(t_{i+1} - t_i),$$

donne une valeur approchée de l'intégrale. Nous voyons aussi que ces sommes sont bien de la forme (6.1.3).

L'idéal serait d'avoir une valeur exacte pour toutes les fonctions susceptibles de nous intéresser, c'est-à-dire en fait pour toutes les fonctions Riemann-intégrables. C'est impossible, mais cette idée donne le principe de construction des méthodes les plus courantes et à partir desquelles beaucoup d'autres sont construites.

On demande ainsi dans un premier temps que les formules de quadrature soient exactes pour des classes de fonctions aussi larges que possible. Comme pour chaque formule on dispose d'un nombre fini de degrés de liberté, on pense vite aux polynômes. En effet, d'une part ils permettent effectivement de construire des formules puisque l'on connaît leurs primitives. D'autre part, ils sont denses dans l'espace des fonctions continues³, ce qui permet d'espérer de bonnes propriétés des formules ainsi construites, pour cette large classe de fonctions.

Lorsque l'on a affaire à une fonction f n'appartenant pas à une classe pour laquelle on dispose de formules exactes, on voudra au moins pouvoir disposer d'une **famille** de formules $(Q_n)_{n \in \mathbb{N}}$ telle que

$$\lim_{n \rightarrow \infty} \sum_{i=0}^n w_{in} f(t_{in}) = \int_a^b f(t) dt.$$

3. ce qui signifie en pratique que toute fonction continue peut être approchée aussi près que l'on veut par un polynôme

Document 6.1

Approchée en
quel sens ?

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Nous construirons plus loin des familles de formule possédant cette propriété pour toute fonction Riemann-intégrable. Nous verrons aussi qu'il existe des familles de formules, qui bien qu'exactes pour les polynômes, ne possèdent pas cette propriété.

[retour au cours](#)

Document 6.1
Approchée en
quel sens ?

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Document 6.2 Estimations a priori et a posteriori de l'erreur de quadrature

L'erreur obtenue pour la méthode du point milieu dans l'exercice (6.3) et celle obtenue pour la méthode des trapèzes dans le paragraphe (Newton-Cotes-formules d'intégration) sont de même nature. Elles montrent par exemple que l'erreur sur un sous-intervalle $[t_i, t_{i+1}]$ tend vers 0 comme h_i^3 , ce qui est intéressant. Cependant la présence d'un terme de la forme $f''(\xi_{M,i})$ (ou $f''(\xi_{T,i})$) interdit toute évaluation de la valeur de l'erreur. Par contre, si nous avons les deux estimations simultanément, à savoir

$$E_M^i(f) = \frac{h_i^3}{24} f''(\xi_{M,i}) \quad \text{et} \quad E_T^i(f) = -\frac{h_i^3}{12} f''(\xi_{T,i})$$

nous remarquons que pour h_i suffisamment petit, nous pouvons écrire

$$E_T^i(f) \simeq -2E_M^i(f)$$

Posons

$$I_{\text{exact}}^i(f) = \int_{t_i}^{t_{i+1}} f(t) dt,$$

$$I_M^i(f) = h_i f\left(\frac{t_i + t_{i+1}}{2}\right), \quad I_T^i(f) = \frac{h_i}{2} (f(t_i) + f(t_{i+1})),$$

Il vient alors

$$I^i(f) - I_T^i \simeq -2 \left(I^i(f) - I_M^i(f) \right).$$

On en déduit une valeur approchée améliorée de $I^i(f)$:

$$I_i(f) \simeq \frac{I_T^i(f) + 2I_M^i(f)}{3} = \frac{h_i}{6} \left(f(t_i) + 4f\left(\frac{t_i + t_{i+1}}{2}\right) + f(t_{i+1}) \right)$$

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

et surtout une estimation de l'erreur, soit

$$E_M^i(f) = I^i(f) - I_M^i(f) \simeq \frac{I_T^i(f) + 2I_M^i(f)}{3} - I_M^i(f) = \frac{I_T^i(f) - I_M^i(f)}{3}.$$

L'estimation ci-dessus est une estimation **a posteriori** : elle donne une 'valeur' calculable de l'erreur dès qu'ont été obtenues les valeurs approchées $I_M^i(f)$ et $I_T^i(f)$. Nous verrons au paragraphe suivant comment utiliser ces estimations pour obtenir des valeurs approchées à une précision donnée.

Concluons ce paragraphe en rappelant que toutes les estimations obtenues n'ont de sens que pour des fonctions à intégrer suffisamment régulières (concrètement de classe C^2 pour les formules de la moyenne et des trapèzes). En outre, les estimations a priori n'ont de sens que pour des h_i suffisamment petits pour que l'on puisse négliger les termes d'ordre supérieur quand on assimile $\xi_{M,i}$ et $\xi_{T,i}$, mais pas trop petits quand même, sinon les erreurs de troncature de la machine deviennent prépondérantes et enlèvent tout sens à ces estimations obtenues sans les prendre en compte.

On donne un intervalle $[a, b]$, une fonction f dont on veut l'intégrale sur $[a, b]$ avec une erreur inférieure à ε . Pour ce faire on évalue une valeur approchée par la méthode du point milieu, puis par la méthode des trapèzes. On obtient ainsi une valeur de l'erreur. Si elle est inférieure à ε c'est terminé et on prend comme valeur approchée la valeur extrapolée. Sinon, on subdivise $[a, b]$ en deux et l'on calcule sur chacun des deux demi-intervalles avec une tolérance de $\varepsilon/2$ pour chacun d'eux, et ainsi de suite.

[retour au cours](#)

Document 6.2

Estimations a priori et a posteriori de l'erreur de quadrature

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercices du chapitre 6

6.1	338
6.2	339
6.3	340
6.4	341
6.5	342
6.6	343
6.7	344

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 6.1

On note $I(f) = \int_a^b f(t) dt$, $J(f) = \sum_{i=0}^n w_i f(t_i)$.

1. On choisit $w_i = \int_a^b \mathcal{L}_i(t) dt$, avec \mathcal{L}_i polynôme de la base de Lagrange associée aux points t_0, t_1, \dots, t_n .

Montrer que alors $I(p) = J(p)$, $\forall p \in \mathcal{P}_n$.

2. Réciproquement, on suppose que $I(p) = J(p)$, $\forall p \in \mathcal{P}_n$, montrer que

$$w_i = \int_a^b \mathcal{L}_i(t) dt.$$

3. Si on note $\{p_0, p_1, \dots, p_n\}$ la base canonique de \mathcal{P}_n , montrer que

$$I(p) = J(p), \forall p \in \mathcal{P}_n \Leftrightarrow I(p_i) = J(p_i), i = 0, \dots, n.$$

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 6.2

On note $I(f) = \int_{-1}^1 f(t) dt$, $J(f) = w_0 f(-1) + w_1 f(0) + w_2 f(1)$.

1. Calculer $(w_i)_{i=0,1,2}$ pour que $I(p) = J(p)$ pour tous les polynômes de \mathcal{P}_n , n le plus grand possible. Quel est ce degré ?
2. En déduire une formule de quadrature pour $\int_0^2 f(t) dt$.

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 6.3

La méthode du point milieu consiste à approcher

$$I(f) = \int_a^b f(t) dt \text{ par } J(f) = (b-a)f(t_M), \text{ où } t_M = \frac{a+b}{2}.$$

1. Vérifier que cette formule est exacte pour les polynômes de degré inférieur ou égal à 1.
2. Quel est l'ordre de cette méthode ?
3. En utilisant la formule de Taylor à l'ordre 2, montrer que l'erreur est donnée par :

$$\exists \eta \in]a, b[, \quad E(f) = I(f) - J(f) = f''(\eta) \int_a^b \frac{(t-t_M)^2}{2} dt = \frac{(b-a)^3}{24} f''(\eta).$$

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 6.4

On approche $I(\varphi) = \int_{-1}^1 \varphi(\xi) d\xi$ par la formule des rectangles $J(\varphi) = w\varphi(t_0)$. Etant donné t_0 , calculer w pour que le degré d'exactitude soit le plus élevé possible. Donner le degré d'exactitude en fonction de t_0 .

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 6.5

Rappeler la formule des trapèzes et donner le degré d'exactitude de cette formule.

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 6.6

Calculer les coefficients w_0 , w_1 , w_2 et w_3 de la formule de Simpson 3/8 qui approche :

$$I(\varphi) = \int_{-1}^1 \varphi(\xi) d\xi \text{ par } J(\varphi) = w_0\varphi(-1) + w_1\varphi(-\frac{1}{3}) + w_2\varphi(\frac{1}{3}) + w_3\varphi(1).$$

Donner le degré d'exactitude de cette formule.

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 6.7

On cherche à approcher $\int_a^b f(t) dt$ à l'aide d'une méthode composée à partir de la méthode de Simpson, on pose $h = \frac{b-a}{2M}$, $t_i = a + ih$. Montrer que

$$J(f) = \frac{h}{3} \left(f(t_0) + 4 \sum_{i=0}^{M-1} f(t_{2i+1}) + 2 \sum_{i=1}^{M-1} f(t_{2i}) + f(t_{2M}) \right),$$

$$E(f) = -\frac{b-a}{180} h^4 f^{(4)}(\eta), \quad \eta \in [a, b].$$

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Chapitre 7

Équations différentielles

7.1	Principes généraux de résolution	346
7.2	Les schémas à un pas	358
7.3	Les schémas multi-pas	376
	Exemples du chapitre 7	387
	Documents du chapitre 7	394
	Exercices du chapitre 7	405

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

7.1 Principes généraux de résolution

7.1.1	Motivations	347
7.1.2	Éléments de théorie des équations différentielles	350
7.1.3	Principe des méthodes numériques	355

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

7.1.1 Motivations

Exercices :

[Exercice 7.1](#)

On considère un pendule initialement au repos et présentant une déviation θ_0 par rapport à la verticale. L'équation différentielle régissant la déviation $\theta(t)$ est la suivante :

$$\begin{cases} \ddot{\theta}(t) = -\frac{g}{L} \sin\theta(t), \\ \theta(0) = \theta_0, \\ \dot{\theta}(0) = 0. \end{cases} \quad (7.1.1)$$

Lorsque θ_0 est petit, on peut se permettre de faire une approximation de $\sin\theta(t)$ au premier ordre, ce qui conduit à poser

$$\sin\theta(t) \approx \theta(t),$$

et dans ce cas on montre aisément que l'on a

$$\theta(t) = \theta_0 \cos\omega t, \quad (7.1.2)$$

où $\omega = \sqrt{g/L}$. Par contre, si θ_0 n'est pas petit, cette approximation n'est plus valable du tout : la figure 7.1.1 montre les solutions obtenues pour $\theta_0 = \frac{63}{64}\pi$. Il est flagrant que la solution provenant du problème linéarisé (en pointillés) donne une idée bien éloignée de ce qu'il se passe en réalité (en trait plein).

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

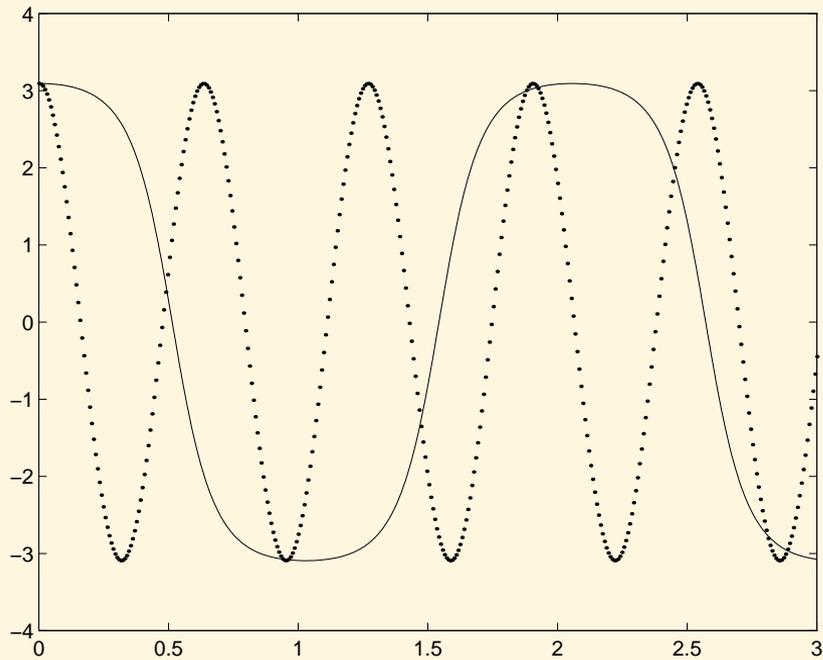
Motivations

FIGURE 7.1.1 – Oscillation d'un pendule linéaire (en pointillés) et non-linéaire (en trait plein)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Nous voyons ainsi que la linéarisation ne donne des résultats exploitables que si l'angle initial θ_0 est petit.

Or on ne sait pas obtenir de solution exacte du problème non linéaire. Il faut donc en rechercher des solutions approchées.

L'objet de ce chapitre est de présenter les principales techniques numériques permettant d'atteindre cet objectif.

D'autres possibilités d'approximation existent et étaient très développées avant l'usage intensif des ordinateurs. Par exemple, on peut chercher une solution du problème non linéaire sous forme d'un développement en série. On détermine alors les coefficients en introduisant le développement dans l'équation différentielle.

Motivations

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

7.1.2 Éléments de théorie des équations différentielles

transform

Exercices :
[Exercice 7.2](#)

Documents :
[Document 7.1](#)

Soit I un intervalle de \mathbb{R} et $f : I \times \mathbb{R} \rightarrow \mathbb{R}$ une fonction continûment différentiable, on notera $f(t, y)$ cette fonction, $\frac{\partial f}{\partial t}$ et $\frac{\partial f}{\partial y}$ les applications dérivées partielles.

Définition 7.1.1. On dit que $\hat{y} : I \rightarrow \mathbb{R}$, fonction continûment dérivable, est solution de l'équation différentielle

$$y' = f(t, y), \quad (7.1.3)$$

si pour tout $t \in I$ on a

$$\hat{y}'(t) = f(t, \hat{y}(t)).$$

Une équation différentielle seule ne suffit pas à définir une solution unique. Classiquement, on lui rajoute une **condition initiale**, c'est-à-dire une condition de la forme :

$$y(t_0) = y_0, \quad \text{où } t_0 \in I \text{ et } y_0 \in \mathbb{R} \text{ sont donnés.} \quad (7.1.4)$$

Dans le document référencé, nous introduisons une condition suffisante classique, pour que l'équation différentielle (7.1.3) avec la condition initiale (7.1.4) admette une solution et une seule dépendant continûment des données.

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Nous ne nous intéresserons qu'à des problèmes de ce type. Cependant, Il faut garder à l'esprit que l'on peut être confronté à d'autres situations. Ainsi, l'équation $y' = f(t, y)$ avec condition initiale $y(t_0) = y_0$, peut admettre, éventuellement, plusieurs solutions. Par exemple l'équation différentielle

$$\begin{cases} y'(t) = \sqrt{y(t)}, & t > 0, \\ y(0) = 0, \end{cases}$$

admet les deux solutions

$$y_1(t) = 0 \text{ et } y_2(t) = \frac{t^2}{4}.$$

Elle en admet même une infinité. En effet, quel que soit le nombre réel $a > 0$, la fonction y_a définie par

$$y_a(t) = \begin{cases} 0, & \text{pour } 0 \leq t \leq a, \\ \frac{(t-a)^2}{4}, & \text{pour } a \leq t, \end{cases}$$

est aussi solution.

Nous nous sommes limités jusqu'ici à une équation du premier ordre. Bien sûr, on rencontre des équations d'ordre supérieur. Ainsi en mécanique rencontre-t-on souvent des équations d'ordre 2 :

$$y''(t) = g(t, y(t), y'(t)), \text{ où } g : I \times \mathbb{R}^2 \rightarrow \mathbb{R}.$$

Une équation d'ordre 2 nécessite deux conditions supplémentaires pour admettre une solution unique.

Éléments de théorie des équations différentielles

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Pour un **problème aux conditions initiales**, qui correspond classiquement au cas où la variable t représente le temps, ces conditions prennent la forme : $y(0) = y_0$ et $y'(0) = y_1$, soit, physiquement, la position et la vitesse à l'instant initial sont données. Nous ne nous intéresserons ici qu'à ce type de problème.

Cependant, il ne faut pas oublier qu'il existe, pour la même équation différentielle, un autre type de problème que nous avons déjà rencontré : le **problème aux limites**. Les deux conditions sont alors partagées entre les deux extrémités de l'intervalle, par exemple : $y(0) = y_0$ et $y(T) = y_T$. Ce deuxième type de problème correspond à des situations physiques complètement différentes et cela se traduit par des méthodes de résolution numériques complètement différentes.

Pour les problèmes aux conditions initiales d'ordre 2, il existe des méthodes numériques spécifiques. Cependant il est toujours possible de les ramener à un système de deux équations différentielles d'ordre 1, couplées et avec conditions initiales. En effet, posons :

$$U(t) = \begin{pmatrix} U_1(t) \\ U_2(t) \end{pmatrix} = \begin{pmatrix} y(t) \\ y'(t) \end{pmatrix}, \quad U_0 = \begin{pmatrix} y_0 \\ y_1 \end{pmatrix}.$$

Il vient :

$$U'(t) = \begin{pmatrix} U_1'(t) \\ U_2'(t) \end{pmatrix} = \begin{pmatrix} y'(t) \\ y''(t) \end{pmatrix} = \begin{pmatrix} U_2(t) \\ g(t, y(t), y'(t)) \end{pmatrix}.$$

Donc en introduisant la fonction à *valeur vectorielle*

$$\begin{aligned} F: I \times \mathbb{R}^2 &\rightarrow \mathbb{R}^2 \\ (t, U) &\rightarrow F(t, U) = \begin{pmatrix} U_2 \\ g(t, U_1, U_2) \end{pmatrix}, \end{aligned} \quad (7.1.5)$$

Éléments de théorie des équations différentielles

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

on s'est ramené au problème :

$$U'(t) = F(t, U(t)) \text{ avec } U(0) = U_0,$$

qui a la même forme que le problème (7.1.4-7.1.3). On a mis l'équation différentielle d'ordre 2 sous forme *normale*.

Ce résultat se généralise aux équations différentielles d'ordre m

$$y^{(m)}(t) = g(t, y(t), y'(t), \dots, y^{(m-1)}(t))$$

qui peuvent se ramener à un système de m équations différentielles d'ordre 1.

Rappel sur la différentiation : soient n, p, q trois entiers > 0 . Soit $f : (x, y) \in \mathbb{R}^n \times \mathbb{R}^p \mapsto f(x, y) \in \mathbb{R}^q$, dont on note $\frac{\partial f}{\partial x}$ et $\frac{\partial f}{\partial y}$ les dérivées partielles par rapport à la première et à la seconde variables. Soit $(x_0, y_0) \in \mathbb{R}^n \times \mathbb{R}^p$. La dérivée partielle $\frac{\partial f}{\partial x}(x_0, y_0)$ est une matrice de taille $(q \times n)$ et $\frac{\partial f}{\partial y}(x_0, y_0)$ est une matrice de taille $(q \times p)$.

Soit I un intervalle de \mathbb{R} . On suppose maintenant que x est une fonction $x : t \in I \mapsto x(t) \in \mathbb{R}^n$. De même, $y : t \in \mathbb{R} \mapsto y(t) \in \mathbb{R}^p$. On prend $t_0 \in I$. On a $x'(t_0) = \frac{dx}{dt}(t_0)$ est un vecteur de \mathbb{R}^n , et $y'(t_0) = \frac{dy}{dt}(t_0) \in \mathbb{R}^p$.

Éléments de théorie des équations différentielles

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Soit enfin la fonction d'une variable $h : t \in I \mapsto h(t) \in \mathbb{R}^q$, définie par $h(t) = f(x(t), y(t))$. Alors sa dérivée au point $t_0 \in I$ s'écrit

$$h'(t_0) = \frac{\partial f}{\partial x}(x(t_0), y(t_0)) \frac{dx}{dt}(t_0) + \frac{\partial f}{\partial y}(x(t_0), y(t_0)) \frac{dy}{dt}(t_0).$$

Noter que les produits entre matrices et vecteurs sont compatibles : $\frac{\partial f}{\partial x}(x(t_0), y(t_0)) \in \mathcal{M}_{q,n}$ et $x'(t_0) \in \mathbb{R}^n$ par exemple.

Éléments de théorie des équations différentielles

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

7.1.3 Principe des méthodes numériques

Exemples :

[Exemple 7.1](#)

Soit $y(t)$ la solution de l'équation différentielle

$$\begin{cases} y'(t) = f(t, y(t)), & t \in [t_0, t_0 + T] \\ y(t_0) = y_0. \end{cases}$$

Le principe général consiste à discrétiser l'intervalle $I = [t_0, t_0 + T]$, en introduisant des points $t_0, t_1, \dots, t_N = t_0 + T$ qui peuvent être équidistants mais ce n'est pas obligatoire. La quantité $h_n = t_{n+1} - t_n$ s'appelle le **pas**. On veut, pour $n = 1, \dots, N$, calculer une approximation de $y(t_n)$, que l'on notera

$$z_n \approx y(t_n),$$

à l'aide d'un procédé itératif.

L'idée la plus simple consiste dans un premier temps à écrire le développement de *Taylor* de $y(t)$ en $t = t_n$: comme la solution est dérivable par rapport à t au moins une fois, on peut écrire que

$$\begin{aligned} y(t_{n+1}) = y(t_n + h) &= y(t_n) + hy'(t_n) + \frac{h^2}{2} y''(\xi), \quad \xi \in [t_n, t_{n+1}] \\ &= y(t_n) + hf(t_n, y(t_n)) + \mathcal{O}(h^2). \end{aligned}$$

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

On a supposé que h_n est constant et égal à $h = \frac{T}{N}$ où N est un entier fixé. Si l'on suppose que h est suffisamment petit, on peut alors proposer le schéma itératif suivant :

$$\begin{cases} z_0 = y_0, \\ z_{n+1} = z_n + hf(t_n, z_n), 0 \leq n \leq N-1. \end{cases}$$

Ce procédé itératif s'appelle **schéma d'Euler simple**. Mais nous verrons d'autres manières de l'introduire.

Une autre solution consiste à pousser le développement de Taylor à l'ordre 2, c'est ce qui est présenté dans l'exemple référencé. On y voit qu'il est alors nécessaire de calculer $y''(t)$ soit

$$\frac{d}{dt}f(t, y(t)) = \frac{\partial f}{\partial t}(t, y(t)) + f(t, y(t)) \frac{\partial f}{\partial y}(t, y(t)).$$

Dans ce même exemple, on présente aussi l'application de ces deux schémas à un exemple très simple.

Enfin, ayant poussé le développement de Taylor jusqu'à l'ordre 2, il est naturel de penser à le développer à l'ordre 3, 4, ... Cependant, il suffit de calculer $(f(t, y(t)))''$ pour voir que cette méthode devient rapidement impraticable dans le cas général.

On distingue deux grandes familles de schémas de résolution numérique des problèmes aux conditions initiales pour les équations différentielles :

- **Les schémas à un pas.** Pour calculer une approximation de la valeur de la fonction cherchée en un point t_{n+1} , on oublie tout ce qui s'est passé avant le point t_n . Le gros avantage est de permettre de changer de pas très facilement au cours du calcul en fonction des estimations d'erreur que l'on obtient en même temps que les valeurs approchées au cours du calcul.

Principe des méthodes numériques

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

- **les schémas multi-pas.** Dans ces méthodes au contraire, pour calculer une approximation de la valeur de la fonction cherchée en un point t_{n+1} , on utilise les valeurs calculées en t_n, t_{n-1}, \dots . Cela permet un coût de calcul moindre, mais rend les valeurs plus interdépendantes de sorte qu'il est plus difficile de changer le pas localement au cours du calcul.

Principe des méthodes numériques

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

7.2 Les schémas à un pas

7.2.1	Schémas d'Euler à partir de l'intégration numérique	359
7.2.2	Schémas prédicteur-correcteurs à un pas	361
7.2.3	Première étude de la méthode d'Euler	363
7.2.4	Ordre et consistance des schémas à un pas	366
7.2.5	Stabilité et convergence des schémas à un pas	368
7.2.6	Les schémas de <i>Runge-Kutta</i>	371

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

7.2.1 Schémas d'Euler à partir de l'intégration numérique

Exercices :

[Exercice 7.4](#)

Cours :

[Principe des méthodes numériques](#)

Pour définir certains schémas numériques de résolution d'équation différentielle, on remarque que la solution exacte $y(t)$ vérifie $y'(t) = f(t, y(t))$, ce qui donne

$$y(t_{n+1}) = y(t_n) + \int_{t_n}^{t_{n+1}} f(t, y(t)) dt.$$

On peut alors penser approcher l'intégrale par une formule utilisant des valeurs de $f(t, y(t))$ sur l'intervalle $[t_n, t_{n+1}]$ bien que $y(t)$ ne soit pas connue sur cet intervalle. Pour simplifier l'écriture, nous supposons que le pas $h_n = t_{n+1} - t_n$ est constant, $h = \frac{T}{N}$ où N est un entier, mais la généralisation à un pas non constant est évidente.

Pour commencer, utilisons la méthode des rectangles "à gauche" pour le calcul approché de l'intégrale, (ceci est équivalent à la formule de Taylor appliquée à $y(t)$ en $t = t_n$ que nous avons introduit dans le paragraphe référencé)

On peut donc proposer le schéma suivant :

$$z_{n+1} = z_n + hf(t_n, z_n), \quad 0 \leq n \leq N-1, \quad (7.2.1)$$

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

qui n'est autre que le **schéma d'Euler simple**. On peut aussi approcher l'intégrale avec la méthode des rectangles à droite. Un calcul équivalent provient de la formule de Taylor appliquée à $y(t)$ en $t = t_{n+1}$:

$$y(t_n) = y(t_{n+1}) - hf(t_{n+1}, y(t_{n+1})) + \frac{h^2}{2} y''(\xi), \quad \xi \in [t_n, t_{n+1}].$$

Ceci conduit au schéma

$$z_{n+1} = z_n + hf(t_{n+1}, z_{n+1}), \quad 0 \leq n \leq N-1, \quad (7.2.2)$$

que l'on nomme **schéma d'Euler rétrograde**. On dit que ce schéma est **implicite** car z_{n+1} est défini implicitement comme solution de l'équation

$$x = z_n + hf(t_{n+1}, x),$$

qui en général est non-linéaire. On fait alors appel à des méthodes de type point fixe ou Newton. Cependant, comme le pas h est petit, le nombre d'itérations nécessaires en pratique est petit : parfois même une seule suffit. Il reste alors à initialiser le processus avec une première estimation de z_{n+1} , qui peut être z_n .

Schémas d'Euler à partir de l'intégration numérique

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

7.2.2 Schémas prédicteur-correcteurs à un pas

Exercices :

[Exercice 7.5](#)

Reprenons le schéma implicite d'Euler rétrograde 7.2.2 défini précédemment :

$$z_{n+1} = z_n + hf(t_{n+1}, z_{n+1}), \quad 0 \leq n \leq N-1.$$

On peut construire un nouveau schéma dit prédicteur-correcteur,

- on détermine une première estimation grossière de z_{n+1} , notée \tilde{z}_{n+1} , on peut avoir recours par exemple à la méthode d'Euler explicite
- on améliore cette estimation en s'inspirant du schéma d'Euler rétrograde.

On obtient le schéma :

$$\begin{cases} \hat{z}_{n+1} = z_n + hf(t_n, z_n), \\ z_{n+1} = z_n + hf(t_{n+1}, \hat{z}_{n+1}). \end{cases} \quad (7.2.3)$$

Dans le langage devenu classique pour ces méthodes, on dit que l'on fait d'abord une **prédiction** (\hat{z}_{n+1}) à l'aide du schéma explicite, puis une **correction** à l'aide du schéma implicite. En outre, on peut être conduit à itérer sur les corrections.

Voyons un autre exemple, si on utilise la méthode des trapèzes pour calculer l'intégrale de

$$y(t_{n+1}) = y(t_n) + \int_{t_n}^{t_{n+1}} f(t, y(t)) dt,$$

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

cela donne l'expression

$$y(t_{n+1}) = y(t_n) + \frac{h}{2} (f(t_n, y(t_n)) + f(t_{n+1}, y(t_{n+1}))) - \frac{h^3}{12} y^{(3)}(\xi), \quad \xi \in [t_n, t_{n+1}].$$

Le terme $y^{(3)}$ correspond à la dérivée seconde par rapport à t de $t \mapsto f(t, y(t))$. Ceci conduit au schéma suivant :

$$z_{n+1} = z_n + \frac{h}{2} (f(t_n, z_n) + f(t_{n+1}, z_{n+1})), \quad 0 \leq n \leq N-1, \quad (7.2.4)$$

qui est implicite, comme le schéma d'Euler rétrograde.

On peut, comme précédemment, construire un nouveau schéma prédicteur-correcteur inspiré de ce schéma implicite, c'est le **schéma prédicteur-correcteur d'Euler-Cauchy** :

$$\begin{cases} \tilde{z}_{n+1} = z_n + hf(t_n, z_n), \\ z_{n+1} = z_n + \frac{h}{2} (f(t_n, z_n) + f(t_{n+1}, \tilde{z}_{n+1})), \end{cases} \quad 0 \leq n \leq N-1. \quad (7.2.5)$$

**Schémas
prédicteur-
correcteurs à un
pas**

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

7.2.3 Première étude de la méthode d'Euler

Exercices :

[Exercice 7.6](#)

Reprenons les schémas d'Euler explicite et implicite [7.2.1](#), [7.2.2](#) :

$$\begin{aligned}z_{n+1} &= z_n + hf(t_n, z_n), \quad n \geq 0, \\z_{n+1} &= z_n + hf(t_{n+1}, z_{n+1}), \quad n \geq 0.\end{aligned}$$

Nous allons maintenant mettre en évidence deux types de difficulté de nature différente qui peuvent apparaître lorsque l'on emploie des méthodes de ce type. Pour cela, il va nous suffire d'appliquer ces schémas au **problème modèle** suivant :

$$\begin{cases} y' = -\lambda y(t), \text{ avec } \lambda > 0, \\ y(0) = y_0. \end{cases}$$

Ce problème admet l'unique solution exacte : $y(t) = y_0 e^{-\lambda t}$.

Le schéma d'Euler explicite construit une suite de valeurs $z_n = (1 - \lambda h)^n y_0$ et le schéma d'Euler implicite $z_n = y_0 / (1 + \lambda h)^n$.

- **Comportement des solutions exactes et approchées lorsque $t \rightarrow +\infty$**

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Nous constatons tout d'abord que $\lim_{t \rightarrow +\infty} y(t) = 0$.

Si nous utilisons le schéma d'Euler pour des valeurs grandes de n , avec pour simplifier les écritures un pas h constant, il faut que l'on ait aussi $\lim_{n \rightarrow +\infty} y_n = 0$. Or, cette condition ne sera remplie que si $|1 - \lambda h| < 1$, soit pour $\lambda > 0$, pour $h < 2/\lambda$. Cette condition peut se révéler très contraignante.

Pour le schéma d'Euler implicite, la suite z_n converge vers 0 quand n tend vers l'infini, sans aucune condition sur h . Ainsi, nous pourrions utiliser cette méthode sans avoir a priori de condition restrictive impérative sur le pas. C'est la précision désirée qui déterminera le pas : plus le pas sera petit, meilleure sera celle-ci.

- Étude de la convergence

Cette fois ci, nous nous plaçons en un point t fixé. Nous choisissons le pas h de manière qu'il existe toujours un entier n tel que $nh = t$. Alors, nous désirons que y_n soit une approximation de $y(t)$. Mieux que cela, nous voulons que, lorsque $n \rightarrow +\infty$, sous la contrainte $nh = t$, alors z_n converge vers $y(t)$. Il est facile de voir que pour les méthodes d'Euler, cette propriété est satisfaite. En effet on montre dans l'exercice 7.6 que

$$\lim_{h \rightarrow 0, nh=t} (1 - \lambda h)^n y_0 = \lim_{h \rightarrow 0, nh=t} y_0 / (1 + \lambda h)^n = y_0 e^{-\lambda(t)}.$$

Insistons bien sur le fait qu'il s'agit de deux points de vue distincts et complémentaires. Le premier n'a d'intérêt que pour t grand. En fait t grand peut être vite atteint. Ainsi, lorsque l'on fait des prévisions météorologiques, $t = 10$ jours est très grand et ce sont justement les questions de stabilité qui limitent en pratique la portée de ces prévisions.

Première étude de la méthode d'Euler

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Si par contre l'on ne s'intéresse qu'à des t 'petits', seule la question de la convergence importera.

Première étude de la méthode d'Euler

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

7.2.4 Ordre et consistance des schémas à un pas

Exercices :

[Exercice 7.7](#)

Etant donnés f, t_0, y_0, T , soit $y(t)$ la solution exacte de

$$\begin{cases} y'(t) = f(t, y(t)), t \in [t_0, t_0 + T] \\ y(t_0) = y_0, \end{cases} \quad (7.2.6)$$

On va supposer, pour simplifier l'exposé, que l'on discrétise avec un pas constant $h = \frac{T}{N}$, on pose $t_n = t_0 + nh$. Les schémas à un pas explicites ou prédicteur-correcteur peuvent se mettre sous la forme générique

$$\begin{cases} z_0 = y_0, \\ z_{n+1} = z_n + h\phi(t_n, z_n, h), 0 \leq n \leq N-1, \end{cases} \quad (7.2.7)$$

On peut se demander comment quantifier l'erreur commise en approchant la solution exacte $y(t)$ par la séquence discrète z_n . La définition suivante donne un début de réponse à cette question.

Définition 7.2.1. On appelle **erreur locale** relative au schéma (7.2.7) la quantité

$$\tau_{n+1}(h) = (y(t_{n+1}) - y(t_n)) - h\phi(t_n, y(t_n), h), \quad n = 0, \dots, N-1,$$

où $y(t)$ est la solution exacte de l'équation différentielle (7.2.6).

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

$\tau_{n+1}(h) = y(t_{n+1}) - z_{n+1}$ si l'on suppose que $z_n = y(t_n)$.

Définition 7.2.2. Lorsque qu'il existe $K > 0$ tel que

$$\max_{1 \leq n \leq N} \left| \frac{\tau_n(h)}{h} \right| \leq Kh^p,$$

le schéma est dit **d'ordre p** .

La définition suivante est propre aux schémas de résolution des équations différentielles et se déduit rapidement de l'ordre.

Définition 7.2.3. Le schéma est dit **consistant** quand

$$\lim_{h \rightarrow 0} \max_{1 \leq n \leq N} \left| \frac{\tau_n(h)}{h} \right| = 0.$$

Un schéma d'ordre strictement positif est donc consistant.

Pour le schéma d'Euler simple, on a

$$\tau_{n+1}(h) = (y(t_{n+1}) - y(t_n)) - hf(t_n, y(t_n)) = \frac{h^2}{2} y''(\xi).$$

Supposons que y'' est bornée sur $[t_n, t_{n+1}]$, soit $|y''(t)| \leq M \forall t \in [t_0, t_0 + T]$, alors on a

$$\tau_{n+1}(h) \leq \frac{M}{2} h^2,$$

et d'après la définition précédente, le schéma d'Euler simple est d'ordre 1. On peut vérifier que le schéma d'Euler-Cauchy est d'ordre 2 voir l'exercice 7.7.

**Ordre et
consistance des
schémas à un
pas**

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

7.2.5 Stabilité et convergence des schémas à un pas

Exemples :

[Exemple 7.2](#)

La consistance ou l'ordre d'un schéma n'est qu'une indication locale de l'erreur. Un moyen plus réaliste de mesurer l'erreur d'approximation de $y(t_i)$ par z_i consiste à considérer l'erreur maximum commise pour $i = 1 \dots N$, et à regarder si cette erreur tend bien vers zéro quand $h \rightarrow 0$.

Définition 7.2.4. *Le schéma*

$$z_{n+1} = z_n + h\phi(t_n, z_n, h), \quad 0 \leq n \leq N-1, \quad z_0 = y_0, \quad h = \frac{T}{N}$$

est dit **convergent** par rapport à l'équation différentielle

$$y'(t) = f(t, y(t)), \quad y(t_0) = y_0, \quad t \in [t_0, t_0 + T]$$

si

$$\lim_{h \rightarrow 0} \max_{1 \leq i \leq N} |y(t_i) - z_i| = 0.$$

La consistance d'un schéma n'implique pas qu'il soit convergent, nous allons voir qu'il s'agit tout au plus d'une condition nécessaire. Une condition supplémentaire fait intervenir la notion de stabilité :

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Définition 7.2.5. *Le schéma*

$$z_{n+1} = z_n + h\phi(t_n, z_n, h), \quad 0 \leq n \leq N-1, \quad z_0 = y_0,$$

est dit **stable** s'il existe une constante M telle que pour tout z_0 , pour tout u_0 , pour tout $h \leq h^*$ et pour toute suite $\{\varepsilon_i\}$, les suites $\{z_i\}$ et $\{u_i\}$ définies par les relations

$$\begin{aligned} z_{i+1} &= z_i + h\phi(t_i, z_i, h), \\ u_{i+1} &= u_i + h\phi(t_i, u_i, h) + \varepsilon_i, \end{aligned}$$

vérifient la condition

$$\forall i = 1 \dots N, |z_i - u_i| \leq M \left(|z_0 - u_0| + \sum_{k=0}^{i-1} |\varepsilon_k| \right).$$

En gros cela signifie qu'un schéma stable n'amplifie ni les erreurs sur la condition initiale, ni les erreurs introduites dans le schéma : il s'agit d'une notion de *continuité*. On peut noter que la stabilité peut éventuellement dépendre de h . La définition ci-dessus montre qu'il peut exister un pas limite h^* au-delà duquel un schéma stable devient instable. On a aussi une condition suffisante de stabilité :

Proposition 7.2.1. *Pour qu'un schéma soit stable, il suffit qu'il existe une constante Λ telle que*

$$\forall t \in [t_0, t_0 + T], \forall z, u \in \mathbb{R}, \forall h \in [0, h^*], |\phi(t, z, h) - \phi(t, u, h)| \leq \Lambda |z - u|.$$

Cela signifie que la fonction ϕ doit vérifier la condition de Lipschitz sur son deuxième argument. En général on montre que f vérifie une condition de Lipschitz pour obtenir cette propriété

**Stabilité et
convergence des
schémas à un
pas**

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

pour ϕ . Dans ce cas, le schéma d'Euler simple est évidemment stable, puisque l'on a $\phi(t, z, h) = f(t, z)$.

Le théorème suivant est simple et essentiel ; il est généralement connu sous le nom de **consistance plus stabilité impliquent convergence**.

Théorème 7.2.6. *Soit ϕ une fonction continue de $t \in [t_0, t_0 + T]$, $z \in \mathbb{R}$ et h , définissant le schéma à un pas*

$$z_{n+1} = z_n + h\phi(t_n, z_n, h), \quad 0 \leq n \leq N-1, \quad z_0 = y_0, \quad h = \frac{T}{N}.$$

Si ce schéma à un pas est consistant et s'il est stable, alors il est convergent par rapport à l'équation différentielle

$$y'(t) = f(t, y(t)), \quad y(t_0) = y_0, \quad t \in [t_0, t_0 + T].$$

L'exemple référencé traite du schéma d'Euler-Cauchy.

**Stabilité et
convergence des
schémas à un
pas**

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

7.2.6 Les schémas de *Runge-Kutta*

Documents :

[Document 7.2](#)

[Document 7.3](#)

Il s'agit de schémas qui permettent de retrouver les bonnes propriétés des schémas de Taylor (ordre élevé), sans en présenter les inconvénients (calcul des dérivées successives de f).

Il existe classiquement deux manières de les construire. Nous en présentons une ici. L'autre est introduite dans le document référencé.

Il s'agit de schémas à un pas, donc se mettant sous la forme

$$\begin{cases} z_0 = y_0, \\ z_{n+1} = z_n + h\phi(t_n, z_n, h), \quad n \geq 0, \end{cases}$$

où ϕ prend la forme particulière suivante :

$$\phi(t, z, h) = \sum_{i=1}^q \gamma_i k_i, \quad (7.2.8)$$

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

et les k_i sont définis récursivement de la façon suivante :

$$\left\{ \begin{array}{l} k_1 = f(t, z), \\ k_2 = f(t + h\alpha_1, z + h\beta_{11}k_1), \\ k_3 = f(t + h\alpha_2, z + h\beta_{21}k_1 + h\beta_{22}k_2), \\ \vdots \\ k_q = f\left(t + h\alpha_{q-1}, z + h\sum_{p=1}^{q-1} \beta_{q-1,p}k_p\right). \end{array} \right. \quad (7.2.9)$$

L'introduction des ces q valeurs intermédiaires a pour but d'approcher un schéma de Taylor d'ordre q . Pour cela il faut déterminer les inconnues qui sont

$$[\gamma_i]_{i=1\dots q}, [\alpha_i]_{i=1\dots q-1}, [\beta_{ij}]_{i=1\dots q-1, j=1\dots i}.$$

Il est difficile d'élaborer une méthode permettant de déterminer ces inconnues pour une valeur quelconque de q , c'est pourquoi nous allons nous contenter de raisonner sur un exemple, ici pour $q = 2$, puis nous donnerons sans démonstration les valeurs obtenues pour $q = 4$, qui est la valeur la plus couramment utilisée dans la pratique.

Un schéma de Taylor d'ordre 2 s'écrit

$$z_{n+1} = z_n + h\psi(t_n, z_n, h),$$

où

$$\psi(t, z, h) = f(t, z) + \frac{h}{2} \left(\frac{\partial f}{\partial t}(t, z) + f(t, z) \frac{\partial f}{\partial y}(t, z) \right), \quad (7.2.10)$$

Les schémas de Runge-Kutta

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Pour $q = 2$ la fonction $\phi(t, z, h)$ définie par (7.2.8)(7.2.9) s'écrit

$$\begin{aligned}\phi(t, z, h) &= \gamma_1 k_1 + \gamma_2 k_2 \\ &= \gamma_1 f(t, z) + \gamma_2 f(t + h\alpha_1, z + h\beta_1 f(t, z)),\end{aligned}$$

où l'on a posé $\beta_1 = \beta_{11}$. On va essayer d'identifier les coefficients inconnus $\gamma_1, \gamma_2, \alpha_1, \beta_1$ de façon à ce que $\phi(t, z, h)$ et $\psi(t, z, h)$ soient les plus proches possibles. Pour cela on développe

$$f(t + h\alpha_1, z + h\beta_1 f(t, z))$$

à l'aide d'un développement de Taylor suivant les deux variables, ce qui donne à l'ordre 2 :

$$\begin{aligned}f(t + u, z + v) &= f(t, z) + u \frac{\partial f}{\partial t}(t, z) + v \frac{\partial f}{\partial y}(t, z) \\ &+ \frac{u^2}{2} \frac{\partial^2 f}{\partial t^2}(\xi, \eta) + uv \frac{\partial^2 f}{\partial t \partial y}(\xi, \eta) + \frac{v^2}{2} \frac{\partial^2 f}{\partial y^2}(\xi, \eta),\end{aligned}$$

où $(\xi, \eta) \in [t, t + u] \times [z, z + v]$. Nous obtenons ainsi

$$f(t + h\alpha_1, z + h\beta_1 f(t, z)) = f(t, z) + h\alpha_1 \frac{\partial f}{\partial t}(t, z) + h\beta_1 f(t, z) \frac{\partial f}{\partial y}(t, z) + \mathcal{O}(h^2),$$

ce qui nous permet d'écrire

$$\phi(t, z, h) = (\gamma_1 + \gamma_2) f(t, z) + \gamma_2 \left(h\alpha_1 \frac{\partial f}{\partial t}(t, z) + h\beta_1 f(t, z) \frac{\partial f}{\partial y}(t, z) \right) + \mathcal{O}(h^2).$$

Les schémas de Runge-Kutta

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Si l'on néglige le reste on peut alors identifier cette expression avec $\psi(t, z, h)$ donné par (7.2.10); on obtient alors les équations suivantes :

$$\begin{cases} \gamma_1 + \gamma_2 &= 1, \\ \gamma_2 \alpha_1 &= \frac{1}{2}, \\ \gamma_2 \beta_1 &= \frac{1}{2}. \end{cases}$$

Comme il y a trois équations pour quatre inconnues, il y a une infinité de possibilités pour le choix des coefficients; nous ne donnerons ici que les plus populaires :

– $\gamma_1 = 0, \gamma_2 = 1, \alpha_1 = \beta_1 = \frac{1}{2}$, qui donnent le schéma

$$z_{n+1} = z_n + hf \left(t_n + \frac{1}{2}h, z_n + \frac{1}{2}hf(t_n, z_n) \right), \quad (7.2.11)$$

appelé **schéma du point milieu**.

– $\gamma_1 = \gamma_2 = \frac{1}{2}, \alpha_1 = \beta_1 = 1$, qui donnent le schéma bien connu

$$z_{n+1} = z_n + \frac{h}{2} (f(t_n, z_n) + f(t_{n+1}, z_n + hf(t_n, z_n))),$$

puisque'il s'agit du schéma d'Euler-Cauchy 7.2.5.

Ces schémas sont bien sûr d'ordre 2 puisque l'on a

$$\phi(t_n, z_n, h) + \mathcal{O}(h^2) = \psi(t_n, z_n, h),$$

ce qui permet de conserver l'ordre de l'erreur locale du schéma de Taylor utilisé.

Les schémas de Runge-Kutta

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Le choix le plus courant est $q = 4$. Cela donne le **schéma de Runge-Kutta d'ordre 4**, ou en abrégé **RK4**, utilisé presque universellement¹ dans les sciences de l'ingénieur :

$$\begin{cases} k_1 = f(t_n, z_n), \\ k_2 = f\left(t_n + \frac{h}{2}, z_n + \frac{h}{2}k_1\right), \\ k_3 = f\left(t_n + \frac{h}{2}, z_n + \frac{h}{2}k_2\right), \\ k_4 = f(t_{n+1}, z_n + hk_3), \end{cases}$$

$$z_{n+1} = z_n + \frac{h}{6}(k_1 + 2k_2 + 2k_3 + k_4).$$

Dans le deuxième document référencé, vous pourrez lire comment on peut adapter le pas du maillage pour augmenter l'efficacité de l'algorithme. On parle alors de maillage adaptatif.

1. parfois même à tort!

Les schémas de Runge-Kutta

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

7.3 Les schémas multi-pas

7.3.1	Les schémas d' <i>Adams-Bashforth</i>	377
7.3.2	Quelques exemples de schémas d' <i>Adams-Bashforth</i>	380
7.3.3	Les schémas implicites d' <i>Adams-Moulton</i>	381
7.3.4	Quelques exemples de schémas d' <i>Adams-Moulton</i>	384
7.3.5	Références bibliographiques	385

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

7.3.1 Les schémas d'*Adams-Bashforth*

Cours :

[Adams-Bashforth - Exemples](#)

La solution exacte $y(t)$ de l'équation différentielle, vérifie

$$y(t_{n+1}) = y(t_n) + \int_{t_n}^{t_{n+1}} f(t, y(t)) dt.$$

Certains schémas sont basés sur une approximation de

$$\int_{t_n}^{t_{n+1}} f(t, y(t)) dt, \quad (7.3.1)$$

par des formules de quadrature à un et deux points. Ici on approche cette intégrale par la quantité

$$\int_{t_n}^{t_{n+1}} p(t) dt, \quad (7.3.2)$$

où $p(t)$ est l'unique polynôme de degré q vérifiant

$$p(t_i) = f(t_i, y(t_i)), \quad i = n, n-1, \dots, n-q.$$

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Si l'on écrit ce polynôme sur la base de Lagrange, on a

$$p(t) = \sum_{k=0}^q \mathcal{L}_k(t) f(t_{n-k}, y(t_{n-k})),$$

et donc

$$\int_{t_n}^{t_{n+1}} p(t) dt = h \sum_{k=0}^q b_k f(t_{n-k}, y(t_{n-k})),$$

avec

$$b_k = \frac{1}{h} \int_{t_n}^{t_{n+1}} \mathcal{L}_k(t) dt. \quad (7.3.3)$$

Soit donc le schéma suivant, appelé **schéma d'Adams-Bashforth** à $q+1$ pas :

$$z_{n+1} = z_n + h \sum_{k=0}^q b_k f(t_{n-k}, z_{n-k}), \quad n \geq q. \quad (7.3.4)$$

Ce schéma à $q+1$ pas fait intervenir $q+1$ points où la fonction f est évaluée (les (t_{n-k}, z_{n-k}) , pour $k = 0, \dots, q$). On dit que c'est un schéma à $q+1$ pas et $q+1$ nœuds.

Ce schéma n'est valable qu'à partir de $n = q$, c'est pourquoi dans la pratique on a besoin d'approcher les q premières valeurs z_1, \dots, z_q par un schéma à un pas d'ordre suffisamment élevé (par exemple Runge-Kutta).

Remarquons que les coefficients b_k donnés par la formule (7.3.3) sont indépendants de h . En effet on a

$$b_k = \frac{1}{h} \int_{t_n}^{t_{n+1}} \prod_{j=0, j \neq k}^q \frac{t - t_{n-j}}{t_{n-k} - t_{n-j}} dt,$$

**Les schémas
d'Adams-
Bashforth**

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

et si l'on pose $s = \frac{t - t_n}{h}$,

$$b_k = \int_0^1 \prod_{j=0, j \neq k}^q \frac{j+s}{j-k} ds,$$

expression qui montre bien que h n'intervient pas.

On peut regarder ce que donne cette approche pour $q = 1$:

$$b_0 = \int_0^1 (s+1) ds = \frac{3}{2} \text{ et } b_1 = - \int_0^1 s ds = -\frac{1}{2},$$

et donc le schéma d'Adams-Bashforth à deux pas prend la forme suivante :

$$z_{n+1} = z_n + h \left[\frac{3}{2} f(t_n, z_n) - \frac{1}{2} f(t_{n-1}, z_{n-1}) \right], \quad n \geq 1.$$

De façon générale, les schémas d'Adams-Bashforth

$$z_{n+1} = z_n + h \sum_{k=0}^q b_k f(t_{n-k}, z_{n-k}), \quad n \geq q.$$

avec les b_k donnés par (7.3.3) sont d'ordre q . Les notions d'ordre, de consistance et de stabilité des schémas multipas sont légèrement différentes de celles concernant les schémas à un pas. Nous n'en parlerons pas dans le cadre de ce cours, elles sont très bien traitées dans l'ouvrage² p. 257-268.

Vous trouverez dans le paragraphe de cours référencé ces schémas pour quelques valeurs de q .

2. BURDEN, R.L. ET FAIRES, J.D. : *Numerical Analysis* - Pws-Kent, Boston

Les schémas d'Adams- Bashforth

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

7.3.2 Quelques exemples de schémas d'Adams-Bashforth

Voici quelques schémas d'Adams-Bashforth correspondant à quelques valeurs de q .

– **Adams-Bashforth à deux pas** (et deux nœuds)

$$z_{n+1} = z_n + h \left[\frac{3}{2}f(t_n, z_n) - \frac{1}{2}f(t_{n-1}, z_{n-1}) \right], \quad n \geq 1,$$

Erreur locale $\tau_{n+1}(h) = \frac{5}{12}y^{(3)}(\eta)h^3$, $\eta \in [t_{n-1}, t_{n+1}]$.

– **Adams-Bashforth à trois pas** (et trois nœuds)

$$z_{n+1} = z_n + \frac{h}{12} [23f(t_n, z_n) - 16f(t_{n-1}, z_{n-1}) + 5f(t_{n-2}, z_{n-2})], \quad n \geq 2,$$

Erreur locale $\tau_{n+1}(h) = \frac{3}{8}y^{(4)}(\eta)h^4$, $\eta \in [t_{n-2}, t_{n+1}]$.

– **Adams-Bashforth à quatre pas** (et quatre nœuds)

$$z_{n+1} = z_n + \frac{h}{24} [55f(t_n, z_n) - 59f(t_{n-1}, z_{n-1}) + 37f(t_{n-2}, z_{n-2}) - 9f(t_{n-3}, z_{n-3})], \quad n \geq 3,$$

Erreur locale $\tau_{n+1}(h) = \frac{251}{720}y^{(5)}(\eta)h^5$, $\eta \in [t_{n-3}, t_{n+1}]$.

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

7.3.3 Les schémas implicites d'Adams-Moulton

Cours :

[Adams-Moulton - Exemples](#)

Nous allons procéder de la même manière que pour les schémas d'Adams-Bashforth. Cette fois-ci cependant, nous allons approcher l'intégrale (7.3.1) par la quantité

$$\int_{t_n}^{t_{n+1}} p(t) dt,$$

où $p(t)$ est l'unique polynôme de degré $q + 1$ vérifiant

$$p(t_i) = f(t_i, y(t_i)), \quad i = n + 1, n, n - 1, \dots, n - q,$$

c'est-à-dire que l'on inclut $y(t_{n+1})$ dans les valeurs "connues", c'est pourquoi nous obtenons ainsi une famille de schémas multi-pas implicites appelés schémas d'**Adams-Moulton** à $q + 1$ pas :

$$z_{n+1} = z_n + h \sum_{k=-1}^q b_k f(t_{n-k}, z_{n-k}), \quad n \geq q. \quad (7.3.5)$$

Ce schéma est un schéma à $q + 1$ pas et $q + 2$ nœuds (les (t_{n-k}, z_{n-k}) , pour $k = -1, \dots, q$).

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Les coefficients b_k , qui se calculent comme pour les schémas d'Adams-Bashforth, sont obtenus par la formule

$$b_k = \int_0^1 \prod_{j=-1, j \neq k}^q \frac{j+s}{j-k} ds, \quad k = -1, 0, \dots, q.$$

Pour $q = 1$, on retrouve le schéma d'Euler-Centré (7.2.4) :

$$z_{n+1} = z_n + \frac{h}{2} [f(t_{n+1}, z_{n+1}) + f(t_n, z_n)], \quad n \geq 0,$$

Erreur locale $\tau_{n+1}(h) = -\frac{1}{12} y^{(3)}(\eta) h^3$, $\eta \in [t_{n-1}, t_{n+1}]$.

Vous trouverez dans le paragraphe référencé ces schémas pour d'autres valeurs de q .

En pratique les schémas d'Adams-Moulton ne sont pas utilisés comme tels, car ils sont implicites, et donc chercher à obtenir z_{n+1} à l'aide d'une méthode de point fixe ou de Newton ferait perdre l'avantage obtenu en réutilisant les valeurs précédentes de f . On utilise plutôt conjointement un schéma d'Adams-Bashforth et un schéma d'Adams-Moulton de même ordre pour construire un schéma prédicteur-correcteur, en utilisant comme valeur de z_{n+1} dans le schéma d'Adams-Moulton la valeur prédite par le schéma Adams-Bashforth. Le prédicteur-correcteur le plus utilisé utilise ces deux schémas à l'ordre 4 :

$$\begin{aligned} \textbf{Prédicteur : } \tilde{z}_{n+1} &= z_n + \frac{h}{24} [55f(t_n, z_n) - 59f(t_{n-1}, z_{n-1}) \\ &+ 37f(t_{n-2}, z_{n-2}) - 9f(t_{n-3}, z_{n-3})], \\ \textbf{Correcteur : } z_{n+1} &= z_n + \frac{h}{24} [9f(t_{n+1}, \tilde{z}_{n+1}) + 19f(t_n, z_n) \\ &- 5f(t_{n-1}, z_{n-1}) + f(t_{n-2}, z_{n-2})]. \end{aligned}$$

Les schémas implicites d'Adams- Moulton

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

On peut montrer que l'ordre de l'erreur locale de chacun des deux schémas (ici 4) est conservé.

**Les schémas
implicites
d'*Adams-
Moulton***

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

7.3.4 Quelques exemples de schémas d'Adams-Moulton

Voici quelques schémas d'Adams-Moulton correspondant à quelques valeurs de q .

– **Adams-Moulton à deux pas** (et trois nœuds)

$$z_{n+1} = z_n + \frac{h}{12} [5f(t_{n+1}, z_{n+1}) + 8f(t_n, z_n) - f(t_{n-1}, z_{n-1})], \quad n \geq 1,$$

Erreur locale $\tau_{n+1}(h) = -\frac{1}{24}y^{(4)}(\eta)h^4$, $\eta \in [t_{n-1}, t_{n+1}]$.

– **Adams-Moulton à trois pas** (et quatre nœuds)

$$\begin{aligned} z_{n+1} = z_n + \frac{h}{24} [9f(t_{n+1}, z_{n+1}) + 19f(t_n, z_n) \\ - 5f(t_{n-1}, z_{n-1}) + f(t_{n-2}, z_{n-2})], \quad n \geq 2, \end{aligned}$$

Erreur locale $\tau_{n+1}(h) = -\frac{19}{720}y^{(5)}(\eta)h^5$, $\eta \in [t_{n-2}, t_{n+1}]$.

– **Adams-Moulton à quatre pas** (et cinq nœuds)

$$\begin{aligned} z_{n+1} = z_n + \frac{h}{720} [251f(t_{n+1}, z_{n+1}) + 646f(t_n, z_n) - 264f(t_{n-1}, z_{n-1}) \\ + 106f(t_{n-2}, z_{n-2}) - 19f(t_{n-3}, z_{n-3})], \quad n \geq 3, \end{aligned}$$

Erreur locale $\tau_{n+1}(h) = -\frac{3}{160}y^{(6)}(\eta)h^6$, $\eta \in [t_{n-3}, t_{n+1}]$.

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

7.3.5 Références bibliographiques

– Références pédagogiques en français

- CROUZEIX, M. ; MIGNOT, A. L. : *Analyse numérique des équations différentielles*, Masson, Paris 1984.
- DEMAILLY, J. P. : *Analyse numérique et équations différentielles*, Presses universitaires de Grenoble, Grenoble, 1991.
- SCHATZMANN, M. : *Analyse numérique, cours et exercices pour la licence*, InterEditions, Paris, 1991.

– Une référence classique historique

- HENRICI, P. : *Discrete variable methods in ordinary differential equations*, John Wiley and sons, New York, 1962.

– Références avec A-stabilité

- GEAR, C. W. : *Numerical initial value problems in ordinary differential equations*, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1971.
- HALL, G. ; WATT, J. M. (editors) : *Modern numerical methods for ordinary differential equations*, Clarendon Press, Oxford, 1976.
- LAMBERT, J. D. : *Computational methods in ordinary differential equations*, John Wiley & Sons, Inc., New York, 1973.

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

– Références historiques

- BASHFORTH, F. and ADAMS, J. C. , *Theory of Capillary Action*, Cambridge U. P. , New-York, 1883.
- DAHLQUIST, G. , 'A special stability problem for linear multistep methods', *BIT*, **3**, 27-43 (1963).
- HEUN, K. 'Neue Methode zur approximativen Integration der Differentialgleichungen einer unabhängigen Veränderlichen', *Z. Math. Physik*, **45**, 23-38 (1900).
- MOULTON, F. R., *New Methods in Exterior Ballistics*, U. of Chicago, Chicago, 1926.

Références bibliographiques

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exemples du chapitre 7

7.1	Applications du développement de Taylor	388
7.2	Applications du développement de Taylor	392

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exemple 7.1 Applications du développement de Taylor

Considérons l'équation différentielle suivante :

$$\begin{cases} y'(t) = -y(t), & t \in [0, 1], \\ y(0) = 1. \end{cases}$$

Ici on a $f(t, y) = -y$, et donc le schéma d'Euler simple donne les itérations suivantes sur z_n :

$$\begin{cases} z_0 = 1, \\ z_{n+1} = (1 - h)z_n, & 0 \leq n \leq N - 1. \end{cases}$$

La figure 7.3.2 permet de comparer la solution exacte $y(t) = \exp(-t)$ avec l'approximation obtenue avec le schéma aux points t_n , avec ici $h = 0.2$.

Il est clair que l'approximation donnée par le schéma d'Euler simple n'est pas très bonne et que l'on pourrait certainement l'améliorer en prenant un h plus petit ou en considérant le schéma correspondant à un développement de Taylor à l'ordre 2 :

$$y(t_{n+1}) = y(t_n) + hy'(t_n) + \frac{h^2}{2}y''(t_n) + \frac{h^3}{6}y''' \xi.$$

Puisque l'on a

$$y''(t) = \frac{d}{dt}f(t, y(t)) = \frac{\partial f}{\partial t}(t, y(t)) + f(t, y(t))\frac{\partial f}{\partial y}(t, y(t)),$$

on peut donc écrire

$$y(t_{n+1}) = y(t_n) + hy'(t_n) + \frac{h^2}{2} \left(\frac{\partial f}{\partial t}(t_n, y(t_n)) + f(t_n, y(t_n))\frac{\partial f}{\partial y}(t_n, y(t_n)) \right) + \mathcal{O}(h^3)$$

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

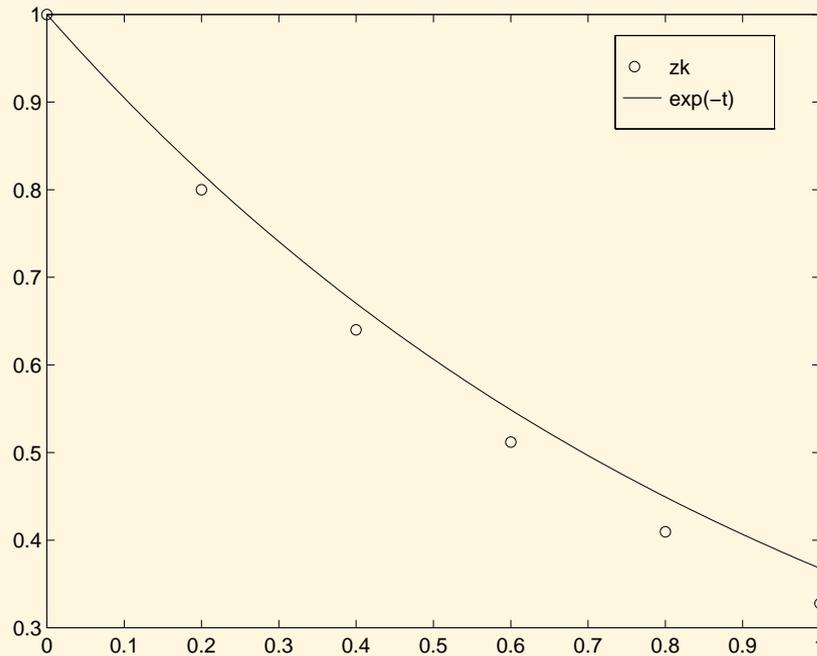


FIGURE 7.3.2 – Approximations obtenues avec le schéma d'Euler Simple ($h = 0.2$)

Exemple 7.1

Applications du développement de Taylor

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

et proposer le schéma suivant :

$$\begin{cases} z_0 = y_0, \\ z_{n+1} = z_n + hf(t_n, z_n) + \frac{h^2}{2} \left(\frac{\partial f}{\partial t}(t_n, z_n) + f(t_n, z_n) \frac{\partial f}{\partial y}(t, z_n) \right), \quad 0 \leq n \leq N-1. \end{cases}$$

Sur l'exemple, ce schéma donne les itérations suivantes :

$$\begin{cases} z_0 = 1, \\ z_{n+1} = (1 - h + \frac{h^2}{2})z_n, \quad 0 \leq n \leq N-1, \end{cases}$$

et les approximations z_k représentées sur la figure 7.3.3. Il est clair que ce schéma est meilleur, à pas égal, que le schéma d'Euler simple, mais ceci au prix du calcul de deux dérivées partielles de f , ce qui n'est pas toujours possible, par exemple quand f est le résultat d'un code de calcul trop complexe. C'est pourquoi les méthodes basées sur le développement de Taylor sont très peu utilisées dans la pratique.

[retour au cours](#)

Exemple 7.1
Applications du
développement
de Taylor

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

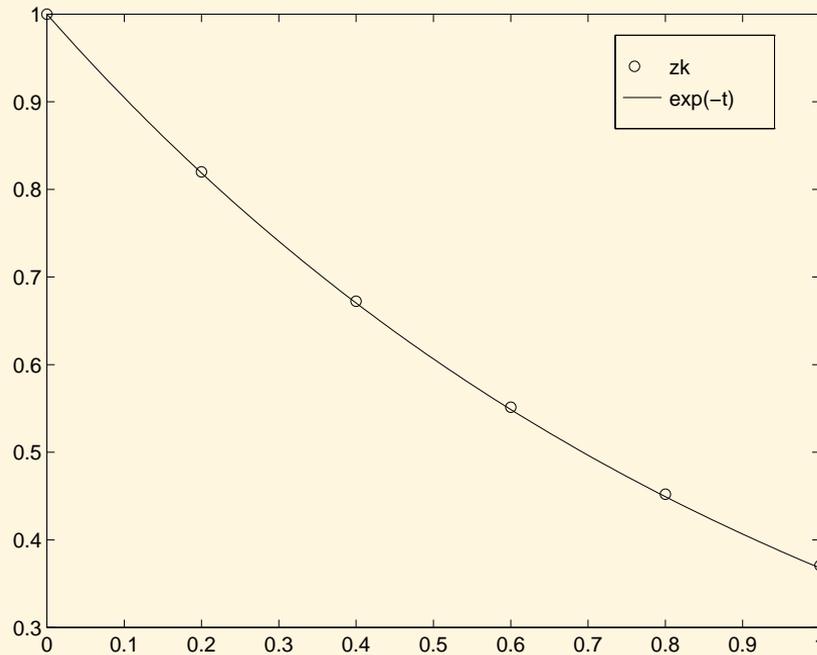


FIGURE 7.3.3 – Approximations obtenues avec le schéma de Taylor d'ordre 2 ($h = 0.2$)

Exemple 7.1

Applications du développement de Taylor

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exemple 7.2 Applications du développement de Taylor

Considérons le schéma d'Euler-Cauchy :

$$\begin{aligned}\tilde{z}_{n+1} &= z_n + hf(t_n, z_n), \\ z_{n+1} &= z_n + \frac{h}{2} (f(t_n, z_n) + f(t_{n+1}, \tilde{z}_{n+1})), \quad n \geq 0.\end{aligned}$$

Pour ce schéma on a

$$\phi(t, z, h) = \frac{1}{2}f(t, z) + \frac{1}{2}f(t+h, z+hf(t, z)),$$

et pour $u \in \mathbb{R}$ on a

$$\begin{aligned}\phi(t, z, h) - \phi(t, u, h) &= \frac{1}{2}(f(t, z) - f(t, u)) \\ &+ \frac{1}{2}(f(t+h, z+hf(t, z)) \\ &- f(t+h, u+hf(t, u))).\end{aligned}$$

Si f vérifie une condition de Lipschitz, on a

$$\begin{aligned}|\phi(t, z, h) - \phi(t, u, h)| &\leq \frac{L}{2}|z-u| + \frac{L}{2}|z+hf(t, z) - u - hf(t, u)|, \\ &\leq L|z-u| + \frac{L}{2}|hf(t, z) - hf(t, u)|, \\ &\leq L|z-u| + \frac{hL^2}{2}|z-u|, \\ &\leq (L + \frac{hL^2}{2})|z-u|.\end{aligned}$$

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

La fonction ϕ vérifie donc la condition de Lipschitz avec

$$\Lambda = L + \frac{hL^2}{2}.$$

Le schéma d'Euler-Cauchy est donc stable. On a démontré qu'il est consistant, il est donc convergent.

[retour au cours](#)

Exemple 7.2
Applications du
développement
de Taylor

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Documents du chapitre 7

7.1	Condition de Lipschitz	395
7.2	Définition des méthodes de Runge-Kutta	397
7.3	Adaptation automatique du pas	401

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Document 7.1 Condition de Lipschitz

Définition 7.3.1. Soit g une fonction définie sur $D \subset \mathbb{R}^n$ à valeurs dans \mathbb{R}^p , on dit que g vérifie une condition de **Lipschitz** (ou qu'elle est **lipschitzienne**) s'il existe une constante $L > 0$ telle que

$$\forall x, y \in D, \quad \|g(x) - g(y)\| \leq L\|x - y\|.$$

Théorème 7.3.2 (Cauchy-Lipschitz). On suppose que la fonction f est lipschitzienne en y , c'est-à-dire qu'il existe une constante $L > 0$ telle que

$$\forall t \in I, \forall x, y \in \mathbb{R}, \quad |f(t, x) - f(t, y)| \leq L|x - y|,$$

alors le problème

$$\begin{cases} y'(t) = f(t, y(t)), & t \in I, \\ y(t_0) = y_0, & t_0 \in I, \end{cases}$$

admet une solution unique.

Soit le problème

$$\begin{cases} y'(t) = \sqrt{y(t)} \\ y(0) = 0 \end{cases}$$

dont la solution n'est pas unique. Alors on voit que f ne satisfait pas la condition de Lipschitz. En effet, il vient

$$f(t, y) - f(t, z) = \sqrt{y} - \sqrt{z} = \frac{y - z}{\sqrt{y} + \sqrt{z}}$$

et le dénominateur est nul en l'origine.

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

[retour au cours](#)

Document 7.1
Condition de
Lipschitz

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Document 7.2 Définition des méthodes de Runge-Kutta

Une méthode à un pas n'utilise que le point (t_n, y_n) pour calculer (t_{n+1}, y_{n+1}) . Il va donc falloir pour obtenir des méthodes d'ordre suffisamment élevé, introduire des points intermédiaires. Soit donc le problème de Cauchy ci-dessous

$$\begin{cases} y' = f(t, y), & t \in [t_0, t_0 + T], \\ y(t_0) = y_0. \end{cases} \quad (7.3.6)$$

On cherche à le discrétiser par rapport à une subdivision $t_0 < t_1 < \dots < t_N = t_0 + T$. L'idée est donc de calculer le point (t_{n+1}, y_{n+1}) à partir du point (t_n, y_n) en utilisant des points intermédiaires $(t_{n,i}, y_{n,i})$. A chacun de ces points, on associe la 'pente'

$$p_{n,i} = f(t_{n,i}, y_{n,i}).$$

qui est en fait une approximation de la pente $y'(t_{n,i})$ de la tangente à la courbe représentative de la fonction y au point $t_{n,i}$.

Ces points intermédiaires vont être les nœuds d'une formule d'intégration numérique (Q) (Q, comme quadrature) :

$$\int_0^1 g(\tau) d\tau \approx \sum_{i=1}^q b_i g(c_i) \quad (Q)$$

qui servira à approcher l'expression exacte

$$y(t_{n+1}) = y(t_n) + \int_{t_n}^{t_{n+1}} f(t, y(t)) dt,$$

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

soit, après passage de l'intervalle de référence $[0, 1]$ à l'intervalle $[t_n, t_{n+1}]$:

$$t_{n,i} = t_n + c_i h_n, \quad c_i \in [0, 1],$$

$$y_{n+1} = y_n + h_n \sum_{i=1}^q b_i f(t_{n,i}, y_{n,i}),$$

Pour obtenir, un schéma numérique, il reste à construire les approximations $y_{n,i}$. Pour ce faire, on a à nouveau recours à l'intégration numérique, mais pour éviter d'introduire indéfiniment de nouveaux points intermédiaires, ces nouvelles formules de quadrature n'utiliseront que les nœuds c_i précédents.

On se donne alors pour chaque point $(t_{n,i}, y_{n,i})$, un méthode d'intégration numérique, que l'on définit sur chaque intervalle de référence, ici $[0, c_i]$:

$$\int_0^{c_i} g(\tau) d\tau \simeq \sum_{j=1}^q a_{ij} g(c_j), \quad (Q_i)$$

Soit alors $y(t)$ une solution exacte de (7.3.6). On a

$$\begin{aligned} y(t_{n,i}) &= y(t_n) + \int_{t_n}^{t_{n,i}} f(t, y(t)) dt \\ &= y(t_n) + h_n \int_0^{c_i} f(t_n + u h_n, y(t_n + u h_n)) du \end{aligned}$$

où l'on a fait le changement de variable $t = t_n + u h_n$. De même,

$$y(t_{n+1}) = y(t_n) + h_n \int_0^1 f(t_n + u h_n, y(t_n + u h_n)) du.$$

Document 7.2

Définition des
méthodes de
Runge-Kutta

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Lorsque l'on applique ces méthodes à $g(u) = f(t_n + u h_n, y(t_n + u h_n))$, il vient

$$\begin{cases} y(t_{n,i}) & \simeq y(t_n) + h_n \sum_{j=1}^q a_{ij} f(t_{n,j}, y(t_{n,j})), & i = 1, \dots, q \\ y(t_{n+1}) & \simeq y(t_n) + h_n \sum_{i=1}^q b_i f(t_{n,i}, y(t_{n,i})), \end{cases}$$

D'où la méthode de Runge-Kutta correspondante :

$$\begin{cases} t_{n,i} & = t_n + \tau_i h_n, \\ y_{n,i} & = y_n + h_n \sum_{j=1}^q a_{ij} p_{n,j}, \\ p_{n,i} & = f(t_{n,i}, y_{n,i}) \\ t_{n+1} & = t_n + h_n, \\ y_{n+1} & = y_n + h_n \sum_{i=1}^q b_i p_{n,i}, \end{cases}$$

On voit qu'il s'agit bien de méthodes à un pas avec

$$\begin{cases} \Phi(t, y, h) = \sum_{i=1}^q b_i f(t + c_i h, y_i), \text{ avec} \\ y_i = y + h \sum_{j=1}^q a_{ij} f(t + c_j h, y_j), \quad 1 \leq i \leq q. \end{cases}$$

les méthodes de Runge-Kutta correspondantes sont *explicites* : le calcul de y_{n+1} se fait à partir de la valeur de y_n sans avoir besoin de résoudre d'équation. Si la sommation sur j se fait de 1 à i , la méthode est dite *semi-implicite*. Si elle va au delà (jusqu'à q , au maximum), elle est dite *implicite*. Alors, il faut résoudre, à chaque pas, une équation, en général non linéaire, pour calculer y_{n+1} à partir de y_n .

Document 7.2

Définition des méthodes de Runge-Kutta

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

[retour au cours](#)

Document 7.2

Définition des
méthodes de
Runge-Kutta

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Document 7.3 Adaptation automatique du pas

La convergence d'un schéma traduit le fait que l'erreur globale $|y(t_n) - w_n|$ tend vers zéro en tous les points de la subdivision, lorsque le nombre de ces points tend vers l'infini. Comme le montre le théorème suivant, elle est directement reliée à l'erreur locale, et c'est cette relation qui justifie la suite de ce paragraphe :

Théorème 7.3.3. *Soit le schéma à un pas*

$$\begin{cases} z_0 = y_0, \\ z_{n+1} = z_n + h\phi(t_n, z_n, h), \quad n \geq 0, \end{cases} \quad (7.3.7)$$

approchant la solution de l'équation différentielle

$$y'(t) = f(t, y), \quad t \in [t_0, t_0 + T], \quad y(t_0) = y_0.$$

Si le schéma est convergent pour $h < h^$, et si l'erreur locale $\tau_{n+1}(h)$ vérifie*

$$\left| \frac{\tau_{n+1}(h)}{h} \right| \leq \epsilon(h), \quad \forall n, \quad \forall h < h^*$$

alors on a la majoration

$$|y(t_n) - z_n| \leq \frac{\epsilon(h)}{L} e^{L(t_n - t_0)}, \quad \forall n,$$

où L est la constante de Lipschitz de la fonction ϕ relative au schéma (7.3.7).

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

On peut donc contrôler l'erreur globale, pour peu que l'on puisse imposer à l'erreur locale de ne pas dépasser un certain seuil de tolérance. Cela permettrait par exemple d'utiliser un minimum de pas de temps, tout en respectant cette tolérance. Malheureusement, et ce n'est pas une surprise, avoir le minimum de pas de temps tout en contrôlant l'erreur globale est un objectif impossible à réaliser si le pas h reste constant.

On peut estimer l'erreur locale en procédant de la façon suivante : supposons que nous disposons d'un schéma à un pas d'ordre p

$$\begin{cases} z_0 = y_0, \\ z_{n+1} = z_n + h\phi(t_n, z_n, h), \quad n \geq 0. \end{cases} \quad (7.3.8)$$

L'erreur locale associée à ce schéma est donnée par la quantité

$$\tau_{n+1}(h) = y(t_{n+1}) - y(t_n) - h\phi(t_n, y(t_n), h) = \mathcal{O}(h^{p+1}),$$

où y désigne la solution de l'équation différentielle originale. On se donne un deuxième schéma

$$\begin{cases} \tilde{z}_0 = y_0, \\ \tilde{z}_{n+1} = \tilde{z}_n + h\tilde{\phi}(t_n, z_n, h), \quad n \geq 0, \end{cases} \quad (7.3.9)$$

d'ordre $p+1$, dont l'erreur locale $\tilde{\tau}_{n+1}(h)$ vérifie donc

$$\tilde{\tau}_{n+1}(h) = \mathcal{O}(h^{p+2}).$$

Supposons que l'on dispose de z_n pour un $n \geq 0$ donné, et appliquons les deux schémas pour obtenir une approximation de $y(t_{n+1})$: on a

$$\begin{aligned} z_{n+1} &= z_n + h\phi(t_n, z_n, h), \\ \tilde{z}_{n+1} &= z_n + h\tilde{\phi}(t_n, z_n, h). \end{aligned}$$

Document 7.3

Adaptation
automatique du
pas

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Si l'on fait l'hypothèse que $z_n = y(t_n)$ (vraie pour $n = 0$!), on peut écrire que

$$\begin{aligned}\tau_{n+1}(h) &= y(t_{n+1}) - z_n - h\phi(t_n, z_n, h), \\ &= y(t_{n+1}) - z_n - (z_{n+1} - z_n) = y(t_{n+1}) - z_{n+1}.\end{aligned}$$

Si on ajoute et retranche \tilde{z}_{n+1} à cette dernière quantité, on obtient

$$\begin{aligned}\tau_{n+1}(h) &= y(t_{n+1}) - \tilde{z}_{n+1} + \tilde{z}_{n+1} - z_{n+1}, \\ &= y(t_{n+1}) - y(t_n) - \tilde{\phi}(t_n, y(t_n), h) + \tilde{z}_{n+1} - z_{n+1}, \\ &= \tilde{\tau}_{n+1}(h) + \tilde{z}_{n+1} - z_{n+1}.\end{aligned}$$

Puisque $\tilde{\tau}_{n+1}(h) = \mathcal{O}(h^{p+2})$ et $\tau_{n+1}(h) = \mathcal{O}(h^{p+1})$ on peut donc négliger $\tilde{\tau}_{n+1}(h)$ devant $\tau_{n+1}(h)$, et estimer $\tau_{n+1}(h)$ par :

$$\hat{\tau}_{n+1}(h) = \tilde{z}_{n+1} - z_{n+1}.$$

On utilise cet estimateur de la façon suivante : on part d'un pas initial, typiquement $h = (T - t_0)/100$, et à chaque itération n , on calcule $\tilde{z}_{n+1} - z_{n+1}$ et donc $\hat{\tau}_{n+1}(h)$. On a ensuite les cas suivants à envisager :

- Si $\frac{|\hat{\tau}_{n+1}(h)|}{h} \in [\varepsilon/10, \varepsilon]$, alors on accepte z_{n+1} , on conserve le même pas pour l'itération suivante.
- Si $\frac{|\hat{\tau}_{n+1}(h)|}{h} \in [0, \varepsilon/10[$ on accepte z_{n+1} et on augmente le pas pour l'itération suivante (par exemple $h \leftarrow 2h$).
- Si $\frac{|\hat{\tau}_{n+1}(h)|}{h} > \varepsilon$, on refuse z_{n+1} , on diminue le pas (par exemple $h \leftarrow h/2$) et on recalcule z_{n+1} , \tilde{z}_{n+1} puis $\hat{\tau}_{n+1}(h)$.

Document 7.3
Adaptation
automatique du
pas

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

En général, on impose une valeur maximale au pas (souvent il s'agit de sa valeur initiale), ainsi qu'une valeur minimale. Lorsque le pas minimum est atteint, cela signifie que soit la valeur minimum est trop grande par rapport à la tolérance ε exigée, soit la solution de l'équation différentielle présente une singularité à l'instant t_n considéré.

La méthode la plus utilisée est basée sur deux schémas de Runge-Kutta, respectivement d'ordre 4 et 5 (c'est la méthode utilisée dans la fonction `ode45` de Matlab).

[retour au cours](#)

Document 7.3

Adaptation
automatique du
pas

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercices du chapitre 7

7.1	406
7.2	407
7.3	408
7.4	409
7.5	410
7.6	411
7.7	412

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 7.1

Résoudre l'équation différentielle du second ordre à coefficients constants

$$\begin{cases} \ddot{\theta}(t) = -\frac{g}{L}\theta(t), \\ \theta(0) = \theta_0, \dot{\theta}(0) = 0. \end{cases}$$

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 7.2

Montrer que pour tout réel $a \geq 0$,

$$y_a(t) = \begin{cases} 0, & \text{pour } 0 \leq t \leq a, \\ \frac{(t-a)^2}{4}, & \text{pour } a \leq t, \end{cases}$$

est solution de

$$\begin{cases} y'(t) = \sqrt{y(t)} \\ y(0) = 0 \end{cases}$$

Bien vérifier que c'est une fonction dérivable en tout point $t \geq 0$. (On s'en assurera en étudiant soigneusement le point de raccord $t = a$.)

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 7.3

Montrer que la résolution de

$$\begin{cases} \ddot{\theta}(t) = -\frac{g}{L} \sin \theta(t), \\ \theta(0) = \theta_0, \dot{\theta}(0) = 0, \end{cases}$$

où θ_0 , g et L sont des réels donnés, se ramène à la résolution d'un système d'équations différentielles du premier ordre

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 7.4

Expliquer comment on obtiendrait le premier itéré de la méthode d'Euler implicite pour résoudre

$$\begin{cases} \ddot{\theta}(t) = -\frac{g}{L} \sin \theta(t), \\ \theta(0) = \theta_0, \dot{\theta}(0) = 0, \end{cases}$$

où θ_0 , g et L sont des réels donnés

[retour au cours](#)

Solution

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 7.5

Expliquer comment on obtiendrait le premier itéré de la méthode d'Euler-Cauchy pour résoudre

$$\begin{cases} \ddot{\theta}(t) = -\frac{g}{L} \sin \theta(t), \\ \theta(0) = \theta_0, \dot{\theta}(0) = 0, \end{cases}$$

où θ_0 , g et L sont des réels donnés.

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 7.6

Montrer que les schémas explicite et implicite d'Euler appliqués à

$$\begin{cases} y' = -\lambda y(t), \text{ avec } \lambda > 0, \\ y(0) = y_0. \end{cases}$$

avec $nh = t$ (pas constant) conduisent chacun à une suite (z_n) telle que

$$\lim_{h \rightarrow 0, nh=t} z_n = y_0 e^{-\lambda(t)}.$$

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 7.7

Calculer l'ordre du schéma d'Euler-Cauchy.

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Chapitre 8

Détermination des valeurs propres et des vecteurs propres

8.1	Rappels et notations	414
8.2	Méthode de la puissance itérée - énoncé et hypothèses	416
8.3	Méthode de la puissance itérée - convergence et remarques	418
8.4	Méthode de la puissance itérée inverse - principe	421
8.5	Méthode de la puissance itérée inverse	423
8.6	Les méthodes de déflation	425
	Exercices du chapitre 8	428

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

8.1 Rappels et notations

On rappelle que, étant donnée une matrice $A \in \mathcal{M}_n(\mathbb{R})$, le problème de calcul des valeurs propres consiste à trouver un scalaire λ tel qu'il existe un vecteur y , $y \neq 0$, tel que

$$Ay = \lambda y.$$

Notations Dans tout ce chapitre on notera $\lambda_1, \lambda_2, \dots, \lambda_n$ les n valeurs propres distinctes ou non de A et on supposera désormais que

$$|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_n|,$$

λ_1 s'appelle la valeur propre *dominante* de A .

On supposera que la matrice A est diagonalisable, il existe donc une base de vecteurs propres que l'on notera $y^{(1)}, y^{(2)}, \dots, y^{(n)}$.

Le problème de la détermination des valeurs propres de A est équivalent à trouver les racines du polynôme caractéristique

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

On pourrait penser que le calcul numérique des valeurs propres se ramène au calcul numérique des racines d'un polynôme. En réalité c'est l'inverse qui se fait. En effet, étant donné un polynôme

$$p(t) = t^n + a_n t^{n-1} + a_{n-1} t^{n-2} + \dots + a_2 t + a_1,$$

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

on peut démontrer (voir exercice de TD) que ce polynôme est le polynôme caractéristique de la matrice suivante

$$A = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 1 \\ -a_1 & -a_2 & \dots & -a_{n-1} & -a_n \end{pmatrix}.$$

Les matrices A qui ont la forme précédente s'appellent des matrices "Compagnon".

Rappels et notations

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

8.2 Méthode de la puissance itérée - énoncé et hypothèses

Exercices :

[Exercice 8.1](#)

La méthode de la puissance itérée est la suivante :

$$\begin{cases} x^{(0)} \text{ donné dans } \mathbb{R}^n, \\ x^{(k+1)} = \frac{Ax^{(k)}}{\|Ax^{(k)}\|}, k \geq 0, \end{cases} \quad (8.1)$$

où $\|\cdot\|$ désigne une norme quelconque.

Hypothèses et notations

- On suppose que

$$|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \dots \geq |\lambda_n|.$$

On a donc que λ_1 est une valeur propre réelle simple (le montrer en exercice).

- On suppose que $x^{(0)}$ n'appartient pas au sous-espace engendré par les vecteurs propres $\{y^{(2)}, y^{(3)}, \dots, y^{(n)}\}$.
- Soit p un indice tel que $y_p^{(1)} \neq 0$

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Théorème 8.4. *Sous les hypothèses précédentes, la suite $(x^{(k)})_{k \in \mathbb{N}}$ générée par les relations (8.1) possède les propriétés suivantes :*

$$\lim_{k \rightarrow \infty} \|Ax^{(k)}\| = |\lambda_1|,$$

on a d'ailleurs plus précisément

$$\lim_{k \rightarrow \infty} \frac{[Ax^{(k)}]_p}{x_p^{(k)}} = \lambda_1.$$

De plus la suite $(x^{(k)})_{k \in \mathbb{N}}$ converge vers un vecteur propre associé à λ_1 de la façon suivante :

$$\lim_{k \rightarrow \infty} [\operatorname{sgn}(\lambda_1)]^k x^{(k)} = \gamma y^{(1)}, \quad (8.2)$$

où γ est une constante réelle.

**Méthode de la
puissance itérée**
- énoncé et
hypothèses

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

8.3 Méthode de la puissance itérée - convergence et remarques

On va maintenant démontrer le théorème 8.4.

On démontre immédiatement que l'on a :

$$x^{(k)} = \frac{A^k x^{(0)}}{\|A^k x^{(0)}\|}. \quad (8.3)$$

En effet cette propriété est vraie pour $k = 1$ par construction même de $x^{(1)}$, si l'on suppose que la propriété est vraie pour k , on a alors

$$x^{(k+1)} = \frac{Ax^{(k)}}{\|Ax^{(k)}\|} = \frac{A \frac{A^k x^{(0)}}{\|A^k x^{(0)}\|}}{\left\| A \frac{A^k x^{(0)}}{\|A^k x^{(0)}\|} \right\|} = \frac{A^{k+1} x^{(0)}}{\|A^{k+1} x^{(0)}\|}.$$

En développant $x^{(0)}$ sur la base des vecteur propres on peut écrire

$$x^{(0)} = \sum_{i=1}^n \xi_i y^{(i)}, \quad \xi_1 \neq 0.$$

On obtient alors

$$A^k x^{(0)} = \sum_{i=1}^n \xi_i A^k y^{(i)} = \sum_{i=1}^n \xi_i \lambda_i^k y^{(i)} = \lambda_1^k \left(\xi_1 y^{(1)} + \sum_{i=2}^n \xi_i \left[\frac{\lambda_i}{\lambda_1} \right]^k y^{(i)} \right) = \lambda_1^k w^{(k)}, \quad (8.4)$$

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

où l'on a posé

$$w^{(k)} = \xi_1 y^{(1)} + \sum_{i=2}^n \xi_i \left[\frac{\lambda_i}{\lambda_1} \right]^k y^{(i)},$$

en utilisant (8.3) et (8.4) on obtient

$$x^{(k)} = \frac{\lambda_1^k}{|\lambda_1^k| \|w^{(k)}\|} w^{(k)} = \operatorname{sgn}(\lambda_1)^k \frac{w^{(k)}}{\|w^{(k)}\|}, \quad (8.5)$$

or $\lim_{k \rightarrow \infty} w^{(k)} = \xi_1 y^{(1)}$, car $|\lambda_i/\lambda_1| < 1, \forall i > 1$, d'où

$$\lim_{k \rightarrow \infty} \operatorname{sgn}(\lambda_1)^k x^{(k)} = \gamma y^{(1)}, \text{ où } \gamma = \frac{\xi_1}{|\xi_1| \|y^{(1)}\|}.$$

De même d'après (8.3) et (8.4)

$$Ax^{(k)} = \frac{A^{k+1}x^{(0)}}{\|A^k x^{(0)}\|} = \frac{\lambda_1^{k+1} w^{(k+1)}}{|\lambda_1^k| \|w^k\|} = \lambda_1 \operatorname{sgn}(\lambda_1)^k \frac{w^{(k+1)}}{\|w^{(k)}\|} \quad (8.6)$$

d'où

$$\|Ax^{(k)}\| = |\lambda_1| \frac{\|w^{(k+1)}\|}{\|w^{(k)}\|},$$

et donc $\lim_{k \rightarrow \infty} \|Ax^{(k)}\| = |\lambda_1|$.

On a de plus en utilisant (8.5) et (8.6)

$$\frac{(Ax^{(k)})_p}{x_p^{(k)}} = \lambda_1 \frac{w_p^{(k+1)}}{w_p^{(k)}}.$$

**Méthode de la
puissance itérée
- convergence et
remarques**

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Or $\lim_{k \rightarrow \infty} w_p^{(k)} = \xi_1 y_p^{(1)}$ (non nul par hypothèse), donc

$$\lim_{k \rightarrow \infty} \frac{[Ax^{(k)}]_p}{x_p^{(k)}} = \lambda_1$$

Quelques remarques :

- Si l'on a par exemple $\lambda_1 = \lambda_2$ et $|\lambda_1| > |\lambda_3| \geq |\lambda_4| \geq \dots \geq |\lambda_n|$, c'est à dire si la valeur propre dominante est double la démonstration est encore valide, il suffit d'écrire

$$x^{(0)} = \xi_1 y^{(1)} + \sum_{i=3}^n \xi_i y^{(i)},$$

où $\xi_1 y^{(1)}$ est la composante de $x^{(0)}$ sur le sous espace propre associé à λ_1 (qui est ici de dimension 2).

Ce serait encore valable si λ_1 était de façon générale multiple. Par contre si on a $\lambda_1 \neq \lambda_2$ et $|\lambda_1| = |\lambda_2|$, alors la démonstration n'est pas valable, c'est ce qui se passe en particulier dans le cas des valeurs propres complexes.

- Dans la formule (8.2) apparaît le signe de λ_1 , ce qui veut dire que, si λ_1 est négatif, alors les vecteurs $x^{(k)}$ oscillent entre les deux vecteurs $+\gamma y^{(1)}$ et $-\gamma y^{(1)}$. On verra en TD (exercice) une variante de la méthode de la puissance itérée qui utilise la norme infinie judicieusement afin d'éviter ces oscillations.

**Méthode de la
puissance itérée
- convergence et
remarques**

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

8.4 Méthode de la puissance itérée inverse - principe

Si λ est valeur propre de A et que A est inversible, alors, de façon équivalente, λ^{-1} est valeur propre de A^{-1} et les vecteurs propres associés sont les mêmes. On a en effet

$$Ay = \lambda y \iff y = \lambda A^{-1}y \iff A^{-1}y = \lambda^{-1}y.$$

Par ailleurs définissons, pour q scalaire donné, la matrice

$$B = A - qI.$$

Cette matrice admet comme valeurs propres $\lambda_i - q$ où λ_i sont les valeurs propres de A , la matrice B^{-1} (si elle est définie, c'est-à-dire si q n'est pas valeur propre de A) admet pour valeurs propres

$$\mu_i = \frac{1}{\lambda_i - q}.$$

En effet si z est vecteur propre de B^{-1} associé à $(\lambda_i - q)^{-1}$, il est vecteur propre de B , donc de A et il est évidemment associé à la valeur propre λ_i . Vérification :

$$\begin{aligned} B^{-1}z = (\lambda_i - q)^{-1}z &\iff z = (\lambda_i - q)^{-1}Bz \iff Bz = (\lambda_i - q)z \iff (A - qI)z = (\lambda_i - q)z \\ &\iff Az - qz = (\lambda_i - q)z \iff Az = \lambda_i z. \end{aligned}$$

Soit A matrice diagonalisable dans \mathbb{R} , soit q un réel et soit λ_j une valeur propre de A qui vérifie

$$0 < |q - \lambda_j| < |q - \lambda_i|, \forall i \neq j, \quad (8.7)$$

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

c'est à dire q n'est pas valeur propre de A et λ_j est la valeur propre la plus proche de q . Il résulte de (8.7) que la matrice B^{-1}

$$B^{-1} = (A - qI)^{-1}$$

admet

$$v_1 = \frac{1}{\lambda_j - q}$$

comme valeur propre dominante. Donc pour q donné on peut s'inspirer de l'algorithme de la puissance itérée appliqué à B^{-1} ce qui permet d'obtenir v_1 , et on retrouve λ_j en posant

$$\lambda_j = \frac{1}{v_1} + q.$$

La méthode s'appelle méthode de la puissance itérée inverse.

Méthode de la puissance itérée inverse - principe

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

8.5 Méthode de la puissance itérée inverse

Posons $B = A - qI$. On donne $x^{(0)}$ arbitraire (ou presque!) et on définit la suite $x^{(k+1)}$ par

$$\begin{cases} Bu^{(k+1)} = x^{(k)} \\ x^{(k+1)} = \frac{u^{(k+1)}}{\|u^{(k+1)}\|}, \quad k \geq 0. \end{cases} \quad (8.8)$$

On peut énoncer le théorème suivant :

Théorème 8.5. Soit $q \in \mathbb{R}$ donné tel que $q \neq \lambda_i, \forall i$, et soit $B = A - qI$. Si le vecteur initial $x^{(0)}$ n'appartient pas au sous-espace engendré par $\{y^{(i)}\}_{i=1, \dots, n, i \neq j}$ alors la suite $(x^{(k)})_{k \in \mathbb{N}}$ générée par la méthode (8.8) possède les propriétés suivantes :

$$\lim_{k \rightarrow \infty} \|u^{(k)}\| = \left| \frac{1}{\lambda_j - q} \right|,$$

où λ_j est la valeur propre de A la plus proche de q . On a d'ailleurs plus précisément si $y_p^{(j)} \neq 0$

$$\lim_{k \rightarrow \infty} \frac{u_p^{(k+1)}}{x_p^{(k)}} = \frac{1}{\lambda_j - q}.$$

De plus la suite $(x^{(k)})_{k \in \mathbb{N}}$ converge vers $y^{(j)}$ le vecteur propre de A associé à λ_j de la façon suivante :

$$\lim_{k \rightarrow \infty} \left[\operatorname{sgn} \left(\frac{1}{\lambda_j - q} \right) \right]^k x^{(k)} = \gamma y^{(j)}, \quad (8.9)$$

où γ est une constante réelle.

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

On peut montrer que la méthode converge d'autant plus vite que $|q - \lambda_j|$ est petit devant $|q - \lambda_i|, \forall i \neq j$.

Méthode de la puissance itérée inverse

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

8.6 Les méthodes de déflation

Les méthodes précédentes permettent de déterminer une valeur propre et le vecteur propre associé, par exemple la méthode de la puissance itérée inverse permet de calculer la valeur propre de plus petite valeur absolue si elle est unique. Donc si l'on a déterminé par une méthode quelconque une valeur propre λ_1 (v.p. dominante par exemple), on peut songer à construire une matrice dans laquelle la valeur propre que l'on vient de déterminer n'est plus dominante et qui, par contre, possède toujours $(\lambda_j), j \neq 1$ comme valeurs propres. La méthode de la puissance (par exemple) permet alors de déterminer une nouvelle valeur propre dominante de la famille $(\lambda_j), j \neq 1$ et ainsi de suite. Cette procédure constitue ce qu'on appelle une méthode de déflation.

– 1^{er} cas : A symétrique.

Supposons que l'on connaisse un vecteur propre y associé à la valeur propre λ avec $\|y\|_2 = 1$. Considérons alors la matrice

$$B = A - \lambda y y^\top.$$

Evidemment on a

$$B y = A y - \lambda y y^\top y = \lambda y - \lambda y = 0.$$

Par ailleurs si z est un autre vecteur propre de A associé à une autre valeur propre μ on a

$$B z = A z - \lambda y y^\top z = \mu z$$

car $y^\top z = 0$ en vertu de l'orthogonalité des vecteurs propres.

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

– 2^{ème} cas : A quelconque (Méthode de *Duncan et Collar*)

Soit A une matrice quelconque et supposons qu'on a obtenu un couple propre (λ, y) et supposons pour simplifier que la première composante de y est égale à 1. Notons \underline{A}_1 la première ligne de la matrice A , par définition du couple (λ, y) on a

$$\underline{A}_1 y = \lambda y_1 = \lambda. \quad (8.10)$$

Définissons alors la matrice B par

$$B = A - y \underline{A}_1.$$

Alors la première ligne de B est nulle, en effet

$$\underline{B}_1 = \underline{A}_1 - y_1 \underline{A}_1 = 0.$$

Par ailleurs, d'après (8.10) :

$$B y = A y - y \underline{A}_1 y = \lambda y - \lambda y = 0. \quad (8.11)$$

Si l'on suppose que les autres vecteurs propres de A ont leur première composante non nulle alors on peut les normaliser de façon à ce qu'elle soit égale à 1. Dans ces conditions si (μ, z) est un autre couple propre de A ($\mu \neq \lambda$) on a $Az = \mu z$, donc

$$\underline{A}_1 z = \mu$$

et

$$B(z - y) = Bz = Az - y \underline{A}_1 z = \mu z - \mu y = \mu(z - y)$$

la première égalité étant une conséquence de (8.11). Les autres valeurs propres de B sont donc identiques à celles de A avec pour vecteurs propres associés $z - y$. Réciproquement

si (μ, u) est un couple propre de B , si on suppose que $\mu \neq 0$ et que $\underline{A}_1 u = \mu - \lambda$ (ceci est toujours possible si $\underline{A}_1 u \neq 0$) alors :

$$A(u + y) = Bu + y\underline{A}_1 u + Ay = \mu u + (\mu - \lambda)y + \lambda y = \mu(u + y)$$

On vient de démontrer que $(u + y)$ est vecteur propre de A associé à la valeur propre μ . Pour obtenir les couples propres de A autres que (λ, y) , on se ramène donc à rechercher les couples propres de B (valeur propre non nulle). Compte tenu de la structure particulière de B (sa première ligne étant nulle) on est ramené à un problème de calcul de valeurs propres dans \mathbb{R}^{n-1} .

Les méthodes de déflation

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercices du chapitre 8

8.1	429
-----	-------	-----

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Exercice 8.1

Montrer que si $A \in \mathcal{M}_{nn}$, si $\lambda_1, \lambda_2, \dots, \lambda_n$ les n valeurs propres de A vérifient $|\lambda_1| > |\lambda_2| \geq \dots \geq |\lambda_n|$, alors λ_1 est une valeur propre réelle et simple.

[retour au cours](#)

[Solution](#)

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Index des concepts

Le gras indique un grain où le concept est défini ; l'italique indique un renvoi à un exercice ou un exemple, le gras italique à un document, et le romain à un grain où le concept est mentionné.

x rang39, 44

A

Adams-Bashforth377
Adams-Bashforth - Exemples377, 380
Adams-Moulton 381
Adams-Moulton - Exemples381, 384
Application linéaire, noyau, image ... 15, 20, 30
Applications linéaires - composition, inverse 17
Approximation d'une fonction - généralités 260

B

Base - dimension 12

C

Convergence des méthodes itératives 209

D

Déflation-méthodes 425
Déterminant - définition par récurrence 33
Déterminant - produit, inverse, calcul pratique
37
Déterminant et colonnes 35, 37
Diagonalisation des matrices 51, 53
Dichotomie 213
Différences divisées
Calcul pratique 276

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)



Définition 272

E

Elimination de Gauss

étape générale..... 117

algorithme 119

deuxième étape 115

première étape 112

Elimination de Gauss

factorisation $A = LU$ 122, 125, 128

Equation différentielle

Exemple 347

Introduction..... 350

Références..... 385

Equations normales 191

Erreur de quadrature 319

Espace vectoriel - Généralités..... 7, 9

Euler-exemples 359

Euler-étude de la méthode 363

F

Factorisation $A = LU$

algorithme de Doolittle 132

calcul direct 130

faisabilité 135

motivations 128

permutation de lignes..... 138

pivot maximal 140

unicité..... 125

Factorisation LDL^T 153

Factorisation $PA = LU$ 142

algorithme 146

applications..... 148

Factorisation $PAQ = LU$ 150

Factorisation de Cholesky

existence 155, 158

mise en oeuvre 158

Famille libre, famille génératrice 10, 12

G

Gauss-Seidel 207

H

Horner 278

I

Intégration numérique

Formules composées..... 323

Introduction..... 312

Sommaire
Concepts

Exemples
Exercices
Documents

Interpolation	
Calcul d'erreur	267
Convergence	280
Introduction.....	262
Intervalle de référence	317

J

Jacobi	205, 207
--------------	-----------------

L

Lagrange-Polynôme d'interpolation.....	265
Lissage	187, 189

M

Matrice de passage.....	28
Matrices - définition.....	20, 23
Matrices - inverse et transposée	26
Matrices - somme et produit	23, 26
Moindres carrés	
formulation matricielle	189

N

Newton	
Convergence	223, 225, 229, 235

Equation non-linéaire.....	221
Système 2 équations non-linéaires ...	229, 235

Newton

Système non linéaire	235
Newton-Cotes-formules d'intégration.....	321
Newton-Polynôme d'interpolation	270

O

Ordre des schémas à un pas	366, 368
----------------------------------	-----------------

P

Point fixe	
Convergence	217, 232
Principe	215, 217, 232
Point fixe	
Système non linéaire	232
Principe des méthodes numériques... ..	355, 359
Puissance itérée.....	416
Puissance itérée inverse-méthode	423
Puissance itérée inverse-principe	421
Puissance itérée-convergence.....	418

Q

Quadrature - formules	315
-----------------------------	------------

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

R

Rang	30
Rang et déterminant	39
Rappels et notations	414
Runge-Kutta	371

S

Schémas prédicteur-correcteurs.....	361
Sous-espace vectoriel	9
Splines	
définition	285
motivations	283
Systèmes linéaires - Matrice carrée	42
Systèmes linéaires - Matrice non carrée	44
Système linéaire - motivations.....	106
Système linéaire triangulaire	110
Sécante	225

V

Valeurs propres - définition	47
Valeurs propres - matrices définies positives	53
Valeurs propres - matrices semblables.....	49

[Sommaire](#)
[Concepts](#)

[Exemples](#)
[Exercices](#)
[Documents](#)

Solution de l'exercice 1.1

Prendre la définition précise du document [1.1](#) et vérifier tous ses éléments en utilisant les propriétés connues des nombres réels.

[Retour à l'exercice ▲](#)

Solution de l'exercice 1.2

Prendre la définition précise du document [1.1](#) et vérifier toutes ses éléments en utilisant les propriétés connues des nombres réels.

[Retour à l'exercice ▲](#)

Solution de l'exercice 1.3

Cet ensemble n'a pas d'élément nul pour l'addition puisque le polynôme nul n'est pas de degré n .

[Retour à l'exercice ▲](#)

Solution de l'exercice 1.4

Il est facile de montrer que la somme de deux éléments de F est un élément de F et que le produit d'un nombre réel par un élément de F est un élément de F . Ainsi, pour le deuxième exemple, on a :

$$(\lambda_1 \vec{x} + \mu_1 \vec{y}) + (\lambda_2 \vec{x} + \mu_2 \vec{y}) = (\lambda_1 + \lambda_2) \vec{x} + (\mu_1 + \mu_2) \vec{y},$$

$$\alpha(\lambda \vec{x} + \mu \vec{y}) = \alpha \lambda \vec{x} + \alpha \mu \vec{y}.$$

[Retour à l'exercice ▲](#)

Solution de l'exercice 1.5

Pour $F \cap G$, il suffit de se rappeler la définition de $F \cap G = \{x \in E, x \in F \text{ et } x \in G\}$ et d'utiliser le fait que F et G sont des sous-espaces vectoriels.

Pour $F \cup G$, il faut exhiber un contre-exemple. Par exemple, si $F = \{\lambda(0, 1), \lambda \in \mathbb{R}\}$ et $G = \{\lambda(1, 0), \lambda \in \mathbb{R}\}$, $F \cup G = \{x \in E, x \in F \text{ ou } x \in G\}$ n'est pas un sous-espace vectoriel car la somme de deux vecteurs de $F \cup G$ tels que $(0, 1)$ et $(1, 0)$ n'est pas dans $F \cup G$.

[Retour à l'exercice ▲](#)

Solution de l'exercice 1.6

Ces questions ne posent pas de difficulté. En ce qui concerne la troisième question, $H = \mathbb{R}^3$ et $F \cup G$ n'est que la réunion des vecteurs de la droite et des vecteurs du plan. Pour l'unicité de la décomposition de la dernière question (comme pour toute démonstration d'unicité), on part de deux décompositions et on démontre qu'elles sont égales.

[Retour à l'exercice ▲](#)

Solution de l'exercice 1.7

Pour donner une famille liée, on peut prendre deux vecteurs dont l'un est un scalaire fois le premier, mais il y a une infinité d'exemples simples possibles...

Pour montrer que la famille $\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$ est libre et génératrice, il suffit d'appliquer les définitions, ce qui correspond au cas particulier $n = 3$ de la dernière question.

[Retour à l'exercice ▲](#)

Solution de l'exercice 1.8

$$\alpha(1, 1, -1) + \beta(1, 1, 1) + \gamma(\sqrt{2}, \sqrt{2}, 3) = 0 \Leftrightarrow \begin{cases} \alpha + \beta + \sqrt{2}\gamma = 0 \\ -\alpha + \beta + 3\gamma = 0 \end{cases} \Leftrightarrow \begin{cases} \alpha + \beta + \sqrt{2}\gamma = 0 \\ +2\beta + (3 + \sqrt{2})\gamma = 0 \end{cases}$$

Il existe des coefficients non tous nuls, par exemple $\gamma = 1$, $\beta = (-\sqrt{2} - 3)/2$, $\alpha = (-\sqrt{2} + 3)/2$, donc la famille est liée.

[Retour à l'exercice ▲](#)

Solution de l'exercice 1.9

On montre que la famille $\{x^0, x, x^2, \dots, x^n\}$ est libre et génératrice, en utilisant en particulier la définition d'un polynôme et plus particulièrement celle du polynôme nul.

[Retour à l'exercice ▲](#)

Solution de l'exercice 1.10

On considère deux décompositions et on utilise le fait que la famille B est libre.

[Retour à l'exercice ▲](#)

Solution de l'exercice 1.11

On choisit un vecteur dont les deux premières composantes sont différentes par exemple $\vec{z} = (1, 0, 0)$, puis on vérifie que les trois vecteurs de la famille ainsi obtenue sont linéairement indépendants. Cela suffit alors puisque l'on connaît la dimension de \mathbb{R}^3 , qui est égale à trois.

[Retour à l'exercice ▲](#)

Solution de l'exercice 1.12

1. $\text{Ker } u_1 = \{x \in \mathbf{R}^n, a_1x_1 + a_2x_2 + \dots + a_nx_n = 0\}$, $\text{Im } u_1 = \mathbf{R}$.
2. $\text{Ker } u_2 = P_1$, l'ensemble des polynômes constants, et $\text{Im } u_2 = P_{k-1}$.

[Retour à l'exercice ▲](#)

Solution de l'exercice 1.13

Vérifier les propriétés de l'application linéaire, puis montrer que $\text{Ker } u = G$ et $\text{Im } u = F$, ce qui donne le résultat de la dernière question.

[Retour à l'exercice ▲](#)

Solution de l'exercice 1.14

Aidez-vous d'une figure et tout est évident.

$$u(\vec{V} + \vec{V}') = u(\vec{V}) + u(\vec{V}'), \quad u(\lambda \vec{V}) = \lambda u(\vec{V}), \quad \text{Ker } u = \{\vec{0}\}, \quad \text{Im } u = \mathbb{R}^2.$$

L'application inverse d'une rotation d'angle θ est une rotation d'angle $-\theta$, c'est donc une application linéaire bijective, puisque vous venez de le démontrer pour la rotation d'angle θ .

[Retour à l'exercice ▲](#)

Solution de l'exercice 1.15

Soit $\{\vec{e}_1, \dots, \vec{e}_n\}$ la base de E . Par définition d'une matrice associée à une application linéaire, la j ème colonne de la matrice associée à i_E est constituée des composantes de $i_E(\vec{e}_j) = \vec{e}_j$ dans la base $\{\vec{e}_1, \dots, \vec{e}_n\}$. Les éléments de cette j ème colonne sont donc tous nuls sauf le j ème élément qui vaut 1. Cette matrice est donc bien la matrice identité.

[Retour à l'exercice ▲](#)

Solution de l'exercice 1.16

$u(1) = 0$, $u(x) = 1$ et $u(x^2) = 2x$. La matrice est donc

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

[Retour à l'exercice ▲](#)

Solution de l'exercice 1.17

1. $AB = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, BA = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$

2. $AB = \begin{pmatrix} -2 & 3 \\ 0 & 0 \\ 2 & -1 \end{pmatrix}$ et le produit BA est impossible car le nombre de colonnes de B est différent du nombre de lignes de A .

[Retour à l'exercice ▲](#)

Solution de l'exercice 1.18

Vous effectuez le produit et vous utilisez les formules trigonométriques bien connues :

$$\cos(\theta_1 + \theta_2) = \cos(\theta_1) \cos(\theta_2) - \sin(\theta_1) \sin(\theta_2),$$

$$\sin(\theta_1 + \theta_2) = \sin(\theta_1) \cos(\theta_2) + \cos(\theta_1) \sin(\theta_2).$$

[Retour à l'exercice ▲](#)

Solution de l'exercice 1.19

$$u(\vec{x}) = u\left(\sum x_j \vec{e}_j\right) = \sum x_j u(\vec{e}_j) = \sum_j x_j \left(\sum_i a_{ij} \vec{f}_i\right) = \sum_i \left(\sum_j a_{ij} x_j\right) \vec{f}_i = \sum_i y_i \vec{f}_i,$$

et l'unicité de la décomposition d'un vecteur sur une base donne

$$y_i = \sum_j a_{ij} x_j$$

soit $Y = AX$.

[Retour à l'exercice ▲](#)

Solution de l'exercice 1.20

En appliquant la proposition 1.3.2, les composantes de $u(\vec{x})$ dans la base canonique de \mathbb{R}^2 sont données par le produit

$$\begin{pmatrix} 3 & 4 & 5 \\ 1 & 2 & 6 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} = \begin{pmatrix} 9 \\ 11 \end{pmatrix}$$

soit $u(\vec{x}) = (9, 11)$.

[Retour à l'exercice ▲](#)

Solution de l'exercice 1.21

On a montré aussi dans cet exercice que son inverse est la rotation d'angle $-\theta$. La matrice inverse est la matrice associée à l'application linéaire inverse, soit

$$A = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}.$$

[Retour à l'exercice ▲](#)

Solution de l'exercice 1.22

Pour montrer l'unicité, on suppose que la matrice possède deux inverses B et C , qui vérifient donc

$$AC = CA = I, AB = BA = I.$$

Calculons alors le produit BAC de deux manières différentes

$$BAC = B(AC) = BI = B, BAC = (BA)C = IC = C,$$

ce qui donne $B = C$.

[Retour à l'exercice ▲](#)

Solution de l'exercice 1.23

Pour que ces produits existent, il faut évidemment que X et Y aient le même nombre n de composantes. Dans ce cas $X^T Y$ est une matrice à une ligne et une colonne, c'est donc un scalaire, plus précisément vous reconnaissez le produit scalaire de deux vecteurs dont les composantes seraient données par les éléments de X et Y . Par contre XY^T est une matrice à n lignes et n colonnes. Les deux matrices $X^T Y$ et XY^T étant de type différent, elles ne peuvent pas être égales !

[Retour à l'exercice ▲](#)

Solution de l'exercice 1.24

Par définition de la matrice de passage (voir le paragraphe [Matrice de passage](#)), on a

$$P = \begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix}.$$

[Retour à l'exercice ▲](#)

Solution de l'exercice 1.25

$$- A = \begin{pmatrix} 1 & 2 \\ 3 & -1 \end{pmatrix}.$$

$$u(\vec{e}'_1) = u(\vec{e}_1 + \vec{e}_2) = u(\vec{e}_1) + u(\vec{e}_2) = \vec{e}_1 + 3\vec{e}_2 + 2\vec{e}_1 - \vec{e}_2 = 3\vec{e}_1 + 2\vec{e}_2.$$

De même

$$u(\vec{e}'_2) = 7\vec{e}_2.$$

- Vous calculez \vec{e}_1 et \vec{e}_2 en fonction de \vec{e}'_1 et \vec{e}'_2 , ce qui donne

$$\vec{e}_1 = \frac{1}{3}(\vec{e}'_1 + \vec{e}'_2), \quad \vec{e}_2 = \frac{1}{3}(2\vec{e}'_1 - \vec{e}'_2).$$

Il suffit alors de remplacer dans le calcul de $u(\vec{e}'_1)$ et $u(\vec{e}'_2)$ de la question précédente.

- Les composantes de $u(\vec{e}'_1)$ et $u(\vec{e}'_2)$ sur \vec{e}'_1 et \vec{e}'_2 obtenues dans la question précédente vous donnent les colonnes de A' .
- Inversez la matrice P calculée dans l'exercice 1.24, vous pouvez alors calculer $P^{-1}AP$, ce qui doit vous redonner la matrice A' .

[Retour à l'exercice ▲](#)

Solution de l'exercice 1.26

- le rang de la matrice de la rotation dans le plan est 2 puisque cette matrice est inversible et de dimension 2.
- Le rang de A est 2, puisque les deux premières colonnes sont linéairement indépendantes et que la troisième colonne est identique à la deuxième.
- Le rang de A est 2 puisque les deux premières colonnes sont linéairement indépendantes et que les colonnes 3 et 4 sont des combinaisons linéaires des deux premières colonnes (lesquelles?).

[Retour à l'exercice ▲](#)

Solution de l'exercice 1.27

Ce sont ds calculs évidents que vous pouvez vérifier avec scilab lorsque les matrices sont numériques. Pour le dernier déterminant, vous pouvez "sortir" λ (pourquoi?).

[Retour à l'exercice ▲](#)

Solution de l'exercice 1.28

Les résultats sont : 1, λ , 1.

[Retour à l'exercice ▲](#)

Solution de l'exercice 1.29

On itère la définition du déterminant sur des matrices dont la dimension diminue jusqu'à ce que l'on arrive sur un scalaire. Il est à noter qu'une matrice diagonale est un cas particulier des matrices triangulaires.

[Retour à l'exercice ▲](#)

Solution de l'exercice 1.30

On peut prendre par exemple $A = I$ et $B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$.

[Retour à l'exercice ▲](#)

Solution de l'exercice 1.31

La réponse est -6 . et s'obtient en développant par rapport une ligne qui a plusieurs 0.

[Retour à l'exercice ▲](#)

Solution de l'exercice 1.32

Le déterminant d'une matrice étant égal au déterminant de sa transposée, on peut appliquer les résultats de l'exercice 1.29.

[Retour à l'exercice ▲](#)

Solution de l'exercice 1.33

Si deux matrices A et B sont semblables, alors il existe une matrice inversible P telle que

$$B = P^{-1}AP.$$

Il suffit alors de calculer les déterminants des deux membres et d'utiliser les règles sur le produit et sur l'inverse des déterminants.

[Retour à l'exercice ▲](#)

Solution de l'exercice 1.34

La réponse est toujours $-6!$ et s'obtient en développant par rapport une ligne (ou une colonne) qui a plusieurs 0.

[Retour à l'exercice ▲](#)

Solution de l'exercice 1.35

Le rang de la matrice est au moins égal à 2 car $\begin{vmatrix} 1 & 3 \\ 2 & 2 \end{vmatrix} \neq 0$. Pour montrer que le rang de A est égal à 3, il faut trouver un déterminant à 3 lignes et 3 colonnes non nul. Or les 4 déterminants de ce type sont nuls (les calculer). Le rang est donc définitivement 2.

[Retour à l'exercice ▲](#)

Solution de l'exercice 1.36

Le déterminant de A est égal à -13 . Alors l'inverse de A est donné par

$$A^{-1} = -\frac{1}{13}B^T,$$

où

$$B = \begin{pmatrix} -7 & 11 & -5 \\ 4 & -10 & 1 \\ -1 & -4 & 3 \end{pmatrix}$$

[Retour à l'exercice ▲](#)

Solution de l'exercice 1.37

$$\begin{aligned}
 x \in \text{Ker } A &\Leftrightarrow \begin{cases} x_1 + x_2 + x_3 = 0 \\ x_1 + 2x_2 - x_3 = 0 \\ x_1 + 3x_3 = 0 \\ 2x_1 + x_2 + 4x_3 = 0 \end{cases} \Leftrightarrow \begin{cases} x_1 + x_2 + x_3 = 0 \\ x_2 - 2x_3 = 0 \\ -x_2 + 2x_3 = 0 \\ -x_2 + 2x_3 = 0 \end{cases} \\
 &\Leftrightarrow \begin{cases} x_1 + x_2 + x_3 = 0 \\ x_2 - 2x_3 = 0 \\ 0 = 0 \\ 0 = 0 \end{cases} \Leftrightarrow \begin{cases} x_1 = -3x_3 \\ x_2 = 2x_3 \end{cases} \Leftrightarrow x = x_3 \begin{pmatrix} -3 \\ 2 \\ 1 \end{pmatrix}
 \end{aligned}$$

La dimension du noyau de A est donc égal à 1 (une base de ce noyau est $\left\{ \begin{pmatrix} -3 \\ 2 \\ 1 \end{pmatrix} \right\}$). Or

$$3 = \text{rang } A + \dim (\text{ker } A),$$

ce qui donne $\text{rang } A = 2$.

Nous venons de démontrer par l'exemple précédent, que la solution d'un système $Ax = 0$, dont la matrice $A \in \mathcal{M}_{n,p}$ est telle que $n > p$, peut avoir une solution non unique. Il est clair que ce système qui a plus d'équations que d'inconnues n'a pas toujours une solution (il est facile de construire des contre-exemples à 3 équations et 2 inconnues).

[Retour à l'exercice ▲](#)

Solution de l'exercice 1.38

On transforme ce système en un système équivalent

$$\left\{ \begin{array}{l} x_1 + x_2 + x_3 = 1 \\ x_2 - 2x_3 = -2 \\ -x_2 + 2x_3 = 2 \\ -x_2 + 2x_3 = 0 \end{array} \right.$$

qui n'a évidemment aucune solution.

[Retour à l'exercice ▲](#)

Solution de l'exercice 1.39

$$sI - A = \begin{pmatrix} s-5 & -2 \\ -2 & s-2 \end{pmatrix} \text{ donc}$$

$$\pi_A(s) = \det(sI - A) = (s-5)(s-2) - 4 = s^2 - 7s + 6 = (s-1)(s-6).$$

On obtient donc 2 valeurs propres réelles $\lambda = 1$ et $\lambda = 6$, on détermine les vecteurs propres associés :

- $\lambda = 1$,

$$(A - \lambda I)Y = 0 \Leftrightarrow \begin{cases} 4y_1 + 2y_2 = 0 \\ 2y_1 + y_2 = 0 \end{cases} \Leftrightarrow y_2 = -2y_1 \Leftrightarrow \begin{cases} Y = \alpha \begin{pmatrix} 1 \\ -2 \end{pmatrix} \\ \alpha \in \mathbb{R} \end{cases}$$

- $\lambda = 6$,

$$(A - \lambda I)Y = 0 \Leftrightarrow \begin{cases} -y_1 + 2y_2 = 0 \\ 2y_1 - 4y_2 = 0 \end{cases} \Leftrightarrow y_1 = 2y_2 \Leftrightarrow \begin{cases} Y = \alpha \begin{pmatrix} 2 \\ 1 \end{pmatrix} \\ \alpha \in \mathbb{R} \end{cases}$$

[Retour à l'exercice ▲](#)

Solution de l'exercice 1.40

- Si une matrice est triangulaire, son déterminant est le produit de ses éléments diagonaux, donc

$$\det(sI - A) = (s - a_{11})(s - a_{22}) \dots (s - a_{nn}).$$

- A non inversible $\iff \det A = 0 \iff 0$ est valeur propre de A .
- Si $B \in \mathcal{M}_{n,n}$, alors $\det B = \det B^T$ et donc $\det(sI - A) = \det(sI - A)^T = \det(sI - A^T)$. Il est facile d'exhiber un contre-exemple pour les vecteurs propres.

[Retour à l'exercice ▲](#)

Solution de l'exercice 1.41

Si deux matrices A et B sont semblables, il existe une matrice P telle que $B = P^{-1}AP$, ce qui donne

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

Si l'on se place dans $\mathcal{M}_{2,2}$, la matrice identité et la matrice triangulaire $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ ont le même déterminant mais ne sont pas semblables. Pour le vérifier, montrer qu'il est impossible de trouver une matrice inversible P telle que $PI = AP$.

[Retour à l'exercice ▲](#)

Solution de l'exercice 1.42

$$A = PDP^{-1} \Leftrightarrow AP = PD \Leftrightarrow AP_i = PD_i = d_{ii}P_i.$$

Ce que l'on vient d'écrire découle uniquement des propriétés du produit matriciel, pour lequel on a bien sûr utilisé le fait que D est diagonale.

On obtient donc que P_i (ième colonne de P) est vecteur propre de A associé à la valeur propre d_{ii} .

[Retour à l'exercice ▲](#)

Solution de l'exercice 1.43

$(B^T B) = B^T B$, ce qui prouve que 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.

[Retour à l'exercice ▲](#)

Solution de l'exercice 1.44

D'après l'exercice précédent, $B^T B$ est symétrique et semi-définie positive.

$$x^T B^T B x = 0 \Rightarrow \|Bx\|^2 = 0 \Rightarrow Bx = 0,$$

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.

[Retour à l'exercice ▲](#)

Solution de l'exercice 2.1

1. On démontre ce résultat par récurrence, c'est trivialement vérifié pour $i = 1$. Supposons que $x_i = ix_1$ pour $i \leq k$, on écrit alors la k ème équation, on obtient :

$$-(k-1)x_1 + 2kx_1 - x_{k+1} = 0 \Leftrightarrow x_{k+1} = (k+1)x_1.$$

Ce qui démontre le résultat.

2. En écrivant la dernière équation, on obtient

$$-(n-1)x_1 + 2nx_1 = 0 \Leftrightarrow x_1 = 0$$

En utilisant la question précédente on a donc $x = 0$.

3. On a vu dans le chapitre 1 qu'une condition nécessaire et suffisante pour que A soit inversible est que son noyau soit réduit à 0, c'est ce que l'on vient de montrer.

[Retour à l'exercice ▲](#)

Solution de l'exercice 2.2

```
1: pour  $i = 1$  jusqu'à  $n$  faire  
2:   si  $|a_{ii}| < \varepsilon$  alors  
3:     Arrêter l'algorithme et donner un message d'erreur  
4:   fin si  
5: fin pour  
6:  $x_1 \leftarrow \frac{b_1}{a_{11}}$   
7: pour  $i = 2$  jusqu'à  $n$  faire  
8:    $x_i \leftarrow \frac{b_i - \sum_{k=1}^{i-1} a_{ik}x_k}{a_{ii}}$   
9: fin pour
```

[Retour à l'exercice ▲](#)

Solution de l'exercice 2.3

Le système linéaire s'écrit

$$\left\{ \begin{array}{l} a_{11}x_1 + a_{12}x_2 + \dots + a_{1,n-1}x_{n-1} + a_{1n}x_n = b_1 \\ a_{22}x_2 + \dots + a_{2,n-1}x_{n-1} + a_{2n}x_n = b_2 \\ \dots = \dots \\ a_{n-1,n-1}x_{n-1} + a_{n-1,n}x_n = b_{n-1} \\ a_{nn}x_n = b_n \end{array} \right.$$

On commence donc par calculer l'inconnue x_n , puis l'inconnue $x_{n-1} = (b_{n-1} - a_{n-1,n}x_n) / a_{n-1,n-1}$ et on remonte ainsi jusqu'à x_1 . Ainsi, lorsque l'on arrive à la i ème équation, on a déjà calculé x_k pour $k = i + 1, \dots, n$. Or cette équation s'écrit

$$a_{ii}x_i + a_{i,i+1}x_{i+1} + \dots + a_{i,n-1}x_{n-1} + a_{in}x_n = b_i$$

ce qui permet de tirer x_i par la formule donnée dans l'énoncé. L'algorithme ne diffère de celui de l'exercice précédent que par les indices. À vous de l'écrire ...

[Retour à l'exercice ▲](#)

Solution de l'exercice 2.4

On élimine le premier élément de la i ème ligne L_i en effectuant une combinaison $L_i - \alpha L_1$, ce qui donne

$$a_{i1} - \alpha a_{11} = 0$$

soit

$$\alpha = \frac{a_{i1}}{a_{11}}.$$

On a alors

$$L_i^{(2)} = L_i - \alpha L_1$$

soit

$$a_{ij}^{(2)} = a_{ij} - \alpha a_{1j}, \text{ pour } j = 1, \dots, n, i = 2, \dots, n.$$

La même combinaison est effectuée sur les composantes du second membre, soit

$$b_i^{(2)} = b_i - \alpha b_1, i = 2, \dots, n.$$

[Retour à l'exercice ▲](#)

Solution de l'exercice 2.5

On élimine le deuxième élément de la i ème ligne $L_i^{(2)}$ en effectuant une combinaison $L_i^{(2)} - \alpha L_2^{(2)}$, ce qui donne

$$a_{i2}^{(2)} - \alpha a_{22}^{(2)} = 0$$

soit

$$\alpha = \frac{a_{i2}^{(2)}}{a_{22}^{(2)}}.$$

On a alors

$$L_i^{(3)} = L_i^{(2)} - \alpha L_2^{(2)}$$

soit

$$a_{ij}^{(3)} = a_{ij}^{(2)} - \alpha a_{2j}^{(2)}, \text{ pour } j = 2, \dots, n, i = 3, \dots, n.$$

La même combinaison est effectuée sur les composantes du second membre, soit

$$b_i^{(3)} = b_i^{(2)} - \alpha b_2^{(2)}, i = 3, \dots, n.$$

[Retour à l'exercice ▲](#)

Solution de l'exercice 2.6

On élimine le k ème élément de la k ème ligne $L_k^{(k)}$ en effectuant une combinaison $L_i^{(k)} - \alpha L_k^{(k)}$, ce qui donne

$$a_{ik}^{(k)} - \alpha a_{kk}^{(k)} = 0$$

soit

$$\alpha = \frac{a_{ik}^{(k)}}{a_{kk}^{(k)}}.$$

On a alors

$$L_i^{(k+1)} = L_i^{(k)} - \alpha L_k^{(k)}$$

soit

$$a_{ij}^{(k+1)} = a_{ij}^{(k)} - \alpha a_{kj}^{(k)}, \text{ pour } j = k, \dots, n, i = k+1, \dots, n.$$

La même combinaison est effectuée sur les composantes du second membre, soit

$$b_i^{(k+1)} = b_i^{(k)} - \alpha b_k^{(k)}, i = k+1, \dots, n.$$

Pour $j = k$ le coefficient α a été déterminé pour que $a_{ik}^{(k+1)} = 0$, donc dans la pratique on affecte directement 0 à ce coefficient sans effectuer le calcul. Les calculs sont donc faits pour i et j variant de $k+1$ à n .

[Retour à l'exercice ▲](#)

Solution de l'exercice 2.7

L'algorithme procède de la manière suivante :

$$\begin{aligned} \begin{pmatrix} 1 & 2 & 1 \\ 2 & 2 & 1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} &= \begin{pmatrix} 4 \\ 5 \\ 3 \end{pmatrix} \Leftrightarrow \begin{pmatrix} 1 & 2 & 1 \\ 0 & -2 & -1 \\ 0 & -1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 4 \\ -3 \\ -1 \end{pmatrix} \\ &\Leftrightarrow \begin{pmatrix} 1 & 2 & 1 \\ 0 & -2 & -1 \\ 0 & 0 & 1/2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 4 \\ -3 \\ 1/2 \end{pmatrix} \end{aligned}$$

La résolution de ce système triangulaire donne :

$$x_3 = 1, x_2 = 1, x_1 = 1.$$

[Retour à l'exercice ▲](#)

Solution de l'exercice 2.8

La démarche est de compter le nombre d'opérations à partir de la boucle la plus intérieure. Nous allons évaluer le nombre de multiplications/divisions, vous laissant le soin d'évaluer le nombre d'additions algébriques. On a ainsi :

- pour $j = k + 1 \rightarrow n$, $a_{ij} \leftarrow a_{ij} - ca_{kj}$: on effectue $n - k$ multiplications,
- calcul de b_i ET c : on effectue 1 multiplication et 1 division
- On effectue les opérations précédentes pour $i = k + 1 \rightarrow n$: on effectue donc $(n - k)(n - k + 2)$ multiplications/divisions
- On effectue ce qui précède pour $k = 1 \rightarrow n - 1$: on effectue donc $\sum_{k=1}^{n-1} (n - k)(n - k + 2)$ multiplications/divisions

soit

$$\sum_{k=1}^{n-1} (n - k)(n - k + 2) = \sum_{p=1}^{n-1} p(p + 2) = \sum_{p=1}^{n-1} p^2 + 2 \sum_{p=1}^{n-1} p = \frac{(n - 1)n(2n - 1)}{6} + 2 \frac{(n - 1)n}{2} \simeq \frac{1}{3}n^3.$$

Dans le résultat, on n'a gardé que les termes de plus haut degré.

[Retour à l'exercice ▲](#)

Solution de l'exercice 2.9

L'élément du produit des matrices L et U est donné par :

$$a_{ij} = \sum_{k=1}^n l_{ik} u_{kj}.$$

Or $l_{ik} = 0$ pour $i < k$ et $u_{kj} = 0$ pour $j < k$. Le produit des éléments sera donc nul lorsque k sera supérieur au plus petit des deux entiers i et j , d'où le résultat.

[Retour à l'exercice ▲](#)

Solution de l'exercice 2.10

On suppose qu'il y a deux décompositions possibles :

$$A = LU = \widehat{L}\widehat{U}.$$

Puisque A est inversible, L , U , \widehat{L} , \widehat{U} sont inversibles. On a alors

$$(\widehat{L})^{-1}L = \widehat{U}U^{-1}.$$

Le produit de deux matrices triangulaires inférieures (resp. supérieures) est une matrice triangulaire inférieure (resp. supérieure). Il en résulte que l'égalité précédente donne une matrice diagonale (car triangulaire inférieure et supérieure). D'autre part, la diagonale de $\widehat{U}U^{-1}$ ne comportant que des 1, la matrice produit est nécessairement la matrice identité. Ainsi

$$(\widehat{L})^{-1}L = \widehat{U}U^{-1} = I,$$

soit

$$L = \widehat{L}, U = \widehat{U}.$$

[Retour à l'exercice ▲](#)

Solution de l'exercice 2.11

La résolution de $Ly = b$ donne

$$y = \begin{pmatrix} 1 \\ 4 \\ -1 \end{pmatrix}$$

puis celle de $Ux = y$ donne

$$x = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

[Retour à l'exercice ▲](#)

Solution de l'exercice 2.12

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

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

$$LU = \begin{pmatrix} 2 & 0 & 0 \\ 4 & \times & 0 \\ -2 & \times & \times \end{pmatrix} \begin{pmatrix} 1 & \times & \times \\ 0 & 1 & \times \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 1 & -2 \\ 4 & 5 & -3 \\ -2 & 5 & 3 \end{pmatrix}.$$

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

$$LU = \begin{pmatrix} 2 & 0 & 0 \\ 4 & \times & 0 \\ -2 & \times & \times \end{pmatrix} \begin{pmatrix} 1 & 1/2 & -1 \\ 0 & 1 & \times \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 1 & -2 \\ 4 & 5 & -3 \\ -2 & 5 & 3 \end{pmatrix}.$$

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

$$LU = \begin{pmatrix} 2 & 0 & 0 \\ 4 & 3 & 0 \\ -2 & 6 & \times \end{pmatrix} \begin{pmatrix} 1 & 1/2 & -1 \\ 0 & 1 & \times \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 1 & -2 \\ 4 & 5 & -3 \\ -2 & 5 & 3 \end{pmatrix}.$$

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

$$LU = \begin{pmatrix} 2 & 0 & 0 \\ 4 & 3 & 0 \\ -2 & 6 & \times \end{pmatrix} \begin{pmatrix} 1 & 1/2 & -1 \\ 0 & 1 & 1/3 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 1 & -2 \\ 4 & 5 & -3 \\ -2 & 5 & 3 \end{pmatrix}.$$

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

$$LU = \begin{pmatrix} 2 & 0 & 0 \\ 4 & 3 & 0 \\ -2 & 6 & -1 \end{pmatrix} \begin{pmatrix} 1 & 1/2 & -1 \\ 0 & 1 & 1/3 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 1 & -2 \\ 4 & 5 & -3 \\ -2 & 5 & 3 \end{pmatrix}.$$

[Retour à l'exercice ▲](#)

Solution de l'exercice 2.13

Le raisonnement s'obtient en échangeant des lignes et les colonnes dans le raisonnement de l'algorithme de Doolittle et les matrices L et U . Ainsi, cela commence par :

En écrivant $A = LU$ et en se souvenant que les matrices L et U sont triangulaires, on obtient

$$a_{ij} = \sum_{k=1}^{j-1} l_{ik}u_{kj} + l_{ij} \text{ pour } i = j, j+1, \dots, n.$$

Ce qui est équivalent à

$$l_{ij} = a_{ij} - \sum_{k=1}^{j-1} l_{ik}u_{kj} \text{ pour } i = j, j+1, \dots, n.$$

Nous voyons que pour calculer les éléments l_{ij} de la j ème colonne de L , il nous faut connaître préalablement les éléments des colonnes 1 à $j-1$ de L ainsi que les éléments des lignes 1 à $j-1$ de U .
etc.... À vous de continuer.

[Retour à l'exercice ▲](#)

Solution de l'exercice 2.14

On découpe les matrices A , L et U :

$$\begin{pmatrix} [A]_k & \dots \\ \dots & \dots \end{pmatrix} = \begin{pmatrix} [L]_k & 0 \\ \dots & \dots \end{pmatrix} \begin{pmatrix} [U]_k & \dots \\ 0 & \dots \end{pmatrix}.$$

En effectuant le produit par blocs, on obtient alors

$$[A]_k = [L]_k [U]_k.$$

Les deux matrices triangulaires $[L]_k$ et $[U]_k$ sont inversibles car les éléments des diagonales des matrices L et U sont non nuls, donc la matrice $[A]_k$ est inversible.

[Retour à l'exercice ▲](#)

Solution de l'exercice 2.15

On rappelle que l'élément p_{ij} correspond à la i ème composante de $g(\vec{e}_j)$. On a donc

$$p_{ij} = (\vec{e}_{\sigma(j)})_i = \delta_{i,\sigma(j)}.$$

Pour montrer que $P^{-1} = P^T$, il suffit de calculer le produit PP^T :

$$(PP^T)_{ij} = \sum_{k=1}^n p_{ik}p_{jk} = \sum_{k=1}^n \delta_{i,\sigma(k)}\delta_{j,\sigma(k)}.$$

Le produit $\delta_{i,\sigma(k)}\delta_{j,\sigma(k)}$ est nul sauf si

$$i = j = \sigma(k)$$

et dans ce cas le produit vaut 1, ce qui montre le résultat.

[Retour à l'exercice ▲](#)

Solution de l'exercice 2.16

1. Le produit de matrices donne

$$a_{ij} = \sum_{k=1}^n (LD)_{ik} l_{jk},$$

soit, puisque la matrice D est diagonale

$$a_{ij} = \sum_{k=1}^n l_{ik} d_k l_{jk}.$$

Or, puisque la matrice L est triangulaire inférieure, le produit $l_{ik} l_{jk}$ est nul pour $k > \inf(i, j)$, soit $k > j$, ce qui donne le résultat.

2. En détaillant la somme de la première question, on obtient ($l_{jj} = 1$) :

$$a_{jj} = \sum_{k=1}^{j-1} d_k l_{jk}^2 + d_j l_{jj}^2 = \sum_{k=1}^{j-1} d_k l_{jk}^2 + d_j,$$

soit

$$d_j = a_{jj} - \sum_{k=1}^{j-1} d_k l_{jk}^2.$$

Et pour $i > j$

$$a_{ij} = \sum_{k=1}^{j-1} d_k l_{ik} l_{jk} + d_j l_{ij} l_{jj} = \sum_{k=1}^{j-1} d_k l_{ik} l_{jk} + d_j l_{ij},$$

soit

$$l_{ij} = \frac{a_{ij} - \sum_{k=1}^{j-1} d_k l_{ik} l_{jk}}{d_j}.$$

[Retour à l'exercice ▲](#)

Solution de l'exercice 2.17

Un raisonnement simple est basé sur le calcul de déterminant. En effet, puisque la matrice B est triangulaire, son déterminant est le produit des éléments de sa diagonale.

$$\det A = \det (BB^T) = \det B \det (B^T) = \prod_{i=1}^n (b_{ii})^2.$$

La matrice A est inversible puisque elle est définie positive (voir les rappels du chapitre 1). Le déterminant de A est donc non nul, d'où l'on déduit que les éléments b_{ii} sont non nuls.

[Retour à l'exercice ▲](#)

Solution de l'exercice 3.1

1. Soit $y = \alpha + \beta t$, l'équation de la droite considérée. Le problème de régression linéaire s'écrit

$$\min_{(\alpha, \beta) \in \mathbb{R}^2} E(\alpha, \beta)$$

où

$$E(\alpha, \beta) = \sum_{i=1}^m (\alpha + \beta t_i - b_i)^2.$$

La solution (α^*, β^*) donne les coefficients de la droite solution du problème de régression linéaire et elle vérifie

$$E(\alpha^*, \beta^*) = \min_{(\alpha, \beta) \in \mathbb{R}^2} E(\alpha, \beta).$$

2. Calculons les deux dérivées partielles

$$\begin{cases} \frac{\partial E}{\partial \alpha}(\alpha, \beta) = \sum_{i=1}^m 2(\alpha + \beta t_i - b_i) \\ \frac{\partial E}{\partial \beta}(\alpha, \beta) = \sum_{i=1}^m 2(\alpha + \beta t_i - b_i) t_i \end{cases}$$

La solution (α^*, β^*) du problème de régression linéaire est donc donnée par la solution des deux équations linéaires obtenues en regroupant les termes

$$\begin{cases} \alpha^* m + \beta^* \sum_{i=1}^m t_i = \sum_{i=1}^m b_i \\ \alpha^* \sum_{i=1}^m t_i + \beta^* \sum_{i=1}^m t_i^2 = \sum_{i=1}^m b_i t_i \end{cases}$$

[Retour à l'exercice ▲](#)

Solution de l'exercice 3.2

Soit

$$p(t) = \alpha_1 + \alpha_2 t + \dots + \alpha_n t^{n-1},$$

un polynôme de degré $n - 1$. Pour que ce polynôme approche les données $(t_i, b_i)_{i=1, \dots, m}$ le plus près possible, il doit minimiser la quantité suivante

$$E(\alpha_1, \alpha_2, \dots, \alpha_n) = \sum_{i=1}^m (p(t_i) - b_i)^2 = \|Ax - b\|_2^2.$$

En effet on peut écrire

$$\begin{pmatrix} p(t_1) - b_1 \\ p(t_2) - b_2 \\ \dots \\ p(t_m) - b_m \end{pmatrix} = \begin{pmatrix} \alpha_1 + \alpha_2 t_1 + \dots + \alpha_n t_1^{n-1} - b_1 \\ \alpha_1 + \alpha_2 t_2 + \dots + \alpha_n t_2^{n-1} - b_2 \\ \dots \\ \alpha_1 + \alpha_2 t_m + \dots + \alpha_n t_m^{n-1} - b_m \end{pmatrix} = A \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \dots \\ \alpha_n \end{pmatrix} - \begin{pmatrix} b_1 \\ b_2 \\ \dots \\ b_m \end{pmatrix} = Ax - b,$$

où les matrices A et x sont définies par :

$$A = \begin{pmatrix} 1 & t_1 & \dots & t_1^{n-1} \\ 1 & t_2 & \dots & t_2^{n-1} \\ \dots & \dots & \dots & \dots \\ 1 & t_m & \dots & t_m^{n-1} \end{pmatrix}, \quad x = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}.$$

Dans les données, les points t_i sont tous distincts, ce qui implique que la matrice V , constituée des n premières lignes de A est inversible, d'où la matrice A est de rang n .

[Retour à l'exercice ▲](#)

Solution de l'exercice 3.3

La matrice A du problème de régression linéaire s'écrit (voir la correction de l'exercice 3.2) :

$$A = \begin{pmatrix} 1 & t_1 \\ 1 & t_2 \\ \dots & \dots \\ 1 & t_m \end{pmatrix}.$$

Les équations normales sont

$$A^T A \hat{x} = A^T b$$

ce qui donne en effectuant les produits matriciels

$$\begin{pmatrix} m & \sum_{i=1}^m t_i \\ \sum_{i=1}^m t_i & \sum_{i=1}^m t_i^2 \end{pmatrix} \hat{x} = \begin{pmatrix} \sum_{i=1}^m b_i \\ \sum_{i=1}^m t_i b_i \end{pmatrix}.$$

On retrouve bien ainsi le système de deux équations à deux inconnues de l'exercice 3.1.

[Retour à l'exercice ▲](#)

Solution de l'exercice 4.1

La définition du déterminant à partir des permutations donne

$$\det A = \sum_{\sigma} a_{\sigma(1)1} a_{\sigma(2)2} \dots a_{\sigma(n)n}$$

où la somme porte sur toutes les permutations de l'ensemble $\{1, 2, \dots, n\}$. Or si la matrice est inversible, son déterminant est non nul et il existe donc au moins un élément de cette somme qui est non nul, cet élément correspond à une permutation $\hat{\sigma}$. On a donc

$$a_{\hat{\sigma}(1)1} a_{\hat{\sigma}(2)2} \dots a_{\hat{\sigma}(n)n} \neq 0,$$

donc $\forall i = 1, \dots, n$ $a_{\hat{\sigma}(i)i} \neq 0$. Si on permute les lignes de A et que l'on note \hat{A} la matrice définie par $\hat{A} = \begin{pmatrix} \underline{A}_{\hat{\sigma}(1)} \\ \underline{A}_{\hat{\sigma}(2)} \\ \vdots \\ \underline{A}_{\hat{\sigma}(n)} \end{pmatrix}$, les

termes diagonaux de \hat{A} qui sont $a_{\hat{\sigma}(1)1}, a_{\hat{\sigma}(2)2}, \dots, a_{\hat{\sigma}(n)n}$ sont tous non nuls.

Il suffit alors de permuter les lignes de la matrice A suivant la permutation $\hat{\sigma}$ pour obtenir le résultat attendu.

[Retour à l'exercice ▲](#)

Solution de l'exercice 4.2

$$D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & 9 \end{pmatrix}.$$

$$E = \begin{pmatrix} 0 & 0 & 0 \\ -6 & 0 & 0 \\ -3 & -4 & 0 \end{pmatrix}.$$

$$F = \begin{pmatrix} 0 & -3 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

[Retour à l'exercice ▲](#)

Solution de l'exercice 4.3

La matrice $D-E$ est triangulaire inférieure. Le nombre de multiplications/divisions est de l'ordre de n^2 pour la résolution d'un système triangulaire, ce qui est beaucoup moins important que les n^3 de la méthode LU. Par contre, la diagonale de A ne doit pas comporter d'éléments nuls.

[Retour à l'exercice ▲](#)

Solution de l'exercice 4.4

On a donc $x^{(k+1)} = M^{-1}Nx^{(k)} + M^{-1}b$, donc en reprenant les notations du paragraphe [Convergence des méthodes itératives](#), on a $C = M^{-1}N$ et $d = M^{-1}b$. On utilise la proposition [4.1.1](#). Si $\|M^{-1}N\| < 1$, alors la suite $x^{(k)}$ converge vers la solution unique de

$$(I - M^{-1}N)x = d \Leftrightarrow x - M^{-1}Nx = M^{-1}b \Leftrightarrow (M - N)x = b \Leftrightarrow Ax = b.$$

On remarque au passage que si M inversible, si $\|M^{-1}N\| < 1$, alors la matrice $A = M - N$ est inversible, en effet il suffit de reprendre la proposition [4.1.1](#), $A = M(I - M^{-1}N)$ est un produit de matrices inversibles.

[Retour à l'exercice ▲](#)

Solution de l'exercice 4.5

La méthode de la dichotomie est basée sur le fait que la solution x^* de $f(x) = 0$ recherchée appartient à une suite de segments emboîtés $[a_k, b_k]$ pour $k \in \mathbb{N}$. Or la longueur de ces segments est donnée par

$$l_k = \frac{b-a}{2^k}.$$

Ceci implique que

$$x^* - a_k \leq l_k.$$

Pour que a_n soit une approximation de x^* avec un précision inférieure à ε , il suffit que

$$l_n \leq \varepsilon$$

soit

$$\frac{b-a}{2^n} \leq \varepsilon \Rightarrow n \leq \log_2 \frac{b-a}{\varepsilon}$$

[Retour à l'exercice ▲](#)

Solution de l'exercice 4.6

Vous pouvez vous aider de SCILAB pour tracer graphiquement la courbe $g(x)$, la bissectrice et les points $(u_k, g(u_k))$.

- Voir sur la figure 8.0.1 un exemple de ce qui vous pourriez obtenir pour g_1

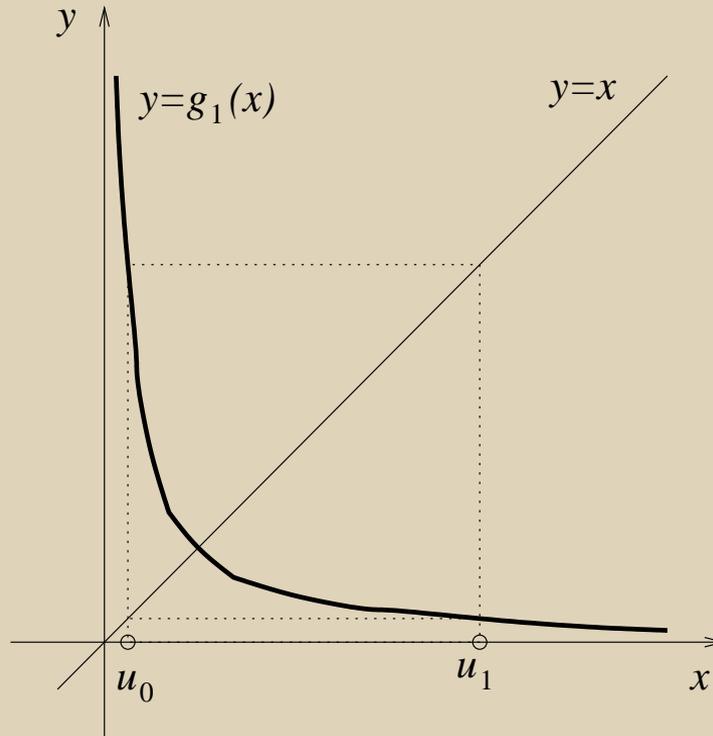


FIGURE 8.0.1 – Point fixe 1

Il est facile de calculer dans le cas de g_1 la suite des itérés. En effet

$$u_0 \text{ donné, } u_1 = \frac{a}{u_0}, u_{2n} = u_0, u_{2n+1} = u_1, n = 1, 2, \dots$$

La suite prend successivement les valeurs u_0 et u_1 et ne converge donc pas (sauf si par hasard $u_0 = \sqrt{a}$!).

- Voir sur la figure 8.0.2 un exemple de ce qui vous pourriez obtenir pour g_2
Le tracé que vous avez effectué pour g_2 vous montre que la suite u_k diverge.
- Voir sur la figure 8.0.3 un exemple de ce qui vous pourriez obtenir pour g_3
Le tracé que vous avez effectué pour g_3 vous montre que la suite u_k se rapproche de la solution. Après avoir étudié les théorèmes sur la convergence d'une telle suite, vous pourrez démontrer qu'elle converge.

[Retour à l'exercice ▲](#)

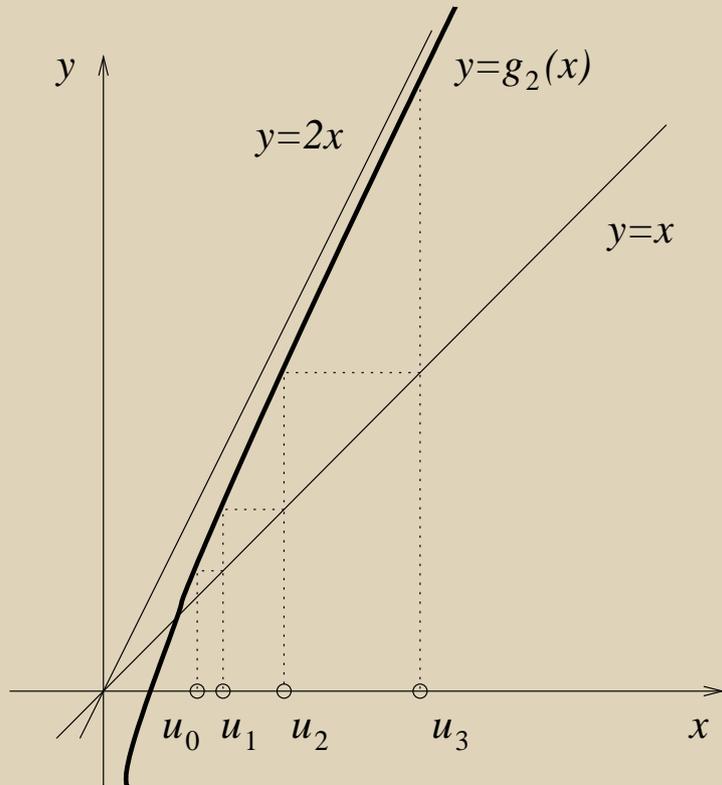


FIGURE 8.0.2 – Point fixe 2

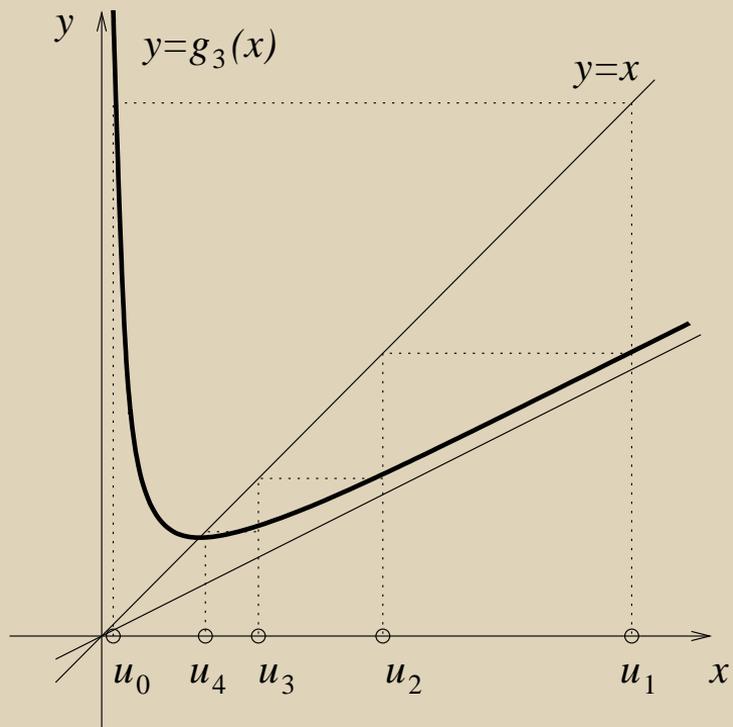


FIGURE 8.0.3 – Point fixe 3

Solution de l'exercice 4.7

Le raisonnement se fait par l'absurde en supposant qu'il existe deux points fixes distincts x^* et \hat{x} distincts dans $[a, b]$. Alors, le théorème des accroissements finis donne

$$g(x^*) - g(\hat{x}) = (x^* - \hat{x})g'(\xi),$$

où ξ est compris entre x^* et \hat{x} et donc appartient à $[a, b]$. Alors

$$|x^* - \hat{x}| = |g(x^*) - g(\hat{x})| = |x^* - \hat{x}| |g'(\xi)| \leq k|x^* - \hat{x}| < |x^* - \hat{x}|$$

ce qui est absurde !

[Retour à l'exercice ▲](#)

Solution de l'exercice 4.8

Pour appliquer la proposition il suffit de démontrer que l'application g va de $[-1, 1]$ dans lui-même et que sa dérivée est bornée par une constante strictement inférieure à 1 sur cet intervalle. Faites un tableau de variation de la fonction g sur $[-1, 1]$ et vous démontrerez la première partie sans problème. Quant à la dérivée $g'(x) = \frac{2}{3}x \dots$

[Retour à l'exercice ▲](#)

Solution de l'exercice 4.9

1. L'inégalité se démontre facilement en utilisant l'inégalité de l'hypothèse pour $i = n, n-1, \dots, 0$.
2. On écrit

$$x_p - x_n = x_p - x_{p-1} + x_{p-1} - \dots - x_{n+1} + x_{n+1} - x_n$$

on prend les valeurs absolues, on utilise l'inégalité triangulaire, puis la majoration précédente.

3. On a

$$k^{p-1} + \dots + k^{n+1} + k^n = k^n(k^{p-1-n} + \dots + k + 1) = k^n \frac{1 - k^{p-n}}{1 - k}$$

ce qui donne

$$|x_p - x_n| \leq k^n \frac{1 - k^{p-n}}{1 - k} |x_1 - x_0|.$$

On passe alors à la limite quand p tend vers $+\infty$ sur les deux membres de l'inégalité (le passage à la limite conserve l'inégalité). Alors, puisque $k < 1$

$$\lim_{p \rightarrow +\infty} k^{p-n} = 0$$

et, par hypothèse, $\lim_{p \rightarrow +\infty} x_p = x^*$, ce qui donne le résultat.

[Retour à l'exercice ▲](#)

Solution de l'exercice 4.10

On utilise les résultats des exercices 4.8 et 4.9, ce qui donne

$$|x^* - x_n| \leq k^n \frac{1}{1-k} |x_1 - x_0|, \text{ avec } k = \frac{2}{3}.$$

On obtient donc

$$x_1 = -\frac{1}{3}, k^n \frac{1}{1-k} |x_1 - x_0| = \left(\frac{2}{3}\right)^n, |x^* - x_n| \leq \left(\frac{2}{3}\right)^n.$$

Il suffit donc d'avoir

$$\left(\frac{2}{3}\right)^n \leq \epsilon \Leftrightarrow n \ln\left(\frac{2}{3}\right) \leq \ln \epsilon \Leftrightarrow n \geq 18.$$

[Retour à l'exercice ▲](#)

Solution de l'exercice 4.11

L'équation de la droite tangente est évidemment

$$y = f'(x_n)(x - x_n) + f(x_n).$$

L'abscisse x_{n+1} de l'intersection de la droite tangente avec l'axe Ox est donné par

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

ce qui permet de retrouver la formule de Newton :

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

Le calcul de x_1 et x_2 donne

$$x_1 = 2 - \frac{2}{4} = \frac{3}{2}, \quad x_2 = \frac{3}{2} - \frac{\frac{1}{4}}{3} = \frac{17}{12}.$$

[Retour à l'exercice ▲](#)

Solution de l'exercice 4.12

Puisque $f(0)f(\frac{\pi}{2}) < 0$ la fonction f admet au moins une racine sur $[0, \frac{\pi}{2}]$. En calculant on obtient

$$x_0 = 0, x_1 = 1, x_2 = 0.7504, x_3 = 0.7391$$

puis

$$x_0 = \frac{\pi}{4}, x_1 = 0.7395, x_2 = 0.7391, x_3 = 0.7391$$

[Retour à l'exercice ▲](#)

Solution de l'exercice 4.13

L'équation de la droite (pour ceux qui auraient un trou de mémoire) est donnée par

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

x_{n+1} est l'abscisse de l'intersection de cette droite avec l'axe Ox , on a donc

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

On retrouve la formule de la sécante.

[Retour à l'exercice ▲](#)

Solution de l'exercice 4.14

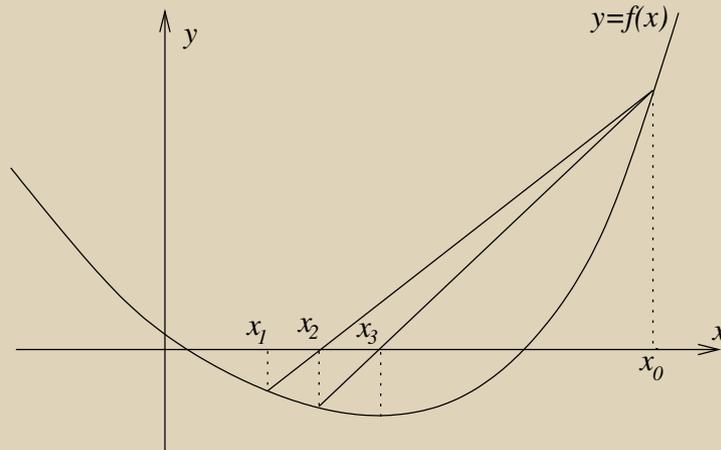


FIGURE 8.0.4 – Méthode de la sécante modifiée

Comme on le voit sur la figure 8.0.4, la suite construite par la méthode de la sécante modifiée est obtenue successivement avec les couples x_0, x_1 puis x_2, x_0 puis x_3, x_0 , chacun des couples encadre la racine.

Par la méthode de la sécante simple, à partir de x_0, x_1 on aurait construit x_2 , (le même !) puis à partir de x_1, x_2 on aurait construit un x'_3 qui n'est pas x_3 .

[Retour à l'exercice ▲](#)

Solution de l'exercice 4.15

Donnons d'abord la fonction f et sa matrice jacobienne

$$f(x) = \begin{pmatrix} x_1^2 + (x_2^2 - 1)^2 - 4 \\ x_1^2 - x_2^2 + 1 \end{pmatrix} \quad Df(x) = \begin{pmatrix} 2x_1 & 2(x_2 - 1) \\ 2x_1 & 2x_2 \end{pmatrix}$$

Alors on peut calculer $x^{(1)}$ en fonction de $x^{(0)}$ de la manière suivante :

$$(Df(x^{(0)}))h = -f(x^{(0)}), \quad x^{(1)} = x^{(0)} + h$$

Par exemple, si on choisit

$$x^{(0)} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \text{ on calcule } Df(x^{(0)}) = \begin{pmatrix} 2 & 0 \\ 2 & -1 \end{pmatrix}, \quad f(x^{(0)}) = \begin{pmatrix} -3 \\ 1 \end{pmatrix}$$

puis après résolution

$$h = \begin{pmatrix} 1.5 \\ 4 \end{pmatrix}, \quad x^{(1)} = \begin{pmatrix} 2.5 \\ 5 \end{pmatrix}, \quad Df(x^{(1)}) = \begin{pmatrix} 5 & 8 \\ 5 & -1 \end{pmatrix}, \quad f(x^{(1)}) = \begin{pmatrix} 18.25 \\ 2.25 \end{pmatrix}$$

Après quelques itérations on obtient la solution $x = \begin{pmatrix} 1.25 \\ 2.56 \end{pmatrix}$.

[Retour à l'exercice ▲](#)

Solution de l'exercice 5.1

Le problème s'écrit :

$$\sum_{k=0}^n a_k t_i^k = f(t_i), \quad i = 0, \dots, n.$$

C'est donc un système linéaire de $n + 1$ équations à $n + 1$ inconnues. Ce problème a une solution unique puisque la matrice M du système (appelée matrice de Van der Monde) est alors inversible.

$$M = \begin{pmatrix} 1 & t_0 & \dots & t_0^n \\ 1 & t_1 & \dots & t_1^n \\ \dots & \dots & \dots & \dots \\ 1 & t_n & \dots & t_n^n \end{pmatrix}$$

[Retour à l'exercice ▲](#)

Solution de l'exercice 5.2

On doit avoir $m \geq 2$ car par trois points non alignés on ne peut faire passer une droite!

Pour $m = 2$, le polynôme d'interpolation s'écrit $p(t) = 1 - t^2$. Remarquons que pour $m > 2$, il passe un infinité de polynômes par trois points.

[Retour à l'exercice ▲](#)

Solution de l'exercice 5.3

1. La droite passant par les points (t_i, z_i) et (t_{i+1}, z_{i+1}) a pour équation $y = g_i(t)$. g_i s'écrit :

$$g_i(t) = z_i \frac{t - t_{i+1}}{t_i - t_{i+1}} + z_{i+1} \frac{t - t_i}{t_{i+1} - t_i}.$$

2. La ligne brisée a pour équation $y = g(t)$ où g est une fonction définie par morceau :

$$g(t) = \begin{cases} g_0(t) & \text{pour } t \in [t_0, t_1] \\ g_1(t) & \text{pour } t \in [t_1, t_2] \\ \dots & \\ g_{n-1}(t) & \text{pour } t \in [t_{n-1}, t_n] \end{cases}$$

On peut remarquer que $g_i(t_{i+1}) = g_{i+1}(t_{i+1}) = z_{i+1}$, la fonction g est une fonction continue sur $[t_0, t_n]$, par contre g n'est pas dérivable aux points t_1, \dots, t_{n-1} , en ces points la courbe présente des points anguleux, les dérivées à droite et à gauche existent mais sont différentes.

On verra plus loin les splines cubiques qui sont également définies par morceaux, mais qui ont plus de régularité.

[Retour à l'exercice ▲](#)

Solution de l'exercice 5.4

1. Si l'on applique le résultat sur le calcul d'erreur, on trouve

$$e(t) = 0$$

car la dérivée d'ordre $n + 1$ d'un polynôme de degré n est nulle. Ce résultat s'explique car par $n + 1$ points il passe un polynôme et un seul de degré n , c'est donc t^n !

2. Si l'on applique le résultat sur le calcul d'erreur, on trouve

$$e(t) = \frac{1}{(n+1)!} (n+1)! \pi_n(t) = \pi_n(t),$$

car la dérivée d'ordre $n + 1$ d'un polynôme de degré t^{n+1} est $(n + 1)!$.

On aurait pu retrouver ce résultat directement. Si l'on note p le polynôme de degré inférieur ou égal à n tel que $p(t_i) = g(t_i)$, $i = 0, 1, \dots, n$, alors $g - p$ est un polynôme de degré inférieur ou égal à $n + 1$ qui vérifie $(g - p)(t_i) = 0$, donc $e(t) = (g - p)(t) = \alpha(t - t_0)(t - t_1)\dots(t - t_n)$, or le coefficient de t^{n+1} dans le polynôme $g - p$ est 1, donc $\alpha = 1$

[Retour à l'exercice ▲](#)

Solution de l'exercice 5.5

Les polynômes étant tous de degré distinct, il est facile de montrer qu'ils sont linéairement indépendants. Or \mathcal{P}_n est un espace vectoriel de dimension $n + 1$. Donc toute famille libre de \mathcal{P}_n de $n + 1$ éléments est une base de \mathcal{P}_n .

[Retour à l'exercice ▲](#)

Solution de l'exercice 5.6

On a

$$p_0(t) = f(t_0) = 1.$$

On a $p_1(t) = c_0 + c_1(t - t_0)$, on écrit que $p_1(t_0) = f(t_0)$, $p_1(t_1) = f(t_1)$, on obtient les coefficients $c_0 = 1$, $c_1 = 2$,

$$p_1(t) = 1 + 2(t - 1).$$

On a $p_2(t) = c'_0 + c'_1(t - t_0) + c'_2(t - t_0)(t - t_1)$, on écrit que $p_2(t_0) = f(t_0)$, $p_2(t_1) = f(t_1)$, $p_2(t_2) = f(t_2)$, on obtient les coefficients $c'_0 = 1$, $c'_1 = 2$, $c'_2 = -\frac{1}{2}$,

$$p_2(t) = 1 + 2(t - 1) - \frac{1}{2}(t - 1)(t - 2).$$

On remarque bien, comme indiqué dans le paragraphe de cours, que les polynômes sont 'emboîtés' et que pour chaque nouveau polynôme il suffit de calculer un seul coefficient.

[Retour à l'exercice ▲](#)

Solution de l'exercice 5.7

1. Par construction de p , il est facile de montrer que $p(t_i) = f(t_i)$ pour $i = 0, \dots, n$ et $p(\theta) = f(\theta)$. Le degré de ce polynôme est évidemment égal à $n + 1$.
2. Les propriétés du polynôme de Newton donne :

$$p(t) = p_n(t) + f[t_0, t_1, \dots, t_n, t]\Pi_n(t).$$

3.

$$e_n(\theta) = f(\theta) - p_n(\theta) = p(\theta) - p_n(\theta) = f[t_0, t_1, \dots, t_n, \theta]\Pi_n(\theta).$$

[Retour à l'exercice ▲](#)

Solution de l'exercice 5.8

Le polynôme s'écrit

$$p(t) = c_0 + c_1(t - t_0) + c_2(t - t_0)(t - t_1).$$

Or

$$p(t_0) = f(t_0) \Rightarrow c_0 = f[t_0]$$

$$p(t_1) = f(t_1) \Rightarrow c_1 = \frac{f(t_1) - f(t_0)}{t_1 - t_0} = \frac{f(t_0) - f(t_1)}{t_0 - t_1} = f[t_0, t_1]$$

$$p(t_2) = f(t_2) \Rightarrow c_2 = \frac{\frac{f(t_2) - f(t_0)}{t_2 - t_0} - f[t_0, t_1]}{t_2 - t_1} = \frac{f[t_2, t_0] - f[t_0, t_1]}{t_2 - t_1} = f[t_2, t_0, t_1] = f[t_0, t_1, t_2]$$

On a utilisé la symétrie des différences divisées.

[Retour à l'exercice ▲](#)

Solution de l'exercice 5.9

Pour calculer les coefficients de $p_3(t)$ dans la base de *Newton*, nous sommes conduits à construire le tableau proposé dans le cours pour $n = 3$, ce qui avec les données de l'exercice nous donne :

	k = 0	k = 1	k = 2	k = 3
$t_0 = 0$	$f[t_0] = \frac{1}{2}$			
$t_1 = 1$	$f[t_1] = 1$	$f[t_0, t_1] = \frac{1}{2}$		
$t_2 = 2$	$f[t_2] = 2$	$f[t_1, t_2] = 1$	$f[t_0, t_1, t_2] = \frac{1}{4}$	
$t_3 = 3$	$f[t_3] = -\frac{1}{2}$	$f[t_2, t_3] = -\frac{5}{2}$	$f[t_1, t_2, t_3] = -\frac{7}{4}$	$f[t_0, t_1, t_2, t_3] = -\frac{2}{3}$

On peut donc écrire $p_3(t)$ de la façon suivante :

$$p_3(t) = \frac{1}{2} + \frac{1}{2}t + \frac{1}{4}t(t-1) - \frac{2}{3}t(t-1)(t-2).$$

[Retour à l'exercice ▲](#)

Solution de l'exercice 5.10

Si vous avez compris les calculs effectués dans le tableau des différences divisées du paragraphe [Différences divisées / Calcul pratique](#), il suffit d'ajouter une ligne à ce tableau pour ajouter un point d'interpolation. Les coefficients $\{c_0, \dots, c_n\}$ sont les mêmes que ceux de p_n le dernier coefficient de la dernière ligne donnera c_{n+1} . Si les calculs ont été faits à la main, vous les avez évidemment gardés. Par contre, si vous avez utilisé l'algorithme donné dans le même paragraphe, une colonne se superpose à la précédente, et le tableau complet n'est donc pas gardé en mémoire. Il faut donc penser à stocker le tableau ...

En ce qui concerne la base des polynômes de Lagrange, chacun d'eux est construit à partir de tous les points d'interpolation, ce qui nécessite de recalculer tous ces polynômes lorsque l'on rajoute un point d'interpolation !

[Retour à l'exercice ▲](#)

Solution de l'exercice 5.11

1. L'algorithme "classique" est le suivant :

1: Les données sont : $c_0, \dots, c_n, t_0, \dots, t_{n-1}, t$

2: $q = t - t_0$

3: $p = c_0 + c_1 * q$

4: **pour** $k = 2$ jusqu'à n **faire**

5: $q = q * (t - t_{k-1})$

6: $p = p + c_k * q$

7: **fin pour**

Ceci correspond à $2n - 1$ multiplications, n additions et n soustractions.

2. Le schéma de Horner, dont l'algorithme est donné dans le cours compte n multiplications, n additions et n soustractions.

[Retour à l'exercice ▲](#)

Solution de l'exercice 5.13

Sur chacun des n intervalles $[t_i, t_{i+1}]$, $i = 0, \dots, n-1$, g est un polynôme de degré inférieur ou égal à 3, on a donc $g(t) = \alpha_i + \beta_i t + \gamma_i t^2 + \delta_i t^3$, il y a donc $4n$ inconnues.

On doit avoir

$$g(t_i^+) = g(t_i^-), g'(t_i^+) = g'(t_i^-), g''(t_i^+) = g''(t_i^-), i = 1, \dots, n-1$$

afin que la fonction g soit 2 fois continûment dérivable. On obtient donc $3(n-1)$ équations.

La fonction g doit interpolier, on doit donc avoir $g(t_i) = y_i$, $i = 0, \dots, n$.

On a donc au total $3n - 3 + n + 1 = 4n - 2$ équations.

Il manque donc deux équations, c'est pourquoi on impose, par exemple, les conditions supplémentaires :

$$g''(t_0) = g''(t_n) = 0$$

[Retour à l'exercice ▲](#)

Solution de l'exercice 6.1

1. Si $p \in \mathcal{P}_n$, alors $\forall t \ p(t) = \sum_{i=0}^n p(t_i) \mathcal{L}_i(t)$, donc

$$I(p) = \int_a^b p(t) dt = \sum_{i=0}^n p(t_i) \int_a^b \mathcal{L}_i(t) dt = \sum_{i=0}^n w_i p(t_i) = J(p).$$

2.

$$I(p) = J(p), \forall p \in \mathcal{P}_n \Rightarrow I(\mathcal{L}_j) = J(\mathcal{L}_j), j = 0, \dots, n,$$

or

$$I(\mathcal{L}_j) = \int_a^b \mathcal{L}_j(t) dt, J(\mathcal{L}_j) = \sum_{i=0}^n w_i \mathcal{L}_j(t_i) = w_j,$$

d'où le résultat.

3. $I(p) = J(p), \forall p \in \mathcal{P}_n \Rightarrow I(p_i) = J(p_i), i = 0, \dots, n$

Réciproquement si $p \in \mathcal{P}_n, p = \sum_{i=0}^n a_i p_i$, donc

$$I(p) = \sum_{i=0}^n a_i \int_a^b p_i(t) dt = \sum_{i=0}^n a_i I(p_i) = \sum_{i=0}^n a_i J(p_i).$$

D'autre part :

$$J(p) = \sum_{j=0}^n w_j p(t_j) = \sum_{j=0}^n w_j \sum_{i=0}^n a_i p_i(t_j) = \sum_{i=0}^n a_i \sum_{j=0}^n w_j a_i p_i(t_j) = \sum_{i=0}^n a_i J(p_i).$$

Ce qui termine de démontrer l'égalité.

[Retour à l'exercice ▲](#)

Solution de l'exercice 6.2

1. Si on note p_i les polynômes de la base canonique $p_i(t) = t^i$, on écrit

$$\begin{cases} I(p_0) = J(p_0) & \Leftrightarrow \int_{-1}^1 dt = w_0 + w_1 + w_2 & \Leftrightarrow w_0 + w_1 + w_2 = 2 \\ I(p_1) = J(p_1) & \Leftrightarrow \int_{-1}^1 t dt = -w_0 + w_2 & \Leftrightarrow w_0 - w_2 = 0 \\ I(p_2) = J(p_2) & \Leftrightarrow \int_{-1}^1 t^2 dt = w_0 + w_2 & \Leftrightarrow w_0 + w_2 = \frac{2}{3} \end{cases} .$$

Pour que ces 3 relations soient vérifiées, il faut que $w_0 = w_2 = \frac{1}{3}$, $w_1 = \frac{4}{3}$

Vérifions avec ces coefficients si $I(p_3) = J(p_3)$, après calcul, on obtient $I(p_3) = J(p_3) = 0$, on continue !

On calcule $I(p_4) = \frac{2}{5}$, $J(p_4) = \frac{2}{3}$.

La formule construite (qui s'appelle la formule de Simpson) est exacte pour tous les polynômes de degré inférieur ou égal à 3.

On aurait pu calculer les coefficients w_i en écrivant $w_i = \int_{-1}^1 \mathcal{L}_i(t) dt$ où \mathcal{L}_i est le polynôme de la base de Lagrange.

2.

$$\int_0^2 f(t) dt = \int_{-1}^1 f(u+1) du, \text{ où } u = t - 1$$

Si on pose $g(u) = f(u+1)$, $\int_{-1}^1 g(u) du$ est approchée par

$$\frac{1}{3}g(-1) + \frac{4}{3}g(0) + \frac{1}{3}g(1) = \frac{1}{3}f(0) + \frac{4}{3}f(1) + \frac{1}{3}f(2)$$

d'où la formule.

[Retour à l'exercice ▲](#)

Solution de l'exercice 6.3

1. On vérifie que $I(p_0) = J(p_0) = b - a$, $I(p_1) = J(p_1) = \frac{b^2 - a^2}{2}$
2. L'ordre est donc au moins égal à 1, on calcule

$$I(p_2) = \frac{b^3 - a^3}{3}, J(p_2) = (b - a) \left(\frac{a + b}{2} \right)^2$$

$I(p_2) \neq J(p_2)$, donc la formule est d'ordre 1.

3. La formule de Taylor est :

$$f(t) = f(t_M) + (t - t_M)f'(t_M) + \frac{(t - t_M)^2}{2}f''(c_t)$$

$$J(f) = (b - a)f(t_M) = \int_a^b f(t_M) dt$$

donc

$$E(f) = I(f) - J(f) = \int_a^b (f(t) - f(t_M)) dt = \int_a^b (t - t_M)f'(t_M) + \frac{(t - t_M)^2}{2}f''(c_t) dt$$

or $\int_a^b (t - t_M) dt = 0$, donc

$$E(f) = \int_a^b \frac{(t - t_M)^2}{2} f''(c_t) dt = f''(\eta) \int_a^b \frac{(t - t_M)^2}{2} dt = f''(\eta) \frac{(b - a)^3}{24}$$

[Retour à l'exercice ▲](#)

Solution de l'exercice 6.4

$I(p_0) = 2$, $J(p_0) = w$, donc il faut que $w = 2$.

$I(p_1) = 0$, $J(p_1) = 2t_0$, donc

– si $t_0 \neq 0$, $I(p_1) \neq J(p_1)$, l'ordre est 0.

– si $t_0 = 0$, $I(p_1) = J(p_1)$, on a de plus $I(p_2) = \frac{2}{3}$, $J(p_2) = 0$, donc l'ordre est 1.

[Retour à l'exercice ▲](#)

Solution de l'exercice 6.5

On approche $I(f) = \int_{-1}^1 f(t) dt$ par $J(f) = f(-1) + f(1)$.

On calcule :

$$I(p_0) = 2, J(p_0) = 2,$$

$$I(p_1) = 0, J(p_1) = 0,$$

$$I(p_2) = \frac{2}{3}, J(p_2) = 2,$$

Donc l'ordre est 1.

[Retour à l'exercice ▲](#)

Solution de l'exercice 6.6

On calcule

$$I(p_0) = 2, J(p_0) = w_0 + w_1 + w_2 + w_3,$$

$$I(p_1) = 0, J(p_1) = -w_0 - \frac{1}{3}w_1 + \frac{1}{3}w_2 + w_3,$$

$$I(p_2) = \frac{2}{3}, J(p_2) = w_0 + \frac{1}{9}w_1 + \frac{1}{9}w_2 + w_3,$$

$$I(p_3) = 0, J(p_3) = -w_0 - \frac{1}{27}w_1 + \frac{1}{27}w_2 + w_3,$$

On obtient un système de 4 équations à 4 inconnues.

En remplaçant la 4ème équation par la 4ème moins la 2ème, on montre que $w_1 = w_2$.

En remplaçant la 3ème équation par la 3ème moins la 1ère, on montre que $w_1 = w_2 = \frac{3}{4}$.

On résout alors les deux premières équations, on obtient :

$$w_0 = w_3 = \frac{1}{4}, w_1 = w_2 = \frac{3}{4}.$$

On calcule $I(p_4)$ et $J(p_4)$, ces deux valeurs sont distinctes, donc l'ordre vaut 3.

[Retour à l'exercice ▲](#)

Solution de l'exercice 6.7

En appliquant la méthode de Simpson sur chacun des intervalles $[t_{2i}, t_{2i+2}]$, on obtient

$$J(f) = \frac{h}{3} (f(t_0) + 4f(t_1) + f(t_2) + f(t_2) + 4f(t_3) + f(t_4) + \dots + f(t_{2M-2}) + 4f(t_{2M-1}) + f(t_{2M})),$$

ce qui donne le résultat en regroupant. De même en sommant les erreurs, on obtient :

$$E(f) = -\frac{h^5}{90} (f^{(4)}(\eta_1) + f^{(4)}(\eta_2) + \dots + f^{(4)}(\eta_M)).$$

On applique le théorème de la valeur intermédiaire pour affirmer qu'il existe η tel que

$$f^{(4)}(\eta) = \frac{(f^{(4)}(\eta_1) + f^{(4)}(\eta_2) + \dots + f^{(4)}(\eta_M))}{M}.$$

On a donc

$$E(f) = -\frac{h^5}{90} M f^{(4)}(\eta) = -\frac{h^4}{180} (b-a) f^{(4)}(\eta).$$

[Retour à l'exercice ▲](#)

Solution de l'exercice 7.1

Revoyez la résolution des équations différentielles linéaires du second ordre à coefficients constants, par exemple dans le polycopié de MT21 chapitre9.

les constantes g et L sont positives, la solution de

$$\begin{cases} \ddot{\theta}(t) = -\frac{g}{L}\theta(t), \\ \theta(0) = \theta_0, \dot{\theta}(0) = 0. \end{cases}$$

est donc

$$\begin{cases} \theta(t) = a \cos \sqrt{\frac{g}{L}} t + b \sin \sqrt{\frac{g}{L}} t \\ a = \theta_0, b = 0. \end{cases}$$

[Retour à l'exercice ▲](#)

Solution de l'exercice 7.2

Pour $t \in]0, a[$, y_a est dérivable et sa dérivée est nulle.

Pour $t \in]a, +\infty[$, y_a est dérivable et $y'_a(t) = \frac{t-a}{2}$.

Donc en $t = a$ y_a admet une dérivée à droite et une dérivée à gauche, ces deux dérivées sont égales donc y_a est dérivable en a et $y'_a(a) = 0$.

On vérifie que

pour $t \in]0, a[$, $\sqrt{y_a(t)} = 0 = y'_a(t)$,

pour $t \in]a, +\infty[$, $\sqrt{y_a(t)} = \frac{|t-a|}{2} = \frac{t-a}{2} = y'_a(t)$.

On a de plus $y_a(0) = 0$, y_a est donc solution du problème de Cauchy.

Le problème de Cauchy admet donc une infinité de solutions

[Retour à l'exercice ▲](#)

Solution de l'exercice 7.3

On pose

$$Y_1(t) = \theta(t), Y_2(t) = \dot{\theta}(t), Y(t) = \begin{pmatrix} \theta(t) \\ \dot{\theta}(t) \end{pmatrix},$$

on a alors

$$\begin{cases} \ddot{\theta}(t) = -\frac{g}{L} \sin \theta(t), \\ \theta(0) = \theta_0, \dot{\theta}(0) = 0, \end{cases} \Leftrightarrow \begin{cases} Y'(t) = \begin{pmatrix} Y_2(t) \\ -\frac{g}{L} \sin Y_1(t) \end{pmatrix} = F(Y(t)), \\ Y(0) = \begin{pmatrix} \theta_0 \\ 0 \end{pmatrix}, \end{cases}$$

[Retour à l'exercice ▲](#)

Solution de l'exercice 7.4

Revoir l'exercice 7.3 qui permet d'obtenir un système d'équations différentielles du premier ordre équivalent.

On choisit un pas h , on a $t_1 = h$, le vecteur $Y^{(1)}$ qui est une approximation de $\begin{pmatrix} \theta(t_1) \\ \dot{\theta}(t_1) \end{pmatrix}$ est obtenu en résolvant $Y^{(1)} = \begin{pmatrix} \theta_0 \\ 0 \end{pmatrix} + hF(Y^{(1)})$, si l'on note u et v les deux composantes du vecteur $Y^{(1)}$, on doit donc résoudre le système de deux équations :

$$Y^{(1)} = \begin{pmatrix} \theta_0 \\ 0 \end{pmatrix} + hF(Y^{(1)}) \Leftrightarrow \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} \theta_0 \\ 0 \end{pmatrix} + h \begin{pmatrix} v \\ -\frac{g}{L} \sin u \end{pmatrix} \Leftrightarrow \begin{cases} u - hv - \theta_0 = 0 \\ v + h\frac{g}{L} \sin u = 0 \end{cases}$$

C'est un système de deux équations non linéaires que l'on peut résoudre par exemple par la méthode de Newton vue dans le chapitre 4. Cette méthode nécessite une valeur initiale pour le vecteur $\begin{pmatrix} u \\ v \end{pmatrix}$, on peut choisir par exemple $Y^{(0)}$.

Pour calculer $Y^{(2)}$ (puis pour les autres itérés), on devra résoudre à nouveau un système de 2 équations

$$Y^{(2)} = Y^{(1)} + hF(Y^{(2)}).$$

[Retour à l'exercice ▲](#)

Solution de l'exercice 7.5

Revoir l'exercice [7.3](#) qui permet d'obtenir un système d'équations différentielles du premier ordre équivalent.

On choisit un pas h , on a $t_1 = h$, on pose $Y^{(0)} = \begin{pmatrix} \theta_0 \\ 0 \end{pmatrix} = Y(0)$. Le vecteur $Y^{(1)}$ qui est une approximation de $Y(t_1) = \begin{pmatrix} \theta(t_1) \\ \dot{\theta}(t_1) \end{pmatrix}$ est obtenu explicitement en écrivant

$$Y^{(1)} = Y^{(0)} + \frac{h}{2} (F(Y^{(0)}) + F(Y^{(0)} + hF(Y^{(0)})))$$

On calcule

$$Y^{(0)} = \begin{pmatrix} \theta_0 \\ 0 \end{pmatrix}, F(Y^{(0)}) = \begin{pmatrix} 0 \\ -\frac{g}{L} \sin \theta_0 \end{pmatrix}, Y^{(0)} + hF(Y^{(0)}) = \begin{pmatrix} \theta_0 \\ -h\frac{g}{L} \sin \theta_0 \end{pmatrix},$$

$$F(Y^{(0)} + hF(Y^{(0)})) = \begin{pmatrix} -h\frac{g}{L} \sin \theta_0 \\ -\frac{g}{L} \sin \theta_0 \end{pmatrix}$$

ce qui permet d'obtenir $Y^{(1)}$.

[Retour à l'exercice ▲](#)

Solution de l'exercice 7.6

Pour le schéma d'Euler simple on a

$$z_0 = y_0, z_i = z_{i-1} - \lambda h z_{i-1} = (1 - \lambda h) z_{i-1} = \dots = (1 - \lambda h)^i z_0$$

On a $ih = t$, donc $i = \frac{t}{h}$, quand h tend vers 0 $(1 - \lambda h)^i$ est indéterminée de la forme 1^∞ , levons l'indétermination.

$$(1 - \lambda h)^i = (1 - \lambda h)^{\frac{t}{h}} = \exp\left(\frac{t}{h} \ln(1 - \lambda h)\right) = \exp\left(\frac{t}{h}(-\lambda h + h\epsilon(h))\right).$$

On a donc

$$\lim_{h \rightarrow 0} (1 - \lambda h)^i = \lim_{h \rightarrow 0} \exp\left(\frac{t}{h}(-\lambda h + h\epsilon(h))\right) = \exp(-\lambda t).$$

Et donc

$$\lim_{h \rightarrow 0} z_i = y_0 \exp(-\lambda t).$$

Pour le schéma d'Euler implicite on a

$$z_0 = y_0, z_i = z_{i-1} - \lambda h z_i \Leftrightarrow z_i = \frac{1}{1 + \lambda h} z_{i-1} = \dots = \left(\frac{1}{1 + \lambda h}\right)^i z_0$$

Là encore, quand h tend vers 0 $\left(\frac{1}{1 + \lambda h}\right)^i$ est indéterminée de la forme 1^∞ , levons l'indétermination.

$$\left(\frac{1}{1 + \lambda h}\right)^i = \left(\frac{1}{1 + \lambda h}\right)^{\frac{t}{h}} = \exp\left(\frac{t}{h} \ln\left(\frac{1}{1 + \lambda h}\right)\right) = \exp\left(-\frac{t}{h} \ln(1 + \lambda h)\right) = \exp\left(-\frac{t}{h}(\lambda h + h\epsilon(h))\right).$$

On a donc

$$\lim_{h \rightarrow 0} \left(\frac{1}{1 + \lambda h}\right)^i = \lim_{h \rightarrow 0} \exp\left(-\frac{t}{h}(\lambda h + h\epsilon(h))\right) = \exp(-\lambda t).$$

Et donc

$$\lim_{h \rightarrow 0} z_i = y_0 \exp(-\lambda t).$$

[Retour à l'exercice ▲](#)

Solution de l'exercice 7.7

Pour obtenir l'ordre du schéma d'Euler-Cauchy, il faut calculer le développement limité de

$$\tau_{n+1}(h) = y(t_{n+1}) - y(t_n) - \frac{h}{2} (f(t_n, y(t_n)) + f(t_n + h, y(t_n) + hf(t_n, y(t_n)))) ,$$

où y est solution de

$$y'(t) = f(t, y(t)).$$

On va supposer que les fonctions f , y sont suffisamment dérivables. En utilisant les résultats sur les dérivées des fonctions composées, on obtient

$$y''(t) = \frac{\partial f}{\partial t}(t, y(t)) + y'(t) \frac{\partial f}{\partial y}(t, y(t))$$

On peut écrire $\tau_{n+1}(h)$ comme la somme de trois termes :

$$A = y(t_{n+1}) - y(t_n), \quad B = -\frac{h}{2} f(t_n, y(t_n)), \quad C = -\frac{h}{2} f(t_n + h, y(t_n) + hf(t_n, y(t_n))).$$

$$A = y(t_{n+1}) - y(t_n) = hy'(t_n) + \frac{h^2}{2} y''(t_n) + \frac{h^3}{3} y'''(c), \quad c \in]t_n, t_{n+1}[.$$

$$B = -\frac{h}{2} f(t_n, y(t_n)) = -\frac{h}{2} y'(t_n).$$

$$C = -\frac{h}{2} f(t_n + h, y(t_n) + hf(t_n, y(t_n))) = -\frac{h}{2} f(t_n + h, y(t_n) + hy'(t_n)) = -\frac{h}{2} g(h),$$

où l'on a noté

$$g(h) = f(t_n + h, y(t_n) + hy'(t_n))$$

en utilisant les dérivées des fonctions composées, on obtient :

$$g'(h) = \frac{\partial f}{\partial t}(t_n + h, y(t_n) + hy'(t_n)) + y'(t_n) \frac{\partial f}{\partial y}(t_n + h, y(t_n) + hy'(t_n))$$

on a donc

$$g(0) = f(t_n, y(t_n)) = y'(t_n), \quad g'(0) = \frac{\partial f}{\partial t}(t_n, y(t_n)) + y'(t_n) \frac{\partial f}{\partial y}(t_n, y(t_n)) = y''(t_n)$$

En utilisant le développement de Taylor, on obtient

$$g(h) = g(0) + hg'(0) + \frac{h^2}{2} g''(d) = y'(t_n) + hy''(t_n) + \frac{h^2}{2} g''(d), \quad d \in]0, h[,$$

d'où

$$C = -\frac{h}{2} y'(t_n) - \frac{h^2}{2} y''(t_n) - \frac{h^3}{4} g''(d).$$

En regroupant

$$\tau_{n+1}(h) = A + B + C = h^3 \left(\frac{y'''(c)}{3} - \frac{g''(d)}{4} \right),$$

si l'on suppose que les fonctions y''' et g'' sont majorées respectivement par M_1 et M_2 , on obtient

$$|\tau_{n+1}(h)| \leq h^3 \left(\frac{M_1}{3} + \frac{M_2}{4} \right), \quad \text{donc} \quad \left| \frac{\tau_{n+1}(h)}{h} \right| \leq Mh^2, \quad 0 \leq n \leq N-1$$

le schéma d'Euler-Cauchy est donc d'ordre 2.

[Retour à l'exercice ▲](#)

Solution de l'exercice 8.1

Si λ_1 était complexe non réelle, alors $\bar{\lambda}_1$ serait une autre valeur propre λ_2 et on aurait $|\lambda_1| = |\lambda_2|$, ce qui est faux, donc λ_1 est réelle. Il est évident que λ_1 est simple.

[Retour à l'exercice ▲](#)