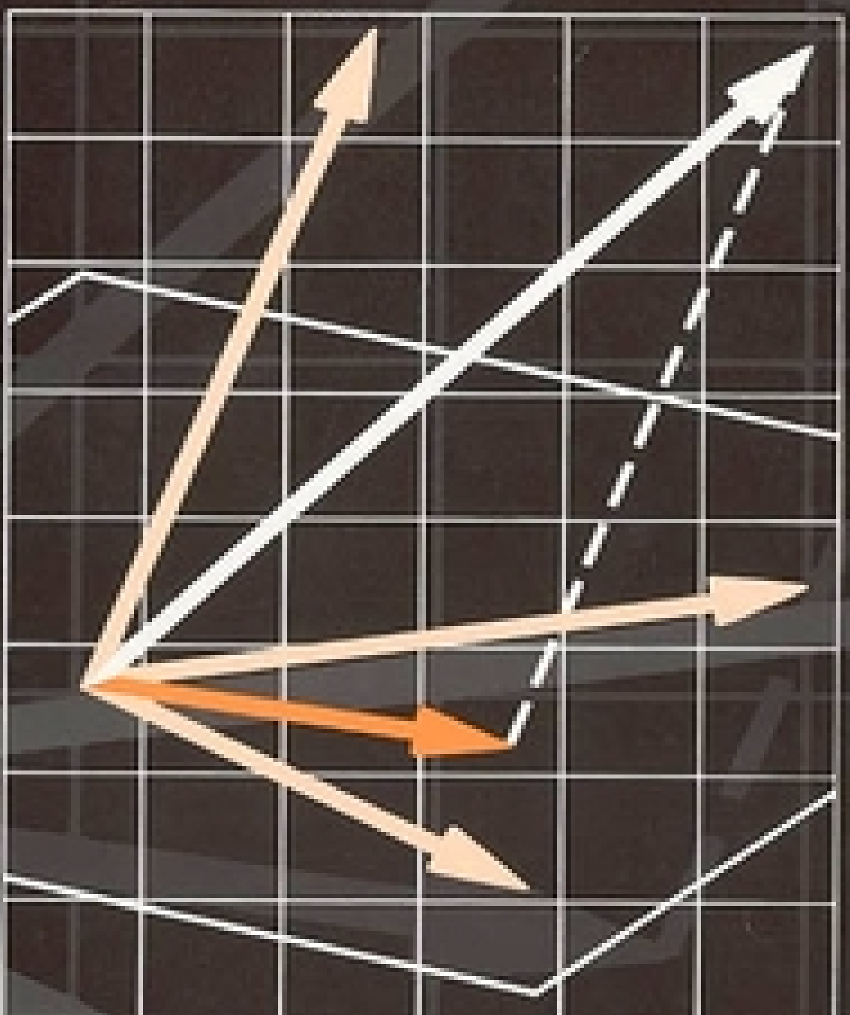
Enseignement des mathématiques

Algèbre linéaire

R. Cairoli



Presses polytechniques et universitaires romandes

Enseignement des mathématiques

Algèbre linéaire

R. Cairoli

Presses polytechniques et universitaires romandes

L'auteur et l'éditeur remercient l'Ecole polytechnique fédérale de Lausanne dont le soutien financier a rendu possible la publication de cet ouvrage.

DANS LA MÊME COLLECTION DIRIGÉE PAR LE PROFESSEUR ROBERT C. DALANG

Analyse
Recueil d'exercices et aide-mémoire vol. 1 et 2
Jacques Douchet

Recherche opérationnelle pour ingénieurs I
Dominique de Werra, Thomas M. Liebling, Jean-François Hêche

Recherche opérationnelle pour ingénieurs II
Jean-François Hêche, Thomas M. Liebling, Dominique de Werra

Introduction à l'analyse numérique
Jacques Rappaz et Marco Picasso

Algèbre linéaire
Aide-mémoire, exercices et applications
Robert C. Dalang et Amel Chaabouni

Analyse avancée pour ingénieurs
Bernard Dacorogna et Chiara Tanteri

Initiation aux probabilités
Sheldon M. Ross

Cours d'Analyse
Srishti D. Chatterji
1 Analyse vectorielle
2 Analyse complexe
3 Equations différentielles

DANS LA COLLECTION «MÉTHODES MATHÉMATIQUES POUR L'INGÉNIEUR»

Introduction à la statistique
Stephan Morgenthaler

Aide-mémoire d'analyse
Heinrich Matzinger

Les Presses polytechniques et universitaires romandes sont une fondation scientifique dont le but est principalement la diffusion des travaux de l'Ecole polytechnique fédérale de Lausanne ainsi que d'autres universités et écolesd'ingénieurs francophones.
Le catalogue de leurs publications peut être obtenu par courrier aux Presses polytechniques et universitaires romandes,
EPFL – Centre Midi, CH-1015 Lausanne, par E-Mail à ppur@epfl.ch,
par téléphone au (0)21 693 41 40, ou par fax au (0)21 693 40 27.

www.ppur.org

Première édition, en deux tomes, 1987
ISBN 2-88074-187-4

Imprimé en Italie

Préface

L'enseignement de l'algèbre linéaire s'est considérablement développé au cours des deux dernières décennies. De nos jours, presque toutes les sections d'études scientifiques et techniques, et notamment les sections d'ingénieurs, incluent l'algèbre linéaire dans la formation de base de leurs étudiants. Un ouvrage qui s'ajoute aux nombreux déjà existants dans la littérature peut donc encore se justifier et prétendre répondre à des exigences restées insatisfaites.

A l'exception de quelques-uns, les sujets développés dans ce livre sont classiques et servent normalement de base à la réalisation de tout livre d'algèbre linéaire de même niveau. Ce qui distingue celui-ci des autres réside dans l'arrangement de la matière et surtout dans sa présentation, qui emprunte beaucoup à la géométrie ordinaire et vise à développer chez le lecteur une compréhension intuitive. Un autre élément distinctif est certainement la place accordée à la notion d'espace affine et à l'étude de la géométrie affine à plusieurs dimensions.

Dans l'élaboration de la matière, l'auteur s'est prévalu de l'expérience de plusieurs années d'enseignement à des sections d'ingénieurs de l'Ecole polytechnique fédérale de Lausanne. C'est à travers cette expérience que l'exposition purement formelle conçue initialement a cédé la place à un discours plus parlé et, oserait-on dire, plus humain.

Conscient du fait que les concepts abstraits ne deviennent clairs qu'à travers les exemples, l'auteur s'est constamment soucié de favoriser la compréhension par des motivations, des commentaires et des illustrations.

Outre créer les conditions propices à une meilleure assimilation des théorèmes et des techniques de calcul de l'algèbre linéaire, cet ouvrage veut inciter le lecteur à un travail de recherche personnel. Quelques-uns des nombreux exercices placés à la fin de chaque chapitre ont pour but d'encourager une telle activité.

Ce livre a été écrit à l'intention des étudiants du premier cycle d'études des écoles d'ingénieurs de niveau universitaire, mais il s'adresse également aux étudiants en mathématique et en physique orientés vers les applications. Il peut en outre venir en aide aux scientifiques à la recherche de méthodes algébriques leur permettant d'apporter des éléments de réponse aux problèmes qu'ils rencontrent, ainsi qu'aux maîtres du degré secondaire désireux de savoir vers quels programmes conduit leur enseignement.

Remerciements

Je remercie vivement Monsieur Jean-Claude Evard de l'intérêt constant qu'il a apporté à la réalisation de cet ouvrage. Ses critiques et ses suggestions m'ont permis d'améliorer la présentation de nombreux points délicats du texte.

Je remercie Monsieur Laurent Perruchoud d'avoir lu plusieurs parties du manuscrit et de m'avoir signalé quelques imprécisions.

Je remercie également Monsieur Claude El-Hayek d'avoir contrôlé la version définitive du manuscrit, Monsieur Klaus-Dieter Semmler d'avoir réalisé les figures représentant les quadriques, Monsieur Jean-François Casteu d'avoir effectué les dessins et Madame Pascale Deppierraz pour la compétence avec laquelle elle s'est occupée des problèmes d'édition.

Conventions

1. Découpage du texte

Ce livre est composé de dix chapitres numérotés de 1 à 10 et d'un appendice repéré par la lettre A. Chaque chapitre est divisé en sections et chaque section en paragraphes. Les sections sont repérées par une double numérotation et les paragraphes par une triple numérotation. Par exemple, 7.2 renvoie à la deuxième section du septième chapitre et 7.2.4 au quatrième paragraphe de cette section. La dernière section de chaque chapitre rassemble les exercices sur la matière traitée dans le chapitre. Ces exercices sont numérotés de la même façon que les paragraphes. Ainsi, 7.4.11 désigne le onzième exercice de la quatrième section du septième chapitre. Les figures sont repérées par l'abréviation Fig. suivie de deux nombres, le premier indiquant le chapitre et le deuxième la figure. Par exemple, Fig. 7.3 désigne la troisième figure du septième chapitre.

2. Conventions sur les nombres

Les lettres $\mathbb{R}$ et $\mathbb{C}$ désigneront respectivement le corps des nombres réels et le corps des nombres complexes. Tout au long des dix chapitres de ce livre, le terme «nombre» aura le sens de «nombre réel». Dans l'appendice consacré à l'extension de certains résultats aux nombres complexes, il sera précisé dans quelles circonstances le terme «nombre» prendra la signification de «nombre complexe».

Les nombres seront généralement désignés par des lettres grecques minuscules telles que α, β, γ, λ, μ, ν, ou par des lettres latines minuscules telles que a, b, c, d. Les nombres entiers seront appelés plus simplement entiers. Ils seront désignés par des lettres latines minuscules telles que n, k, i, j, l, m.

Nous dirons qu'un nombre est *positif* (*négatif*) s'il est supérieur (inférieur) à zéro. Nous dirons qu'il est *non négatif* s'il est supérieur ou égal à zéro.

3. Avertissement concernant l'emploi des adjectifs numéraux

Tout au long de ce livre, par deux, trois, ..., n objets, nous entendrons (sauf mention explicite du contraire) deux, trois, ..., n objets distincts.

4. Familles d'éléments d'un ensemble

Dans ce livre, nous appellerons *famille finie* d'éléments d'un ensemble E tout n-uplet (ordonné) d'éléments de E, c'est-à-dire tout élément du produit cartésien

$(x_1, x_2, ..., x_n)$ et dirons que x_i est le *i-ième terme* ou le *terme d'indice i*. Nous dirons en outre que l'ensemble $\{1, 2, ..., n\}$ est l'*ensemble d'indices*.

On remarquera que notre définition n'exclut pas que $x_i = x_j$ pour des indices i et j différents, ni même que $x_i = x$ pour tout indice i, x désignant un élément de E. On remarquera encore que $(x_1, x_2, ..., x_n) = (x'_1, x'_2, ..., x'_n)$ si et seulement si $x_1 = x'_1, x_2 = x'_2, ..., x_n = x'_n$.

Nous appellerons les familles à deux et à trois termes respectivement *couples* et *triplets*.

Les *familles infinies* (ou *suites*) sont définies similairement en prenant comme ensemble d'indices l'ensemble des entiers positifs. Elles seront désignées par le symbole $(x_1, x_2, ...)$. Précisons toutefois qu'elles apparaîtront très rarement dans ce livre.

Ajoutons quelques lignes de commentaire à la notion de famille finie. Dans l'étude de nombreuses questions, telles l'indépendance linéaire ou la génération de sous-espaces ou l'orthogonalité, l'incorporation d'un ordre à la définition de famille finie est inutile. A ce type de questions, les familles «non ordonnées», c'est-à-dire les applications d'un ensemble fini I dans E (notées habituellement par $(x_i)_{i \in I}$) conviennent parfaitement. Par contre, dans d'autres contextes, notamment dans certaines relations avec les matrices (en raison de leur représentation sous la forme de tableaux) ou dans des questions concernant l'orientation, l'ordre de disposition des termes de la famille joue un rôle important. Dans ce livre, par souci de simplicité, nous n'avons pas jugé nécessaire d'utiliser deux notions de famille finie. Le lecteur intéressé pourra facilement déceler de lui-même les parties qui s'énoncent plus naturellement en termes de familles finies «non ordonnées».

Table des matières

CHAPITRE 1

Espaces vectoriels et espaces affines

1.1 UN MODÈLE D'ESPACE VECTORIEL

1.1.1 Introduction

Des notions physiques telles que la force ou la vitesse sont caractérisées par une direction, un sens et une intensité. Ce triple caractère est mis en évidence par les flèches. Celles-ci sont à l'origine de la notion de vecteur et en constituent l'exemple le plus suggestif. Bien que leur nature soit essentiellement géométrique, c'est leur aptitude à se lier les unes aux autres, donc leur comportement algébrique, qui retiendra principalement notre attention. Partagé en classes d'équivalence et muni de deux opérations appelées addition et multiplication par un scalaire, l'ensemble qu'elles forment représente le modèle classique d'un espace vectoriel. Un de nos premiers objectifs est la description détaillée de ce modèle.

1.1.2 Notion de flèche

Nous désignerons par $\mathscr{G}$ l'espace ordinaire de la géométrie élémentaire et par P, Q, ... ses points. Nous appellerons *flèche* tout segment de droite orienté. La flèche d'origine P et d'extrémité Q sera notée PQ (fig. 1.1). Il est évident que toute flèche est caractérisée par sa direction, son sens, son intensité ou grandeur et son origine.

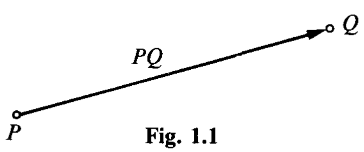


Fig. 1.1

1.1.3 Ensemble des vecteurs

Nous dirons que deux flèches sont équivalentes si elles ont la même direction, le même sens et la même intensité. Partageons l'ensemble des flèches en classes d'équivalence: deux flèches appartiennent à une même classe si et seulement si elles sont équivalentes. Nous dirons que chacune de ces classes est un *vecteur*. Rangeons, en outre, les flèches dégénérées (c'est-à-dire de la forme PP) en une classe distinguée que nous appellerons *vecteur nul* et noterons $\mathbf{0}$. L'ensemble des vecteurs ainsi définis sera désigné par V. Il faut souligner que les éléments de V sont des classes de flèches et non pas des flèches individuelles. Il est cependant clair qu'une flèche quelconque suffit à déterminer la classe à laquelle elle appartient et il est donc naturel de l'appeler *représentant* de la classe ou du vecteur.

Dans ce livre, les vecteurs seront désignés par des lettres latines minuscules imprimées en caractère gras ou par des couples de lettres latines majuscules surmontées d'une flèche (par exemple, $\overrightarrow{PQ}$ désigne le vecteur déterminé par la flèche PQ). Dans les figures, les flèches seront toutefois désignées par le symbole du vecteur qu'elles représentent.

1.1.4 Addition de vecteurs

Traçons le représentant d'un vecteur $\mathbf{y}$ à partir de l'extrémité d'un représentant d'un vecteur $\mathbf{x}$. La flèche dont l'origine est celle du représentant de $\mathbf{x}$ et l'extrémité celle du représentant de $\mathbf{y}$ détermine un vecteur que nous noterons $\mathbf{x} + \mathbf{y}$ et appelerons *somme* de $\mathbf{x}$ et $\mathbf{y}$. L'opération qui associe à tout couple de vecteurs leur somme s'appelle *addition vectorielle* (fig. 1.2).

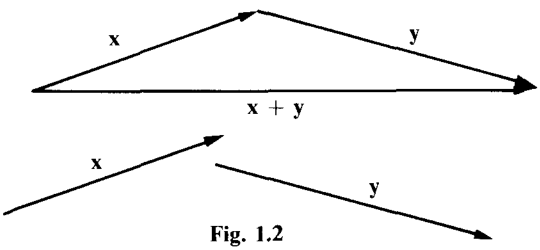


Fig. 1.2

A l'aide d'une figure, il est facile de montrer que l'opération d'addition vectorielle est *associative* et *commutative*, autrement dit, que

$$(\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z})$$

et

$$\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}.$$

Il est en outre évident que le vecteur nul $\mathbf{0}$ est l'élément neutre de l'addition vectorielle, autrement dit, que

$$\mathbf{x} + \mathbf{0} = \mathbf{0} + \mathbf{x} = \mathbf{x},$$

et que

$$\mathbf{x} + (-\mathbf{x}) = \mathbf{0},$$

où $-\mathbf{x}$ désigne le *vecteur opposé* de $\mathbf{x}$, c'est-à-dire le vecteur dont les représentants ont la même direction et la même intensité que ceux de $\mathbf{x}$, mais le sens opposé (fig. 1.3).

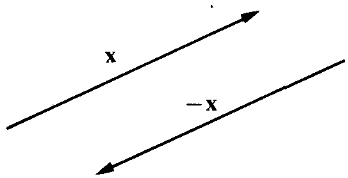


Fig. 1.3

1.1.5 Soustraction de vecteurs

L'opération inverse de l'addition vectorielle est la *soustraction vectorielle*. Soustraire un vecteur revient à additionner le vecteur opposé (fig. 1.4):

$$\mathbf{x} - \mathbf{y} = \mathbf{x} + (-\mathbf{y}).$$

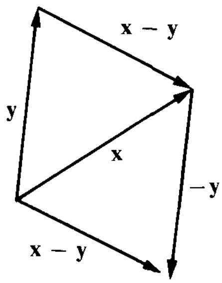


Fig. 1.4

1.1.6 Remarque

L'addition s'étend, par récurrence, au cas d'une famille finie quelconque de vecteurs $(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k)$:

$$((\mathbf{x}_1 + \mathbf{x}_2) + \mathbf{x}_3) + \ldots .$$

En vertu de l'associativité, ces additions successives peuvent être effectuées dans n'importe quel ordre, ce qui justifie l'écriture sans parenthèses

$$\mathbf{x}_1 + \mathbf{x}_2 + \ldots + \mathbf{x}_k.$$

1.1.7 Multiplication par un scalaire

Dans ce livre «scalaire» sera synonyme de «nombre». Rappelons en outre que, sauf indication contraire, «nombre» signifie «nombre réel». Le vecteur $\alpha\mathbf{x}$, appelé *produit du nombre* α *par* $\mathbf{x}$, est défini de la manière suivante: prenons une flèche représentative de $\mathbf{x}$ et construisons une flèche de même direction, de même sens ou de sens opposé, suivant que α est positif ou négatif, et d'intensité $|\alpha|$ fois l'intensité de la flèche initiale; la flèche ainsi obtenue est un représentant du vecteur $\alpha\mathbf{x}$; si $\alpha = 0$ ou $\mathbf{x} = \mathbf{0}$, nous posons $\alpha\mathbf{x} = \mathbf{0}$. L'opération qui consiste à effectuer le produit d'un nombre par un vecteur est appelée *multiplication par un scalaire* (fig. 1.5). Les deux cas particuliers suivants méritent d'être relevés:

$$1\mathbf{x} = \mathbf{x}, \quad (-1)\mathbf{x} = -\mathbf{x}.$$

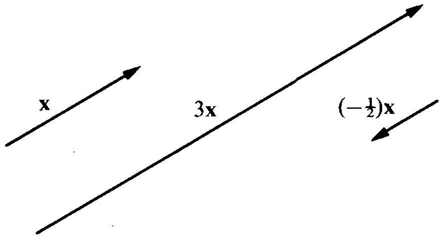


Fig. 1.5

On vérifie aisément que la multiplication par un scalaire est *associative* et *distributive* par rapport à l'addition numérique et à l'addition vectorielle, autrement dit, que

$$\alpha(\beta\mathbf{x}) = (\alpha\beta)\mathbf{x},$$
$$(\alpha + \beta)\mathbf{x} = \alpha\mathbf{x} + \beta\mathbf{x},$$
$$\alpha(\mathbf{x} + \mathbf{y}) = \alpha\mathbf{x} + \alpha\mathbf{y}.$$

1.1.8 Espace vectoriel géométrique

L'addition vectorielle et la multiplication par un scalaire définissent dans V une structure d'espace vectoriel. Nous appellerons V *espace vectoriel géométrique.*

1.2 DÉFINITION DE LA NOTION D'ESPACE VECTORIEL

1.2.1 Introduction

Nous allons maintenant aborder l'étude des espaces vectoriels sous une forme abstraite. En gros, un espace vectoriel sera un ensemble d'éléments pouvant être additionnés et multipliés par les nombres conformément aux règles de calcul mises en évidence dans la section 1.1.

1.2.2 Espaces vectoriels

On appelle *espace vectoriel* un ensemble E d'éléments désignés par $\mathbf{x}$, $\mathbf{y}$, ... et appelés *vecteurs*, muni d'une structure algébrique définie par la donnée de deux opérations:

- l'*addition vectorielle*: à tout couple $(\mathbf{x}, \mathbf{y})$ de vecteurs correspond un vecteur désigné par $\mathbf{x} + \mathbf{y}$ et appelé *somme* de $\mathbf{x}$ et $\mathbf{y}$;
- la *multiplication par un scalaire*: à tout couple $(\alpha, \mathbf{x})$ formé d'un nombre α et d'un vecteur $\mathbf{x}$ correspond un vecteur désigné par $\alpha\mathbf{x}$ et appelé *produit* de α par $\mathbf{x}$.

Ces deux opérations satisfont aux conditions suivantes:

(a) $(\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z})$ (loi associative).
(b) $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$ (loi commutative).
(c) Il existe un vecteur, noté $\mathbf{0}$ et appelé *vecteur nul*, tel que $\mathbf{x} + \mathbf{0} = \mathbf{x}$ pour tout vecteur $\mathbf{x}$.
(d) Pour tout vecteur $\mathbf{x}$, il existe un vecteur, noté $-\mathbf{x}$ et appelé *vecteur opposé* de $\mathbf{x}$, tel que $\mathbf{x} + (-\mathbf{x}) = \mathbf{0}$.
(e) $\alpha(\beta\mathbf{x}) = (\alpha\beta)\mathbf{x}$ (loi associative).
(f) $(\alpha + \beta)\mathbf{x} = \alpha\mathbf{x} + \beta\mathbf{x}$
(g) $\alpha(\mathbf{x} + \mathbf{y}) = \alpha\mathbf{x} + \alpha\mathbf{y}$ (lois distributives).
(h) $1\mathbf{x} = \mathbf{x}$.

1.2.3 Remarques et règles élémentaires

Voici quelques observations découlant directement de la définition.

(1) La remarque 1.1.6 s'applique dans le cas présent.

(2) Les lois distributives (f) et (g) s'étendent par récurrence à un nombre fini quelconque de termes:

$$(\alpha_1 + \ldots + \alpha_k)\mathbf{x} = \alpha_1\mathbf{x} + \ldots + \alpha_k\mathbf{x},$$
$$\alpha(\mathbf{x}_1 + \ldots + \mathbf{x}_k) = \alpha\mathbf{x}_1 + \ldots + \alpha\mathbf{x}_k.$$

(3) Le vecteur nul dont l'existence est assurée par la condition (c) est unique. En effet, si un vecteur $\mathbf{0}'$ satisfait aussi à cette condition, alors $\mathbf{0} = \mathbf{0} + \mathbf{0}' \underset{\text{par (b)}}{=} \mathbf{0}' + \mathbf{0} = \mathbf{0}'$, ce qui montre que $\mathbf{0} = \mathbf{0}'$. De manière analogue, on montre que l'opposé d'un vecteur est unique.

(4) *Soustraction vectorielle.* Soustraire le vecteur $\mathbf{y}$ du vecteur $\mathbf{x}$ signifie additionner $-\mathbf{y}$ à $\mathbf{x}$. Le vecteur ainsi obtenu est appelé *différence* de $\mathbf{x}$ et $\mathbf{y}$ et est noté $\mathbf{x} - \mathbf{y}$. Nous laissons au lecteur le soin de vérifier, en partant des conditions de la définition, que $\mathbf{x} - \mathbf{y}$ est l'unique vecteur $\mathbf{z}$ tel que $\mathbf{z} + \mathbf{y} = \mathbf{x}$.

(5) Si $\alpha = 0$ ou $\mathbf{x} = \mathbf{0}$, alors $\alpha\mathbf{x} = \mathbf{0}$. Il suffit de poser $\beta = 0$ dans (f) si $\alpha = 0$, $\mathbf{y} = \mathbf{0}$ dans (g) si $\mathbf{x} = \mathbf{0}$, et de soustraire aux deux membres de l'égalité $0\mathbf{x}$ dans le premier cas et $\alpha\mathbf{0}$ dans le deuxième cas.

(6) Si $\alpha\mathbf{x} = \mathbf{0}$, alors $\alpha = 0$ ou $\mathbf{x} = \mathbf{0}$. En effet, si $\alpha \neq 0$, en utilisant successivement les conditions (h), (e) et la règle (5), nous obtenons

$$\mathbf{x} = 1\mathbf{x} = (\frac{1}{\alpha}\alpha)\mathbf{x} = \frac{1}{\alpha}(\alpha\mathbf{x}) = \frac{1}{\alpha}\mathbf{0} = \mathbf{0}.$$

(7) $(-\alpha)\mathbf{x} = -(\alpha\mathbf{x})$, donc en particulier $(-1)\mathbf{x} = -\mathbf{x}$. En effet, l'opposé d'un vecteur étant unique d'après (3), il suffit de montrer que $\alpha\mathbf{x} + (-\alpha)\mathbf{x} = \mathbf{0}$. Or, par la condition (f), $\alpha\mathbf{x} + (-\alpha)\mathbf{x} = (\alpha - \alpha)\mathbf{x} = 0\mathbf{x}$ et, par (5), $0\mathbf{x} = \mathbf{0}$.

1.3 EXEMPLES D'ESPACES VECTORIELS

1.3.1 Espaces vectoriels géométriques

L'espace vectoriel géométrique V étudié dans la section 1.1 est un premier exemple d'espace vectoriel selon la définition 1.2.2. Un deuxième exemple est le plan vectoriel géométrique, c'est-à-dire l'ensemble des classes de flèches équivalentes du plan usuel de la géométrie élémentaire, muni des deux opérations introduites dans 1.1.4 et 1.1.7. Nous le désignerons également par V, mais s'il faut le distinguer du premier, nous utiliserons les symboles V^3 pour l'espace et V^2 pour le plan.

1.3.2 Vectorialisé de $\mathscr{G}$.

Soit O un point arbitrairement choisi et fixé de l'espace ponctuel $\mathscr{G}$ introduit dans 1.1.2. Définissons l'opération d'addition des points P et Q de $\mathscr{G}$ par la règle du parallélogramme: $P + Q$ est le sommet opposé à O du parallélogramme construit sur O, P, Q (fig. 1.6). Cette opération peut également être définie à l'aide des flèches: $P + Q$ est l'extrémité de la flèche d'origine P équivalente à la flèche OQ. Définissons similairement la multiplication de P par un nombre α (fig. 1.7). Muni de ces deux opérations, $\mathscr{G}$ devient un espace vectoriel appelé *vectorialisé* de $\mathscr{G}$ relativement à O. Nous désignerons cet espace par $\mathscr{G}_O$ et appellerons le point O *origine*. En identifiant chaque point P à la flèche OP, nous pouvons considérer $\mathscr{G}_O$ comme étant formé des flèches d'origine commune O.

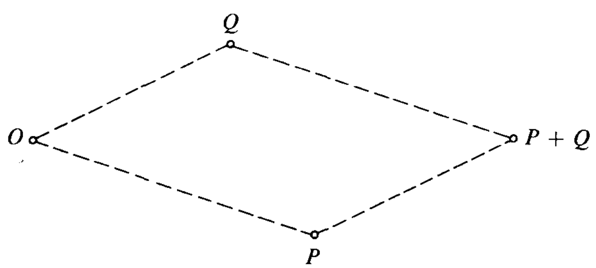


Fig. 1.6

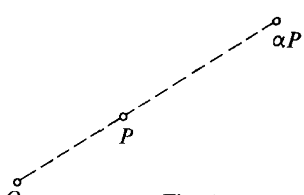


Fig. 1.7

On remarquera que si O' est une deuxième origine, $\mathscr{G}_{O'}$ et $\mathscr{G}_O$ sont égaux en tant qu'ensembles, mais différents en tant qu'espaces vectoriels (fig. 1.8).

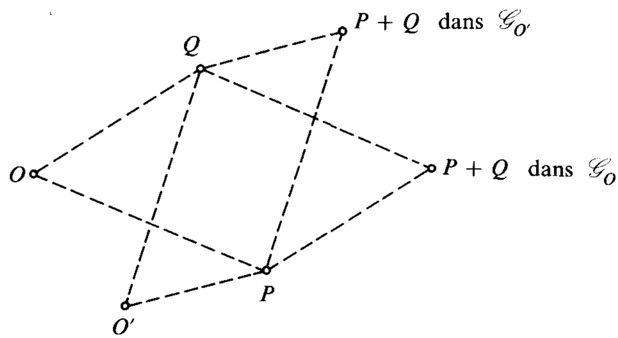


Fig. 1.8

1.3.3 Espaces $\mathbb{R}^n$

Pour tout entier positif n, $\mathbb{R}^n$ désignera l'ensemble des n-uplets de nombres disposés en colonne:

$$\begin{pmatrix} a_1 \\ a_2 \\ \cdot \\ \cdot \\ \cdot \\ a_n \end{pmatrix}.$$

Munissons $\mathbb{R}^n$ des opérations d'addition et de multiplication par un scalaire définies au moyen des formules suivantes:

$$\begin{pmatrix} a_1 \\ a_2 \\ \cdot \\ \cdot \\ \cdot \\ a_n \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \\ \cdot \\ \cdot \\ \cdot \\ b_n \end{pmatrix} = \begin{pmatrix} a_1 + b_1 \\ a_2 + b_2 \\ \cdot \\ \cdot \\ \cdot \\ a_n + b_n \end{pmatrix}, \quad \alpha \begin{pmatrix} a_1 \\ a_2 \\ \cdot \\ \cdot \\ \cdot \\ a_n \end{pmatrix} = \begin{pmatrix} \alpha a_1 \\ \alpha a_2 \\ \cdot \\ \cdot \\ \cdot \\ \alpha a_n \end{pmatrix}. \tag{1.1}$$

Ces deux opérations satisfont, de toute évidence, aux conditions (a)–(h) de la définition 1.2.2 et confèrent donc à $\mathbb{R}^n$ une structure d'espace vectoriel. Les vecteurs de cet espace seront appelés *vecteurs-colonnes*. Ils seront souvent désignés plus brièvement par $(a_i)_{1 \leqslant i \leqslant n}$ ou simplement par (a_i). Le nombre a_i sera appelé *terme d'indice i* de (a_i).

On remarquera que le vecteur nul de $\mathbb{R}^n$ est le n-uplet

$$\begin{pmatrix} 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{pmatrix}$$

et que

$$\begin{pmatrix} -a_1 \\ -a_2 \\ \cdot \\ \cdot \\ \cdot \\ -a_n \end{pmatrix} \quad \text{est l'opposé de} \quad \begin{pmatrix} a_1 \\ a_2 \\ \cdot \\ \cdot \\ \cdot \\ a_n \end{pmatrix}.$$

$\mathbb{R}^1$ sera identifié à $\mathbb{R}$.

1.3.4 Un espace vectoriel fonctionnel

Soit $C_{[a,\,b]}$ l'ensemble des fonctions réelles continues définies dans l'intervalle fermé $[a, b]$. Nous désignerons les éléments de cet ensemble par les lettres **f**, **g**, ... La valeur de **f** au point t sera notée $\mathbf{f}(t)$. Dire que $\mathbf{f} = \mathbf{g}$ équivaudra donc à dire que $\mathbf{f}(t) = \mathbf{g}(t)$ pour tout t de l'intervalle $[a, b]$. De manière abrégée, nous écrirons $\mathbf{f}(t) \equiv \mathbf{g}(t)$, le signe $\equiv$ indiquant ainsi que les deux membres sont égaux pour tout t de l'intervalle $[a, b]$. Considérons les deux opérations suivantes:

- $\mathbf{f} + \mathbf{g}$, définie par la formule $(\mathbf{f} + \mathbf{g})(t) \equiv \mathbf{f}(t) + \mathbf{g}(t)$,
- $\alpha\mathbf{f}$, définie par la formule $(\alpha\mathbf{f})(t) \equiv \alpha\mathbf{f}(t)$.

Ces deux opérations satisfont aux conditions (a)–(h) de la définition 1.2.2 et munissent $C_{[a,\,b]}$ d'une structure d'espace vectoriel. Le vecteur nul de cet espace est la fonction nulle et l'opposé de **f** est la fonction $-\mathbf{f}$ définie par $(-\mathbf{f})(t) \equiv -\mathbf{f}(t)$.

Il est intéressant de constater que $C_{[a,\,b]}$, en tant qu'espace vectoriel, est une généralisation naturelle de $\mathbb{R}^n$ au cas continu. On peut en effet concevoir tout vecteur (a_i) de $\mathbb{R}^n$ sous la forme d'une fonction réelle définie dans l'ensemble $\{1, 2, ..., n\}$: la valeur de cette fonction au point i est tout simplement a_i.

1.3.5 Autres espaces vectoriels fonctionnels

Voici quelques autres exemples d'espaces vectoriels fonctionnels. Les opérations d'addition et de multiplication par un scalaire sont définies comme dans 1.3.4.

(1) L'espace vectoriel $C^k_{(a,\,b)}$ formé des fonctions réelles k fois continûment dérivables, définies dans l'intervalle ouvert (a, b).

(2) L'espace vectoriel des fonctions réelles définies dans un intervalle.

(3) L'espace vectoriel des polynômes à une indéterminée.

(4) L'espace vectoriel des polynômes à une indéterminée de degré inférieur ou égal à n.

1.4 COMBINAISONS LINÉAIRES, SOUS-ESPACES VECTORIELS, FAMILLES GÉNÉRATRICES

1.4.1 Avertissement

Dorénavant, sauf mention explicite du contraire, les vecteurs seront les éléments d'un espace vectoriel donné E.

1.4.2 Combinaisons linéaires

On appelle *combinaison linéaire* des vecteurs $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$ tout vecteur de la forme $\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 + \ldots + \alpha_k\mathbf{x}_k$, où $\alpha_1, \alpha_2, \ldots, \alpha_k$ sont des nombres appelés *coefficients* de la combinaison linéaire.

1.4.3 Exemples

(1) Le vecteur nul est combinaison linéaire de $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$. Pour voir cela, il suffit de prendre $\alpha_1 = \alpha_2 = \ldots = \alpha_k = 0$. Dans ce cas, la combinaison linéaire est appelée *combinaison linéaire triviale*.

(2) Les combinaisons linéaires d'un vecteur $\mathbf{x}$ sont appelées *multiples* de $\mathbf{x}$. Un multiple de $\mathbf{x}$ est donc un vecteur de la forme $\alpha\mathbf{x}$. On notera que le vecteur nul est multiple de tout vecteur.

(3) *Combinaisons convexes*. On appelle *combinaison convexe* toute combinaison linéaire dont les coefficients sont non négatifs et de somme égale à 1. L'ensemble des combinaisons convexes de deux points P et Q de $\mathscr{G}_O$ est le segment de droite joignant P et Q. Pour s'en rendre compte, il suffit d'écrire

$$\alpha P + (1 - \alpha)Q = Q + \alpha(P - Q),$$

de faire varier α de 0 à 1 et de constater que tous les points du segment sont ainsi obtenus (fig. 1.9).

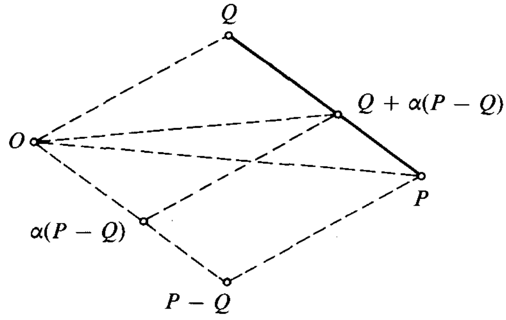


Fig. 1.9

Nous laissons au lecteur le soin de vérifier que l'ensemble des combinaisons convexes de trois points est le triangle, éventuellement dégénéré, dont les sommets sont ces trois points.

D'une manière générale, on peut montrer que l'ensemble des combinaisons convexes de $k > 3$ points est le plus petit polyèdre convexe (polygone convexe si $\mathscr{G}_O$ désigne le plan), éventuellement dégénéré, comprenant ces points.

(4) On dit que le vecteur-colonne (x_i^0) de $\mathbb{R}^3$ est solution du système d'équations linéaires

$$\begin{aligned} 3x_1 - 2x_2 + 4x_3 &= 0 \\ -x_1 + x_2 - 5x_3 &= 0 \end{aligned} \tag{1.2}$$

si ses termes x_1^0, x_2^0, x_3^0, substitués aux inconnues x_1, x_2, x_3, vérifient les deux équations.

Toute combinaison linéaire $\alpha_1(x_i^1) + \alpha_2(x_i^2)$ de solutions (x_i^1) et (x_i^2) du système (1.2) est encore une solution de ce système.

(5) Soit $\mathbf{f}_1$ et $\mathbf{f}_2$ les deux fonctions définies par

$$\mathbf{f}_1(t) \equiv \cos t \quad \text{et} \quad \mathbf{f}_2(t) \equiv \sin t. \tag{1.3}$$

Toute combinaison linéaire de $\mathbf{f}_1$ et $\mathbf{f}_2$ est solution de l'équation différentielle

$$\ddot{\mathbf{f}} + \mathbf{f} = \mathbf{0},$$

où $\ddot{\mathbf{f}}$ désigne la deuxième dérivée de $\mathbf{f}$.

1.4.4 Combinaisons linéaires itérées

Si le vecteur $\mathbf{x}$ est combinaison linéaire des vecteurs $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$ et chacun de ces vecteurs est combinaison linéaire des vecteurs $\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_l$, alors $\mathbf{x}$ est combinaison linéaire de $\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_l$.

1.4.5 Sous-espaces vectoriels

On appelle *sous-espace vectoriel* de E tout sous-ensemble de E qui est lui-même un espace vectoriel pour les opérations d'addition et de multiplication par un scalaire définies dans E.

Un sous-espace vectoriel, en tant qu'espace vectoriel, ne peut être vide, puisqu'il comprend au moins un vecteur, à savoir son vecteur nul, celui-ci étant d'ailleurs forcément le vecteur nul de E, en vertu de la règle (5) de 1.2.3. En outre, en même temps que les vecteurs $\mathbf{x}$ et $\mathbf{y}$, il comprend toutes leurs combinaisons linéaires $\alpha\mathbf{x} + \beta\mathbf{y}$. Inversement, on voit aussitôt que tout sous-ensemble jouissant de ces propriétés est un sous-espace vectoriel. Nous avons ainsi établi la proposition suivante:

1.4.6 Proposition. Caractérisation des sous-espaces vectoriels

Un sous-ensemble S de E est un sous-espace vectoriel de E si et seulement si S est non vide et $\alpha\mathbf{x} + \beta\mathbf{y}$ *appartient à S pour tout couple* $(\mathbf{x}, \mathbf{y})$ *de vecteurs de S et tout couple* (α, β) *de nombres.*

La proposition suivante en découle aisément:

1.4.7 Proposition

Soit $(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k)$ *une famille de vecteurs. L'ensemble des combinaisons linéaires de* $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k$ *est un sous-espace vectoriel S de E, plus précisément le plus petit sous-espace vectoriel de E (au sens de l'inclusion) comprenant* $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k$.

1.4.8 Générateurs, familles génératrices

Les vecteurs $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k$ de la proposition 1.4.7 sont appelés *générateurs* de S et la famille $(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k)$ *famille génératrice* de S. On dit aussi que ces vecteurs ou cette famille *engendrent* S.

1.4.9 Somme et intersection de sous-espaces vectoriels

Soit S et T des sous-espaces vectoriels de E. On appelle *somme* de S et T, et on note $S + T$, l'ensemble des vecteurs de la forme $\mathbf{s} + \mathbf{t}$, où $\mathbf{s}$ est un vecteur de S et $\mathbf{t}$ un vecteur de T. A l'aide de la proposition 1.4.7, le lecteur constatera facilement que la somme $S + T$ et l'intersection $S \cap T$ sont des sous-espaces vectoriels de E. Au contraire, la réunion $S \cup T$ n'est pas un sous-espace vectoriel de E, à moins que S ne soit contenu dans T, ou réciproquement. En fait, le plus petit sous-espace vectoriel de E contenant $S \cup T$ est la somme $S + T$ (cf. exercice 1.12.14).

Les résultats de ce paragraphe s'étendent, de manière évidente, au cas d'une famille quelconque $(S_1, S_2, ..., S_k)$ de sous-espaces vectoriels de E: la somme $S_1 + S_2 + ... + S_k$, c'est-à-dire l'ensemble des vecteurs de la forme $\mathbf{s}_1 + \mathbf{s}_2 + ... + \mathbf{s}_k$, où $\mathbf{s}_i$ est un vecteur de S_i pour $i = 1, 2, ..., k$, et l'intersection $S_1 \cap S_2 \cap ... \cap S_k$ sont des sous-espaces vectoriels de E.

1.4.10 Exemples

(1) $\{\mathbf{0}\}$ et E sont des sous-espaces vectoriels de E.

(2) Le sous-espace vectoriel engendré par un vecteur non nul est formé de tous les multiples de ce vecteur. On appelle un tel sous-espace *droite vectorielle.* Un sous-espace vectoriel engendré par deux vecteurs non multiples l'un de l'autre est appelé *plan vectoriel.* Dans $\mathscr{G}_O$ une droite et un plan vectoriels sont effectivement une droite et un plan passant par l'origine O.

(3) $\mathbb{R}^4$ est engendré par les vecteurs-colonnes

$$\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

En effet,

$$\begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix} = a_1 \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + a_2 \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} + a_3 \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} + a_4 \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}. \tag{1.4}$$

(4) Les fonctions $\mathbf{f}_1$ et $\mathbf{f}_2$ définies dans (1.3) engendrent le sous-espace vectoriel de $C^2_{(a,b)}$ formé des solutions de l'équation différentielle $\ddot{\mathbf{f}} + \mathbf{f} = \mathbf{0}$. Ce résultat d'analyse est énoncé sans démonstration.

(5) Etant donné cinq nombres $t_0, \ldots, t_4$ arbitrairement choisis, définissons les cinq polynômes du quatrième degré $\mathbf{p}_0, \ldots, \mathbf{p}_4$ par la formule

$$\mathbf{p}_i(t) \equiv \frac{(t - t_0) \ldots (t \hat{-} t_i) \ldots (t - t_4)}{(t_i - t_0) \ldots (t_i \hat{-} t_i) \ldots (t_i - t_4)}, \tag{1.5}$$

où l'accent circonflexe indique l'absence des facteurs d'indice i. Il est évident que

$$\mathbf{p}_i(t_i) = 1 \quad \text{et} \quad \mathbf{p}_i(t_j) = 0 \text{ si } i \neq j. \tag{1.6}$$

Nous allons établir que ces cinq polynômes engendrent l'espace vectoriel de tous les polynômes de degré inférieur ou égal à 4, en démontrant que si $\mathbf{p}$ est un tel polynône, alors

$$\mathbf{p}(t) \equiv \mathbf{p}(t_0)\,\mathbf{p}_0(t) + \ldots + \mathbf{p}(t_4)\,\mathbf{p}_4(t) \quad \textit{(formule de Lagrange)}. \tag{1.7}$$

A cet effet, désignons par $\tilde{\mathbf{p}}$ le polynôme défini par le second membre de (1.7). D'après (1.6),

$$\begin{aligned} \tilde{\mathbf{p}}(t_0) &= \mathbf{p}(t_0)\cdot 1 + 0 + 0 + 0 + 0 = \mathbf{p}(t_0) \\ &\ldots\ldots \\ \tilde{\mathbf{p}}(t_4) &= 0 + 0 + 0 + 0 + \mathbf{p}(t_4)\cdot 1 = \mathbf{p}(t_4), \end{aligned}$$

ce qui montre que $\mathbf{p}$ et $\tilde{\mathbf{p}}$ prennent les mêmes valeurs en $t = t_0, \ldots, t = t_4$, donc que le polynôme $\tilde{\mathbf{p}} - \mathbf{p}$ a au moins cinq zéros. Puisqu'il est de degré inférieur ou égal à 4, nous en concluons qu'il est nul, donc que $\tilde{\mathbf{p}} = \mathbf{p}$. La formule (1.7) est ainsi démontrée.

(6) L'ensemble des polynômes de degré inférieur ou égal à n est un sous-espace vectoriel de l'espace de tous les polynômes. Il est engendré par les monômes

$$\mathbf{p}_0\colon \mathbf{p}_0(t) \equiv 1\,, \quad \mathbf{p}_1\colon \mathbf{p}_1(t) \equiv t\,, \quad \ldots, \quad \mathbf{p}_n\colon \mathbf{p}_n(t) \equiv t^n. \tag{1.8}$$

(7) Si $\mathbf{x}$ n'est pas multiple de $\mathbf{y}$, $S = \{\mathbf{z} : \mathbf{z} = \mathbf{x} + \alpha\mathbf{y}, \alpha \in \mathbb{R}\}$ n'est pas un sous-espace vectoriel de E, ne serait-ce que par le fait que $\mathbf{0}$ n'est pas élément de S. En particulier, une droite ne passant pas par l'origine n'est pas un sous-espace vectoriel de $\mathscr{G}_O$. (Pour plus de détails, voir les sections 1.8 et 1.9.)

1.5 DÉPENDANCE ET INDÉPENDANCE LINÉAIRES

1.5.1 Caractérisation de l'absence de parallélisme à un même plan

Si $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ sont trois vecteurs de V^3 dont les représentants ne sont pas parallèles à un même plan (par convention, une flèche d'intensité nulle est parallèle à tout plan), alors tout vecteur $\mathbf{x}$ de V^3 s'écrit de manière unique sous la forme

$$\mathbf{x} = \alpha_1\mathbf{e}_1 + \alpha_2\mathbf{e}_2 + \alpha_3\mathbf{e}_3,$$

où α_1, α_2, α_3 sont des nombres (fig. 1.10).

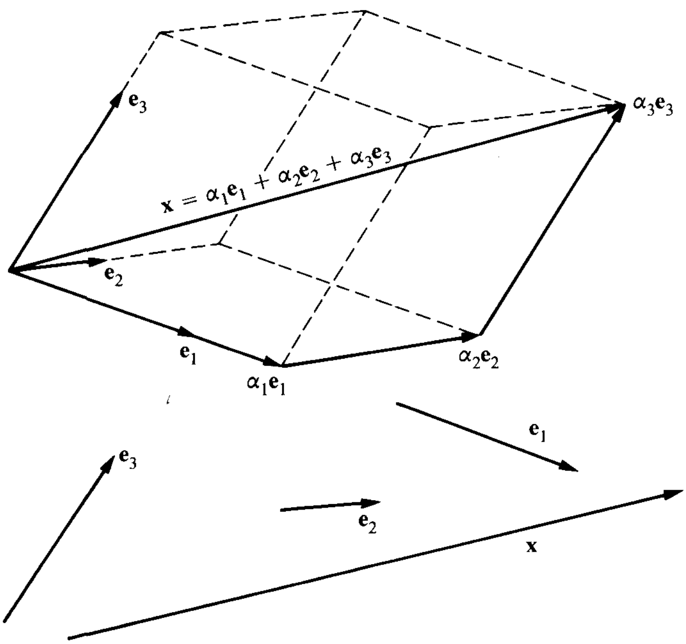


Fig. 1.10

En particulier, la seule possibilité d'obtenir le vecteur nul comme combinaison linéaire de $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ est d'attribuer la valeur 0 à α_1, α_2, α_3.

Réciproquement, si pour trois vecteurs $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ de V^3 la relation $\alpha_1\mathbf{e}_1 + \alpha_2\mathbf{e}_2 + \alpha_3\mathbf{e}_3 = \mathbf{0}$ implique $\alpha_1 = \alpha_2 = \alpha_3 = 0$, aucun de ces vecteurs ne peut être combinaison linéaire des deux autres, autrement dit, leurs représentants ne sont pas parallèles à un même plan.

Sur la base de ces observations, nous allons étendre la notion d'absence de parallélisme à un même plan au cas d'un nombre quelconque de vecteurs d'un espace vectoriel E.

1.5.2 Indépendance et dépendance linéaires, familles libres et liées

On dit que les vecteurs $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k$ sont *linéairement indépendants* si la relation $\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 + ... + \alpha_k\mathbf{x}_k = \mathbf{0}$ implique $\alpha_1 = \alpha_2 = ... = \alpha_k = 0$, autrement dit, si la combinaison linéaire triviale est la seule combinaison linéaire de $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k$ qui soit nulle. Dans le cas contraire, on dit que les vecteurs $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k$ sont *linéairement dépendants.*

Si l'attention est fixée sur la famille $(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k)$ plutôt que sur les termes dont elle est constituée, on dit que celle-ci est *libre* ou *liée* suivant que les vecteurs $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k$ sont linéairement indépendants ou dépendants.

1.5.3 Formulation de la dépendance linéaire

Selon la définition 1.5.2, les vecteurs $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k$ sont linéairement dépendants s'il existe des nombres non tous nuls $\alpha_1, \alpha_2, ..., \alpha_k$ tels que $\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 + ... + \alpha_k\mathbf{x}_k = \mathbf{0}$.

1.5.4 Exemples

Voici quelques cas où la définition 1.5.2 s'applique directement.

(1) Une famille réduite à un seul terme $\mathbf{x}$ est libre ou liée suivant que $\mathbf{x}$ est non nul ou nul.

(2) Pour qu'un couple $(\mathbf{x}, \mathbf{y})$ soit lié, il faut et il suffit que l'un des vecteurs $\mathbf{x}$ ou $\mathbf{y}$ soit multiple de l'autre. On remarquera que le couple $(\mathbf{x}, \mathbf{0})$ est lié, mais que $\mathbf{x}$ n'est pas multiple de $\mathbf{0}$ si $\mathbf{x} \neq \mathbf{0}$.

(3) Si un des vecteurs $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k$ est nul, ces vecteurs sont linéairement dépendants, car, en supposant $\mathbf{x}_i = \mathbf{0}$, nous voyons que la combinaison linéaire $0\mathbf{x}_1 + ... + 1\mathbf{x}_i + ... + 0\mathbf{x}_k$ est nulle sans être triviale.

(4) Si $\mathbf{x}_i = \mathbf{x}_j$ pour un couple d'indices i et j différents, alors les vecteurs $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k$ sont linéairement dépendants, car, étant admis par exemple que $i < j$, la combinaison linéaire $0\mathbf{x}_1 + ... + 1\mathbf{x}_i + ... + (-1)\mathbf{x}_j + ... + 0\mathbf{x}_k$ est nulle sans être triviale. Cet exemple et le précédent se généralisent de la façon suivante:

(5) Si les vecteurs $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k$ sont linéairement dépendants, alors les vecteurs $\mathbf{x}_1, ..., \mathbf{x}_k, \mathbf{x}_{k+1}, ..., \mathbf{x}_l$ le sont aussi, quels que soient $\mathbf{x}_{k+1}, ..., \mathbf{x}_l$ $(l > k)$. En d'autres termes, toute famille finie de vecteurs admettant une sous-famille liée est elle-même liée.

(6) Si les vecteurs $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k$ sont linéairement indépendants, les vecteurs $\mathbf{x}_{i_1}, \mathbf{x}_{i_2}, ..., \mathbf{x}_{i_l}$ $(1 \leqslant i_1 < i_2 < ... < i_l \leqslant k)$ le sont également. En d'autres termes, toute sous-famille d'une famille libre est elle-même libre.

La proposition suivante généralise l'exemple (2):

1.5.5 Proposition. Caractérisation de la dépendance linéaire

Pour qu'une famille de vecteurs $(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k)$ $(k > 1)$ *soit liée, il faut et il suffit qu'un vecteur* $\mathbf{x}_i$ *soit combinaison linéaire des vecteurs* $\mathbf{x}_j$ *avec* $j \neq i$.

DÉMONSTRATION

Si la famille $(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k)$ est liée, il existe des nombres non tous nuls α_1, $\alpha_2, ..., \alpha_k$ tels que $\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 + ... + \alpha_k\mathbf{x}_k = \mathbf{0}$. En supposant α_i non nul et en résolvant par rapport à $\mathbf{x}_i$, nous obtenons

$$\mathbf{x}_i = \frac{1}{\alpha_i}(-\alpha_1\mathbf{x}_1 - ... \widehat{\alpha_i\mathbf{x}_i} - ... - \alpha_k\mathbf{x}_k),$$

où l'accent circonflexe indique l'absence du terme d'indice i. Cela montre que $\mathbf{x}_i$ est combinaison linéaire des $\mathbf{x}_j$ avec $j \neq i$.

Inversement, si $\mathbf{x}_i = \alpha_1\mathbf{x}_1 + ... + \widehat{\alpha_i\mathbf{x}_i} + ... + \alpha_k\mathbf{x}_k$, alors $\alpha_1\mathbf{x}_1 + ... + (-1)\mathbf{x}_i + ... + \alpha_k\mathbf{x}_k = \mathbf{0}$, ce qui montre que la famille $(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k)$ est liée, car au moins un coefficient de la combinaison linéaire est non nul.

1.5.6 Proposition. Critère d'indépendance linéaire

Pour qu'une famille de vecteurs $(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k)$ *soit libre, il faut et il suffit qu'aucun vecteur* $\mathbf{x}$ *ne puisse s'écrire de deux manières sous la forme d'une combinaison linéaire des vecteurs* $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k$.

DÉMONSTRATION

Supposons que la famille $(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k)$ soit libre. S'il existait un vecteur $\mathbf{x}$ tel que

$$\mathbf{x} = \alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 + ... + \alpha_k\mathbf{x}_k = \alpha_1'\mathbf{x}_1 + \alpha_2'\mathbf{x}_2 + ... + \alpha_k'\mathbf{x}_k,$$

avec $\alpha_i \neq \alpha_i'$ pour au moins un indice i, la combinaison linéaire

$$(\alpha_1 - \alpha_1')\mathbf{x}_1 + (\alpha_2 - \alpha_2')\mathbf{x}_2 + ... + (\alpha_k - \alpha_k')\mathbf{x}_k$$

serait nulle sans être triviale, ce qui contredirait l'hypothèse.

Réciproquement, supposons qu'aucun vecteur $\mathbf{x}$ ne puisse être écrit de deux manières sous la forme d'une combinaison linéaire des vecteurs $\mathbf{x}_1, \mathbf{x}_2, .., \mathbf{x}_k$. Alors, en particulier, le vecteur nul ne pourra l'être, autrement dit, la combinaison linéaire triviale est la seule combinaison linéaire des vecteurs $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k$ qui s'annule, donc la famille $(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k)$ est libre.

1.6 BASES D'UN ESPACE VECTORIEL

1.6.1 Introduction

Les familles génératrices libres jouent un rôle important en algèbre linéaire. Elles seront étudiées dans la présente section et la suivante.

Comme précédemment, E désignera un espace vectoriel.

1.6.2 Bases d'un espace vectoriel

On dit qu'une famille finie de vecteurs est une *base* de E si elle est libre et engendre E.

D'après cette définition, toute famille libre $(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k)$ est une base du sous-espace vectoriel qu'elle engendre.

1.6.3 Proposition

Pour qu'une famille de vecteurs $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ *soit une base de E, il faut et il suffit que tout vecteur* $\mathbf{x}$ *de E s'exprime de manière unique sous la forme d'une combinaison linéaire des vecteurs* $\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n$*:*

$$\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + ... + x_n\mathbf{e}_n. \tag{1.9}$$

L'expression (1.9) est appelée *décomposition* de $\mathbf{x}$ suivant la base $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$.

DÉMONSTRATION

Par définition d'une base, les vecteurs $\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n$ engendrent E, donc tout vecteur $\mathbf{x}$ s'écrit sous la forme (1.9). D'autre part, puisque $\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n$ sont linéairement indépendants, la décomposition (1.9) est unique, d'après la proposition 1.5.6.

Réciproquement, si tout vecteur $\mathbf{x}$ s'écrit de manière unique sous la forme (1.9), les vecteurs $\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n$ engendrent E et, en outre, ils sont linéairement indépendants, encore d'après la proposition 1.5.6, donc $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ est une base de E.

1.6.4 Composantes d'un vecteur

Les coefficients $x_1, x_2, ..., x_n$ de la décomposition d'un vecteur $\mathbf{x}$ suivant une base sont appelés *composantes* de $\mathbf{x}$ dans cette base.

En présence d'une base, tout vecteur est donc entièrement déterminé par ses composantes.

1.6.5 Proposition

Si $x_1, x_2, ..., x_n$ *sont les composantes de* $\mathbf{x}$ *et* $y_1, y_2, ..., y_n$ *celles de* $\mathbf{y}$, *alors* $x_1 + y_1, x_2 + y_2, ..., x_n + y_n$ *sont les composantes de* $\mathbf{x} + \mathbf{y}$ *et* $\alpha x_1, \alpha x_2, ..., \alpha x_n$ *celles de* $\alpha\mathbf{x}$.

En d'autres termes, additionner deux vecteurs revient à additionner leurs composantes et multiplier un vecteur par α revient à multiplier ses composantes par α. La base est donc un outil de calcul important, car elle permet d'effectuer les opérations sur les vecteurs au moyen d'opérations sur les nombres.

1.6.6 Notation

De toute évidence, les composantes d'un vecteur dépendent du choix de la base, mais une fois que ce choix est fait, il n'y aura aucun danger d'ambiguïté lorsqu'on écrira $\mathbf{x}(x_1, x_2, ..., x_n)$ pour exprimer que $x_1, x_2, ..., x_n$ sont les composantes de $\mathbf{x}$. (Une écriture mieux appropriée au cas où des changements de base interviennent sera introduite dans 6.5.6.)

1.6.7 Exemples

(1) La donnée d'une base dans $\mathscr{G}_O$ équivaut à celle d'un système d'axes de référence d'origine O dans $\mathscr{G}$. Les composantes d'un point de $\mathscr{G}_O$ sont, dans ce cas, les coordonnées de ce point par rapport au système d'axes.

(2) Deux (trois) vecteurs linéairement indépendants de V^2 (V^3) forment une base.

(3) Les vecteurs-colonnes de $\mathbb{R}^n$

$$\begin{bmatrix} 1 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{bmatrix}, ..., \begin{bmatrix} 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 1 \end{bmatrix} \tag{1.10}$$

forment une base que l'on appelle *base canonique* de $\mathbb{R}^n$. Les composantes du vecteur-colonne (a_i) dans cette base sont $a_1, a_2, ..., a_n$.

(4) Les polynômes $\mathbf{p}_0, ..., \mathbf{p}_4$ définis dans (1.5) forment une base de l'espace vectoriel des polynômes de degré inférieur ou égal à 4. En effet, ils engendrent cet espace et, de plus, ils sont linéairement indépendants, car la relation $\alpha_0\mathbf{p}_0 + ... + \alpha_4\mathbf{p}_4 = \mathbf{0}$ implique $\alpha_0\mathbf{p}_0(t_i) + ... + \alpha_4\mathbf{p}_4(t_i) = 0$ et donc, par (1.6), $\alpha_i = 0$ pour $i = 0, ..., 4$. D'après (1.7), les composantes dans cette base d'un polynôme $\mathbf{p}$ de degré inférieur ou égal à 4 sont $\mathbf{p}(t_0), ..., \mathbf{p}(t_4)$.

(5) Les monômes $\mathbf{p}_0, ..., \mathbf{p}_n$ définis dans (1.8) forment une base de l'espace vectoriel des polynômes de degré inférieur ou égal à n. En effet, ils engendrent cet

espace et, en outre, ils sont linéairement indépendants, car la relation $\alpha_0\mathbf{p}_0 + \ldots + \alpha_n\mathbf{p}_n = \mathbf{0}$ s'écrit aussi sous la forme $\alpha_0 + \alpha_1 t + \ldots + \alpha_n t^n \equiv 0$ et implique donc $\alpha_0 = \alpha_1 = \ldots = \alpha_n = 0$ (cf. exercice 1.12.5). Les composantes d'un polynôme de degré inférieur ou égal à n dans cette base sont les coefficients de ce polynôme.

1.7 DIMENSION D'UN ESPACE VECTORIEL

1.7.1 Introduction

Au long de ce livre, nous nous occuperons principalement d'espaces vectoriels admettant une base. Un des objectifs de cette section est de caractériser ces espaces.

Dans cette section, E désignera encore un espace vectoriel.

1.7.2 Dimension finie et infinie

On dit que E est de *dimension finie* s'il est engendré par une famille finie de vecteurs. Dans le cas contraire, on dit que E est de *dimension infinie.*

1.7.3 Théorème. Prolongement d'une famille libre

Soit $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k)$ *une famille libre et* $(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m)$ *une famille génératrice de E. Si* $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k)$ *n'est pas une base de E, on peut extraire une sous-famille* $(\mathbf{v}_{i_1}, \mathbf{v}_{i_2}, \ldots, \mathbf{v}_{i_l})$ *de* $(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m)$ *de telle manière que la famille* $(\mathbf{x}_1, \ldots, \mathbf{x}_k, \mathbf{v}_{i_1}, \ldots, \mathbf{v}_{i_l})$ *soit une base de E.*

DÉMONSTRATION

Au moins un des vecteurs $\mathbf{v}_i$ n'est pas combinaison linéaire des vecteurs $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$, sinon $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k)$ engendrerait E, en raison de 1.4.4, et serait donc une base de E. Notons ce vecteur $\mathbf{v}_{i_1}$. La famille $(\mathbf{x}_1, \ldots, \mathbf{x}_k, \mathbf{v}_{i_1})$ est alors libre. En effet, la relation $\alpha_1\mathbf{x}_1 + \ldots + \alpha_k\mathbf{x}_k + \beta_1\mathbf{v}_{i_1} = \mathbf{0}$ implique d'abord $\beta_1 = 0$, autrement $\mathbf{v}_{i_1}$ serait combinaison linéaire des vecteurs $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$, et ensuite $\alpha_1 = \alpha_2 = \ldots = \alpha_k = 0$, puisque les vecteurs $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$ sont linéairement indépendants. Si la famille $(\mathbf{x}_1, \ldots, \mathbf{x}_k, \mathbf{v}_{i_1})$ engendre E, elle est une base de E et le théorème est alors démontré. Dans le cas contraire, le même raisonnement nous assure de l'existence d'un vecteur $\mathbf{v}_{i_2}$ tel que $(\mathbf{x}_1, \ldots, \mathbf{x}_k, \mathbf{v}_{i_1}, \mathbf{v}_{i_2})$ est une famille libre. Si cette famille n'engendre pas E, le procédé d'extraction de vecteurs $\mathbf{v}_i$ de $(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m)$ se poursuit. Lorsqu'il s'arrête, nous aurons obtenu un prolongement de $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k)$ en une famille libre engendrant E, c'est-à-dire en une base de E.

1.7.4 Corollaire. Existence et extraction d'une base

Tout espace vectoriel de dimension finie et non réduit au vecteur nul admet une base. En fait, de toute famille génératrice d'un tel espace on peut extraire une base.

DÉMONSTRATION

Soit $(\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_m)$ une famille génératrice de E. Si E n'est pas réduit à $\{\mathbf{0}\}$, au moins un vecteur $\mathbf{v}_i$ n'est pas nul. Désignons ce vecteur par $\mathbf{x}$. Le corollaire résulte alors du théorème 1.7.3 appliqué aux familles $(\mathbf{x})$ et $(\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_m)$.

1.7.5 Théorème

Si $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ *est une base de E, toute famille de vecteurs* $(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k)$ *dont le nombre de termes k est supérieur à n est liée.*

DÉMONSTRATION

Nous raisonnons par l'absurde en supposant que la famille $(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k)$ soit libre. Considérons la décomposition de $\mathbf{x}_1$ suivant la base $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$:

$$\mathbf{x}_1 = \alpha_1\mathbf{e}_1 + \alpha_2\mathbf{e}_2 + ... + \alpha_n\mathbf{e}_n.$$

Comme $\mathbf{x}_1$ n'est pas nul, au moins un des coefficients $\alpha_1, \alpha_2, ..., \alpha_n$ n'est pas nul. Quitte à énumérer autrement les termes de la base, nous pouvons admettre que ce coefficient est α_1. Dans ce cas,

$$\mathbf{e}_1 = \frac{1}{\alpha_1}\mathbf{x}_1 - \frac{\alpha_2}{\alpha_1}\mathbf{e}_2 - ... - \frac{\alpha_n}{\alpha_1}\mathbf{e}_n$$

et donc $(\mathbf{x}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ est une famille génératrice de E, puisque le sous-espace vectoriel qu'elle engendre comprend les vecteurs $\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n$. Exprimons alors $\mathbf{x}_2$ sous la forme d'une combinaison linéaire de $\mathbf{x}_1, \mathbf{e}_2, ..., \mathbf{e}_n$:

$$\mathbf{x}_2 = \beta_1\mathbf{x}_1 + \beta_2\mathbf{e}_2 + ... + \beta_n\mathbf{e}_n.$$

Les coefficients $\beta_2, ..., \beta_n$ ne sont pas tous nuls, sinon $\mathbf{x}_1$ et $\mathbf{x}_2$ seraient linéairement dépendants. Sans restreindre la généralité, nous pouvons admettre que β_2 n'est pas nul, ce qui nous permet d'écrire

$$\mathbf{e}_2 = -\frac{\beta_1}{\beta_2}\mathbf{x}_1 + \frac{1}{\beta_2}\mathbf{x}_2 - \frac{\beta_3}{\beta_2}\mathbf{e}_3 - ... - \frac{\beta_n}{\beta_2}\mathbf{e}_n$$

et donc de conclure que $(\mathbf{x}_1, \mathbf{x}_2, \mathbf{e}_3, ..., \mathbf{e}_n)$ est une famille génératrice de E, puisque le sous-espace vectoriel qu'elle engendre comprend les vecteurs $\mathbf{x}_1, \mathbf{e}_2, ..., \mathbf{e}_n$. La suite de la démonstration poursuit ce procédé d'échange: $\mathbf{e}_3, ..., \mathbf{e}_n$ sont remplacés successivement par $\mathbf{x}_3, ..., \mathbf{x}_n$. Le résultat montre que la famille $(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_n)$ engendre E. Mais alors $\mathbf{x}_k$ est combinaison linéaire des vecteurs $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_n$, ce qui contredit l'hypothèse, en raison de la proposition 1.5.5.

1.7.6 Corollaire

Si $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ *et* $(\mathbf{e}'_1, \mathbf{e}'_2, ..., \mathbf{e}'_k)$ *sont deux bases de* E, *alors* $n = k$.

DÉMONSTRATION

Du théorème 1.7.5 nous déduisons que $k \leqslant n$, ainsi que $n \leqslant k$, par un échange du rôle des deux bases. Il s'ensuit que $n = k$.

1.7.7 Dimension d'un espace vectoriel

Au moyen des corollaires 1.7.4 et 1.7.6, il est maintenant possible d'attribuer une dimension à tout espace vectoriel de dimension finie. Soit E un tel espace. On appelle *dimension* de E le nombre de termes d'une quelconque de ses bases. Si E se réduit au seul vecteur nul, on dit que sa dimension est nulle. La dimension de E sera notée $\dim E$.

1.7.8 Exemples

La dimension de V^2 est 2, celle de V^3 est 3 et celle de $\mathbb{R}^n$ est n. D'après l'exemple (5) de 1.6.7, la dimension de l'espace vectoriel des polynômes de degré inférieur ou égal à n est $n + 1$. L'espace vectoriel de tous les polynômes et, par conséquent, les espaces vectoriels $C_{[a,b]}$ et $C^k_{(a,b)}$ sont de dimension infinie. En effet, ces espaces admettent des familles libres arbitrairement grandes, par exemple $(\mathbf{p}_0, \mathbf{p}_1, ..., \mathbf{p}_n)$, où les $\mathbf{p}_i$ sont les monômes définis dans (1.8), n étant pris arbitrairement grand; par le théorème 1.7.5, ces espaces n'admettent aucune base, donc leur dimension est infinie, d'après le corollaire 1.7.4.

1.7.9 Proposition. Caractérisations d'une base

Supposons que E *soit de dimension finie non nulle* n.

(a) *Toute famille libre à* n *termes est une base de* E.

(b) *Toute famille génératrice à* n *termes est une base de* E.

DÉMONSTRATION

Assertion (a). Une famille libre à n termes qui ne serait pas une base se prolongerait en une base, d'après le théorème 1.7.3, et la dimension de E serait alors supérieure à n.

Assertion (b). D'une famille génératrice à n termes qui ne serait pas une base on pourrait extraire une base, d'après le corollaire 1.7.4, et la dimension de E serait alors inférieure à n.

1.7.10 Isomorphie de E et $\mathbb{R}^n$

Tout espace vectoriel E de dimension finie non nulle n peut être mis en correspondance biunivoque avec $\mathbb{R}^n$. Il suffit de choisir une base de E et de faire correspondre à tout vecteur $\mathbf{x}$ de E le vecteur-colonne dont les termes sont les composantes de $\mathbf{x}$ dans la base choisie:

$$\begin{array}{ccc} E & & \mathbb{R}^n \\ \mathbf{x}(x_1, x_2, \ldots, x_n) & \longleftrightarrow & \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix} . \end{array} \qquad (1.11)$$

D'après la proposition 1.6.5, cette correspondance conserve les opérations d'addition vectorielle et de multiplication par un scalaire; en d'autres termes, elle permet, comme nous l'avons déjà fait remarquer, d'effectuer les opérations sur les vecteurs par des opérations sur les nombres. On dit que E et $\mathbb{R}^n$ sont *isomorphes*, ou que la correspondance est un *isomorphisme* (cf. 6.3.7). Evidemment, cet isomorphisme dépend de la base de E choisie.

Il importe de noter qu'un espace vectoriel E de dimension n n'admet, en général, aucune base privilégiée, contrairement à $\mathbb{R}^n$; par conséquent, sauf si le choix d'une base de E a été fait, il faudra éviter de considérer E et $\mathbb{R}^n$ comme identiques.

1.8 RETOUR AUX SOUS-ESPACES VECTORIELS, SOMMES DIRECTES

1.8.1 Introduction

Dans cette section, nous reprenons l'étude des sous-espaces vectoriels d'un espace vectoriel donné E et introduisons les notions de rang d'une famille finie de vecteurs, ainsi que celle de somme directe de sous-espaces vectoriels.

1.8.2 Rang d'une famille finie de vecteurs

On appelle *rang* d'une famille finie de vecteurs la dimension du sous-espace vectoriel de E qu'elle engendre.

1.8.3 Proposition

Le rang d'une famille de vecteurs $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k)$ *est inférieur ou égal à k. Il est égal à k si et seulement si cette famille est libre.*

DÉMONSTRATION

Ecartons le cas trivial où le rang de la famille $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k)$ est nul. D'après le corollaire 1.7.4, on peut alors extraire de cette famille une base du sous-espace vectoriel qu'elle engendre. Le rang est donc inférieur à k ou égal à k suivant que $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k)$ est une famille liée ou libre.

1.8.4 Proposition. Comparaison de deux rangs

Pour que le rang d'une famille de vecteurs $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k)$ *soit égal au rang de la famille augmentée* $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k, \mathbf{y})$, *il faut et il suffit que le vecteur* $\mathbf{y}$ *soit combinaison linéaire des vecteurs* $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$.

DÉMONSTRATION

Notons S et T les sous-espaces vectoriels engendrés respectivement par $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k)$ et $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k, \mathbf{y})$. Si $\mathbf{y}$ est combinaison linéaire des vecteurs $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$, alors $S = T$, donc les deux rangs sont égaux. Réciproquement, supposons que les deux rangs soient égaux et montrons que $S = T$, ce qui nous permettra de conclure que $\mathbf{y}$ est combinaison linéaire des vecteurs $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$. Si les deux rangs sont nuls, il est clair que $S = T = \{\mathbf{0}\}$. Sinon, une base de S est également une base de T, puisque S est inclus dans T, ce qui entraîne que $S = T$.

1.8.5 Proposition. Sous-espaces vectoriels d'un espace vectoriel de dimension finie

Si E est de dimension finie et S est un sous-espace vectoriel de E, alors S est de dimension finie et $\dim S \leqslant \dim E$. *En outre,* $\dim S = \dim E$ *si et seulement si* $S = E$.

DÉMONSTRATION

Désignons la dimension de E par n et supposons que celle de S soit infinie ou finie et supérieure à n. Par récurrence, nous allons démontrer l'existence d'une famille libre $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{n+1})$ de vecteurs de S, ce qui contredit le théorème 1.7.5, vu la définition 1.7.7. Comme S n'est pas de dimension nulle, il comprend au moins un vecteur non nul $\mathbf{x}_1$, donc $(\mathbf{x}_1)$ est une famille libre. Supposons que $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k)$ est une famille libre de vecteurs de S. Si k est inférieur ou égal à n, au moins un des vecteurs de S n'est pas combinaison linéaire de $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$, sinon S serait de dimension k, contrairement à l'hypothèse. Désignons ce vecteur par $\mathbf{x}_{k+1}$. En vertu de la proposition précédente, la famille $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{k+1})$ de vecteurs de S est alors libre.

Pour démontrer la seconde assertion, il suffit de poser $T = E$ dans la dernière partie de la démonstration précédente.

1.8.6 Hyperplans vectoriels

Si E est de dimension finie non nulle n, un sous-espace vectoriel de E de dimension $n - 1$ est appelé *hyperplan vectoriel.* Par exemple, un hyperplan vectoriel se réduit au vecteur nul si $n = 1$, est une droite vectorielle si $n = 2$, un plan vectoriel si $n = 3$.

1.8.7 Sommes directes

On dit que la somme $S + T$ de deux sous-espaces vectoriels S et T de E est *directe* si $S \cap T = \{\mathbf{0}\}$. Dans ce cas, on la note $S \oplus T$.

Par exemple, si E est de dimension 2 et $(\mathbf{e}_1, \mathbf{e}_2)$ est une base de E, alors $E = D_1 \oplus D_2$, où D_1 et D_2 sont les droites vectorielles engendrées respectivement par $\mathbf{e}_1$ et $\mathbf{e}_2$.

Tout vecteur d'une somme directe $S \oplus T$ s'écrit d'une seule manière sous la forme $\mathbf{s} + \mathbf{t}$, où $\mathbf{s}$ est un vecteur de S et $\mathbf{t}$ un vecteur de T. En effet, si

$$\mathbf{s} + \mathbf{t} = \mathbf{u} + \mathbf{v},$$

où $\mathbf{s}$, $\mathbf{u}$ sont des vecteurs de S et $\mathbf{t}$, $\mathbf{v}$ des vecteurs de T, alors

$$\mathbf{s} - \mathbf{u} = \mathbf{v} - \mathbf{t}$$

est un vecteur de $S \cap T = \{\mathbf{0}\}$, donc $\mathbf{s} - \mathbf{u} = \mathbf{v} - \mathbf{t} = \mathbf{0}$, ce qui entraîne $\mathbf{s} = \mathbf{u}$ et $\mathbf{t} = \mathbf{v}$.

Inversement, si tout vecteur d'une somme $S + T$ s'écrit d'une seule manière sous la forme $\mathbf{s} + \mathbf{t}$, où $\mathbf{s}$ est un vecteur de S et $\mathbf{t}$ un vecteur de T, alors $S \cap T = \{\mathbf{0}\}$ et donc $S + T$ est une somme directe. En effet, un vecteur non nul $\mathbf{u}$ de $S \cap T$ permettrait au vecteur nul d'avoir deux décompositions, à savoir $\mathbf{0} = \mathbf{0} + \mathbf{0}$ et $\mathbf{0} = \mathbf{u} + (-\mathbf{u})$.

1.8.8 Proposition. Existence d'un sous-espace complémentaire

Supposons que E soit de dimension finie. Pour tout sous-espace vectoriel S de E, il existe un sous-espace vectoriel T de E (non unique) tel que E soit somme directe de S et T. On dit que T est un sous-espace complémentaire *de S dans E.*

DÉMONSTRATION

Ecartons les cas triviaux où $S = \{\mathbf{0}\}$ et $S = E$. Le sous-espace vectoriel S admet alors une base $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_k)$, où k est inférieur à la dimension n de E. Par le théorème 1.7.3, cette base se prolonge en une base $(\mathbf{e}_1, ..., \mathbf{e}_k, \mathbf{e}_{k+1}, ..., \mathbf{e}_n)$ de E. Soit T le sous-espace vectoriel engendré par la famille $(\mathbf{e}_{k+1}, \mathbf{e}_{k+2}, ..., \mathbf{e}_n)$. Si $\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + ... + x_n\mathbf{e}_n$ est un vecteur quelconque de E, alors $\mathbf{x} = \mathbf{s} + \mathbf{t}$, où $\mathbf{s} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + ... + x_k\mathbf{e}_k$ est un vecteur de S et $\mathbf{t} = x_{k+1}\mathbf{e}_{k+1} + x_{k+2}\mathbf{e}_{k+2} + ...$

$+ \, x_n\mathbf{e}_n$ est un vecteur de T. En outre, $S \cap T = \{\mathbf{0}\}$ car aucun vecteur, excepté le vecteur nul, ne peut être combinaison linéaire des vecteurs $\mathbf{e}_1$, $\mathbf{e}_2$, ..., $\mathbf{e}_k$ et des vecteurs $\mathbf{e}_{k+1}$, $\mathbf{e}_{k+2}$, ..., $\mathbf{e}_n$. Nous en concluons que $E = S \oplus T$.

Par exemple, si S est un hyperplan, tout vecteur n'appartenant pas à S engendre une droite vectorielle D telle que $E = S \oplus D$.

1.8.9 Proposition. Dimension d'une somme directe

Si E est somme directe de deux sous-espaces vectoriels S et T de dimension finie, alors E est de dimension finie et

$$\dim E = \dim S + \dim T. \tag{1.12}$$

DÉMONSTRATION

Ecartons le cas trivial où un des sous-espaces S et T se réduit à $\{\mathbf{0}\}$. Soit $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_k)$ une base de S et $(\mathbf{e}_{k+1}, \mathbf{e}_{k+2}, ..., \mathbf{e}_n)$ une base de T. Il suffit de démontrer que $(\mathbf{e}_1, ..., \mathbf{e}_k, \mathbf{e}_{k+1}, ..., \mathbf{e}_n)$ est une base de E, ou, ce qui revient au même, que tout vecteur de E s'écrit de manière unique sous la forme d'une combinaison linéaire des vecteurs $\mathbf{e}_1, ..., \mathbf{e}_k, \mathbf{e}_{k+1}, ..., \mathbf{e}_n$. Or, cela est immédiat, car tout vecteur de E s'écrit de manière unique sous la forme $\mathbf{s} + \mathbf{t}$, où $\mathbf{s}$ est un vecteur de S et $\mathbf{t}$ un vecteur de T.

1.8.10 Extension

On dit que la somme $S_1 + S_2 + ... + S_k$ des sous-espaces vectoriels S_1, S_2, ..., S_k de E est *directe* si, pour $i = 1, 2, ..., k$, $S_i \cap T_i = \{\mathbf{0}\}$, où T_i est la somme des S_j d'indice j différent de i. On note cette somme directe $S_1 \oplus S_2 \oplus ... \oplus S_k$.

De même qu'en 1.8.7, on démontre qu'une somme $S_1 + S_2 + ... + S_k$ est directe si et seulement si chacun de ses vecteurs s'écrit d'une seule manière sous la forme $\mathbf{s}_1 + \mathbf{s}_2 + ... + \mathbf{s}_k$, où $\mathbf{s}_i$ est un vecteur de S_i pour $i = 1, 2, ..., k$. Il est évident que cette condition est remplie si et seulement si la seule décomposition du vecteur nul est celle dans laquelle tous les $\mathbf{s}_i$ sont nuls.

Si les dimensions de S_1, S_2, ..., S_k sont finies, la relation (1.12) se généralise et devient

$$\dim(S_1 \oplus S_2 \oplus ... \oplus S_k) = \dim S_1 + \dim S_2 + ... + \dim S_k. \tag{1.13}$$

1.9 ESPACES AFFINES

1.9.1 Introduction

L'espace $\mathscr{G}$ de la géométrie élémentaire est à la fois le modèle usuel et la source de la notion d'espace affine que nous allons introduire. Cet espace $\mathscr{G}$ est associé à l'espace vectoriel géométrique V par la correspondance entre flèches et vecteurs étudiée dans la section 1.1. La définition suivante ne fait que mettre en évidence les traits dominants de cette correspondance.

1.9.2 Espaces affines

Soit $\mathscr{E}$ un ensemble non vide d'éléments que nous appellerons *points* et désignerons par les lettres P, Q, ... ; soit en outre E un espace vectoriel. Supposons qu'à tout couple de points (P, Q) corresponde un vecteur noté $\overrightarrow{PQ}$. On dit que $\mathscr{E}$ est un *espace affine* d'*espace directeur* ou *direction E* si les conditions suivantes sont satisfaites:

(a) Pour tout point P fixé, la correspondance entre couples (P, Q) et vecteurs $\mathbf{x}$ est biunivoque, autrement dit, pour tout vecteur $\mathbf{x}$ il existe un point Q et un seul tel que $\mathbf{x} = \overrightarrow{PQ}$.

(b) Pour tout triplet de points (P, Q, R), $\overrightarrow{PQ} + \overrightarrow{QR} = \overrightarrow{PR}$ (*relation de Chasles*).

1.9.3 Notation

Si P est un point et $\mathbf{x}$ un vecteur, pour exprimer que Q est l'unique point tel que $\mathbf{x} = \overrightarrow{PQ}$, nous écrirons

$$Q = P + \mathbf{x}. \tag{1.14}$$

Bien qu'un peu abusive, cette écriture est commode à l'usage et suggère bien le sens de l'opération qu'elle désigne (fig. 1.11).

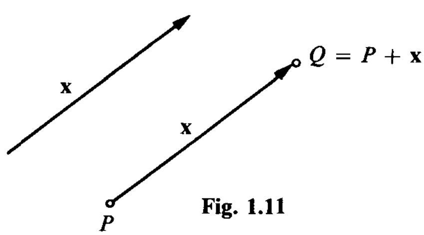


Fig. 1.11

On remarquera que

$$P + (\mathbf{x} + \mathbf{y}) = (P + \mathbf{x}) + \mathbf{y}.$$

1.9.4 Dimension d'un espace affine

On appelle *dimension* d'un espace affine la dimension de son espace directeur.

1.9.5 Règles de calcul dans les espaces affines

Les règles suivantes découlent directement de la définition 1.9.2.

(1) Pour tout point P, $\overrightarrow{PP} = \mathbf{0}$. Cela résulte de la condition (b) appliquée au cas où $P = Q = R$.

(2) Si $\overrightarrow{PQ} = \mathbf{0}$, alors $P = Q$. Cela résulte de la condition (a), compte tenu de la règle (1).

(3) $\overrightarrow{PQ} = -\overrightarrow{QP}$. Il suffit de poser $R = P$ dans la condition (b).

(4) *Règle du parallélogramme.* $\overrightarrow{PP'} = \overrightarrow{QQ'}$ si et seulement si $\overrightarrow{P'Q'} = \overrightarrow{PQ}$. En effet, d'après la condition (b), $\overrightarrow{PP'} + \overrightarrow{P'Q'} = \overrightarrow{PQ'} = \overrightarrow{PQ} + \overrightarrow{QQ'}$ (fig. 1.12).

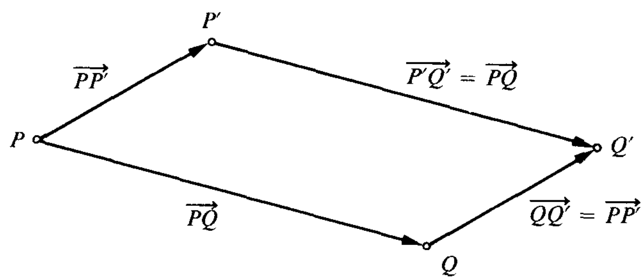


Fig. 1.12

1.9.6 Vectorialisé d'un espace affine

Dans 1.3.2, nous avons montré comment l'espace $\mathscr{S}$ peut être muni d'une structure d'espace vectoriel. Dans le cas général d'un espace affine $\mathscr{E}$, le procédé est le même. On choisit un point quelconque O de $\mathscr{E}$. La correspondance entre couples (O, P) et vecteurs de l'espace directeur E étant alors biunivoque (par (a)), tout comme celle entre couples (O, P) et points P, on définit l'addition de points et la multiplication d'un point par un scalaire par les opérations correspondantes sur les vecteurs de E. Muni de ces deux opérations, $\mathscr{E}$ devient un espace vectoriel, appelé *vectorialisé* de $\mathscr{E}$ relativement à O. Nous désignerons cet espace par $\mathscr{E}_O$ et appellerons O *origine*.

Vu la manière dont les opérations ont été définies, il résulte que $\mathscr{E}_O$ est isomorphe (cf. 1.7.10) à l'espace directeur E:

$$\begin{array}{ccc} \mathscr{E}_O & & E \\ P = O + \mathbf{x} & \longleftrightarrow & \mathbf{x} \end{array}$$

Toutefois, cet isomorphisme dépend du choix de l'origine O et en pratique cette origine est choisie sur la base de données inhérentes aux problèmes posés. Par exemple, si une transformation affine admet un point invariant, nous verrons qu'il y a avantage à choisir ce point comme origine.

1.9.7 Exemples

(1) Il est dit dans 1.9.1 que l'espace $\mathscr{G}$ de la géométrie élémentaire est un espace affine. En effet, sa direction est l'espace géométrique V et les conditions (a) et (b) de la définition 1.9.2 sont satisfaites. Il faut bien noter qu'au couple de points (P, Q) est associé le vecteur $\overrightarrow{PQ}$ et non pas la flèche PQ. En fait, la flèche pouvant être identifée au couple de points, nous voyons que ce que postule la définition 1.9.2 n'est rien d'autre qu'une forme abstraite de correspondance entre flèches et vecteurs.

(2) Tout espace vectoriel E peut être considéré comme un espace affine de direction E lui-même si au couple de vecteurs $(\mathbf{x}, \mathbf{y})$ est associé le vecteur $\mathbf{y} - \mathbf{x}$. En effet, les conditions (a) et (b) de la définition 1.9.2 sont satisfaites.

(3) Soit E un espace vectoriel, S un sous-espace vectoriel de E distinct de E et $\mathbf{x}$ un vecteur de E. Nous désignerons par $\mathbf{x} + S$ l'ensemble des vecteurs $\mathbf{z}$ de la forme $\mathbf{x} + \mathbf{y}$, où $\mathbf{y}$ parcourt S. Si $\mathbf{x}$ n'appartient pas à S, $\mathbf{x} + S$ n'est pas un sous-espace vectoriel de E, car le vecteur nul n'appartient pas à $\mathbf{x} + S$. Par contre, $\mathbf{x} + S$ devient un espace affine de direction S, lorsqu'on y introduit la correspondance entre couples et vecteurs définie dans l'exemple (2). En effet, la différence de deux vecteurs de $\mathbf{x} + S$ est un vecteur de S et les conditions (a) et (b) de la définition 1.9.2 sont satisfaites.

(4) Pour illustrer l'exemple (3), considérons le système d'équations linéaires

$$\begin{aligned} 3x_1 - 2x_2 + 4x_3 &= 3 \\ -x_1 + x_2 - 5x_3 &= -4. \end{aligned} \tag{1.15}$$

Comme dans l'exemple (4) de 1.4.3, nous appellerons solution de ce système tout vecteur-colonne (x_i^0) de $\mathbb{R}^3$ dont les termes x_1^0, x_2^0, x_3^0 vérifient les deux équations. Prenons une solution particulière de ce système, par exemple

$$\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}.$$

Nous obtenons alors la solution générale en additionnant à cette solution particulière la solution générale du système (1.2) (cf. 3.5.5). En d'autres termes, l'ensemble des solutions de (1.15) s'écrit sous la forme

$$\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} + S,$$

où S est le sous-espace vectoriel de $\mathbb{R}^3$ formé des solutions du système (1.2).

1.10 SOUS-ESPACES AFFINES, PARALLÉLISME

1.10.1 Introduction

L'exemple (3) de 1.9.7 et son illustration (4) nous suggèrent la notion de sous-espace affine que nous allons introduire. Les sous-espaces affines de l'espace affine $\mathscr{G}$ de la géométrie élémentaire sont les points, les droites, les plans et $\mathscr{G}$ lui-même.

Dans ce qui suit, $\mathscr{E}$ désignera un espace affine de direction E.

1.10.2 Sous-espaces affines

On dit qu'un sous-ensemble $\mathscr{S}$ de $\mathscr{E}$ est un *sous-espace affine* s'il existe un point P_0 de $\mathscr{E}$ et un sous-espace vectoriel S de E tels que

$$\mathscr{S} = \{P\colon \overrightarrow{P_0P} \in S\} = \{P\colon P = P_0 + \mathbf{x},\ \ \mathbf{x} \in S\}. \tag{1.16}$$

En d'autres termes, $\mathscr{S}$ est un sous-espace affine si, pour un point P_0 de $\mathscr{E}$, $\mathscr{S}$ est un sous-espace vectoriel de $\mathscr{E}_{P_0}$ (fig. 1.13).

On notera que $\mathscr{S}$ ainsi défini n'est pas vide, car il comprend au moins le point P_0.

Suivant l'exemple (3) de 1.9.7, nous désignerons le dernier membre de (1.16) plus brièvement par $P_0 + S$, ce qui nous permettra de considérer un sous-espace affine comme un sous-ensemble $\mathscr{S}$ de $\mathscr{E}$ de la forme

$$\mathscr{S} = P_0 + S, \tag{1.17}$$

où P_0 est un point de $\mathscr{E}$ et S un sous-espace vectoriel de E.

On prendra garde à bien distinguer les différentes significations du signe $+$: addition dans E, «addition» d'un point et d'un vecteur, «addition» d'un point et d'un sous-espace vectoriel.

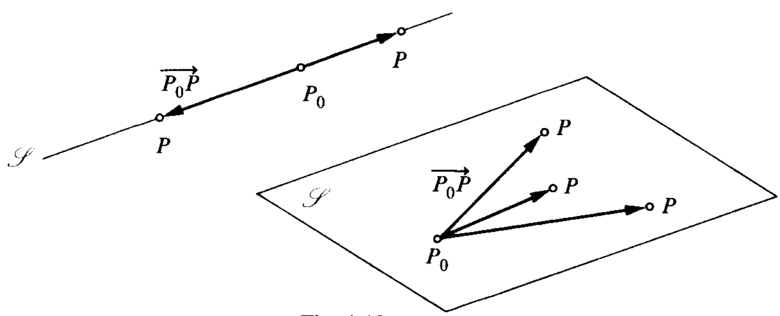


Fig. 1.13

1.10.3 Proposition

Soit $\mathscr{S}$ le sous-espace affine défini par P_0 et S.
(a) *Si P est un point quelconque de $\mathscr{S}$, alors $\mathscr{S} = P + S$.*
(b) $S = \{\mathbf{x}: \mathbf{x} = \overrightarrow{PQ},\ P, Q \in \mathscr{S}\}$.

DÉMONSTRATION

Assertion (a). Posons $\mathscr{S}' = P + S$. Il suffit de montrer que $\mathscr{S}'$ est inclus dans $\mathscr{S}$, car l'argument symétrique nous fournira l'inclusion opposée et donc la conclusion $\mathscr{S} = \mathscr{S}'$. Soit Q un point de $\mathscr{S}'$, c'est-à-dire un point tel que $\overrightarrow{PQ}$ appartienne à S. Puisque P est un point de $\mathscr{S}$, $\overrightarrow{P_0P}$ appartient à S. Mais $\overrightarrow{P_0Q} = \overrightarrow{P_0P} + \overrightarrow{PQ}$, par la condition (b) de la définition 1.9.2, donc $\overrightarrow{P_0Q}$ appartient aussi à S et, par conséquent, Q est un point de $\mathscr{S}$.

Assertion (b). En vertu de (a), pour tout couple (P, Q) de points de $\mathscr{S}$, $\overrightarrow{PQ}$ est un vecteur de S. Réciproquement, si $\mathbf{x}$ est un vecteur de S, choisissons un point quelconque P de $\mathscr{S}$ et posons $Q = P + \mathbf{x}$. D'après (a), Q est un point de $\mathscr{S}$, donc $\mathbf{x}$ est bien un vecteur de la forme $\overrightarrow{PQ}$, avec P et Q des points de $\mathscr{S}$.

1.10.4 Espace directeur

D'après la proposition 1.10.3, le sous-espace affine $\mathscr{S}$ défini par (1.16) ne dépend pas du choix de l'«origine» P_0 et détermine le sous-espace vectoriel S. On dit que S est l'*espace directeur* ou la *direction* de $\mathscr{S}$.

1.10.5 Sous-espaces affines comme espaces affines

Soit $\mathscr{S}$ un sous-espace affine de $\mathscr{E}$ de direction S. A l'aide de la proposition 1.10.3, nous voyons que la correspondance héritée de $\mathscr{E}$ entre couples de points de $\mathscr{S}$ et vecteurs de S fait de $\mathscr{S}$ un espace affine dans le sens de la définition 1.9.2.

1.10.6 Dimension d'un sous-espace affine

On appelle *dimension* d'un sous-espace affine $\mathscr{S}$ de $\mathscr{E}$ la dimension de $\mathscr{S}$ en tant qu'espace affine, c'est-à-dire la dimension de l'espace directeur de $\mathscr{S}$.

1.10.7 Intersection de sous-espaces affines

Soit $\mathscr{S}$ et $\mathscr{T}$ deux sous-espaces affines de $\mathscr{E}$. Si $\mathscr{S}$ et $\mathscr{T}$ ont au moins un point commun P, leur intersection est un sous-espace affine de $\mathscr{E}$ de direction l'intersection de leurs directions. En effet, $\mathscr{S} \cap \mathscr{T} = \{Q: \overrightarrow{PQ} \in S\} \cap \{Q: \overrightarrow{PQ} \in T\} = \{Q: \overrightarrow{PQ} \in S \cap T\}$, où S et T sont les directions respectives de $\mathscr{S}$ et de $\mathscr{T}$.

1.10.8 Sous-espaces affines engendrés

Soit $\mathscr{S}$ un sous-espace affine de $\mathscr{E}$ de direction S et $P_1, P_2, \ldots, P_l$ des points de $\mathscr{E}$. On appelle sous-espace affine *engendré* par $\mathscr{S}$ et $P_1, P_2, \ldots, P_l$ le plus petit sous-espace affine $\mathscr{T}$ contenant $\mathscr{S}$ et comprenant $P_1, P_2, \ldots, P_l$. Si S admet une base $(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k)$, on voit aisément que la direction de $\mathscr{T}$ est le sous-espace vectoriel engendré par les vecteurs $\mathbf{v}_1, \ldots, \mathbf{v}_k, \overrightarrow{P_0P_1}, \ldots, \overrightarrow{P_0P_l}$, où P_0 est un point quelconque de $\mathscr{S}$. Par exemple, le sous-espace affine de $\mathscr{G}$ engendré par une droite et un point est le plan passant par cette droite et ce point (supposé hors de la droite).

1.10.9 Exemples

(1) Comme déjà annoncé et illustré (fig. 1.13), les sous-espaces affines de $\mathscr{G}$ sont les points, les droites, les plans et $\mathscr{G}$ lui-même.

(2) L'espace affine $\mathscr{E}$ est un sous-espace affine de lui-même. Pour s'en rendre compte, il suffit de mettre E à la place de S dans (1.16).

(3) Pour tout point P_0 de $\mathscr{E}$, $\{P_0\}$ est un sous-espace affine. On voit cela en posant $S = \{\mathbf{0}\}$ dans (1.16).

(4) Un sous-espace affine de dimension nulle se réduit à un seul point. Un sous-espace affine de dimension 1 est appelé *droite affine* ou simplement *droite*; un de dimension 2 *plan affine* ou simplement *plan*. Il est clair qu'une droite est déterminée par deux de ses points et un plan par trois de ses points non alignés, c'est-à-dire n'appartenant pas à une même droite.

1.10.10 Hyperplans affines

Si $\mathscr{E}$ est de dimension finie non nulle n, un sous-espace affine de $\mathscr{E}$ de dimension $n - 1$ est appelé *hyperplan affine* ou simplement *hyperplan*. Un hyperplan est donc un sous-espace affine de direction un hyperplan vectoriel. Par exemple, un hyperplan est un point si $n = 1$, une droite si $n = 2$ et un plan si $n = 3$.

1.10.11 Parallélisme

Soit $\mathscr{S}$ et $\mathscr{T}$ des sous-espaces affines de $\mathscr{E}$ de directions respectives S et T. On dit que $\mathscr{S}$ et $\mathscr{T}$ sont *parallèles* si l'une des directions S ou T est incluse dans l'autre. Si $S = T$, on dit aussi que $\mathscr{S}$ et $\mathscr{T}$ sont *parallèles au sens étroit*.

On remarquera que ces deux notions de parallélisme s'équivalent lorsqu'elles sont appliquées à des sous-espaces affines de même dimension finie (en raison de la seconde partie de la proposition 1.8.5). En les appliquant au cas particulier de $\mathscr{G}$, nous retrouvons les notions usuelles de parallélisme entre droites, entre plans et entre droites et plans.

1.10.12 Deux résultats

Formulées en termes de droites et de plans de $\mathscr{G}$, les deux assertions suivantes deviennent des énoncés bien connus de la géométrie élémentaire.

(1) *Généralisation du cinquième postulat d'Euclide.* Soit P un point de $\mathscr{E}$ et $\mathscr{S}$ un sous-espace affine de $\mathscr{E}$. Il existe un unique sous-espace affine $\mathscr{T}$ de $\mathscr{E}$ parallèle à $\mathscr{S}$ au sens étroit et comprenant P. En effet, ce sous-espace affine est tout simplement $\mathscr{T} = \{Q: \overrightarrow{PQ} \in S\}$, où S est la direction de $\mathscr{S}$.

(2) *Position relative de deux sous-espaces affines parallèles.* Si deux sous-espaces affines sont parallèles, soit ils sont disjoints, soit l'un d'entre eux est inclus dans l'autre. S'ils sont parallèles au sens étroit, soit ils sont disjoints, soit confondus. Il suffit en effet de choisir comme point P_0 des représentations (1.16) des deux sous-espaces affines un point de leur intersection (dans le cas où celle-ci n'est pas vide), pour que la conclusion découle de la définition 1.10.11.

1.11 REPÈRES, REPRÉSENTATION PARAMÉTRIQUE, GÉOMÉTRIE ANALYTIQUE AFFINE

1.11.1 Introduction

Comme pour le choix d'une base dans un espace vectoriel, le choix d'un repère dans un espace affine de dimension n permettra d'identifier cet espace à $\mathbf{R}^n$ et, par ce moyen, de traiter les problèmes géométriques par des calculs sur les coordonnées (géométrie analytique).

Dans cette section, $\mathscr{E}$ désignera un espace affine de dimension finie non nulle n et de direction E.

1.11.2 Repères

On appelle *repère* de $\mathscr{E}$ tout ensemble $(O; \mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ formé d'un point O, appelé *origine*, et d'une base $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ de E.

1.11.3 Remarque

On peut concevoir un repère sous la forme d'une famille de points $(O; P_1, P_2, ..., P_n)$ telle que $(\overrightarrow{OP_1}, \overrightarrow{OP_2}, ..., \overrightarrow{OP_n})$ est une base de E.

1.11.4 Coordonnées

On appelle *coordonnées cartésiennes* ou simplement *coordonnées* d'un point P dans un repère $(O; \mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ les composantes du vecteur $\overrightarrow{OP}$ dans la base $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$.

Si $x_1, x_2, ..., x_n$ sont les coordonnées d'un point P et $y_1, y_2, ..., y_n$ celles d'un point Q, les composantes du vecteur $\overrightarrow{PQ} = \overrightarrow{OQ} - \overrightarrow{OP}$ sont $y_1 - x_1, y_2 - x_2, ..., y_n - x_n$. Inversement, si $x_1, x_2, ..., x_n$ sont les coordonnées d'un point P et $z_1, z_2, ..., z_n$ les composantes d'un vecteur $\mathbf{z}$, les coordonnées du point $Q = P + \mathbf{z}$ sont $x_1 + z_1, x_2 + z_2, ..., x_n + z_n$, car P et Q sont liés par la relation $\overrightarrow{PQ} = \overrightarrow{OQ} - \overrightarrow{OP} = \mathbf{z}$.

Par la suite, un point P défini par ses coordonnées $x_1, x_2, ..., x_n$ sera désigné par $P(x_1, x_2, ..., x_n)$.

1.11.5 Représentation paramétrique d'un sous-espace affine

Soit $\mathscr{S}$ un sous-espace affine de $\mathscr{E}$ de direction S et de dimension non nulle k. Soit en outre P_0 un point de $\mathscr{S}$ et $(\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k)$ une base de S. D'après (a) de la proposition 1.10.3, un point P appartient à $\mathscr{S}$ si et seulement si $P = P_0 + \mathbf{x}$, où $\mathbf{x}$ est un vecteur de S, c'est-à-dire un vecteur de la forme $\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + ... + \alpha_k\mathbf{v}_k$. Il s'ensuit que $\mathscr{S}$ est l'ensemble des points P satisfaisant à la relation

$$P = P_0 + \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + ... + \alpha_k\mathbf{v}_k, \tag{1.18}$$

où $\alpha_1, \alpha_2, ..., \alpha_k$ sont des variables parcourant $\mathbb{R}$. Cette relation est appelée *représentation paramétrique* de $\mathscr{S}$. Les vecteurs $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k$ sont appelés *vecteurs directeurs* de $\mathscr{S}$ et les variables $\alpha_1, \alpha_2, ..., \alpha_k$ *paramètres* de la représentation. On notera que le nombre de paramètres est égal à la dimension de $\mathscr{S}$ (fig. 1.14).

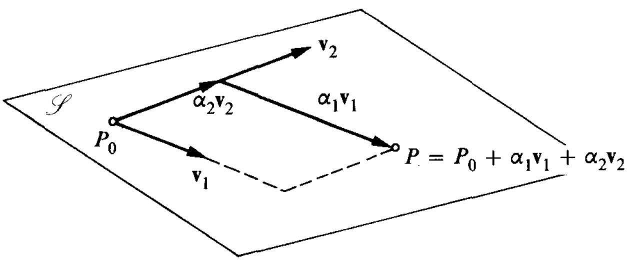


Fig. 1.14

1.11.6 Equations paramétriques

Supposons maintenant que $\mathscr{E}$ soit muni d'un repère $(O; \mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$. En désignant par $x_1, x_2, ..., x_n$ et $x_1^0, x_2^0, ..., x_n^0$ les coordonnées respectives de P et P_0 et par $v_{1i}, v_{2i}, ..., v_{ni}$ les composantes de $\mathbf{v}_i$, nous voyons que la relation (1.18) s'écrit, de manière équivalente, sous la forme

$$\begin{aligned} x_1 &= x_1^0 + \alpha_1 v_{11} + \alpha_2 v_{12} + ... + \alpha_k v_{1k} \\ x_2 &= x_2^0 + \alpha_1 v_{21} + \alpha_2 v_{22} + ... + \alpha_k v_{2k} \\ &\cdots\cdots \\ x_n &= x_n^0 + \alpha_1 v_{n1} + \alpha_2 v_{n2} + ... + \alpha_k v_{nk}. \end{aligned} \tag{1.19}$$

Ces équations sont appelées *équations paramétriques* de $\mathscr{S}$.

1.11.7 Cas particuliers

Une droite est déterminée par un de ses points et un vecteur directeur, un plan par un de ses points et deux vecteurs directeurs.

	Droite:	Plan:	
Représentation paramétrique:	$\mathbf{P} = \mathbf{P}_0 + \alpha\mathbf{v}$	$\mathbf{P} = \mathbf{P}_0 + \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2$	
Equations paramétriques dans le cas $n = 3$:	$x_1 = x_1^0 + \alpha v_1$ $x_2 = x_2^0 + \alpha v_2$ $x_3 = x_3^0 + \alpha v_3$	$x_1 = x_1^0 + \alpha_1 v_{11} + \alpha_2 v_{12}$ $x_2 = x_2^0 + \alpha_1 v_{21} + \alpha_2 v_{22}$ $x_3 = x_3^0 + \alpha_1 v_{31} + \alpha_2 v_{32}$.	(1.20)

1.11.8 Equation cartésienne d'un hyperplan

Supposons que n soit supérieur à 1. Si $k = n - 1$, les équations paramétriques (1.19) sont celles d'un hyperplan. Par le procédé d'élimination (cf. exemple (2) de 3.1.6), $n - 1$ équations peuvent être utilisées pour éliminer les $n - 1$ paramètres $\alpha_1, \alpha_2, \ldots, \alpha_{n-1}$ de l'équation restante. Le résultat sera une relation linéaire entre les coordonnées $x_1, x_2, \ldots, x_n$, plus exactement une équation de la forme

$$\nu_1 x_1 + \nu_2 x_2 + \ldots + \nu_n x_n = \delta, \tag{1.21}$$

où $\nu_1, \nu_2, \ldots, \nu_n$ sont des nombres non tous nuls et δ est un nombre pouvant s'annuler, auquel cas l'hyperplan passe par l'origine O. Cette équation est appelée *équation cartésienne* ou simplement *équation de l'hyperplan*. Les coordonnées d'un point P satisfont à (1.21) si et seulement si P appartient à l'hyperplan.

Réciproquement, toute équation de la forme (1.21) est l'équation d'un hyperplan, à condition, bien entendu, que les coefficients $\nu_1, \nu_2, \ldots, \nu_n$ ne soient pas tous nuls. En effet, si par exemple ν_n est non nul, les solutions de (1.21) peuvent être écrites sous la forme

$$\begin{aligned}
x_1 &= \alpha_1\\
x_2 &= \alpha_2\\
&\cdots\cdots\\
x_{n-1} &= \alpha_{n-1}\\
x_n &= \frac{\delta}{\nu_n} - \frac{\nu_1}{\nu_n}\alpha_1 - \ldots - \frac{\nu_{n-1}}{\nu_n}\alpha_{n-1},
\end{aligned}$$

où $\alpha_1, \alpha_2, \ldots, \alpha_{n-1}$ sont des variables parcourant $\mathbb{R}$. Ces équations sont de la forme (1.19) et représentent donc un hyperplan.

Dans le cas particulier où $n = 2$,

$$\nu_1 x_1 + \nu_2 x_2 = \delta$$

est l'équation cartésienne d'une droite et dans celui où $n = 3$,

$$\nu_1 x_1 + \nu_2 x_2 + \nu_3 x_3 = \delta$$

est l'équation cartésienne d'un plan.

1.11.9 Problèmes de géométrie analytique affine

La géométrie analytique affine traite les problèmes de la géométrie affine (parallélisme, incidence, ...) par des calculs dans $\mathbb{R}^n$. Voici quelques problèmes résolus.

(1) Trouver les équations paramétriques du sous-espace affine engendré par les points $P_0(1, 0, 3, -1)$, $P_1(-1, 2, 0, 3)$, $P_2(0, 0, -1, 2)$, $P_3(-3, 5, 1, 2)$.

Solution. Trois vecteurs directeurs de ce sous-espace affine sont

$$\begin{aligned}
\mathbf{v}_1 &= \overrightarrow{P_0P_1}(-2, 2, -3, 4)\\
\mathbf{v}_2 &= \overrightarrow{P_0P_2}(-1, 0, -4, 3)\\
\mathbf{v}_3 &= \overrightarrow{P_0P_3}(-4, 5, -2, 3)\,.
\end{aligned}$$

Ses équations paramétriques sont donc

$$\begin{aligned}
x_1 &= 1 - 2\alpha_1 - \alpha_2 - 4\alpha_3\\
x_2 &= 2\alpha_1 + 5\alpha_3\\
x_3 &= 3 - 3\alpha_1 - 4\alpha_2 - 2\alpha_3\\
x_4 &= -1 + 4\alpha_1 + 3\alpha_2 + 3\alpha_3\,.
\end{aligned}$$

(2) Trouver l'intersection des deux plans d'équations paramétriques

$$\begin{aligned}
x_1 &= 1 + \alpha_1 + 2\alpha_2 & \qquad x_1 &= 1 - \alpha_1'\\
x_2 &= 1 - \alpha_1 + \alpha_2 & \qquad x_2 &= 1 + \alpha_1' + \alpha_2'\\
x_3 &= -1 + 2\alpha_1 - \alpha_2 & \qquad x_3 &= -1 + 2\alpha_2'\\
x_4 &= -2 + \alpha_1 + \alpha_2\,, & \qquad x_4 &= -2 + 2\alpha_1' + 3\alpha_2'\,.
\end{aligned}$$

Solution. Deux vecteurs directeurs du premier plan sont $\mathbf{v}_1(1, -1, 2, 1)$ et $\mathbf{v}_2(2, 1, -1, 1)$, deux du deuxième $\mathbf{v}_1'(-1, 1, 0, 2)$ et $\mathbf{v}_2'(0, 1, 2, 3)$. Ces quatre vecteurs sont linéairement indépendants, donc aucun vecteur, sauf le vecteur nul, ne peut être en même temps combinaison linéaire de $\mathbf{v}_1$ et $\mathbf{v}_2$ et de $\mathbf{v}_1'$ et $\mathbf{v}_2'$. Cela entraîne que l'intersection des deux espaces directeurs est $\{\mathbf{0}\}$. D'autre part, en posant $\alpha_1 = \alpha_2 = \alpha_1' = \alpha_2' = 0$ dans les équations paramétriques, nous voyons que le point $P_0(1, 1, -1, -2)$ appartient aux deux plans. Nous en concluons que l'intersection cherchée se réduit au point P_0.

(3) Deux droites $\mathscr{D}$ et $\mathscr{D}'$ sont données par leurs équations paramétriques

$$\begin{aligned}
x_1 &= -\tfrac{2}{3} - 2\alpha & \qquad x_1 &= \alpha'\\
x_2 &= \tfrac{3}{2} + 3\alpha & \qquad x_2 &= -1 + 2\alpha'\\
x_3 &= 6\alpha\,, & \qquad x_3 &= 4 + \alpha'.
\end{aligned}$$

Déterminer les équations paramétriques de la droite passant par le point $P_0(1, 1, -1)$ et rencontrant $\mathscr{D}$ et $\mathscr{D}'$.

Solution. Les points d'intersection P de $\mathscr{D}$ et la droite cherchée et P' de $\mathscr{D}'$ et cette même droite satisfont à la relation $\overrightarrow{P_0P} = \beta\overrightarrow{P_0P'}$, où β est un nombre inconnu non nul. En composantes cette relation s'exprime par les trois équations

$$\begin{aligned} -\tfrac{5}{3} - 2\alpha &= \beta(-1 + \alpha') \\ \tfrac{1}{2} + 3\alpha &= \beta(-2 + 2\alpha') \\ 1 + 6\alpha &= \beta(\ 5 + \alpha'). \end{aligned}$$

En multipliant la deuxième équation par 2 et en la soustrayant de la troisième, nous obtenons aussitôt $\alpha' = 3$, ce qui nous permet de calculer les coordonnées de P', à savoir 3, 5, 7. En prenant alors pour vecteur directeur de la droite cherchée le vecteur $\frac{1}{2}\overrightarrow{P_0P'}$, nous voyons que les équations paramétriques de celle-ci s'écrivent

$$\begin{aligned} x_1 &= 1 + \alpha \\ x_2 &= 1 + 2\alpha \\ x_3 &= -1 + 4\alpha . \end{aligned}$$

Une autre manière de résoudre ce problème consiste à déterminer l'intersection des deux plans définis par P_0 et chacune des deux droites $\mathscr{D}$ et $\mathscr{D}'$.

(4) Déterminer la projection du point $P(1, -3, -2, 1)$ sur l'hyperplan d'équation

$$x_1 - x_2 + 3x_3 - 4x_4 = 4,$$

parallèlement à une droite de vecteur directeur $\mathbf{v}(1, 4, 3, -1)$.

Solution. Les équations paramétriques de la droite passant par P de vecteur directeur $\mathbf{v}$ sont

$$\begin{aligned} x_1 &= 1 + \alpha \\ x_2 &= -3 + 4\alpha \\ x_3 &= -2 + 3\alpha \\ x_4 &= 1 - \alpha . \end{aligned}$$

La projection cherchée est le point d'intersection de cette droite et l'hyperplan donné. Ses coordonnées s'obtiennent en déterminant la valeur de α pour laquelle les seconds membres des équations paramétriques vérifient l'équation cartésienne de l'hyperplan. Cette valeur est $\alpha = 1$, donc les coordonnées de la projection de P sont 2, 1, 1, 0.

1.12 EXERCICES

1.12.1 Montrer que les vecteurs-colonnes

$$\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}$$

engendrent $\mathbb{R}^3$ et exprimer le vecteur-colonne $\begin{pmatrix} -3 \\ 4 \\ -1 \end{pmatrix}$ comme combinaison linéaire de ces vecteurs

1.12.2 Montrer que toute fonction de la forme $\mathbf{f}(t) \equiv \alpha \sin(\lambda t + \mu)$ $(\alpha, \lambda \geqslant 0, \mu \in [0, 2\pi))$ est combinaison linéaire des fonctions $\mathbf{s}(t) \equiv \sin(\lambda t)$ et $\mathbf{c}(t) \equiv \cos(\lambda t)$ et, inversement, que toute combinaison linéaire de $\mathbf{s}$ et $\mathbf{c}$ est une fonction de la même forme que $\mathbf{f}$.

1.12.3 Exprimer le polynôme $\mathbf{p}(t) \equiv 1 + t^4$ sous la forme d'une combinaison linéaire des polynômes $\mathbf{p}_0$, $\mathbf{p}_1$, $\mathbf{p}_2$, $\mathbf{p}_3$ et $\mathbf{p}_4$ définis dans (1.5).

1.12.4 Démontrer que les vecteurs-colonnes

$$\begin{pmatrix} a_1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} b_1 \\ b_2 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ 0 \end{pmatrix}, \begin{pmatrix} d_1 \\ d_2 \\ d_3 \\ d_4 \end{pmatrix}$$

sont linéairement indépendants si et seulement si a_1, b_2, c_3 et d_4 sont différents de 0.

1.12.5 En dérivant n fois la relation $\alpha_0 + \alpha_1 t + \ldots + \alpha_n t^n \equiv 0$, montrer que les monômes $\mathbf{p}_0$, $\mathbf{p}_1$, ..., $\mathbf{p}_n$ définis dans (1.8) sont linéairement indépendants.

1.12.6 Montrer que les fonctions $\mathbf{f}_1(t) \equiv 1$, $\mathbf{f}_2(t) \equiv e^t$ et $\mathbf{f}_3(t) \equiv e^{2t}$ sont linéairement indépendantes.

1.12.7 Soit $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k)$ une famille libre de vecteurs d'un espace vectoriel E. Soit $(\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_k)$ une deuxième famille de vecteurs de E. On suppose que chaque vecteur $\mathbf{x}_i$ soit combinaison linéaire des vecteurs $\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_k$. Montrer que la famille $(\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_k)$ est libre.

1.12.8 Soit $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4)$ une base d'un espace vectoriel E de dimension 4.
(a) Montrer que les vecteurs $\mathbf{v}_1 = \mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_4$, $\mathbf{v}_2 = \mathbf{e}_1 - \mathbf{e}_2 + \mathbf{e}_3 - 2\mathbf{e}_4$ et $\mathbf{v}_3 = -\mathbf{e}_1 + \mathbf{e}_2 - \mathbf{e}_3 + \mathbf{e}_4$ sont linéairement indépendants.
(b) Prolonger la famille $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ en une base de E.

1.12.9 Soit $(\mathbf{e}_1, \mathbf{e}_2)$ une base d'un espace vectoriel E de dimension 2. Montrer que les familles de vecteurs $(\mathbf{e}_1 + \mathbf{e}_2, \mathbf{e}_2)$, $(\mathbf{e}_1 + \mathbf{e}_2, \mathbf{e}_1 - \mathbf{e}_2)$ et $(\mathbf{e}_1, \mathbf{e}_2 + \alpha\mathbf{e}_1)$ (α étant un nombre arbitraire) sont des bases de E. Calculer, en outre, les composantes de $\mathbf{e}_1$ et $\mathbf{e}_2$ dans chacune de ces bases.

1.12.10 Dans chacun des cas suivants, dire si la famille $(\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3)$ est une base de l'espace vectoriel des polynômes de degré inférieur ou égal à 2. Si oui, trouver les composantes dans cette base des monômes $\mathbf{p}_0(t) \equiv 1$, $\mathbf{p}_1(t) \equiv t$ et $\mathbf{p}_2(t) \equiv t^2$.
(a) $\mathbf{f}_1(t) \equiv 1 + t^2$, $\mathbf{f}_2(t) \equiv 2 - t + t^2$, $\mathbf{f}_3(t) \equiv -6 + 3t - 3t^2$.
(b) $\mathbf{f}_1(t) \equiv 1 + 2t + t^2$, $\mathbf{f}_2(t) \equiv 1 - 2t + t^2$, $\mathbf{f}_3(t) \equiv t$.
(c) $\mathbf{f}_1(t) \equiv t$, $\mathbf{f}_2(t) \equiv t + t^2$, $\mathbf{f}_3(t) \equiv 1 + t + t^2$.

1.12.11 Les sous-ensembles suivants sont-ils des sous-espaces vectoriels de $\mathbb{R}^5$?

$\{(a_i)\colon a_1 = 0\}$, $\{(a_i)\colon a_2 = 1\}$, $\{(a_i)\colon 2a_1 - a_4 = 0\}$, $\{(a_i)\colon a_1a_4 = a_5\}$, $\{(a_i)\colon a_1 \text{ rationnel}\}$.

1.12.12 Soit $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_6)$ une base d'un espace vectoriel E de dimension 6. Soit S le sous-espace vectoriel de E engendré par les vecteurs $\mathbf{e}_1 - \mathbf{e}_2$, $\mathbf{e}_3 + \mathbf{e}_4$ et $\mathbf{e}_5 + \mathbf{e}_6$. Quelles conditions doivent satisfaire les composantes d'un vecteur $\mathbf{x}$ de E pour qu'il appartienne à S?

1.12.13 Soit E un espace vectoriel de dimension 4, muni d'une base $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4)$.
(a) Montrer que l'ensemble S des vecteurs $\mathbf{x}$ de E dont les composantes x_1, x_2, x_3, x_4 satisfont à la condition $x_1 - 2x_2 + x_3 - x_4 = 0$ est un hyperplan vectoriel.
(b) Exhiber une base de S.
(c) Trouver un sous-espace complémentaire de S dans E.

1.12.14 Soit S et T deux sous-espaces vectoriels d'un espace vectoriel E.
(a) Démontrer que la somme $S + T$ et l'intersection $S \cap T$ sont des sous-espaces vectoriels de E.
(b) Démontrer que $S + T$ est l'intersection de tous les sous-espaces vectoriels U de E tels que $U \supset S \cup T$.

1.12.15 Soit S et T les sous-espaces vectoriels de $\mathbb{R}^4$ engendrés par les familles respectives

$$\left(\begin{bmatrix} 1 \\ -1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -2 \\ 5 \\ -1 \end{bmatrix}\right) \text{ et } \left(\begin{bmatrix} 1 \\ -1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 4 \\ -2 \\ 3 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ -1 \\ 1 \\ -1 \end{bmatrix}\right).$$

Déterminer les dimensions de S et de T, ainsi que celles de $S + T$ et de $S \cap T$.

1.12.16 Soit S, T et U trois sous-espaces vectoriels d'un espace vectoriel E.
(a) Montrer que $S + (T \cap U) \subset (S + T) \cap (S + U)$.
(b) Montrer que si $S \subset T$, les deux membres de l'inclusion sont égaux.
(c) Trouver un exemple montrant que l'égalité n'est généralement pas vraie.

1.12.17 Soit S et T deux sous-espaces vectoriels de dimension finie d'un espace vectoriel E. Montrer que

$$\dim(S + T) + \dim(S \cap T) = \dim S + \dim T.$$

En déduire la relation (1.12).

Exercices sur les espaces affines

1.12.18 Soit $\mathscr{S}$ l'hyperplan d'équation $3x_1 - x_2 + x_3 - 2x_4 = 1$ et $\mathscr{D}$ le lieu géométrique défini par les équations $\dfrac{x_1 - 1}{-2} = \dfrac{x_2 - 1}{2} = x_3 - 2 = \dfrac{x_4 + 1}{-4}$.
(a) Montrer que $\mathscr{D}$ est une droite et écrire ses équations paramétriques.
(b) Calculer les coordonnées du point d'intersection de $\mathscr{D}$ et $\mathscr{S}$.

1.12.19 Déterminer la projection parallèlement à une droite de vecteur directeur $\mathbf{v}(-1, 2, 1, 2, -2)$ de la droite d'équations $x_1 - 2 = \dfrac{x_2 + 1}{2} = \dfrac{x_3 - 5}{3} = \dfrac{x_4 + 2}{-1} = \dfrac{x_5 - 1}{4}$ sur l'hyperplan d'équation $2x_1 + 3x_2 - x_5 = 2$.

1.12.20 Dans chacun des cas suivants, étudier la position relative des sous-espaces affines $\mathscr{S}$ et $\mathscr{T}$.
(a) $\mathscr{S}$ est l'hyperplan d'équation $x_1 - x_2 + x_4 = 1$ et $\mathscr{T}$ le plan passant par les points $P_1(1, 1, 2, 2)$, $P_2(2, 2, 4, 2)$ et $P_3(1, 3, 1, 4)$.
(b) $\mathscr{S}$ et $\mathscr{T}$ sont les plans d'équations paramétriques respectives

$$\begin{aligned} x_1 &= 1 + \alpha_1 - \alpha_2 \\ x_2 &= 2 + \alpha_1 + \alpha_2 \\ x_3 &= -3 + \alpha_1 - \alpha_2 \\ x_4 &= 1 + \alpha_1 + \alpha_2, \end{aligned} \qquad \begin{aligned} x_1 &= 2 - \alpha_1' - \alpha_2' \\ x_2 &= 3 + \alpha_1' - \alpha_2' \\ x_3 &= -2 + \alpha_1' + \alpha_2' \\ x_4 &= 4 - \alpha_1' + \alpha_2'. \end{aligned}$$

(c) $\mathscr{S}$ est l'hyperplan d'équation $3x_1 - x_3 + 2x_4 = 2$ et $\mathscr{T}$ la droite passant par le point $P_0(1, 5, 3, 1)$ et de vecteur directeur $\mathbf{v}(1, 0, 1, -1)$.

(d) $\mathscr{S}$ est l'hyperplan d'équation $x_1 - 3x_2 + x_4 = 3$ et $\mathscr{T}$ le plan passant par le point $P_0(1, 0, 3, -2)$ et de vecteurs directeurs $\mathbf{v}_1(-1, 0, 1, 1)$ et $\mathbf{v}_2(1, 1, -2, 1)$.

(e) $\mathscr{S}$ est l'hyperplan d'équation $x_1 - 3x_2 + x_3 - 2x_4 = 4$ et $\mathscr{T}$ l'hyperplan d'équations paramétriques

$$\begin{aligned} x_1 &= 1 + \alpha_1 - 2\alpha_2 + 2\alpha_3 \\ x_2 &= 1 + \alpha_1 + \alpha_2 + \alpha_3 \\ x_3 &= 2 - \alpha_2 - \alpha_3 \\ x_4 &= -2 - \alpha_1 - 3\alpha_2 - \alpha_3. \end{aligned}$$

1.12.21 Soit $\mathscr{E}_O$ le vectorialisé relativement à une origine O d'un espace affine $\mathscr{E}$ de dimension finie supérieure à 1 ou infinie. Etudier le lieu géométrique des points de la forme $(\alpha - \beta)P + \beta Q$, où P et Q sont des points distincts de $\mathscr{E}_O$, $\alpha \in (-\infty, 1]$ et $\beta \in \mathbb{R}$.

1.12.22 Soit $\mathscr{E}_O$ le vectorialisé relativement à une origine O d'un espace affine $\mathscr{E}$ de dimension 3. Soit P_1, P_2, P_3 et P_4 quatre points n'appartenant pas à un même plan. Montrer que l'ensemble des combinaisons convexes de P_1, P_2, P_3 et P_4 est le tétraèdre dont les sommets sont ces quatre points.

1.12.23 Soit $\mathscr{E}$ un espace affine muni d'une origine O. Soit $(P_1, P_2, ..., P_k)$ une famille de points de $\mathscr{E}$ et $(\alpha_1, \alpha_2, ..., \alpha_k)$ une famille de nombres telle que $\alpha = \alpha_1 + \alpha_2 + ... + \alpha_k \neq 0$. On appelle le point G de $\mathscr{E}$ défini par la relation $\overrightarrow{OG} = \frac{1}{\alpha}(\alpha_1\overrightarrow{OP_1} + \alpha_2\overrightarrow{OP_2} + ... + \alpha_k\overrightarrow{OP_k})$ *barycentre* des points $P_1, P_2, ..., P_k$ affectés des coefficients respectifs $\alpha_1, \alpha_2, ..., \alpha_k$. Lorsque $\alpha_1 = \alpha_2 = ... = \alpha_k = \frac{1}{k}$, on appelle G *centre de gravité* des points $P_1, P_2, ..., P_k$.

(a) Montrer que la définition de G ne dépend pas du choix de l'origine O.

(b) Situer le centre de gravité des sommets d'un triangle.

1.12.24 Suite de l'exercice précédent. Soit $N_1, N_2, \ldots, N_l$ les ensembles d'une partition de $\{1, 2, \ldots, k\}$. Pour $i = 1, 2, \ldots, l$, on suppose que $\beta_i = \sum_{j \in N_i} \alpha_j$ soit non nul et on désigne par G_i le barycentre des points P_j avec $j \in N_i$, affectés des coefficients α_j.

(a) Montrer que G (défini dans l'exercice précédent) est le barycentre des points $G_1, G_2, \ldots, G_l$ affectés des coefficients respectifs $\beta_1, \beta_2, \ldots, \beta_l$.

(b) A l'aide de (a), situer le centre de gravité des sommets d'un tétraèdre.

1.12.25 Soit $\mathscr{S}$, $\mathscr{S}'$ et $\mathscr{S}''$ trois hyperplans parallèles d'un espace affine de dimension finie supérieure à 1. Soit en outre $\mathscr{D}_1$, $\mathscr{D}_2$, ... des droites du même espace, non parallèles à $\mathscr{S}$. On désigne par P_i, P_i' et P_i'' les points d'intersection respectifs de $\mathscr{S}$, $\mathscr{S}'$, $\mathscr{S}''$ et $\mathscr{D}_i$. Montrer que $\overrightarrow{P_iP_i''} = \alpha\overrightarrow{P_iP_i'}$, avec α indépendant de i (*théorème de Thalès*).

CHAPITRE 2

Espaces vectoriels euclidiens et espaces affines euclidiens

2.1 PRODUIT SCALAIRE DANS L'ESPACE VECTORIEL GÉOMÉTRIQUE

2.1.1 Introduction

Dans le présent chapitre, nous étudierons les notions qui dérivent de la donnée d'une nouvelle opération sur les vecteurs, celle de produit scalaire. Il s'agit de notions métriques telles que la norme ou l'orthogonalité. Dans l'espace vectoriel géométrique V, on peut définir un produit scalaire en partant des idées de longueur et d'angle. Muni de ce produit scalaire, V devient un exemple de ce que nous appellerons *espace vectoriel euclidien* (cf. définition 2.2.3).

Tout au long de cette section, les vecteurs seront les éléments de V.

2.1.2 Longueur, angle

En choisissant une unité de longueur, nous pouvons mesurer l'intensité de chaque flèche, autrement dit, déterminer sa *longueur*. Nous pouvons aussi mesurer l'écart angulaire de deux flèches quelconques d'origine commune (non nécessairement distinctes) en prenant comme unité de mesure par exemple le radian. La mesure de cet écart est alors un nombre compris entre 0 et π, appelé *angle* des deux flèches. Si les deux flèches ont même direction et même sens, leur angle est nul et si elles ont même direction et sens opposé, ce même angle est π.

2.1.3 Norme, angle de deux vecteurs

Les flèches représentatives d'un même vecteur $\mathbf{x}$ ont toutes la même longueur. Nous désignerons cette longueur par $\|\mathbf{x}\|$ et l'appellerons *norme* de $\mathbf{x}$. Il est clair qu'un vecteur est nul si et seulement si sa norme est nulle. Nous dirons qu'un vecteur est *unitaire* si sa norme est 1. Si $\mathbf{x}$ est un vecteur non nul, $\dfrac{1}{\|\mathbf{x}\|}\mathbf{x}$ est un vecteur unitaire, puisque la norme de $\alpha\mathbf{x}$ est $|\alpha|\,\|\mathbf{x}\|$.

Nous appellerons *angle des vecteurs non nuls* $\mathbf{x}$ *et* $\mathbf{y}$ l'angle de deux flèches d'origine commune représentant l'une $\mathbf{x}$ et l'autre $\mathbf{y}$.

2.1.4 Produit scalaire

On appelle *produit scalaire* des vecteurs non nuls $\mathbf{x}$ et $\mathbf{y}$ le nombre

$$\| \mathbf{x} \| \| \mathbf{y} \| \cos\theta, \tag{2.1}$$

où θ est l'angle de $\mathbf{x}$ et $\mathbf{y}$. Si $\mathbf{x}$ ou $\mathbf{y}$ sont nuls, le produit scalaire est nul par convention.

Nous noterons le produit scalaire de $\mathbf{x}$ et $\mathbf{y}$ par $(\mathbf{x} \mid \mathbf{y})$. D'autres symboles couramment utilisés sont $(\mathbf{x}, \mathbf{y})$, $\langle \mathbf{x}, \mathbf{y} \rangle$, $\mathbf{x} \cdot \mathbf{y}$.

On remarquera que $(\mathbf{x} \mid \mathbf{x}) = \| \mathbf{x} \|^2$.

2.1.5 Orthogonalité

On dit que des vecteurs $\mathbf{x}$ et $\mathbf{y}$ sont *orthogonaux* s'ils sont non nuls et leur angle est $\frac{\pi}{2}$, ou si l'un d'entre eux est nul.

En vertu de la définition 2.1.4, $\mathbf{x}$ et $\mathbf{y}$ sont donc orthogonaux si et seulement si $(\mathbf{x} \mid \mathbf{y}) = 0$.

2.1.6 Projection orthogonale

La *projection orthogonale* d'un vecteur non nul $\mathbf{x}$ sur la droite vectorielle engendrée par un vecteur non nul $\mathbf{v}$ est le vecteur

$$\| \mathbf{x} \| \cos\theta \left(\frac{1}{\| \mathbf{v} \|}\mathbf{v}\right),$$

où θ est l'angle de $\mathbf{x}$ et $\mathbf{v}$. Nous la noterons $\mathrm{proj}_{\mathbf{v}}\mathbf{x}$ et l'appellerons aussi projection orthogonale de $\mathbf{x}$ sur $\mathbf{v}$ (fig. 2.1).

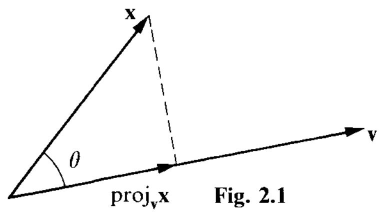


Fig. 2.1

A l'aide du produit scalaire, le vecteur $\mathrm{proj}_{\mathbf{v}}\mathbf{x}$ peut être exprimé autrement: il suffit de substituer $\frac{(\mathbf{x} \mid \mathbf{v})}{\| \mathbf{x} \| \| \mathbf{v} \|}$ à $\cos\theta$ pour obtenir

$$\mathrm{proj}_{\mathbf{v}}\mathbf{x} = \frac{(\mathbf{x} \mid \mathbf{v})}{\| \mathbf{v} \|^2}\mathbf{v} = \frac{(\mathbf{x} \mid \mathbf{v})}{(\mathbf{v} \mid \mathbf{v})}\mathbf{v}. \tag{2.2}$$

Cette expression vaut également dans le cas où $\mathbf{x}$ est nul, à condition d'admettre que la projection orthogonale du vecteur nul est le vecteur nul. La norme de $\mathrm{proj}_{\mathbf{v}}\mathbf{x}$ s'écrit

$$\| \mathrm{proj}_{\mathbf{v}}\mathbf{x} \| = \frac{|(\mathbf{x} \mid \mathbf{v})|}{\| \mathbf{v} \|^2} \| \mathbf{v} \| = \frac{|(\mathbf{x} \mid \mathbf{v})|}{\| \mathbf{v} \|}. \tag{2.3}$$

Si $\mathbf{v}$ est unitaire, (2.2) et (2.3) se simplifient et deviennent

$$\mathrm{proj}_{\mathbf{v}}\mathbf{x} = (\mathbf{x} \mid \mathbf{v})\mathbf{v}, \quad \| \mathrm{proj}_{\mathbf{v}}\mathbf{x} \| = |(\mathbf{x} \mid \mathbf{v})|. \tag{2.4}$$

Par des considérations géométriques élémentaires, il est facile de se rendre compte que

$$\mathrm{proj}_{\mathbf{v}}(\mathbf{x} + \mathbf{y}) = \mathrm{proj}_{\mathbf{v}}\mathbf{x} + \mathrm{proj}_{\mathbf{v}}\mathbf{y}, \quad \mathrm{proj}_{\mathbf{v}}\alpha\mathbf{x} = \alpha\,\mathrm{proj}_{\mathbf{v}}\mathbf{x}. \tag{2.5}$$

2.1.7 Propriétés

Le produit scalaire jouit de trois propriétés qui constitueront le point de départ d'une formulation abstraite de cette opération. Les voici:

(a) $(\mathbf{x} \mid \mathbf{y}) = (\mathbf{y} \mid \mathbf{x})$.
(b) $(\alpha\mathbf{x} + \beta\mathbf{y} \mid \mathbf{z}) = \alpha(\mathbf{x} \mid \mathbf{z}) + \beta(\mathbf{y} \mid \mathbf{z})$.
(c) $(\mathbf{x} \mid \mathbf{x}) > 0$ si $\mathbf{x} \neq \mathbf{0}$.

La seule qui demande une vérification est la deuxième. Si $\mathbf{z}$ est nul, les trois produits scalaires sont nuls et l'égalité est vraie. Si $\mathbf{z}$ n'est pas nul,

$$\mathrm{proj}_{\mathbf{z}}(\alpha\mathbf{x} + \beta\mathbf{y}) = \alpha\,\mathrm{proj}_{\mathbf{z}}\mathbf{x} + \beta\,\mathrm{proj}_{\mathbf{z}}\mathbf{y},$$

d'après (2.5), ce qui entraîne, grâce à (2.2),

$$\frac{(\alpha\mathbf{x} + \beta\mathbf{y} \mid \mathbf{z})}{(\mathbf{z} \mid \mathbf{z})}\mathbf{z} = \alpha\frac{(\mathbf{x} \mid \mathbf{z})}{(\mathbf{z} \mid \mathbf{z})}\mathbf{z} + \beta\frac{(\mathbf{y} \mid \mathbf{z})}{(\mathbf{z} \mid \mathbf{z})}\mathbf{z},$$

d'où la propriété (b) s'ensuit.

2.1.8 Bases orthonormales

Une base $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ de V est dite orthonormale si les vecteurs $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ sont orthogonaux deux à deux et unitaires.

2.1.9 Décomposition suivant une base orthonormale

Par le raisonnement géométrique, on voit facilement que chaque vecteur est la somme de ses projections orthogonales sur les vecteurs d'une base orthonormale, autrement dit, si $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ est une base orthonormale,

$$\mathbf{x} = (\mathbf{x} \mid \mathbf{e}_1)\mathbf{e}_1 + (\mathbf{x} \mid \mathbf{e}_2)\mathbf{e}_2 + (\mathbf{x} \mid \mathbf{e}_3)\mathbf{e}_3. \tag{2.6}$$

Cette décomposition s'obtient également par (b) de 2.1.7. En effet, x_1, x_2, x_3 étant les composantes de $\mathbf{x}$,

$$(\mathbf{x} \mid \mathbf{e}_1) = (x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3 \mid \mathbf{e}_1) = x_1(\mathbf{e}_1 \mid \mathbf{e}_1) + x_2(\mathbf{e}_2 \mid \mathbf{e}_1) + x_3(\mathbf{e}_3 \mid \mathbf{e}_1) = x_1,$$

puisque $(\mathbf{e}_1 \mid \mathbf{e}_1) = 1$ et $(\mathbf{e}_2 \mid \mathbf{e}_1) = (\mathbf{e}_3 \mid \mathbf{e}_1) = 0$; de même,

$$(\mathbf{x} \mid \mathbf{e}_2) = x_2, \quad (\mathbf{x} \mid \mathbf{e}_3) = x_3,$$

d'où la décomposition.

2.1.10 Produit scalaire en fonction des composantes

Soit x_1, x_2, x_3 et y_1, y_2, y_3 les composantes respectives des vecteurs $\mathbf{x}$ et $\mathbf{y}$ dans une base orthonormale $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$. Grâce à (a) et (b) de 2.1.7, le produit scalaire

$$(\mathbf{x} \mid \mathbf{y}) = (x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3 \mid y_1\mathbf{e}_1 + y_2\mathbf{e}_2 + y_3\mathbf{e}_3)$$

se développe en une somme de neuf termes de la forme $x_iy_j(\mathbf{e}_i \mid \mathbf{e}_j)$; or, par l'orthonormalité de la base, $(\mathbf{e}_i \mid \mathbf{e}_j)$ est nul si $i \neq j$ et vaut 1 si $i = j$, ce qui entraîne

$$(\mathbf{x} \mid \mathbf{y}) = x_1y_1 + x_2y_2 + x_3y_3.$$

2.2 ESPACES VECTORIELS EUCLIDIENS

2.2.1 Introduction

Le produit scalaire dans V a été défini au moyen des notions de norme et d'angle. Il jouit des propriétés (a), (b) et (c) (mises en évidence dans 2.1.7) qui en résument les caractères. Pour étendre la notion de produit scalaire aux espaces vectoriels abstraits, nous suivrons le procédé inverse; plus exactement, nous admettrons les trois propriétés comme données de départ, en déduirons les notions de norme, d'orthogonalité et d'angle et aboutirons à un certain nombre de résultats applicables à des situations les plus variées, notamment aux espaces vectoriels fonctionnels. La géométrie aura ainsi laissé la place à l'algèbre, en demeurant toutefois l'inspiratrice des idées et des méthodes utilisées.

2.2.2 Produit scalaire

Soit E un espace vectoriel. On appelle *produit scalaire* dans E toute opération qui fait correspondre à chaque couple $(\mathbf{x}, \mathbf{y})$ de vecteurs de E un nombre, noté $(\mathbf{x} \mid \mathbf{y})$ et appelé *produit scalaire* de $\mathbf{x}$ et $\mathbf{y}$, satisfaisant aux conditions suivantes:

(a) $(\mathbf{x} \mid \mathbf{y}) = (\mathbf{y} \mid \mathbf{x})$ (symétrie ou commutativité).
(b) $(\alpha\mathbf{x} + \beta\mathbf{y} \mid \mathbf{z}) = \alpha(\mathbf{x} \mid \mathbf{z}) + \beta(\mathbf{y} \mid \mathbf{z})$ (linéarité).
(c) $(\mathbf{x} \mid \mathbf{x}) > 0$ si $\mathbf{x} \neq \mathbf{0}$ (positivité).

2.2.3 Espaces vectoriels euclidiens

On appelle *espace vectoriel euclidien* tout espace vectoriel muni d'un produit scalaire.

Il est clair que tout sous-espace vectoriel d'un espace vectoriel euclidien E devient lui-même un espace vectoriel euclidien s'il est muni du produit scalaire hérité de E.

2.2.4 Remarques

(1) La condition (b) exprime la *linéarité à gauche* du produit scalaire (cf. 6.2.2). La *linéarité à droite* découle de l'union de (a) et (b):

$$(\mathbf{x} \mid \beta\mathbf{y} + \gamma\mathbf{z}) = \beta(\mathbf{x} \mid \mathbf{y}) + \gamma(\mathbf{x} \mid \mathbf{z}).$$

(2) En posant $\alpha = \beta = 0$ dans (b), ou $\beta = \gamma = 0$ ci-dessus, nous voyons que le produit scalaire $(\mathbf{x} \mid \mathbf{y})$ est nul si $\mathbf{x}$ ou $\mathbf{y}$ sont nuls.

(3) La linéarité à gauche et à droite du produit scalaire s'étend par récurrence aux combinaisons linéaires de plus de deux termes.

2.2.5 Norme d'un vecteur

Soit E un espace vectoriel euclidien. On appelle *norme* d'un vecteur $\mathbf{x}$ de E, et l'on note $\|\mathbf{x}\|$, le nombre $\sqrt{(\mathbf{x} \mid \mathbf{x})}$.

En vertu de la condition (c), la norme $\|\mathbf{x}\|$ est positive si $\mathbf{x}$ est non nul; d'après (2) de 2.2.4, elle est nulle si $\mathbf{x}$ est nul.

On remarquera que $\sqrt{(\alpha\mathbf{x} \mid \alpha\mathbf{x})} = \sqrt{\alpha^2(\mathbf{x} \mid \mathbf{x})} = |\alpha| \sqrt{(\mathbf{x} \mid \mathbf{x})}$, ce qui montre que

$$\|\alpha\mathbf{x}\| = |\alpha| \|\mathbf{x}\|. \tag{2.7}$$

On dit qu'un vecteur est *unitaire* si sa norme est 1. D'après (2.7), si $\mathbf{x}$ est non nul, $\dfrac{1}{\|\mathbf{x}\|}\mathbf{x}$ est un vecteur unitaire.

2.2.6 Exemples

(1) Vu les propriétés 2.1.7, l'opération définie dans 2.1.4 est un produit scalaire dans V conforme à la définition 2.2.2.

(2) L'opération $(\cdot|\cdot)$ définie par la formule

$$\left(\begin{pmatrix} a_1 \\ a_2 \\ \cdot \\ \cdot \\ \cdot \\ a_n \end{pmatrix} \middle| \begin{pmatrix} b_1 \\ b_2 \\ \cdot \\ \cdot \\ \cdot \\ b_n \end{pmatrix}\right) = a_1b_1 + a_2b_2 + \ldots + a_nb_n \tag{2.8}$$

est un produit scalaire dans $\mathbb{R}^n$ que l'on appelle *produit scalaire canonique*.

Par la suite, sauf mention explicite du contraire, $\mathbb{R}^n$ sera considéré comme muni de ce produit scalaire.

Deux autres exemples de produits scalaires dans $\mathbb{R}^3$ sont définis par les formules

$$\left(\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \middle| \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}\right) = 2a_1b_1 + 3a_2b_2 + 4a_3b_3,$$

$$\left(\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \middle| \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}\right) = 7a_1b_1 + 2a_1b_2 + 2a_2b_1 + 6a_2b_2 + 2a_2b_3 + 2a_3b_2 + 5a_3b_3. \tag{2.9}$$

La preuve que l'opération définie par (2.9) satisfait à la condition (c) n'est pas immédiate. Nous laissons toutefois au lecteur la tâche de l'effectuer, car nous reviendrons sur cette question dans la section 9.2.

(3) Le champ d'application privilégié de la théorie abstraite du produit scalaire est constitué par les espaces vectoriels fonctionnels (cf. 1.3.4 et 1.3.5). On appelle *produit scalaire canonique* dans $C_{[a,\,b]}$ l'opération $(\cdot|\cdot)$ définie par la formule

$$(\mathbf{f} \mid \mathbf{g}) = \int_a^b \mathbf{f}(t)\mathbf{g}(t)\mathrm{d}t. \tag{2.10}$$

Cette opération définit bien un produit scalaire, car les conditions (a) et (b) de la définition 2.2.2 sont manifestement satisfaites et, en outre, l'intégrale

$$\int_a^b \mathbf{f}^2(t)\mathrm{d}t$$

est positive si la fonction continue $\mathbf{f}$ n'est pas identiquement nulle.

Par la suite, sauf mention explicite du contraire, $C_{[a,\,b]}$ sera considéré comme muni du produit scalaire canonique.

(4) Un autre exemple de produit scalaire dans $C_{[a,\,b]}$ est fourni par la formule

$$(\mathbf{f} \mid \mathbf{g}) = \int_a^b \mathbf{f}(t)\mathbf{g}(t)\mathbf{p}(t)\mathrm{d}t, \tag{2.11}$$

où $\mathbf{p}$ est une fonction continue à valeurs positives.

2.3 ORTHOGONALITÉ

2.3.1 Introduction

Il a été remarqué dans 2.1.5 que deux vecteurs de V sont orthogonaux si et seulement si leur produit scalaire est nul. Cette équivalence tiendra lieu de définition de la notion abstraite d'orthogonalité.

Dans la suite de ce chapitre, E désignera un espace vectoriel euclidien. Comme déjà convenu, sauf indication contraire, les vecteurs seront les éléments de E.

2.3.2 Orthogonalité

On dit que les vecteurs $\mathbf{x}$ et $\mathbf{y}$ sont *orthogonaux*, ou que $\mathbf{x}$ est orthogonal à $\mathbf{y}$, si $(\mathbf{x} \mid \mathbf{y}) = 0$. On dit qu'une famille finie ou infinie de vecteurs $(\mathbf{x}_1, \mathbf{x}_2, ...)$ est *orthogonale*, ou que les vecteurs $\mathbf{x}_1$, $\mathbf{x}_2$, ... sont *orthogonaux deux à deux*, si $(\mathbf{x}_i \mid \mathbf{x}_j) = 0$ pour tout couple (i, j) tel que $i \neq j$. Si, de plus, tous les vecteurs $\mathbf{x}_i$ sont unitaires, on dit que la famille est *orthonormale*.

Comme déjà noté dans la remarque 2.2.4, le vecteur nul est orthogonal à tout vecteur. En fait, vu la condition (c) de la définition 2.2.2, il est le seul vecteur qui possède cette propriété.

2.3.3 Exemples

(1) La base canonique de $\mathbb{R}^n$ est manifestement orthonormale.

(2) Considérons les fonctions $\mathbf{c}_0$, $\mathbf{c}_k$ et $\mathbf{s}_k$ $(k > 0)$ de $C_{[-\pi, \pi]}$ définies par

$$\mathbf{c}_0(t) \equiv \frac{1}{\sqrt{2\pi}}, \quad \mathbf{c}_k(t) \equiv \frac{1}{\sqrt{\pi}}\cos(kt), \quad \mathbf{s}_k(t) \equiv \frac{1}{\sqrt{\pi}}\sin(kt). \tag{2.12}$$

La famille infinie $(\mathbf{c}_0, \mathbf{c}_1, \mathbf{s}_1, \mathbf{c}_2, \mathbf{s}_2, ...)$ est appelée *système trigonométrique*. Cette famille est orthonormale, car

$$\int_{-\pi}^{\pi} \cos(kt)\cos(lt)\mathrm{d}t = \tfrac{1}{2}\int_{-\pi}^{\pi} (\cos((k+l)t) + \cos((k-l)t))\mathrm{d}t = \begin{cases} 0 & \text{si } k \neq l, \\ \pi & \text{si } k = l \neq 0, \\ 2\pi & \text{si } k = l = 0; \end{cases}$$

$$\int_{-\pi}^{\pi} \sin(kt)\sin(lt)\mathrm{d}t = \tfrac{1}{2}\int_{-\pi}^{\pi} (\cos((k-l)t) - \cos((k+l)t))\mathrm{d}t = \begin{cases} 0 & \text{si } k \neq l, \\ \pi & \text{si } k = l \neq 0; \end{cases}$$

$$\int_{-\pi}^{\pi} \sin(kt)\cos(lt)\mathrm{d}t = \tfrac{1}{2}\int_{-\pi}^{\pi} (\sin((k+l)t) + \sin((k-l)t))\mathrm{d}t = 0.$$

2.3.4 Orthogonalité à un sous-espace vectoriel

On dit qu'un vecteur $\mathbf{x}$ est orthogonal à un sous-espace vectoriel S de E s'il est orthogonal à tout vecteur de S.

Si $(\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k)$ est une famille génératrice de S, pour qu'un vecteur $\mathbf{x}$ soit orthogonal à S, il suffit qu'il soit orthogonal à tous les vecteurs $\mathbf{v}_i$. En effet, tout vecteur $\mathbf{y}$ de S s'écrit sous la forme $\mathbf{y} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + ... + \alpha_k\mathbf{v}_k$, donc, par la linéarité à droite du produit scalaire,

$$\begin{aligned}(\mathbf{x} \mid \mathbf{y}) &= (\mathbf{x} \mid \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + ... + \alpha_k\mathbf{v}_k)\\ &= \alpha_1(\mathbf{x} \mid \mathbf{v}_1) + \alpha_2(\mathbf{x} \mid \mathbf{v}_2) + ... + \alpha_k(\mathbf{x} \mid \mathbf{v}_k) = 0,\end{aligned}$$

ce qui démontre l'assertion.

2.3.5 Sous-espaces vectoriels orthogonaux

On dit que des sous-espaces vectoriels S et T de E sont orthogonaux, ou que S est orthogonal à T, si chaque vecteur de S est orthogonal à T (ou, ce qui revient au même, chaque vecteur de T est orthogonal à S).

2.3.6 Proposition. Orthogonalité et indépendance linéaire

Toute famille orthogonale finie de vecteurs non nuls est libre. En particulier, toute famille orthonormale finie est libre.

DÉMONSTRATION

Soit $(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k)$ une famille orthogonale de vecteurs non nuls. La relation $\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 + ... + \alpha_k\mathbf{x}_k = \mathbf{0}$ entraîne, pour $i = 1, 2, ..., k$, grâce à la linéarité à gauche du produit scalaire,

$$\begin{aligned}0 &= (\mathbf{0} \mid \mathbf{x}_i) = (\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 + ... + \alpha_k\mathbf{x}_k \mid \mathbf{x}_i)\\ &= \alpha_1(\mathbf{x}_1 \mid \mathbf{x}_i) + \alpha_2(\mathbf{x}_2 \mid \mathbf{x}_i) + ... + \alpha_k(\mathbf{x}_k \mid \mathbf{x}_i) = \alpha_i(\mathbf{x}_i \mid \mathbf{x}_i),\end{aligned}$$

d'où nous concluons que $\alpha_i = 0$, puisque $(\mathbf{x}_i \mid \mathbf{x}_i) \neq 0$.

2.3.7 Théorème de Pythagore

La relation

$$(\mathbf{x} + \mathbf{y} \mid \mathbf{x} + \mathbf{y}) = (\mathbf{x} \mid \mathbf{x}) + (\mathbf{y} \mid \mathbf{y}) + 2(\mathbf{x} \mid \mathbf{y}) \tag{2.13}$$

résulte aisément de la linéarité à gauche et à droite du produit scalaire. Cette relation s'écrit aussi sous la forme

$$\| \mathbf{x} + \mathbf{y} \|^2 = \| \mathbf{x} \|^2 + \| \mathbf{y} \|^2 + 2(\mathbf{x} \mid \mathbf{y}) \tag{2.14}$$

et devient, lorsque $\mathbf{x}$ et $\mathbf{y}$ sont orthogonaux, le *théorème de Pythagore*:

$$\| \mathbf{x} + \mathbf{y} \|^2 = \| \mathbf{x} \|^2 + \| \mathbf{y} \|^2. \tag{2.15}$$

L'extension de (2.15) au cas d'une famille orthogonale $(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k)$ de plus de deux termes se fait par récurrence:

$$\| \mathbf{x}_1 + \mathbf{x}_2 + ... + \mathbf{x}_k \|^2 = \| \mathbf{x}_1 \|^2 + \| \mathbf{x}_2 \|^2 + ... + \| \mathbf{x}_k \|^2. \tag{2.16}$$

2.3.8 Projection orthogonale sur une droite vectorielle

On appelle *projection orthogonale d'un vecteur* $\mathbf{x}$ sur la droite vectorielle engendrée par un vecteur non nul $\mathbf{v}$, ou simplement projection orthogonale de $\mathbf{x}$ sur $\mathbf{v}$, le vecteur

$$\mathrm{proj}_{\mathbf{v}}\mathbf{x} = \frac{(\mathbf{x} \mid \mathbf{v})}{(\mathbf{v} \mid \mathbf{v})}\mathbf{v}. \tag{2.17}$$

On remarquera que cette projection est l'unique vecteur $\alpha\mathbf{v}$ de la droite en question tel que $\mathbf{x} - \alpha\mathbf{v}$ est orthogonal à $\mathbf{v}$. En effet, l'unique solution de l'équation $(\mathbf{x} - \alpha\mathbf{v} \mid \mathbf{v}) = 0$ est

$$\alpha = \frac{(\mathbf{x} \mid \mathbf{v})}{(\mathbf{v} \mid \mathbf{v})}.$$

2.3.9 Orthogonalisation

Le procédé que nous allons présenter est connu sous le nom de *procédé d'orthogonalisation de Gram-Schmidt.* A partir d'une famille libre $(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k)$, il montre par quelles opérations sur les vecteurs $\mathbf{x}_i$ on peut construire une famille orthogonale $(\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k)$ engendrant le même sous-espace vectoriel que $(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k)$.

On pose successivement

$$\begin{aligned}
\mathbf{v}_1 &= \mathbf{x}_1, \\
\mathbf{v}_2 &= \mathbf{x}_2 - \frac{(\mathbf{x}_2 \mid \mathbf{v}_1)}{(\mathbf{v}_1 \mid \mathbf{v}_1)}\mathbf{v}_1, \\
\mathbf{v}_3 &= \mathbf{x}_3 - \frac{(\mathbf{x}_3 \mid \mathbf{v}_1)}{(\mathbf{v}_1 \mid \mathbf{v}_1)}\mathbf{v}_1 - \frac{(\mathbf{x}_3 \mid \mathbf{v}_2)}{(\mathbf{v}_2 \mid \mathbf{v}_2)}\mathbf{v}_2, \\
&\cdots\cdots \\
\mathbf{v}_k &= \mathbf{x}_k - \frac{(\mathbf{x}_k \mid \mathbf{v}_1)}{(\mathbf{v}_1 \mid \mathbf{v}_1)}\mathbf{v}_1 - \frac{(\mathbf{x}_k \mid \mathbf{v}_2)}{(\mathbf{v}_2 \mid \mathbf{v}_2)}\mathbf{v}_2 - \dots - \frac{(\mathbf{x}_k \mid \mathbf{v}_{k-1})}{(\mathbf{v}_{k-1} \mid \mathbf{v}_{k-1})}\mathbf{v}_{k-1}.
\end{aligned} \tag{2.18}$$

En d'autres termes, $\mathbf{v}_i$ $(i > 1)$ est la différence de $\mathbf{x}_i$ et la somme des projections orthogonales de $\mathbf{x}_i$ sur les vecteurs $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_{i-1}$.

Nous allons montrer, par récurrence, que les vecteurs $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k$ sont orthogonaux deux à deux et non nuls. A cet effet, supposons que les vecteurs $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_i$ $(i < k)$ le soient. Alors, pour $j = 1, 2, ..., i$,

$$\begin{aligned}(\mathbf{v}_{i+1} \mid \mathbf{v}_j) &= (\mathbf{x}_{i+1} - \sum_{l=1}^{i} \frac{(\mathbf{x}_{i+1} \mid \mathbf{v}_l)}{(\mathbf{v}_l \mid \mathbf{v}_l)} \mathbf{v}_l \mid \mathbf{v}_j) \\ &= (\mathbf{x}_{i+1} \mid \mathbf{v}_j) - \sum_{l=1}^{i} \frac{(\mathbf{x}_{i+1} \mid \mathbf{v}_l)}{(\mathbf{v}_l \mid \mathbf{v}_l)} (\mathbf{v}_l \mid \mathbf{v}_j) = (\mathbf{x}_{i+1} \mid \mathbf{v}_j) - \frac{(\mathbf{x}_{i+1} \mid \mathbf{v}_j)}{(\mathbf{v}_j \mid \mathbf{v}_j)} (\mathbf{v}_j \mid \mathbf{v}_j) = 0,\end{aligned}$$

ce qui démontre que $\mathbf{v}_{i+1}$ est orthogonal à $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_i$. En outre, $\mathbf{v}_{i+1}$ n'est pas nul, sinon $\mathbf{x}_{i+1}$ serait combinaison linéaire des vecteurs $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_i$ et donc des vecteurs $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_i$, en contradiction avec l'hypothèse d'indépendance linéaire des vecteurs $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k$. Les vecteurs $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_{i+1}$ sont donc orthogonaux deux à deux et non nuls, ce qui établit notre assertion.

On notera que, pour chaque i, les familles $(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_i)$ et $(\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_i)$ engendrent le même sous-espace vectoriel de E.

2.3.10 Exemples

(1) Par le procédé de Gram-Schmidt (2.18), nous allons orthogonaliser le triplet de $\mathbb{R}^4$

$$\left(\begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix} \right).$$

En effectuant les calculs, nous obtenons

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix},$$

$$\mathbf{v}_2 = \begin{pmatrix} 1 \\ -1 \\ 1 \\ 1 \end{pmatrix} - \frac{1}{3}\begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 2 \\ -4 \\ 2 \\ 3 \end{pmatrix} \quad \text{(le facteur } \frac{1}{3} \text{ peut être omis),}$$

$$\mathbf{v}_3 = \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix} - \frac{1}{3}\begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix} - \frac{-7}{33}\begin{pmatrix} 2 \\ -4 \\ 2 \\ 3 \end{pmatrix} = -\frac{2}{11}\begin{pmatrix} -6 \\ 1 \\ 5 \\ 2 \end{pmatrix} \quad \text{(le facteur } -\frac{2}{11} \text{ peut être omis).}$$

(2) Soit $(\mathbf{p}_0, \mathbf{p}_1, \mathbf{p}_2)$ la famille de vecteurs de $C_{[-1,1]}$ définie par $\mathbf{p}_i(t) \equiv t^i$ (cf. (1.8) et exemple (5) de 1.6.7). Le procédé d'orthogonalisation de Gram-Schmidt (2.18) appliqué à cette famille entraîne:

$$\begin{aligned} \mathbf{v}_0(t) &\equiv 1, \\ \mathbf{v}_1(t) &\equiv t - \tfrac{1}{2}\int_{-1}^{1} s\,\mathrm{d}s \equiv t, \\ \mathbf{v}_2(t) &\equiv t^2 - \tfrac{1}{2}\int_{-1}^{1} s^2\mathrm{d}s - \tfrac{3}{2}\int_{-1}^{1} s^3\mathrm{d}s\cdot t \equiv t^2 - \tfrac{1}{3}. \end{aligned}$$

Les polynômes $\mathbf{v}_0$, $\mathbf{v}_1$ et $\mathbf{v}_2$ ainsi obtenus sont les trois premiers termes d'une famille orthogonale infinie connue sous le nom de *système de Legendre*.

2.4 INÉGALITÉS, ANGLES

2.4.1 Inégalité de Cauchy-Schwarz

On appelle *inégalité de Cauchy-Schwarz* l'inégalité, valable pour tout choix des vecteurs $\mathbf{x}$ et $\mathbf{y}$,

$$|(\mathbf{x} \mid \mathbf{y})| \leqslant \|\mathbf{x}\|\,\|\mathbf{y}\|. \tag{2.19}$$

Dans le cas particulier où $\mathbf{x}$ et $\mathbf{y}$ sont des vecteurs de V et le produit scalaire est défini par (2.1), cette inégalité est évidente, puisque $|\cos\theta| \leqslant 1$. Dans le cas général, elle nécessite une preuve. Si $\mathbf{y}$ est nul, les deux membres de (2.19) sont nuls et l'inégalité est en fait une égalité. Si $\mathbf{y}$ n'est pas nul, posons $\mathbf{e} = \dfrac{1}{\|\mathbf{y}\|}\mathbf{y}$ et écrivons

$$\mathbf{x} = (\mathbf{x} \mid \mathbf{e})\mathbf{e} + \mathbf{x} - (\mathbf{x} \mid \mathbf{e})\mathbf{e}. \tag{2.20}$$

Comme $(\mathbf{x} \mid \mathbf{e})\mathbf{e}$ est la projection orthogonale de $\mathbf{x}$ sur $\mathbf{e}$, $\mathbf{x} - (\mathbf{x} \mid \mathbf{e})\mathbf{e}$ est orthogonal à $\mathbf{e}$, donc à $(\mathbf{x} \mid \mathbf{e})\mathbf{e}$, ce qui nous permet d'appliquer (2.15) au second membre de (2.20) et d'obtenir

$$\|\mathbf{x}\|^2 = \|(\mathbf{x} \mid \mathbf{e})\mathbf{e}\|^2 + \|\mathbf{x} - (\mathbf{x} \mid \mathbf{e})\mathbf{e}\|^2 \geqslant (\mathbf{x} \mid \mathbf{e})^2\|\mathbf{e}\|^2 = (\mathbf{x} \mid \mathbf{e})^2. \tag{2.21}$$

Il ne reste alors plus qu'à multiplier les racines carrées des deux membres extrêmes de (2.21) par $\|\mathbf{y}\|$ et à substituer $\mathbf{y}$ à $\|\mathbf{y}\|\mathbf{e}$ pour établir (2.19).

On remarquera que l'inégalité de Cauchy-Schwarz est une égalité si et seulement si $\mathbf{x}$ et $\mathbf{y}$ sont linéairement dépendants. En effet, l'inégalité dans (2.21) est une égalité si et seulement si $\|\mathbf{x} - (\mathbf{x} \mid \mathbf{e})\mathbf{e}\|$ est nul, autrement dit, $\mathbf{x}$ est multiple de $\mathbf{e}$ ou, ce qui revient au même, de $\mathbf{y}$.

2.4.2 Exemples d'application

(1) Lorsque E est $\mathbb{R}^n$, l'inégalité de Cauchy-Schwarz s'écrit

$$\Big|\sum_{i=1}^{n} a_i b_i\Big| \leqslant \sqrt{\sum_{i=1}^{n} a_i^2 \sum_{i=1}^{n} b_i^2}.$$

Dans le cas particulier où $b_1 = b_2 = \ldots = b_n = 1$, elle devient

$$\left(\sum_{i=1}^{n} a_i\right)^2 \leqslant n\sum_{i=1}^{n} a_i^2,$$

ou encore

$$\left(\frac{1}{n}\sum_{i=1}^{n} a_i\right)^2 \leqslant \frac{1}{n}\sum_{i=1}^{n} a_i^2,$$

ce qui montre que le carré de la moyenne arithmétique est inférieur ou égal à la moyenne arithmétique des carrés.

(2) Lorsque E est $C_{[a,\,b]}$, l'inégalité de Cauchy-Schwarz s'écrit

$$\left|\int_a^b \mathbf{f}(t)\mathbf{g}(t)\mathrm{d}t\right| \leqslant \sqrt{\int_a^b \mathbf{f}^2(t)\mathrm{d}t\int_a^b \mathbf{g}^2(t)\mathrm{d}t}.$$

En prenant $\mathbf{g}(t) \equiv 1$ et $|\mathbf{f}|$ à la place de $\mathbf{f}$, nous en déduisons l'inégalité

$$\left(\int_a^b |\mathbf{f}(t)|\mathrm{d}t\right)^2 \leqslant (b-a)\int_a^b \mathbf{f}^2(t)\mathrm{d}t.$$

Par exemple,

$$\left(\int_0^\pi \sqrt{\sin t}\,\mathrm{d}t\right)^2 \leqslant \pi\int_0^\pi \sin t\,\mathrm{d}t = 2\pi,$$

ou encore

$$\int_0^\pi \sqrt{\sin t}\,\mathrm{d}t \leqslant \sqrt{2\pi}.$$

2.4.3 Inégalité triangulaire

En majorant $2(\mathbf{x} \mid \mathbf{y})$ par $2\|\mathbf{x}\|\,\|\mathbf{y}\|$ (inégalité de Cauchy-Schwarz) dans (2.14), nous voyons que

$$\|\mathbf{x} + \mathbf{y}\|^2 \leqslant \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 + 2\|\mathbf{x}\|\,\|\mathbf{y}\| = (\|\mathbf{x}\| + \|\mathbf{y}\|)^2,$$

ce qui entraîne *l'inégalité triangulaire*:

$$\|\mathbf{x} + \mathbf{y}\| \leqslant \|\mathbf{x}\| + \|\mathbf{y}\|. \tag{2.22}$$

En l'appliquant une fois aux vecteurs $\mathbf{x}$ et $\mathbf{y} - \mathbf{x}$ et une autre fois aux vecteurs $\mathbf{x} - \mathbf{y}$ et $\mathbf{y}$, nous déduisons la variante

$$\big|\,\|\mathbf{x}\| - \|\mathbf{y}\|\,\big| \leqslant \|\mathbf{x} - \mathbf{y}\|.$$

2.4.4 Angles. Théorème du cosinus

Supposons que les vecteurs **x** et **y** soient non nuls. En vertu de l'inégalité de Cauchy-Schwarz,

$$-1 \leqslant \frac{(\mathbf{x} \mid \mathbf{y})}{\|\mathbf{x}\| \|\mathbf{y}\|} \leqslant 1.$$

Il existe donc un et un seul nombre θ de l'intervalle $[0, \pi]$ tel que

$$\cos\theta = \frac{(\mathbf{x} \mid \mathbf{y})}{\|\mathbf{x}\| \|\mathbf{y}\|} \tag{2.23}$$

Ce nombre est appelé *angle des vecteurs* **x** *et* **y**. On notera les cas particuliers suivants:

$$\mathbf{x} \text{ et } \mathbf{y} \text{ orthogonaux: } \frac{(\mathbf{x} \mid \mathbf{y})}{\|\mathbf{x}\| \|\mathbf{y}\|} = 0, \text{ ce qui équivaut à } \theta = \frac{\pi}{2};$$

$$\mathbf{x} = \beta\mathbf{y}\text{: } \frac{(\mathbf{x} \mid \mathbf{y})}{\|\mathbf{x}\| \|\mathbf{y}\|} = \frac{\beta\|\mathbf{y}\|^2}{|\beta|\|\mathbf{y}\|^2} = \begin{cases} 1 \text{ si } \beta > 0, \text{ ce qui équivaut à } \theta = 0; \\ -1 \text{ si } \beta < 0, \text{ ce qui équivaut à } \theta = \pi. \end{cases}$$

Si **x** et **y** sont non nuls, la relation (2.14), ou celle qui en dérive par substitution de $-\mathbf{y}$ à **y**, s'écrit, à l'aide de (2.23), sous la forme équivalente

$$\|\mathbf{x} \pm \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 \pm 2\|\mathbf{x}\| \|\mathbf{y}\| \cos\theta, \tag{2.24}$$

où θ est l'angle de **x** et **y**. On appelle (2.24) *théorème du cosinus*.

2.5 ESPACES VECTORIELS EUCLIDIENS DE DIMENSION FINIE

2.5.1 Introduction

Cette section a pour objet l'étude de quelques questions liées à l'existence d'une base orthonormale.

Nous supposerons que l'espace vectoriel euclidien E est de dimension finie non nulle n.

2.5.2 Proposition. Prolongement d'une famille orthonormale

Soit $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_k)$ *une famille de vecteurs orthonormale. Si k est inférieur à la dimension n de E, cette famille se prolonge en une base orthonormale de E.*

DÉMONSTRATION

D'après le théorème 1.7.3, la famille $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_k)$ se prolonge en une base $(\mathbf{e}_1, ..., \mathbf{e}_k, \mathbf{x}_{k+1}, ..., \mathbf{x}_n)$ de E. Appliqué à cette base, le procédé d'orthogonalisation de Gram-Schmidt (2.18) nous fournit une base orthogonale $(\mathbf{e}_1, ..., \mathbf{e}_k, \mathbf{v}_{k+1}, ..., \mathbf{v}_n)$ de E. Posons

$$\mathbf{e}_{k+1} = \frac{1}{\| \mathbf{v}_{k+1} \|}\mathbf{v}_{k+1}, ..., \mathbf{e}_n = \frac{1}{\| \mathbf{v}_n \|}\mathbf{v}_n.$$

La famille $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ est alors une base orthonormale de E.

2.5.3 Corollaire. Existence d'une base orthonormale

Tout espace vectoriel euclidien de dimension finie non nulle admet une base orthonormale.

DÉMONSTRATION

Un tel espace comprend au moins un vecteur non nul $\mathbf{e}_1$, que nous pouvons supposer unitaire. Il suffit alors d'appliquer la proposition 2.5.2 à la famille $(\mathbf{e}_1)$.

2.5.4 Composantes d'un vecteur dans une base orthonormale

Soit x_1 x_2, ..., x_n les composantes d'un vecteur $\mathbf{x}$ dans une base orthonormale $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ de E. Par la linéarité à gauche du produit scalaire, pour $i = 1, 2, ..., n$,

$$\begin{aligned}(\mathbf{x} \mid \mathbf{e}_i) &= (x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + ... + x_n\mathbf{e}_n \mid \mathbf{e}_i)\\ &= x_1(\mathbf{e}_1 \mid \mathbf{e}_i) + x_2(\mathbf{e}_2 \mid \mathbf{e}_i) + ... + x_n(\mathbf{e}_n \mid \mathbf{e}_i)\\ &= x_i(\mathbf{e}_i \mid \mathbf{e}_i) = x_i,\end{aligned}$$

ce qui montre que $x_i\mathbf{e}_i$ est la projection orthogonale de $\mathbf{x}$ sur $\mathbf{e}_i$. La décomposition de $\mathbf{x}$ suivant la base $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ s'écrit donc

$$\mathbf{x} = (\mathbf{x} \mid \mathbf{e}_1)\mathbf{e}_1 + (\mathbf{x} \mid \mathbf{e}_2)\mathbf{e}_2 + ... + (\mathbf{x} \mid \mathbf{e}_n)\mathbf{e}_n. \tag{2.25}$$

Cela généralise la décomposition (2.6).

2.5.5 Expression du produit scalaire en fonction des composantes

Soit $x_1, x_2, ..., x_n$ et $y_1, y_2, ..., y_n$ les composantes respectives de $\mathbf{x}$ et de $\mathbf{y}$ dans une base $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ de E. Par la linéarité à gauche et à droite du produit scalaire,

$$(\mathbf{x} \mid \mathbf{y}) = \left(\sum_{i=1}^{n} x_i\mathbf{e}_i \mid \sum_{j=1}^{n} y_j\mathbf{e}_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} x_iy_j(\mathbf{e}_i \mid \mathbf{e}_j) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_iy_j, \tag{2.26}$$

où $a_{ij} = (\mathbf{e}_i \mid \mathbf{e}_j)$. En particulier,

$$(\mathbf{x} \mid \mathbf{x}) = \| \mathbf{x} \|^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j. \tag{2.27}$$

Nous reviendrons sur les expressions ainsi obtenues dans le chapitre 9. Pour l'instant, limitons-nous à observer que si la base $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ est orthonormale, (2.26) et (2.27) se simplifient et se réduisent aux expressions qui définissent le produit scalaire canonique dans $\mathbb{R}^n$ et le carré de la norme correspondante:

$$(\mathbf{x} \mid \mathbf{y}) = x_1 y_1 + x_2 y_2 + ... + x_n y_n, \tag{2.28}$$

$$(\mathbf{x} \mid \mathbf{x}) = \| \mathbf{x} \|^2 = x_1^2 + x_2^2 + ... + x_n^2. \tag{2.29}$$

2.5.6 Isomorphie de E et $\mathbb{R}^n$

D'après (2.28), lorsque la base de E est orthonormale, l'isomorphisme (1.11) entre E et $\mathbb{R}^n$ conserve le produit scalaire, ce qui signifie que le produit scalaire de $\mathbf{x}$ et $\mathbf{y}$ dans E est égal au produit scalaire dans $\mathbb{R}^n$ des vecteurs-colonnes correspondants (x_i) et (y_i).

2.5.7 Produit scalaire défini par une base

Tout espace vectoriel de dimension finie non nulle n peut être rendu euclidien. Il suffit de choisir une base $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ de cet espace et de définir le produit scalaire par la formule (2.28). Pour ce produit scalaire, la base $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ est orthonormale.

Il est évident que tout espace vectoriel réduit au vecteur nul peut également être considéré comme euclidien.

2.6 PROJECTION ORTHOGONALE ET MEILLEURE APPROXIMATION

2.6.1 Introduction

Dans cette section, nous définirons la notion de projection orthogonale d'un vecteur sur un sous-espace vectoriel et utiliserons cette notion pour calculer la meilleure approximation d'un vecteur par des vecteurs d'un sous-espace vectoriel de dimension finie.

L'espace vectoriel euclidien E sera de nouveau de dimension finie ou infinie.

2.6.2 Complémentaire orthogonal

Soit S un sous-espace vectoriel de E. Désignons par $S^\perp$ l'ensemble des vecteurs orthogonaux à S. Cet ensemble n'est pas vide, car il comprend au moins le vecteur nul. En outre, par la linéarité du produit scalaire, en même temps que $\mathbf{x}$ et $\mathbf{y}$ il comprend toutes les combinaisons linéaires $\alpha\mathbf{x} + \beta\mathbf{y}$. D'après la proposition 1.4.6, $S^\perp$ est donc un sous-espace vectoriel de E. Ce sous-espace est orthogonal à S, selon la définition 2.3.5.

D'après (c) de la définition 2.2.2, $S \cap S^\perp = \{\mathbf{0}\}$. Cependant, E n'est pas, en général, somme directe de S et $S^\perp$ (cf. exercice 2.9.13). Il l'est, toutefois, s'il existe un sous-espace vectoriel T de E, orthogonal à S et tel que E soit somme directe de S et T, car alors T est nécessairement $S^\perp$. En effet, grâce à l'existence d'un tel T, tout vecteur $\mathbf{t}'$ de $S^\perp$ s'écrit sous la forme $\mathbf{t}' = \mathbf{s} + \mathbf{t}$, où $\mathbf{s}$ est un vecteur de S et $\mathbf{t}$ un vecteur de T; cela entraîne que $\mathbf{t}' - \mathbf{t}$ est un vecteur de S orthogonal à S, donc que $\mathbf{t}' - \mathbf{t} = \mathbf{0}$ et, par conséquent, que $\mathbf{t}'$ $(= \mathbf{t})$ appartient à T, ce qui montre que $S^\perp$ est contenu dans T. D'autre part, par définition de $S^\perp$, T est contenu dans $S^\perp$, donc $T = S^\perp$.

Il s'avère ainsi que les conditions suivantes sont équivalentes:

(a) Il existe un sous-espace vectoriel T de E, orthogonal à S et tel que E soit somme directe de S et T.
(b) E est somme directe de S et $S^\perp$.
(c) Tout vecteur $\mathbf{x}$ de E s'écrit (nécessairement de manière unique) sous la forme

$$\mathbf{x} = \mathbf{s} + \mathbf{t}, \tag{2.30}$$

où $\mathbf{s}$ est un vecteur de S et $\mathbf{t}$ un vecteur orthogonal à S.
(d) Pour tout vecteur $\mathbf{x}$ de E, il existe un vecteur $\mathbf{s}$ de S (nécessairement unique) tel que $\mathbf{x} - \mathbf{s}$ est orthogonal à S.

On dit que *S admet un complémentaire orthogonal dans E*, ou que *$S^\perp$ est le complémentaire orthogonal de S dans E*, si l'une des conditions équivalentes (a)–(d) est satisfaite.

D'après ce que nous venons d'établir, si $S^\perp$ est le complémentaire orthogonal de S dans E, alors S est le complémentaire orthogonal de $S^\perp$ dans E et

$$(S^\perp)^\perp = S. \tag{2.31}$$

2.6.3 Projection orthogonale

Soit S un sous-espace vectoriel de E et $\mathbf{x}$ un vecteur de E. Si S admet un complémentaire orthogonal dans E, on appelle le vecteur $\mathbf{s}$ de la décomposition (2.30) *projection orthogonale de* $\mathbf{x}$ *sur* S. Nous désignerons cette projection par $\mathrm{proj}_S\mathbf{x}$.

2.6.4 Proposition. Existence du complémentaire orthogonal d'un sous-espace vectoriel de dimension finie

Tout sous-espace vectoriel de E de dimension finie admet un complémentaire orthogonal dans E.

DÉMONSTRATION

Soit S un sous-espace vectoriel de E de dimension finie. Soit en outre $\mathbf{x}$ un vecteur quelconque de E. Nous allons montrer qu'il existe un vecteur $\mathbf{s}$ de S tel que $\mathbf{x} - \mathbf{s}$ est orthogonal à S. Si S se réduit à $\{\mathbf{0}\}$, $\mathbf{s}$ est le vecteur nul. Si S est de dimension non nulle k, choisissons une base orthogonale quelconque $(\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k)$ de S et posons

$$\mathbf{s} = \mathrm{proj}_{\mathbf{v}_1}\mathbf{x} + ... + \mathrm{proj}_{\mathbf{v}_k}\mathbf{x} = \frac{(\mathbf{x} \mid \mathbf{v}_1)}{(\mathbf{v}_1 \mid \mathbf{v}_1)}\mathbf{v}_1 + ... + \frac{(\mathbf{x} \mid \mathbf{v}_k)}{(\mathbf{v}_k \mid \mathbf{v}_k)}\mathbf{v}_k.$$

Compte tenu de 2.3.8, nous voyons que $\mathbf{x} - \mathbf{s}$ est orthogonal à tous les vecteurs $\mathbf{v}_i$. D'après 2.3.4, $\mathbf{x} - \mathbf{s}$ est donc orthogonal à S.

2.6.5 Calcul de la projection orthogonale

Lorsqu'on doit calculer la projection orthogonale d'un vecteur $\mathbf{x}$ de E sur un sous-espace vectoriel S de E défini par une de ses bases $(\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k)$, il faut avoir en vue les deux cas suivants:

(a) La base $(\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k)$ est orthogonale. Dans ce cas, d'après ce qui précède,

$$\mathrm{proj}_S\mathbf{x} = \frac{(\mathbf{x} \mid \mathbf{v}_1)}{(\mathbf{v}_1 \mid \mathbf{v}_1)}\mathbf{v}_1 + \frac{(\mathbf{x} \mid \mathbf{v}_2)}{(\mathbf{v}_2 \mid \mathbf{v}_2)}\mathbf{v}_2 + ... + \frac{(\mathbf{x} \mid \mathbf{v}_k)}{(\mathbf{v}_k \mid \mathbf{v}_k)}\mathbf{v}_k. \tag{2.32}$$

(b) La base $(\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k)$ n'est pas orthogonale. Dans ce cas, soit on l'orthogonalise par le procédé de Gram-Schmidt et on applique (2.32) à la base orthogonale $(\mathbf{v}'_1, \mathbf{v}'_2, ..., \mathbf{v}'_k)$ obtenue par ce procédé, soit on détermine les coefficients de la combinaison linéaire $\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + ... + \alpha_k\mathbf{v}_k$ de manière que le vecteur $\mathbf{x} - (\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + ... + \alpha_k\mathbf{v}_k)$ soit orthogonal à $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k$; cette condition se traduit par les k équations

$$\alpha_1(\mathbf{v}_1 \mid \mathbf{v}_i) + \alpha_2(\mathbf{v}_2 \mid \mathbf{v}_i) + ... + \alpha_k(\mathbf{v}_k \mid \mathbf{v}_i) = (\mathbf{x} \mid \mathbf{v}_i), \quad i = 1, 2, ..., k,$$

d'où l'on tire les valeurs de $\alpha_1, \alpha_2, ..., \alpha_k$ qui produisent la combinaison linéaire égale à $\mathrm{proj}_S\mathbf{x}$.

2.6.6 Vecteur normal à un hyperplan vectoriel

Soit S un hyperplan vectoriel de E (supposé de dimension finie non nulle). D'après la proposition 2.6.4, S admet un complémentaire orthogonal dans E. D'après (1.12), ce complémentaire est une droite vectorielle. On appelle *vecteur normal* à S tout vecteur non nul de cette droite vectorielle.

2.6.7 Meilleure approximation

Soit S un sous-espace vectoriel de E de dimension finie. On appelle *meilleure approximation* d'un vecteur $\mathbf{x}$ par des vecteurs de S l'unique vecteur de S qui minimise la norme $\| \mathbf{x} - \mathbf{y} \|$, lorsque $\mathbf{y}$ parcourt S. Dans le cas particulier où E est l'espace vectoriel géométrique et S une droite ou un plan vectoriel de cet espace, on voit facilement que la meilleure approximation de $\mathbf{x}$ est la projection orthogonale de $\mathbf{x}$ sur S. Nous allons montrer qu'il en va de même dans le cas général, c'est-à-dire que

$$\| \mathbf{x} - \mathrm{proj}_S \mathbf{x} \| < \| \mathbf{x} - \mathbf{y} \|, \tag{2.33}$$

pour tout vecteur $\mathbf{y}$ de S distinct de $\mathrm{proj}_S \mathbf{x}$. Considérons un tel vecteur $\mathbf{y}$ et écrivons

$$\mathbf{x} - \mathbf{y} = (\mathbf{x} - \mathrm{proj}_S \mathbf{x}) + (\mathrm{proj}_S \mathbf{x} - \mathbf{y}).$$

Comme $\mathbf{x} - \mathrm{proj}_S \mathbf{x}$ est orthogonal à tout vecteur de S, donc à $\mathrm{proj}_S \mathbf{x} - \mathbf{y}$, la relation (2.15) entraîne

$$\| \mathbf{x} - \mathbf{y} \|^2 = \| \mathbf{x} - \mathrm{proj}_S \mathbf{x} \|^2 + \| \mathrm{proj}_S \mathbf{x} - \mathbf{y} \|^2 > \| \mathbf{x} - \mathrm{proj}_S \mathbf{x} \|^2,$$

ce qui établit (2.33).

2.6.8 Exemples d'application

La meilleure approximation est utilisée notamment dans des problèmes de prédiction et d'interpolation. Les deux exemples que nous présentons concernent le système de Legendre et le système trigonométrique.

(1) La meilleure approximation de la fonction $\mathbf{f}(t) \equiv |t|$ de $C_{[-1,1]}$ par des polynômes de degré inférieur ou égal à 2 est, d'après 2.6.7 et (2.32), le polynôme

$$\mathbf{p} = \frac{(\mathbf{f} \mid \mathbf{v}_0)}{(\mathbf{v}_0 \mid \mathbf{v}_0)}\mathbf{v}_0 + \frac{(\mathbf{f} \mid \mathbf{v}_1)}{(\mathbf{v}_1 \mid \mathbf{v}_1)}\mathbf{v}_1 + \frac{(\mathbf{f} \mid \mathbf{v}_2)}{(\mathbf{v}_2 \mid \mathbf{v}_2)}\mathbf{v}_2,$$

où $\mathbf{v}_0$, $\mathbf{v}_1$ et $\mathbf{v}_2$ sont les polynômes orthogonaux obtenus dans l'exemple (2) de 2.3.10. Le calcul des produits scalaires donne:

$$(\mathbf{f} \mid \mathbf{v}_0) = \int_{-1}^{1} |t| \mathrm{d}t = 1, \quad (\mathbf{v}_0 \mid \mathbf{v}_0) = \int_{-1}^{1} 1^2 \mathrm{d}t = 2;$$

$$(\mathbf{f} \mid \mathbf{v}_1) = \int_{-1}^{1} |t| t \mathrm{d}t = 0 \quad \text{(le calcul de } (\mathbf{v}_1 \mid \mathbf{v}_1) \text{ est donc inutile);}$$

$$(\mathbf{f} \mid \mathbf{v}_2) = \int_{-1}^{1} |t|(t^2 - \tfrac{1}{3}) \mathrm{d}t = \tfrac{1}{6}, \quad (\mathbf{v}_2 \mid \mathbf{v}_2) = \int_{-1}^{1} (t^2 - \tfrac{1}{3})^2 \mathrm{d}t = \tfrac{8}{45}.$$

Il en résulte donc que

$$\mathbf{p}(t) \equiv \tfrac{1}{2} + \tfrac{1}{6} \cdot \tfrac{45}{8}(t^2 - \tfrac{1}{3}) \equiv \tfrac{3}{16}(5t^2 + 1) \quad \text{(fig. 2.2).}$$

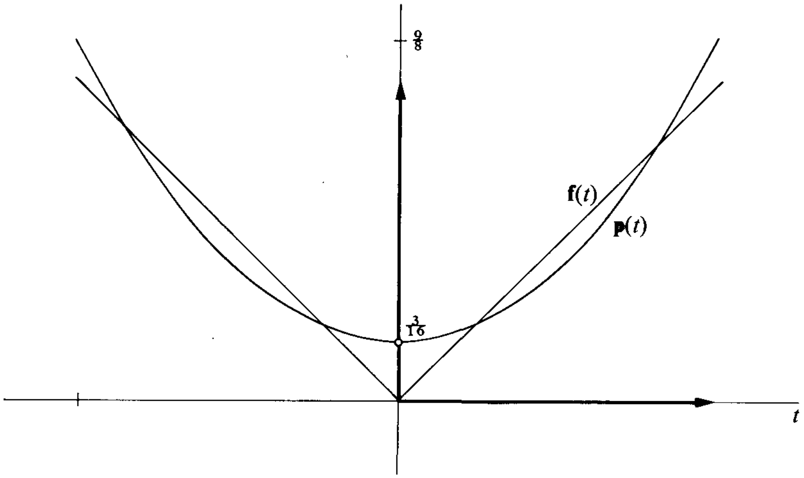


Fig. 2.2

(2) La meilleure approximation de la fonction $\mathbf{f}(t) \equiv t$ de $C_{[-\pi,\pi]}$ par des *polynômes trigonométriques* d'ordre inférieur ou égal à k, c'est-à-dire des combinaisons linéaires des $2k + 1$ premiers termes du système trigonométrique défini dans 2.3.3, est, d'après 2.6.7 et (2.32), le polynôme trigonométrique

$$\mathbf{f}_k = (\mathbf{f} \mid \mathbf{c}_0)\mathbf{c}_0 + (\mathbf{f} \mid \mathbf{c}_1)\mathbf{c}_1 + (\mathbf{f} \mid \mathbf{s}_1)\mathbf{s}_1 + \ldots + (\mathbf{f} \mid \mathbf{c}_k)\mathbf{c}_k + (\mathbf{f} \mid \mathbf{s}_k)\mathbf{s}_k.$$

Les produits scalaires $(\mathbf{f} \mid \mathbf{c}_j)$ $(j \geqslant 0)$ et $(\mathbf{f} \mid \mathbf{s}_j)$ $(j \geqslant 1)$ sont appelés *coefficients de Fourier* de la fonction $\mathbf{f}$. Comme $\mathbf{f}\cdot\mathbf{c}_j$ est une fonction impaire,

$$(\mathbf{f} \mid \mathbf{c}_j) = \frac{1}{\sqrt{(2)\pi}} \int_{-\pi}^{\pi} t\cos(jt)\mathrm{d}t = 0.$$

D'autre part,

$$(\mathbf{f} \mid \mathbf{s}_j) = \frac{1}{\sqrt{\pi}} \int_{-\pi}^{\pi} t\sin(jt)\mathrm{d}t = \frac{2\sqrt{\pi}}{j}(-1)^{j+1},$$

donc le polynôme trigonométrique $\mathbf{f}_k$ s'écrit sous la forme

$$\mathbf{f}_k(t) \equiv \sum_{j=1}^{k} \frac{2}{j}(-1)^{j+1}\sin(jt).$$

L'étude de la convergence de $\mathbf{f}_k(t)$, lorsque k tend vers l'infini, dépasse le cadre de ce livre et sera donc omise. Signalons néanmoins qu'elle a lieu et que la limite est $t = \mathbf{f}(t)$ si $-\pi < t < \pi$ et 0 si $t = \pm\pi$ (car $\sin(j\pi) = 0$). D'une manière générale, on peut démontrer que $\mathbf{f}_k$ *converge en norme* vers $\mathbf{f}$, c'est-à-dire que

$$\lim_{k\to\infty} \| \mathbf{f} - \mathbf{f}_k \| = 0. \tag{2.34}$$

2.7 PRODUIT VECTORIEL ET PRODUIT MIXTE

2.7.1 Introduction

Le produit vectoriel de deux vecteurs est une opération propre à la dimension 3. Pour l'introduire, il faut préalablement orienter l'espace destiné à le recevoir. L'orientation étant définie au moyen de la notion de déterminant, nous commencerons par une brève introduction à l'étude de cette notion. Cette étude sera reprise dans le chapitre 5.

Dans cette section, nous supposerons que l'espace vectoriel euclidien E est de dimension 3.

2.7.2 Déterminants d'ordre 2 et 3

On appelle *déterminant* des vecteurs-colonnes de $\mathbb{R}^2$

$$\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}, \begin{pmatrix} b_1 \\ b_2 \end{pmatrix},$$

et on note

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}, \tag{2.35}$$

le nombre

$$a_1 b_2 - a_2 b_1. \tag{2.36}$$

On appelle *déterminant* des vecteurs-colonnes de $\mathbb{R}^3$

$$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}, \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}, \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix},$$

et on note

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}, \tag{2.37}$$

le nombre

$$a_1 \begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix} - a_2 \begin{vmatrix} b_1 & c_1 \\ b_3 & c_3 \end{vmatrix} + a_3 \begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix} \tag{2.38}$$

$$= a_1 b_2 c_3 + a_2 b_3 c_1 + a_3 b_1 c_2 - a_1 b_3 c_2 - a_2 b_1 c_3 - a_3 b_2 c_1.$$

La fonction qui associe à tout couple de vecteurs-colonnes de $\mathbb{R}^2$ (à tout triplet de vecteurs-colonnes de $\mathbb{R}^3$) son déterminant est appelée déterminant d'ordre 2 (d'ordre 3).

2.7.3 Propriétés des déterminants

Les propriétés les plus importantes des déterminants seront établies dans la section 5.1. En anticipant sur cette section, nous énonçons ici les deux propriétés qui nous servent à étudier le produit vectoriel et le produit mixte.

(a) Le déterminant est multiplié par -1 si l'un de ses vecteurs-colonnes est remplacé par son opposé ou si deux de ses vecteurs-colonnes sont échangés.
(b) Le déterminant est non nul si et seulement si ses vecteurs-colonnes sont linéairement indépendants.

Ces deux propriétés sont évidentes dans le cas du déterminant d'ordre 2. La propriété (a) l'est également dans le cas du déterminant d'ordre 3. La vérification de la propriété (b) dans ce même cas est laissée comme exercice. S'il le désire, le lecteur pourra se reporter au paragraphe 5.2.2, où cette vérification est faite en toute généralité.

2.7.4 Déterminants de passage

On appelle *déterminant de passage* d'une base $(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$ à une base $(\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3)$ le déterminant

$$\begin{vmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{vmatrix}, \tag{2.39}$$

où p_{1j}, p_{2j}, p_{3j} sont les composantes du vecteur $\mathbf{y}_j$ dans la base $(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$.

2.7.5 Orientation de E

On définit l'orientation de l'espace vectoriel E par le choix d'une de ses bases. Ce choix étant fait, on dit que E est *orienté* ou *doté d'une orientation.* Soit $(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$ la base définissant l'orientation de E. On dit alors qu'une base $(\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3)$ est *directe* ou *positivement orientée* si le déterminant de passage (2.39) est positif et *indirecte* ou *négativement orientée* si ce même déterminant est négatif.

2.7.6 Remarque

Nous verrons dans 6.6.10 que le déterminant de passage entre bases de même orientation est positif, tandis qu'entre bases d'orientations différentes ce déterminant est négatif. Par les déterminants de passage, les bases de E peuvent donc être rangées en deux classes définissant chacune l'une des deux orientations possibles de E.

2.7.7 Effet d'une transposition et d'une permutation cyclique

Les deux assertions suivantes découlent directement de (a) de 2.7.3.

(1) L'orientation d'une base $(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$ change si l'un des vecteurs $\mathbf{x}_i$ est remplacé par son opposé, ou si deux vecteurs $\mathbf{x}_i$ et $\mathbf{x}_j$ sont échangés.

(2) L'orientation d'une base $(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$ reste inchangée si les vecteurs $\mathbf{x}_i$ sont permutés cycliquement, c'est-à-dire si $\mathbf{x}_1$ prend la place de $\mathbf{x}_2$, $\mathbf{x}_2$ celle de $\mathbf{x}_3$ et $\mathbf{x}_3$ celle de $\mathbf{x}_1$ (ou, ce qui revient au même, si $\mathbf{x}_1$ et $\mathbf{x}_3$ et ensuite $\mathbf{x}_2$ et $\mathbf{x}_1$ sont échangés).

A l'aide de (1) et (2), un examen des six dispositions possibles de $\mathbf{x}_1$, $\mathbf{x}_2$ et $\mathbf{x}_3$ nous permet de conclure que les trois triplets

$$(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3), \quad (\mathbf{x}_3, \mathbf{x}_1, \mathbf{x}_2), \quad (\mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_1)$$

sont de même orientation, tandis que les trois autres

$$(\mathbf{x}_2, \mathbf{x}_1, \mathbf{x}_3), \quad (\mathbf{x}_3, \mathbf{x}_2, \mathbf{x}_1), \quad (\mathbf{x}_1, \mathbf{x}_3, \mathbf{x}_2)$$

sont d'orientation opposée à celle des trois précédents.

2.7.8 Choix d'une base

Dans la suite de cette section, $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ est une base orthonormale définissant l'orientation de E.

2.7.9 Produit vectoriel

Soit x_1, x_2, x_3 et y_1, y_2, y_3 les composantes respectives des vecteurs $\mathbf{x}$ et $\mathbf{y}$ dans la base $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$. On appelle *produit vectoriel* de $\mathbf{x}$ et $\mathbf{y}$, et on note $\mathbf{x} \times \mathbf{y}$, le vecteur

$$(x_2y_3 - x_3y_2)\mathbf{e}_1 + (x_3y_1 - x_1y_3)\mathbf{e}_2 + (x_1y_2 - x_2y_1)\mathbf{e}_3. \tag{2.40}$$

Cette définition dépend du choix de la base $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$, mais nous verrons dans 2.7.13 qu'elle n'en dépend pas si ce choix est fait parmi les bases orthonormales directes.

La i-ième composante de $\mathbf{x} \times \mathbf{y}$ est le déterminant des deux colonnes

$$\begin{matrix} x_1 & y_1 \\ x_2 & y_2 \\ x_3 & y_3 \end{matrix} \tag{2.41}$$

privées de leur i-ième terme, le deuxième déterminant étant cependant pris avec le signe $-$:

$$\mathbf{x} \times \mathbf{y} = \begin{vmatrix} x_2 & y_2 \\ x_3 & y_3 \end{vmatrix}\mathbf{e}_1 - \begin{vmatrix} x_1 & y_1 \\ x_3 & y_3 \end{vmatrix}\mathbf{e}_2 + \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix}\mathbf{e}_3. \tag{2.42}$$

Cas particuliers:

$$\mathbf{e}_1 \times \mathbf{e}_1 = \mathbf{e}_2 \times \mathbf{e}_2 = \mathbf{e}_3 \times \mathbf{e}_3 = \mathbf{0},$$
$$\mathbf{e}_1 \times \mathbf{e}_2 = \mathbf{e}_3, \quad \mathbf{e}_2 \times \mathbf{e}_3 = \mathbf{e}_1, \quad \mathbf{e}_3 \times \mathbf{e}_1 = \mathbf{e}_2.$$

2.7.10 Propriétés du produit vectoriel

Le produit vectoriel jouit des propriétés suivantes:

(a) $\mathbf{x} \times \mathbf{y} = -(\mathbf{y} \times \mathbf{x})$ (antisymétrie).
(b) $(\alpha\mathbf{x} + \beta\mathbf{y}) \times \mathbf{z} = \alpha(\mathbf{x} \times \mathbf{z}) + \beta(\mathbf{y} \times \mathbf{z})$ (linéarité).
(c) $\mathbf{x} \times \mathbf{y} \neq \mathbf{0}$ si et seulement si $\mathbf{x}$ et $\mathbf{y}$ sont linéairement indépendants.

Les deux premières découlent directement de la définition. Nous établirons la troisième en montrant que $\mathbf{x} \times \mathbf{y}$ est nul si et seulement si $\mathbf{x}$ et $\mathbf{y}$ sont linéairement dépendants. Sous cette condition, l'une des deux colonnes dans (2.41) est multiple de l'autre, donc les trois déterminants dans (2.42) sont nuls, ce qui entraîne que le produit vectoriel $\mathbf{x} \times \mathbf{y}$ est nul. Réciproquement, supposons que $\mathbf{x} \times \mathbf{y}$ soit nul. Alors les trois déterminants dans (2.42) sont nuls, c'est-à-dire

$$x_2y_3 = x_3y_2, \quad x_3y_1 = x_1y_3, \quad x_1y_2 = x_2y_1. \tag{2.43}$$

Si tous les y_i sont nuls, les vecteurs $\mathbf{x}$ et $\mathbf{y}$ sont linéairement dépendants. Si un des y_i, par exemple y_3, est non nul, il existe α tel que $x_3 = \alpha y_3$, ce qui entraîne, par substitution de αy_3 à x_3 dans (2.43),

$$x_2y_3 = \alpha y_3y_2, \quad \alpha y_3y_1 = x_1y_3$$

et donc

$$x_2 = \alpha y_2, \quad x_1 = \alpha y_1.$$

La première colonne dans (2.41) est donc le produit de α par la deuxième, ce qui implique la dépendance linéaire de $\mathbf{x}$ et $\mathbf{y}$.

La propriété (b) exprime la *linéarité à gauche* du produit vectoriel (cf. 6.2.2). Cette propriété, jointe à l'antisymétrie, entraîne la *linéarité à droite*:

$$\mathbf{x} \times (\beta\mathbf{y} + \gamma\mathbf{z}) = \beta(\mathbf{x} \times \mathbf{y}) + \gamma(\mathbf{x} \times \mathbf{z}). \tag{2.44}$$

On remarquera, en particulier, que

$$\alpha(\mathbf{x} \times \mathbf{y}) = (\alpha\mathbf{x}) \times \mathbf{y} = \mathbf{x} \times (\alpha\mathbf{y}), \tag{2.45}$$

ce qui nous permet d'omettre les parenthèses et d'écrire $\alpha\mathbf{x} \times \mathbf{y}$ sans danger d'ambiguïté.

2.7.11 Propriétés métriques du produit vectoriel

Les vecteurs $\mathbf{x}$ et $\mathbf{y}$ étant supposés non nuls, le produit vectoriel de $\mathbf{x}$ et $\mathbf{y}$ est

(a) orthogonal à $\mathbf{x}$ et à $\mathbf{y}$,
(b) de norme $\|\mathbf{x}\|\,\|\mathbf{y}\|\sin\theta$, où θ est l'angle de $\mathbf{x}$ et $\mathbf{y}$.

En effet, d'après (2.42), (2.28), 2.7.2 et (b) de 2.7.3,

$$(\mathbf{x} \mid \mathbf{x} \times \mathbf{y}) = x_1 \begin{vmatrix} x_2 & y_2 \\ x_3 & y_3 \end{vmatrix} - x_2 \begin{vmatrix} x_1 & y_1 \\ x_3 & y_3 \end{vmatrix} + x_3 \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} = \begin{vmatrix} x_1 & x_1 & y_1 \\ x_2 & x_2 & y_2 \\ x_3 & x_3 & y_3 \end{vmatrix} = 0,$$

ce qui montre que $\mathbf{x} \times \mathbf{y}$ est orthogonal à $\mathbf{x}$ et donc aussi à $\mathbf{y}$, puisque $\mathbf{x} \times \mathbf{y} = -\mathbf{y} \times \mathbf{x}$. D'autre part, d'après (2.40), (2.29), (2.28) et (2.23),

$$\begin{aligned} \| \mathbf{x} \times \mathbf{y} \|^2 &= (x_2y_3 - x_3y_2)^2 + (x_3y_1 - x_1y_3)^2 + (x_1y_2 - x_2y_1)^2 \\ &= (x_1^2 + x_2^2 + x_3^2)(y_1^2 + y_2^2 + y_3^2) - (x_1y_1 + x_2y_2 + x_3y_3)^2 \\ &= \| \mathbf{x} \|^2 \| \mathbf{y} \|^2 \sin^2\theta, \end{aligned}$$

ce qui démontre (b).

On remarquera que dans le cas où E est l'espace vectoriel géométrique V^3, la norme $\| \mathbf{x} \| \| \mathbf{y} \| \sin\theta$ du produit vectoriel représente l'aire du parallélogramme construit sur des représentants de $\mathbf{x}$ et $\mathbf{y}$ d'origine commune (fig. 2.3).

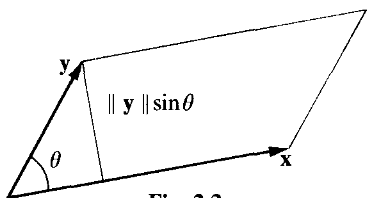


Fig. 2.3

2.7.12 Orientation

Si $\mathbf{x}$ et $\mathbf{y}$ sont linéairement indépendants, le triplet $(\mathbf{x} \times \mathbf{y}, \mathbf{x}, \mathbf{y})$ et donc aussi le triplet $(\mathbf{x}, \mathbf{y}, \mathbf{x} \times \mathbf{y})$ sont directs.

En effet, z_1, z_2, z_3 étant les composantes de $\mathbf{x} \times \mathbf{y}$ (dans la base $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$), le déterminant de passage de $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ à $(\mathbf{x} \times \mathbf{y}, \mathbf{x}, \mathbf{y})$ s'écrit

$$\begin{vmatrix} z_1 & x_1 & y_1 \\ z_2 & x_2 & y_2 \\ z_3 & x_3 & y_3 \end{vmatrix} = z_1^2 + z_2^2 + z_3^2. \tag{2.46}$$

Ce déterminant est donc positif, puisqu'au moins un des z_i n'est pas nul, d'après (c) de 2.7.10.

2.7.13 Indépendance du produit vectoriel du choix de la base orthonormale directe

L'expression du produit vectoriel (2.40) est invariante par changement de base orthonormale directe. Soit en effet $(\mathbf{e}_1', \mathbf{e}_2', \mathbf{e}_3')$ une telle base. Selon les propriétés métriques 2.7.11, $\mathbf{e}_1' \times \mathbf{e}_2'$ est soit $\mathbf{e}_3'$, soit $-\mathbf{e}_3'$. D'après 2.7.12, le triplet $(\mathbf{e}_1', \mathbf{e}_2', \mathbf{e}_1' \times \mathbf{e}_2')$ est direct, donc $\mathbf{e}_1' \times \mathbf{e}_2' = \mathbf{e}_3'$, vu l'assertion (1) de 2.7.7. Par le même raisonnement, compte tenu de l'assertion (2) de 2.7.7, nous voyons que $\mathbf{e}'_2 \times \mathbf{e}'_3 = \mathbf{e}'_1$ et $\mathbf{e}'_3 \times \mathbf{e}'_1 = \mathbf{e}'_2$. Soit maintenant x'_1, x'_2, x'_3 et y'_1, y'_2, y'_3 les composantes respectives de $\mathbf{x}$ et $\mathbf{y}$ dans la base $(\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3)$. Grâce à (a) et (b) de 2.7.10, nous pouvons écrire

$$\begin{aligned}\mathbf{x} \times \mathbf{y} &= \left(\sum_{i=1}^{3} x'_i \mathbf{e}'_i\right) \times \left(\sum_{j=1}^{3} y'_j \mathbf{e}'_j\right) = \sum_{i=1}^{3}\sum_{j=1}^{3} x'_i y'_j \mathbf{e}'_i \times \mathbf{e}'_j \\ &= (x'_2 y'_3 - x'_3 y'_2)\mathbf{e}'_1 + (x'_3 y'_1 - x'_1 y'_3)\mathbf{e}'_2 + (x'_1 y'_2 - x'_2 y'_1)\mathbf{e}'_3,\end{aligned}$$

ce qui montre l'invariance de (2.40).

2.7.14 Produit mixte

On appelle *produit mixte* des vecteurs $\mathbf{x}$, $\mathbf{y}$ et $\mathbf{z}$ le double produit $(\mathbf{x} \mid \mathbf{y} \times \mathbf{z})$.

D'après (2.23), si $\mathbf{x}$ et $\mathbf{y} \times \mathbf{z}$ sont non nuls, le produit mixte de $\mathbf{x}$, $\mathbf{y}$ et $\mathbf{z}$ s'écrit aussi sous la forme

$$\| \mathbf{x} \| \, \| \mathbf{y} \times \mathbf{z} \| \cos\theta, \tag{2.47}$$

où θ est l'angle de $\mathbf{x}$ et $\mathbf{y} \times \mathbf{z}$.

On remarquera que dans le cas où E est l'espace vectoriel géométrique V^3, le produit mixte, écrit sous la forme (2.47), symbolise le volume (orienté) du parallélépipède construit sur des représentants de $\mathbf{x}$, $\mathbf{y}$ et $\mathbf{z}$ d'origine commune (fig. 2.4).

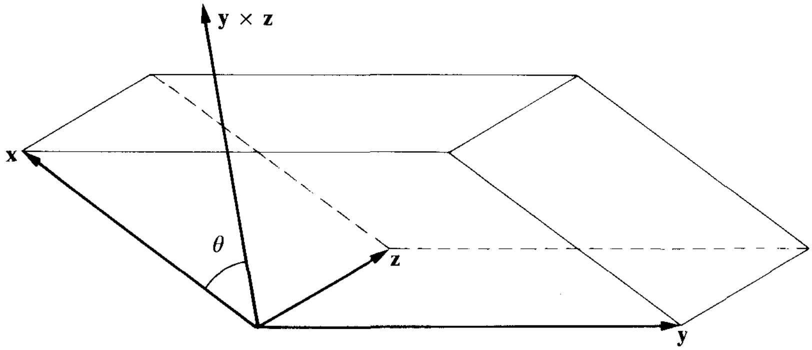


Fig. 2.4

2.7.15 Produit mixte en fonction des composantes

Soit x_1, x_2, x_3, y_1, y_2, y_3 et z_1, z_2, z_3 les composantes respectives de $\mathbf{x}$, $\mathbf{y}$ et $\mathbf{z}$ dans une base orthonormale directe. D'après (2.42), (2.28) et 2.7.2,

$$(\mathbf{x} \mid \mathbf{y} \times \mathbf{z}) = x_1 \begin{vmatrix} y_2 & z_2 \\ y_3 & z_3 \end{vmatrix} - x_2 \begin{vmatrix} y_1 & z_1 \\ y_3 & z_3 \end{vmatrix} + x_3 \begin{vmatrix} y_1 & z_1 \\ y_2 & z_2 \end{vmatrix} = \begin{vmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{vmatrix}, \tag{2.48}$$

donc l'expression du produit mixte en fonction des composantes est le déterminant des vecteurs-colonnes des composantes. Il s'ensuit, grâce à la commutativité du produit scalaire et à (a) de 2.7.3, que

$$(\mathbf{x} \times \mathbf{y} \mid \mathbf{z}) = (\mathbf{z} \mid \mathbf{x} \times \mathbf{y}) = (\mathbf{x} \mid \mathbf{y} \times \mathbf{z}), \tag{2.49}$$

ce qui montre que l'ordre dans lequel les deux produits sont effectués est indifférent et justifie la notation suivante:

2.7.16 Notation

Désormais le produit mixte des vecteurs $\mathbf{x}$, $\mathbf{y}$ et $\mathbf{z}$ sera noté $[\mathbf{x}, \mathbf{y}, \mathbf{z}]$.

2.7.17 Propriétés du produit mixte

Le produit mixte jouit des propriétés suivantes:

(a) $[\mathbf{x}, \mathbf{y}, \mathbf{z}] = [\mathbf{z}, \mathbf{x}, \mathbf{y}] = [\mathbf{y}, \mathbf{z}, \mathbf{x}] = -[\mathbf{y}, \mathbf{x}, \mathbf{z}] = -[\mathbf{z}, \mathbf{y}, \mathbf{x}] = -[\mathbf{x}, \mathbf{z}, \mathbf{y}]$.
(b) $[\alpha\mathbf{x} + \beta\mathbf{y}, \mathbf{z}, \mathbf{v}] = \alpha[\mathbf{x}, \mathbf{z}, \mathbf{v}] + \beta[\mathbf{y}, \mathbf{z}, \mathbf{v}]$.
(c) $[\mathbf{x}, \mathbf{y}, \mathbf{z}] \neq 0$ si et seulement si $\mathbf{x}$, $\mathbf{y}$ et $\mathbf{z}$ sont linéairement indépendants.

La première et la troisième de ces propriétés découlent de la relation (2.48) jointe respectivement à (a) et à (b) de 2.7.3. La deuxième résulte de la définition 2.7.14 jointe à (b) de 2.2.2.

En particulier, la propriété (a) montre que le produit mixte est invariant par permutation cyclique.

Par (2.48), la propriété (b) résulte aussi immédiatement de (2.38). Elle exprime la *linéarité à gauche* du produit mixte (cf. 6.2.2). La *linéarité au centre* et *à droite* suit de l'union de (a) et (b).

2.7.18 Quelques identités vectorielles

Voici trois identités vectorielles d'utilité pratique:

$$\mathbf{x} \times (\mathbf{y} \times \mathbf{z}) = (\mathbf{x} \mid \mathbf{z})\mathbf{y} - (\mathbf{x} \mid \mathbf{y})\mathbf{z}, \tag{2.50}$$
$$(\mathbf{x} \times \mathbf{y} \mid \mathbf{z} \times \mathbf{v}) = (\mathbf{x} \mid \mathbf{z})(\mathbf{y} \mid \mathbf{v}) - (\mathbf{x} \mid \mathbf{v})(\mathbf{y} \mid \mathbf{z}), \tag{2.51}$$
$$(\mathbf{x} \times \mathbf{y}) \times (\mathbf{z} \times \mathbf{v}) = [\mathbf{x}, \mathbf{y}, \mathbf{v}]\mathbf{z} - [\mathbf{x}, \mathbf{y}, \mathbf{z}]\mathbf{v}. \tag{2.52}$$

Le lecteur pourra vérifier aisément (2.50) au moyen des composantes dans une base orthonormale directe. Par (2.49), la première croix et la barre peuvent être échangées dans le premier membre de (2.51); grâce à (2.50), il s'ensuit que

$$\begin{aligned}(\mathbf{x} \times \mathbf{y} \mid \mathbf{z} \times \mathbf{v}) &= (\mathbf{x} \mid \mathbf{y} \times (\mathbf{z} \times \mathbf{v}))\\ &= (\mathbf{x} \mid (\mathbf{y} \mid \mathbf{v})\mathbf{z} - (\mathbf{y} \mid \mathbf{z})\mathbf{v}) = (\mathbf{x} \mid \mathbf{z})(\mathbf{y} \mid \mathbf{v}) - (\mathbf{x} \mid \mathbf{v})(\mathbf{y} \mid \mathbf{z}),\end{aligned}$$

ce qui établit (2.51). Pour obtenir (2.52), il suffit d'appliquer (2.50) au double produit vectoriel $\mathbf{u} \times (\mathbf{z} \times \mathbf{v})$, où $\mathbf{u} = \mathbf{x} \times \mathbf{y}$, car cela donne, compte tenu de (2.49),

$$(\mathbf{x} \times \mathbf{y}) \times (\mathbf{z} \times \mathbf{v}) = (\mathbf{x} \times \mathbf{y} \mid \mathbf{v})\mathbf{z} - (\mathbf{x} \times \mathbf{y} \mid \mathbf{z})\mathbf{v} = [\mathbf{x}, \mathbf{y}, \mathbf{v}]\mathbf{z} - [\mathbf{x}, \mathbf{y}, \mathbf{z}]\mathbf{v},$$

ce qui prouve (2.52).

L'identité (2.50) montre, en particulier, que le produit vectoriel n'est pas associatif, c'est-à-dire qu'en général

$$\mathbf{x} \times (\mathbf{y} \times \mathbf{z}) \neq (\mathbf{x} \times \mathbf{y}) \times \mathbf{z}.$$

2.8 ESPACES AFFINES EUCLIDIENS

2.8.1 Introduction

Le concept d'espace affine défini dans la section 1.9 met en relation un espace ponctuel avec la structure linéaire d'un espace vectoriel. Dans un espace affine, des notions comme la distance ou l'orthogonalité n'ont pas de sens. Pour leur en donner un, il faut munir l'espace vectoriel directeur d'un produit scalaire. Le but de cette section est précisément d'examiner les conséquences qu'implique l'introduction d'une telle opération.

2.8.2 Espaces affines euclidiens

Soit $\mathscr{E}$ un espace affine de direction E. On dit que $\mathscr{E}$ est un *espace affine euclidien* si E est un espace vectoriel euclidien.

En vertu de 1.10.5 et 2.2.3, tout sous-espace affine d'un espace affine euclidien est lui-même un espace affine euclidien.

2.8.3 Exemples

(1) L'espace $\mathscr{G}$ de la géométrie élémentaire est un espace affine euclidien, car l'espace directeur V qui lui est associé est un espace vectoriel euclidien.

(2) Dans l'exemple (2) de 1.9.7, il est montré que tout espace vectoriel E peut être considéré comme un espace affine d'espace directeur E lui-même. Cet espace affine est euclidien si E est euclidien.

(3) Tout espace affine de dimension finie peut être rendu euclidien par le procédé indiqué dans 2.5.7.

(4) L'ensemble des solutions du système (1.15) est un sous-espace affine euclidien de $\mathbb{R}^3$.

2.8.4 Distance et ses propriétés

Dans la suite de cette section, $\mathscr{E}$ désignera un espace affine euclidien. Le lecteur désireux de se familiariser de nouveau avec la notation est prié de se reporter à la section 1.9.

On appelle *distance des points P et Q* le nombre $\delta(P, Q)$ défini par

$$\delta(P, Q) = \| \overrightarrow{PQ} \|. \tag{2.53}$$

Des règles (1), (2) et (3) de 1.9.5 et des propriétés de la norme vues dans 2.2.5, il résulte que $\delta(P, Q) = \delta(Q, P)$, que $\delta(P, Q)$ est positif si les points P et Q sont distincts et nul s'ils sont confondus. De la condition (b) de la définition 1.9.2 et de (2.22), il résulte en outre que tout triplet de points (P, Q, R) vérifie l'inégalité triangulaire:

$$\delta(P, R) \leqslant \delta(P, Q) + \delta(Q, R). \tag{2.54}$$

Cette inégalité est une égalité si et seulement si P, Q et R sont alignés, c'est-à-dire appartiennent à une même droite.

2.8.5 Sous-espaces affines orthogonaux

On dit que deux sous-espaces affines de $\mathscr{E}$ sont orthogonaux, ou que l'un d'entre eux est orthogonal à l'autre, si leurs espaces directeurs sont orthogonaux.

2.8.6 Vecteur normal à un hyperplan

Lorsque $\mathscr{E}$ est de dimension finie non nulle, on appelle *vecteur normal* à un hyperplan $\mathscr{S}$ de $\mathscr{E}$ tout vecteur normal à la direction S de $\mathscr{S}$. Ainsi, une droite orthogonale à un hyperplan $\mathscr{S}$ a pour vecteur directeur un vecteur normal à $\mathscr{S}$.

2.8.7 Projection orthogonale d'un point

Soit $\mathscr{S}$ un sous-espace affine de $\mathscr{E}$ de direction S et de dimension finie. Soit en outre P un point de $\mathscr{E}$. Nous allons démontrer qu'il existe un point R de $\mathscr{S}$ et un seul tel que $\overrightarrow{RP}$ soit orthogonal à S. Ce point R s'appelle *projection orthogonale de P sur* $\mathscr{S}$ et se note $\operatorname{proj}_{\mathscr{S}} P$.

Soit P_0 un point quelconque de $\mathscr{S}$. Posons $R = P_0 + \operatorname{proj}_S \overrightarrow{P_0P}$ et montrons que R est la projection cherchée (fig. 2.5). D'après (1.16), R est un point de $\mathscr{S}$; en outre, d'après (d) de 2.6.2, le vecteur

$$\overrightarrow{P_0P} - \operatorname{proj}_S \overrightarrow{P_0P} = \overrightarrow{P_0P} - \overrightarrow{P_0R} = \overrightarrow{RP}$$

est orthogonal à S. D'autre part, si Q est un point de $\mathscr{S}$ tel que $\overrightarrow{QP}$ est orthogonal à S, en écrivant

$$\overrightarrow{QR} = \overrightarrow{QP} - \overrightarrow{RP},$$

nous voyons que $\overrightarrow{QR}$ est un vecteur de S orthogonal à S, donc à lui-même. Par conséquent, $\overrightarrow{QR} = \mathbf{0}$, ce qui entraîne $Q = R$.

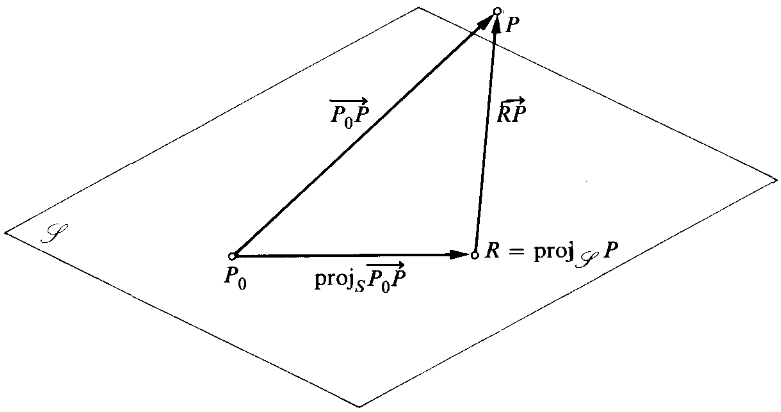


Fig. 2.5

2.8.8 Perpendiculaire par un point à un sous-espace affine

Soit $\mathscr{S}$ un sous-espace affine de $\mathscr{E}$ de dimension finie. Soit en outre P un point de $\mathscr{E}$ n'appartenant pas à $\mathscr{S}$. On appelle *perpendiculaire* par P à $\mathscr{S}$ la droite déterminée par P et $\mathrm{proj}_{\mathscr{S}} P$.

2.8.9 Distance d'un point à un sous-espace affine

Soit $\mathscr{S}$ un sous-espace affine de $\mathscr{E}$ de dimension finie. On appelle *distance d'un point P à $\mathscr{S}$*, et l'on note $\mathrm{dist}(P, \mathscr{S})$, le minimum de $\delta(P, Q)$, Q parcourant $\mathscr{S}$. Nous nous proposons de démontrer que

$$\mathrm{dist}(P, \mathscr{S}) = \delta(P, \mathrm{proj}_{\mathscr{S}} P). \tag{2.55}$$

Soit Q un point de $\mathscr{S}$ distinct de $\mathrm{proj}_{\mathscr{S}} P$. Comme $\overrightarrow{(\mathrm{proj}_{\mathscr{S}} P)Q}$ est orthogonal à $\overrightarrow{P(\mathrm{proj}_{\mathscr{S}} P)}$, par (2.15) nous voyons que

$$\begin{aligned}\delta^2(P, Q) &= \|\overrightarrow{PQ}\|^2 = \|\overrightarrow{P(\mathrm{proj}_{\mathscr{S}} P)} + \overrightarrow{(\mathrm{proj}_{\mathscr{S}} P)Q}\|^2\\ &= \|\overrightarrow{P(\mathrm{proj}_{\mathscr{S}} P)}\|^2 + \|\overrightarrow{(\mathrm{proj}_{\mathscr{S}} P)Q}\|^2 > \|\overrightarrow{P(\mathrm{proj}_{\mathscr{S}} P)}\|^2\\ &= \delta^2(P, \mathrm{proj}_{\mathscr{S}} P),\end{aligned}$$

ce qui prouve (2.55).

Lorsque P n'appartient pas à $\mathcal{S}$, la connaissance d'un vecteur directeur $\mathbf{n}$ de la perpendiculaire par P à $\mathcal{S}$ permet de calculer la distance de P à $\mathcal{S}$ au moyen de la formule

$$\operatorname{dist}(P, \mathcal{S}) = \| \operatorname{proj}_{\mathbf{n}} \overrightarrow{P_0P} \| = \frac{|(\overrightarrow{P_0P} \mid \mathbf{n})|}{\| \mathbf{n} \|}, \tag{2.56}$$

où P_0 est un point quelconque de $\mathcal{S}$. Pour justifier cette formule, il suffit d'observer que $\overrightarrow{P_0P} - \overrightarrow{(\operatorname{proj}_{\mathcal{S}} P)P} = \overrightarrow{P_0(\operatorname{proj}_{\mathcal{S}} P)}$ est un vecteur orthogonal à $\mathbf{n}$, ce qui nous permet de conclure que

$$\overrightarrow{(\operatorname{proj}_{\mathcal{S}} P)P} = \operatorname{proj}_{\mathbf{n}} \overrightarrow{P_0P} = \frac{(\overrightarrow{P_0P} \mid \mathbf{n})}{(\mathbf{n} \mid \mathbf{n})} \mathbf{n}.$$

Si $\mathcal{E}$ est de dimension finie non nulle et $\mathcal{S}$ est un hyperplan de $\mathcal{E}$, tout vecteur $\mathbf{n}$ normal à $\mathcal{S}$ est un vecteur directeur de la perpendiculaire par P à $\mathcal{S}$. Lorsque la dimension de $\mathcal{E}$ est 3, un tel vecteur est fourni par le produit vectoriel $\mathbf{v}_1 \times \mathbf{v}_2$, où $\mathbf{v}_1$, $\mathbf{v}_2$ sont des vecteurs directeurs de $\mathcal{S}$ (pour le cas général, voir l'exercice 5.5.13).

2.8.10 Perpendiculaire commune à deux droites gauches

Soit $\mathcal{D}$ et $\mathcal{D}'$ deux droites gauches de $\mathcal{E}$, c'est-à-dire deux droites non parallèles et d'intersection vide. Nous allons démontrer qu'il existe un unique point P de $\mathcal{D}$ et un unique point P' de $\mathcal{D}'$ tels que la droite déterminée par P et P' soit orthogonale à $\mathcal{D}$ et à $\mathcal{D}'$. Cette droite est appelée *perpendiculaire commune* à $\mathcal{D}$ et à $\mathcal{D}'$ (fig. 2.6).

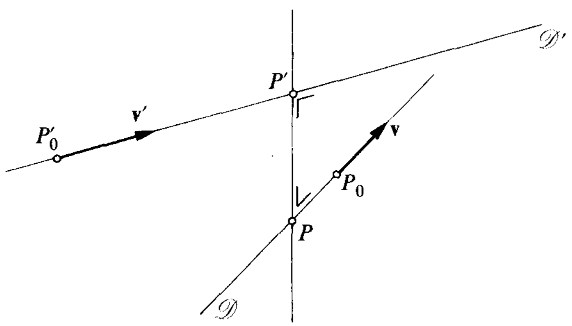


Fig. 2.6

Soit P_0 un point quelconque de $\mathcal{D}$ et P'_0 un point quelconque de $\mathcal{D}'$. Soit en outre $\mathbf{v}$ un vecteur directeur de $\mathcal{D}$ et $\mathbf{v}'$ un vecteur directeur de $\mathcal{D}'$. Des représentations paramétriques de $\mathcal{D}$ et de $\mathcal{D}'$

$$P = P_0 + \alpha \mathbf{v} \quad \text{et} \quad P' = P'_0 + \alpha' \mathbf{v}', \tag{2.57}$$

nous déduisons la relation

$$\overrightarrow{PP'} = \overrightarrow{P_0P'_0} + \alpha'\mathbf{v}' - \alpha\mathbf{v}.$$

Posons

$$\mathbf{u} = \mathbf{v} - \mathrm{proj}_{\mathbf{v}'}\mathbf{v} = \mathbf{v} - \frac{(\mathbf{v} \mid \mathbf{v}')}{(\mathbf{v}' \mid \mathbf{v}')}\mathbf{v}'. \tag{2.58}$$

On notera que $\mathbf{u}$ n'est pas nul, car $\mathscr{D}$ et $\mathscr{D}'$ ne sont pas parallèles. Supposons que la droite déterminée par P et P' soit orthogonale à $\mathscr{D}$ et à $\mathscr{D}'$. Alors le vecteur $\overrightarrow{PP'}$ est orthogonal à $\mathbf{v}$ et à $\mathbf{v}'$, donc à $\mathbf{u}$. En d'autres termes,

$$(\overrightarrow{P_0P'_0} \mid \mathbf{u}) + \alpha'(\mathbf{v}' \mid \mathbf{u}) - \alpha(\mathbf{v} \mid \mathbf{u}) = 0. \tag{2.59}$$

Mais

$$(\mathbf{v}' \mid \mathbf{u}) = (\mathbf{v}' \mid \mathbf{v}) - \frac{(\mathbf{v} \mid \mathbf{v}')}{(\mathbf{v}' \mid \mathbf{v}')}(\mathbf{v}' \mid \mathbf{v}') = 0$$

et

$$(\mathbf{v} \mid \mathbf{u}) = (\mathbf{v} \mid \mathbf{v}) - \frac{(\mathbf{v} \mid \mathbf{v}')}{(\mathbf{v}' \mid \mathbf{v}')}(\mathbf{v} \mid \mathbf{v}') = (\mathbf{u} \mid \mathbf{u}),$$

donc l'équation (2.59) se réduit à

$$(\overrightarrow{P_0P'_0} \mid \mathbf{u}) - \alpha(\mathbf{u} \mid \mathbf{u}) = 0.$$

En posant

$$\mathbf{u}' = \mathbf{v}' - \mathrm{proj}_{\mathbf{v}}\mathbf{v}' = \mathbf{v}' - \frac{(\mathbf{v}' \mid \mathbf{v})}{(\mathbf{v} \mid \mathbf{v})}\mathbf{v}, \tag{2.60}$$

nous obtenons, de manière analogue, l'équation

$$(\overrightarrow{P_0P'_0} \mid \mathbf{u}') + \alpha'(\mathbf{u}' \mid \mathbf{u}') = 0.$$

Nous en concluons que

$$\alpha = \frac{(\overrightarrow{P_0P'_0} \mid \mathbf{u})}{(\mathbf{u} \mid \mathbf{u})} \quad \text{et} \quad \alpha' = -\frac{(\overrightarrow{P_0P'_0} \mid \mathbf{u}')}{(\mathbf{u}' \mid \mathbf{u}')}, \tag{2.61}$$

où $\mathbf{u}$ et $\mathbf{u}'$ sont définis dans (2.58) et (2.60). Par les représentations (2.57), ces valeurs de α et de α' déterminent les points P et P' cherchés. En effet, le vecteur $\overrightarrow{PP'}$ est orthogonal à $\mathbf{u}$ et à $\mathbf{u}'$, donc à $\mathbf{v}$ et à $\mathbf{v}'$, vu que ces deux vecteurs appartiennent au plan vectoriel engendré par $\mathbf{u}$ et $\mathbf{u}'$.

2.8.11 Distance de deux droites gauches

Soit $\mathscr{S}$ et $\mathscr{S}'$ des sous-espaces affines de $\mathscr{E}$ de dimension finie. On appelle *distance de* $\mathscr{S}$ *et* $\mathscr{S}'$, et l'on note $\mathrm{dist}(\mathscr{S}, \mathscr{S}')$, le minimum de $\delta(P_0, P'_0)$, P_0 et P'_0 parcourant respectivement $\mathscr{S}$ et $\mathscr{S}'$. Nous nous limiterons à examiner le cas où $\mathscr{S}$ et $\mathscr{S}'$ sont deux droites gauches, que nous noterons $\mathscr{D}$ et $\mathscr{D}'$. Soit $\mathbf{n}$ un vecteur directeur de leur perpendiculaire commune, P et P' respectivement le point de $\mathscr{D}$

et le point de $\mathscr{D}'$ appartenant à cette perpendiculaire. Si P_0 est un point quelconque de $\mathscr{D}$ et P_0' un point quelconque de $\mathscr{D}'$, alors

$$\operatorname{dist}(\mathscr{D}, \mathscr{D}') = \delta(P, P') = \| \operatorname{proj}_{\mathbf{n}} \overrightarrow{P_0P_0'} \| = \frac{|(\overrightarrow{P_0P_0'} \mid \mathbf{n})|}{\| \mathbf{n} \|}. \tag{2.62}$$

La vérification de cette formule est analogue à celle de la formule (2.56); c'est pourquoi nous la laissons comme exercice au lecteur.

Dans le cas particulier où la dimension de $\mathscr{E}$ est 3, on posera $\mathbf{n} = \mathbf{v} \times \mathbf{v}'$, où $\mathbf{v}$ est un vecteur directeur de $\mathscr{D}$ et $\mathbf{v}'$ un vecteur directeur de $\mathscr{D}'$.

2.8.12 Angle déterminé par trois points

Supposons que les points P et R soient distincts d'un point Q. On appelle *angle déterminé par* P, Q, R, et l'on note $P\hat{Q}R$, l'angle des vecteurs $\overrightarrow{QP}$ et $\overrightarrow{QR}$ (fig. 2.7).

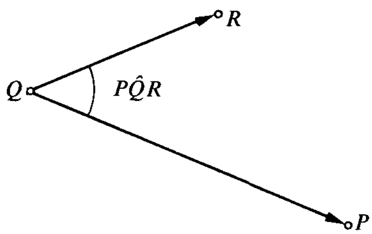


Fig. 2.7

2.8.13 Théorèmes de Pythagore et du cosinus

Soit P, Q, et R comme dans 2.8.12. D'après (2.24),

$$\begin{aligned} \| \overrightarrow{RP} \|^2 &= \| \overrightarrow{QP} - \overrightarrow{QR} \|^2 \\ &= \| \overrightarrow{QP} \|^2 + \| \overrightarrow{QR} \|^2 - 2 \| \overrightarrow{QP} \| \| \overrightarrow{QR} \| \cos(P\hat{Q}R), \end{aligned}$$

ce qui s'écrit aussi sous la forme

$$\delta^2(P, R) = \delta^2(P, Q) + \delta^2(Q, R) - 2\delta(P, Q)\delta(Q, R)\cos(P\hat{Q}R). \tag{2.63}$$

Dans le cas particulier où $P\hat{Q}R = \frac{\pi}{2}$, cela devient

$$\delta^2(P, R) = \delta^2(P, Q) + \delta^2(Q, R). \tag{2.64}$$

Les relations (2.63) et (2.64) sont appelées respectivement *théorème du cosinus* et *théorème de Pythagore*. On remarquera que (2.64) est encore vraie si $P = Q$ ou $R = Q$.

2.8.14 Angle d'une droite et un sous-espace affine

Soit $\mathscr{D}$ une droite de $\mathscr{E}$ de vecteur directeur $\mathbf{v}$ et $\mathscr{S}$ un sous-espace affine de $\mathscr{E}$ de direction S et de dimension finie non nulle k. Si $\mathscr{D}$ n'est pas orthogonale à $\mathscr{S}$, on appelle *angle de $\mathscr{D}$ et $\mathscr{S}$* l'angle θ des vecteurs $\mathbf{v}$ et $\mathrm{proj}_S\mathbf{v}$ (fig. 2.8). En choisissant une base orthogonale quelconque $(\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k)$ de S et en exprimant $\mathrm{proj}_S\mathbf{v}$ sous la forme (2.32), nous voyons aussitôt que $\cos\theta$ est positif, donc que θ appartient à l'intervalle $[0, \frac{\pi}{2})$. Si $\mathscr{D}$ est orthogonale à $\mathscr{S}$, l'angle de $\mathscr{D}$ et $\mathscr{S}$ est, par convention, $\frac{\pi}{2}$ (c'est-à-dire l'angle de $\mathbf{v}$ et n'importe quel vecteur non nul de $\mathscr{S}$).

On remarquera que l'intersection de $\mathscr{D}$ et $\mathscr{S}$ peut être vide. En particulier, notre définition d'angle s'applique à deux droites gauches.

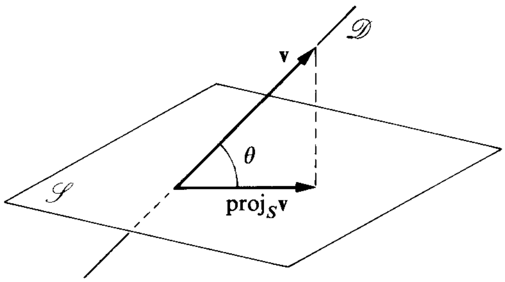


Fig. 2.8

2.8.15 Angle de deux hyperplans

Si $\mathscr{E}$ est de dimension finie supérieure à 1, on appelle *angle des hyperplans $\mathscr{S}$ et $\mathscr{S}'$* de $\mathscr{E}$ l'angle θ de deux droites $\mathscr{D}$ et $\mathscr{D}'$ orthogonales respectivement à $\mathscr{S}$ et à $\mathscr{S}'$ (fig. 2.9). Il est évident que $\cos\theta = \dfrac{|(\mathbf{n} \mid \mathbf{n}')|}{\|\mathbf{n}\| \, \|\mathbf{n}'\|}$, où $\mathbf{n}$ et $\mathbf{n}'$ sont des vecteurs normaux respectivement à $\mathscr{S}$ et à $\mathscr{S}'$. Il est également évident que lorsque $\mathscr{E}$ est de dimension 2, la présente définition de θ coïncide avec celle du paragraphe précédent.

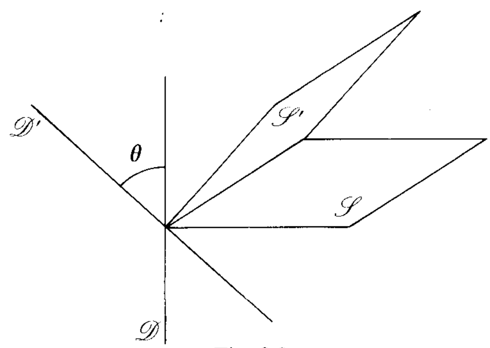


Fig. 2.9

2.8.16 Repères orthonormaux

Supposons que $\mathscr{E}$ soit de dimension finie non nulle n. On dit qu'un repère $(O; \mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ de $\mathscr{E}$ est *orthonormal* si la base $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ est orthonormale.

2.8.17 Retour à l'équation cartésienne d'un hyperplan

Encore sous l'hypothèse que $\mathscr{E}$ est de dimension finie non nulle, considérons un hyperplan $\mathscr{S}$ de $\mathscr{E}$ de direction S, un vecteur $\mathbf{n}$ normal à $\mathscr{S}$ et un point quelconque P_0 de $\mathscr{S}$. D'après (1.16), P est un point de $\mathscr{S}$ si et seulement si le vecteur $\overrightarrow{P_0P}$ appartient à S. D'autre part, en vertu de (2.31), ce vecteur appartient à S si et seulement s'il est orthogonal à $\mathbf{n}$. L'hyperplan $\mathscr{S}$ est donc constitué des points P qui vérifient l'équation

$$(\mathbf{n} \mid \overrightarrow{P_0P}) = 0. \tag{2.65}$$

Munissons $\mathscr{E}$ d'un repère orthonormal $(O; \mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ et désignons par $x_1^0, x_2^0, ..., x_n^0$ les coordonnées de P_0, par $x_1, x_2, ..., x_n$ celles de P et par $v_1, v_2, ..., v_n$ les composantes de $\mathbf{n}$. L'équation (2.65) s'écrit, de manière équivalente, sous la forme

$$v_1(x_1 - x_1^0) + v_2(x_2 - x_2^0) + ... + v_n(x_n - x_n^0) = 0, \tag{2.66}$$

ou encore

$$v_1x_1 + v_2x_2 + ... + v_nx_n = \delta, \tag{2.67}$$

où

$$\delta = v_1x_1^0 + v_2x_2^0 + ... + v_nx_n^0 = (\mathbf{n} \mid \overrightarrow{OP_0}). \tag{2.68}$$

L'équation (2.67) n'est rien d'autre que l'équation cartésienne de l'hyperplan $\mathscr{S}$ (cf. 1.11.8). Elle est ici obtenue par un procédé propre aux espaces affines euclidiens. Ce procédé montre que les coefficients du premier membre de cette équation sont les composantes d'un vecteur normal à l'hyperplan. Grâce à (2.56), il fournit en outre une interprétation du second membre:

$$|\delta| = \|\mathbf{n}\| \operatorname{dist}(O, \mathscr{S}). \tag{2.69}$$

2.8.18 Problèmes de géométrie analytique euclidienne

La géométrie analytique euclidienne traite les problèmes de la géométrie euclidienne (distance, orthogonalité, angles, ...) par des calculs dans $\mathbb{R}^n$. Dans les exemples suivants, l'espace affine euclidien est supposé muni d'un repère orthonormal.

(1) Déterminer la distance du point $P(0, 1, -1, 2)$ à l'hyperplan $\mathscr{S}$ d'équation

$$x_1 - x_2 + 3x_3 + x_4 = 2.$$

Solution. En appliquant la formule (2.56) à $P(0, 1, -1, 2)$, $P_0(0, 0, 0, 2)$ et $\mathbf{n}(1, -1, 3, 1)$, nous obtenons

$$\operatorname{dist}(P, \mathscr{S}) = \frac{4}{\sqrt{12}} = \frac{2}{\sqrt{3}}.$$

(2) Calculer la distance des deux droites gauches $\mathscr{D}$ et $\mathscr{D}'$ d'équations paramétriques respectives

$$\begin{aligned} x_1 &= 1 \\ x_2 &= 1 + \alpha \\ x_3 &= 1 + 2\alpha, \end{aligned} \qquad \begin{aligned} x_1 &= -1 - 2\alpha' \\ x_2 &= 2 - \alpha' \\ x_3 &= - 2\alpha'. \end{aligned}$$

Solution. $\mathbf{v}(0, 1, 2)$ est un vecteur directeur de $\mathscr{D}$ et $\mathbf{v}'(-2, -1, -2)$ un vecteur directeur de $\mathscr{D}'$. En appliquant la formule (2.62) à $P_0(1, 1, 1)$, $P_0'(-1, 2, 0)$ et $\mathbf{n} = \mathbf{v} \times \mathbf{v}'$, nous obtenons

$$\operatorname{dist}(\mathscr{D}, \mathscr{D}') = \frac{|-3|}{\sqrt{5}} = \frac{3}{\sqrt{5}}.$$

(3) Les données étant celles du problème précédent, déterminer les points d'intersection P, P' des droites $\mathscr{D}$, $\mathscr{D}'$ et la perpendiculaire commune à $\mathscr{D}$ et à $\mathscr{D}'$.

Solution. En appliquant les formules (2.61) à $P_0(1, 1, 1)$, $P_0'(-1, 2, 0)$, $\mathbf{v}(0, 1, 2)$ et $\mathbf{v}'(-2, -1, -2)$, nous obtenons

$$\alpha = \frac{4}{5}, \ \alpha' = -1.$$

Les coordonnées des points cherchés P et P' se calculent alors par (2.57):

$$P(1, \frac{9}{5}, \frac{13}{5}), \ P'(1, 3, 2).$$

(4) Trouver la projection orthogonale du point $P(-1, 1, 2, 3)$ sur le plan $\mathscr{S}$ passant par le point $P_0(0, -1, 0, -1)$ et de vecteurs directeurs $\mathbf{v}_1(1, 1, 0, 0)$ et $\mathbf{v}_2(1, 1, 1, 1)$. Déterminer, en outre, la distance de P à $\mathscr{S}$.

Solution. On voit aussitôt que $\mathbf{v}_1$ et $\mathbf{v}_2 - \mathbf{v}_1$ sont des vecteurs directeurs orthogonaux de $\mathscr{S}$. D'après 2.8.7 et (2.32),

$$\operatorname{proj}_{\mathscr{S}} P = P_0 + \operatorname{proj}_S \overrightarrow{P_0P} = P_0 + \frac{(\overrightarrow{P_0P} \mid \mathbf{v}_1)}{(\mathbf{v}_1 \mid \mathbf{v}_1)}\mathbf{v}_1 + \frac{(\overrightarrow{P_0P} \mid \mathbf{v}_2 - \mathbf{v}_1)}{(\mathbf{v}_2 - \mathbf{v}_1 \mid \mathbf{v}_2 - \mathbf{v}_1)}(\mathbf{v}_2 - \mathbf{v}_1),$$

ce qui nous permet de calculer les coordonnées de $\operatorname{proj}_{\mathscr{S}} P$, soit $\frac{1}{2}$, $-\frac{1}{2}$, 3, 2.

La distance de P à $\mathscr{S}$ est égale à la norme du vecteur $\overrightarrow{P(\operatorname{proj}_{\mathscr{S}} P)}$, c'est-à-dire à $\sqrt{13/2}$.

(5) Trouver l'équation du lieu géométrique des points équidistants des points $P_1(1, 0, 2, -1)$ et $P_2(0, 1, -1, 1)$.

Solution. Un point $P(x_1, x_2, x_3, x_4)$ est situé à égale distance de P_1 et de P_2 si et seulement si

$$(x_1 - 1)^2 + x_2^2 + (x_3 - 2)^2 + (x_4 + 1)^2 = x_1^2 + (x_2 - 1)^2 + (x_3 + 1)^2 + (x_4 - 1)^2.$$

En développant les carrés, on voit aisément que cette équation se réduit à

$$2x_1 - 2x_2 + 6x_3 - 4x_4 = 3.$$

C'est l'équation d'un hyperplan, appelé *hyperplan médiateur* du segment de droite P_1P_2.

(6) Déterminer l'équation du cône de révolution de sommet $P(3, 0, 3)$ et dont l'intersection avec le plan $\mathscr{S}$ d'équation

$$2x_1 - x_2 + 2x_3 = 3$$

est un cercle de rayon 3.

Solution. En appliquant la formule (2.56) à $P(3, 0, 3)$, $P_0(0, -3, 0)$ et $\mathbf{n}(2, -1, 2)$, nous obtenons

$$\operatorname{dist}(P, \mathscr{S}) = \frac{|-9|}{3} = 3.$$

La tangente de l'angle θ des génératrices et l'axe du cône est donc

$$\operatorname{tg}\theta = \frac{3}{3} = 1,$$

ce qui entraîne $\theta = \frac{\pi}{4}$. Soit $Q(x_1, x_2, x_3)$ un point quelconque du cône distinct de P. L'angle du vecteur directeur $\mathbf{n}(2, -1, 2)$ de l'axe du cône et $\overrightarrow{PQ}(x_1 - 3, x_2, x_3 - 3)$ étant $\frac{\pi}{4}$ ou $\frac{3\pi}{4}$, il s'ensuit, grâce à (2.23), que

$$2(x_1 - 3) - x_2 + 2(x_3 - 3) = ((x_1 - 3)^2 + x_2^2 + (x_3 - 3)^2)^{1/2} \cdot 3 \cdot \left(\pm\frac{\sqrt{2}}{2}\right).$$

On notera que les coordonnées de P vérifient également cette équation. Il ne reste alors plus qu'à élever les deux membres au carré et à développer ces carrés pour obtenir l'équation demandée:

$$x_1^2 + 7x_2^2 + x_3^2 + 8x_1x_2 - 16x_1x_3 + 8x_2x_3 + 42x_1 - 48x_2 + 42x_3 = 126.$$

2.9 EXERCICES

2.9.1 Lesquelles des opérations définies ci-dessous sont des produits scalaires dans $\mathbb{R}^3$?

(a) $(\mathbf{x} \mid \mathbf{y}) = x_1y_1 + x_2y_2 + x_2y_3 + x_3y_2 + 2x_3y_3$.
(b) $(\mathbf{x} \mid \mathbf{y}) = x_1y_1 + x_2y_2 - 3x_2y_3 + x_3y_2 + x_3y_3$.
(c) $(\mathbf{x} \mid \mathbf{y}) = x_1y_1 + x_2y_2 + x_3^2y_2$.

Même question rapportée aux espaces vectoriels respectifs $C_{[-1,1]}$, $C_{[0,\infty)}$ et $C_{\mathbb{R}}$:

(d) $(\mathbf{f} \mid \mathbf{g}) = \int_{-1}^{1} t^2\mathbf{f}(t)\mathbf{g}(t)\mathrm{d}t$.

(e) $(\mathbf{f} \mid \mathbf{g}) = \int_{0}^{\infty} t\mathbf{f}^2(t)\mathbf{g}(t)\mathrm{d}t$.

(f) $(\mathbf{f} \mid \mathbf{g}) = \int_{-\infty}^{\infty} \mathrm{e}^{-t}\mathbf{f}(t)\mathbf{g}(t)\mathrm{d}t$.

2.9.2 Pour tout couple de polynômes $\mathbf{p}(t) \equiv \sum_{i=0}^{m} \alpha_i t^i$ et $\mathbf{q}(t) \equiv \sum_{i=0}^{n} \beta_i t^i$, posons

$$(\mathbf{p} \mid \mathbf{q}) = \sum_{i=0}^{\min(m,n)} \alpha_i\beta_i.$$

(a) Montrer que l'opération ainsi définie est un produit scalaire dans l'espace vectoriel des polynômes.
(b) Exhiber une famille infinie de polynômes orthonormale pour ce produit scalaire.

2.9.3 Montrer que pour tout couple $(\mathbf{x}, \mathbf{y})$ de vecteurs d'un espace vectoriel euclidien,

$$\|\mathbf{x} + \mathbf{y}\|^2 + \|\mathbf{x} - \mathbf{y}\|^2 = 2\|\mathbf{x}\|^2 + 2\|\mathbf{y}\|^2 \quad (\textit{identité du parallélogramme}).$$

2.9.4 Montrer que dans un espace vectoriel euclidien la somme et la différence de deux vecteurs de même norme sont orthogonales.

2.9.5 Par le procédé de Gram-Schmidt, orthogonaliser le triplet de vecteurs de $\mathbb{R}^4$

$$\left(\begin{bmatrix}1\\0\\2\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\\-1\end{bmatrix}, \begin{bmatrix}0\\3\\-1\\0\end{bmatrix}\right).$$

2.9.6 Soit $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4)$ une base orthonormale d'un espace vectoriel euclidien de dimension 4. Par le procédé de Gram-Schmidt, orthogonaliser le triplet de vecteurs $(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$, où $\mathbf{x}_1 = \mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3 + \mathbf{e}_4$, $\mathbf{x}_2 = \mathbf{e}_2 + \mathbf{e}_3 + \mathbf{e}_4$ et $\mathbf{x}_3 = \mathbf{e}_3 + \mathbf{e}_4$.

2.9.7 A l'aide de l'inégalité de Cauchy-Schwarz, montrer que

$$\int_0^1 \sqrt{1 + t^3}\,\mathrm{d}t < \frac{\sqrt{5}}{2}.$$

2.9.8 Exprimer le vecteur-colonne $\begin{bmatrix} 3 \\ -1 \\ 1 \end{bmatrix}$ sous la forme d'une combinaison linéaire des vecteurs-colonnes

$$\frac{1}{\sqrt{3}}\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}, \quad \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad \frac{1}{\sqrt{6}}\begin{bmatrix} 1 \\ -2 \\ -1 \end{bmatrix}.$$

2.9.9 Soit $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_k)$ une famille orthonormale de vecteurs d'un espace vectoriel euclidien E. Montrer que pour tout vecteur $\mathbf{x}$ de E,

$$\sum_{i=1}^{k} (\mathbf{x} \mid \mathbf{e}_i)^2 \leqslant \| \mathbf{x} \|^2.$$

Dans quel cas les deux membres sont-ils égaux?

2.9.10 Soit E un espace vectoriel euclidien de dimension 4, muni d'une base orthonormale. Déterminer la projection orthogonale du vecteur $\mathbf{x}(3, 2, -2, -1)$ sur le plan vectoriel de E engendré par les vecteurs $\mathbf{v}_1(1, 0, -1, 1)$ et $\mathbf{v}_2(1, 0, 0, 1)$.

2.9.11 Soit E un espace vectoriel euclidien de dimension 4, muni d'une base orthonormale. Soit S le sous-espace vectoriel de E engendré par les vecteurs $\mathbf{v}_1(1, 1, 1, 1)$ $\mathbf{v}_2(-3, 1, 1, 1)$ et $\mathbf{v}_3(0, -2, 1, 1)$. Déterminer le complémentaire orthogonal de S dans E.

2.9.12 Calculer la meilleure approximation de la fonction $\mathbf{f}(t) \equiv t^2$ de $C_{[-\pi, \pi]}$ par des polynômes trigonométriques d'ordre inférieur ou égal à k.

2.9.13 Soit S le sous-espace vectoriel de $C_{[-\pi, \pi]}$ formé de tous les polynômes trigonométriques.

(a) A l'aide de (2.34), montrer que $S^\perp = \{\mathbf{0}\}$. (Prendre un $\mathbf{f}$ de $S^\perp$ et s'assurer que $\mathbf{f}_k = \mathbf{0}$.)

(b) En déduire que S n'admet aucun complémentaire orthogonal dans $C_{[-\pi, \pi]}$ (autrement dit, que $S \oplus S^\perp \neq C_{[-\pi, \pi]}$).

2.9.14 Soit E un espace vectoriel euclidien, orienté et de dimension 3. A quelle condition doivent satisfaire deux vecteurs donnés $\mathbf{y}$ et $\mathbf{z}$ de E, pour que l'équation vectorielle $\mathbf{x} \times \mathbf{y} = \mathbf{z}$ ait au moins une solution $\mathbf{x}_0$ dans E? Quel est alors l'ensemble des solutions?

2.9.15 Soit E comme dans l'exercice précédent. Démontrer les identités vectorielles suivantes:

(a) $(\mathbf{x} \times (\mathbf{y} \times \mathbf{z})) + (\mathbf{y} \times (\mathbf{z} \times \mathbf{x})) + (\mathbf{z} \times (\mathbf{x} \times \mathbf{y})) = \mathbf{0}$. (Utiliser (2.50).)

(b) $(\mathbf{x} \times \mathbf{y} \mid (\mathbf{y} \times \mathbf{z}) \times (\mathbf{z} \times \mathbf{x})) = [\mathbf{x}, \mathbf{y}, \mathbf{z}]^2$. (Utiliser (2.51) ou (2.52).)

(c) $\mathbf{v} = \dfrac{[\mathbf{v}, \mathbf{y}, \mathbf{z}]}{[\mathbf{x}, \mathbf{y}, \mathbf{z}]}\mathbf{x} + \dfrac{[\mathbf{x}, \mathbf{v}, \mathbf{z}]}{[\mathbf{x}, \mathbf{y}, \mathbf{z}]}\mathbf{y} + \dfrac{[\mathbf{x}, \mathbf{y}, \mathbf{v}]}{[\mathbf{x}, \mathbf{y}, \mathbf{z}]}\mathbf{z}$ $([\mathbf{x}, \mathbf{y}, \mathbf{z}] \neq 0)$. (Utiliser (2.52).)

Exercices sur les espaces affines euclidiens

Dans les exercices suivants, lorsqu'il est sous-entendu que l'espace affine euclidien est muni d'un repère, celui-ci est supposé orthonormal.

2.9.16 Calculer la distance du point $P(0, 1, -1, 2)$ à l'hyperplan d'équation $x_1 - x_2 + 3x_3 + x_4 = 2$.

2.9.17 Trouver l'équation de l'hyperplan médiateur du segment d'extrémités $P_1(3, -2, 5, -1)$ et $P_2(-3, 1, 2, 4)$.

2.9.18 Trouver les équations paramétriques des bissectrices des deux droites passant par le point $P_0(-1, 2, 1, 3)$, de vecteurs directeurs respectifs $\mathbf{v}(1, -3, 2, 2)$ et $\mathbf{v}'(4, 0, 1, 1)$.

2.9.19 Déterminer la projection orthogonale du point $P(-1, 1, 1, 1)$ sur le plan passant par le point $P_0(0, 1, 2, -1)$, de vecteurs directeurs $\mathbf{v}_1(-1, 0, 2, 1)$ et $\mathbf{v}_2(0, 1, 1, -1)$. Calculer en outre la distance de P à ce plan.

2.9.20 Dans un espace affine euclidien, orienté et de dimension 3, on considère un point P et la droite $\mathscr{D}$ passant par le point P_0 et de vecteur directeur $\mathbf{v}$. Montrer que

$$\operatorname{dist}(P, \mathscr{D}) = \frac{\| \overrightarrow{P_0P} \times \mathbf{v} \|}{\| \mathbf{v} \|}.$$

2.9.21 Déterminer la perpendiculaire commune et la distance des deux droites d'équations paramétriques respectives

$$\begin{aligned} x_1 &= \alpha \\ x_2 &= 1 \\ x_3 &= -1 + \alpha \\ x_4 &= 2 - \alpha, \end{aligned} \qquad \begin{aligned} x_1 &= \alpha' \\ x_2 &= 1 - \alpha' \\ x_3 &= 0 \\ x_4 &= -1 + 2\alpha'. \end{aligned}$$

2.9.22 Calculer l'angle de la droite passant par le point $P_0(1, 1, 1, 1)$, de vecteur directeur $\mathbf{v}(0, 1, -1, 0)$, et le plan passant par le point $Q_0(1, 0, -1, 1)$, de vecteurs directeurs $\mathbf{v}_1(1, 1, 0, 0)$ et $\mathbf{v}_2(0, 1, 1, -1)$.

2.9.23 Déterminer l'angle que forment la droite et l'hyperplan de l'exercice 1.12.18.

2.9.24 Déterminer l'angle des deux hyperplans d'équations respectives $x_1 - 3x_2 + x_3 - 2x_4 - x_5 = 2$ et $2x_1 - 2x_2 - 3x_3 + 2x_4 + 2x_5 = -1$.

2.9.25 Soit $\mathscr{E}$ un espace affine euclidien de dimension finie supérieure à 1.

(a) Montrer que les équations des hyperplans bissecteurs des deux hyperplans de $\mathscr{E}$ d'équations respectives $(\mathbf{n}_1 \mid \mathbf{x}) = \delta_1$ et $(\mathbf{n}_2 \mid \mathbf{x}) = \delta_2$ sont données par

$$\frac{(\mathbf{n}_1 \mid \mathbf{x}) - \delta_1}{\| \mathbf{n}_1 \|} = \pm \frac{(\mathbf{n}_2 \mid \mathbf{x}) - \delta_2}{\| \mathbf{n}_2 \|}.$$

(b) Appliquer ce résultat à $\mathbf{n}_1(1, -1, 3, 1, 2)$, $\mathbf{n}_2(-1, 2, 4, 0, -2)$, $\delta_1 = -1$ et $\delta_2 = 2$.

2.9.26 Dans un espace affine euclidien, orienté et de dimension 3, on considère trois points P_1, P_2, P_3 non alignés et on pose $\theta_1 = P_3\hat{P}_1P_2$, $\theta_2 = P_1\hat{P}_2P_3$ et $\theta_3 = P_2\hat{P}_3P_1$. A l'aide du produit vectoriel, démontrer que

$$\frac{\sin\theta_1}{\delta(P_2, P_3)} = \frac{\sin\theta_2}{\delta(P_1, P_3)} = \frac{\sin\theta_3}{\delta(P_1, P_2)} \quad (\textit{théorème du sinus}).$$

2.9.27 Les quatre sommets d'un tétraèdre sont $P_1(1, 1, 1)$, $P_2(-1, 2, 1)$, $P_3(5, 2, 3)$ et P_4. Trouver les coordonnées du point P_4, sachant qu'il appartient à la droite passant par $P_0(1, 0, 1)$, de vecteur directeur $\mathbf{v}(-2, -1, 3)$, et que le volume du tétraèdre est 5.

2.9.28 Trouver l'équation cartésienne du plan passant par le point P_0 $(5, 3, -1)$, perpendiculaire au plan $\mathscr{S}$ d'équations paramétriques

$$\begin{aligned} x_1 &= 3 + 4\alpha_1 + \alpha_2 \\ x_2 &= -1 - \alpha_1 - \alpha_2 \\ x_3 &= -2 + 2\alpha_1 - \alpha_2 \end{aligned}$$

et parallèle à la droite $\mathscr{D}$ d'équations paramétriques

$$\begin{aligned} x_1 &= 17 - \alpha' \\ x_2 &= 9 + 3\alpha' \\ x_3 &= -11 - 2\alpha'. \end{aligned}$$

(Deux hyperplans sont dits *perpendiculaires* si leurs vecteurs normaux sont orthogonaux (cf. 2.8.15).)

CHAPITRE 3

Systèmes linéaires

3.1 DÉFINITIONS ET EXEMPLES

3.1.1 Introduction

Nous avons déjà eu l'occasion de constater que la résolution de certains problèmes liés aux notions de sous-espace vectoriel et de sous-espace affine se réduit à la recherche des solutions d'un système d'équations linéaires. D'un point de vue plus général, on voit souvent que des problèmes d'origines les plus diverses se traduisent en termes d'équations linéaires. Le présent chapitre exposera les principaux résultats de la théorie des systèmes d'équations linéaires. L'idée à la base de notre étude est simple: par une succession d'opérations élémentaires sur les équations, le système est transformé en un système échelonné équivalent que l'on peut résoudre facilement. Le procédé de transformation est connu sous le nom de *méthode de Gauss* ou *méthode d'élimination des inconnues.*

3.1.2 Systèmes linéaires

On appelle *système linéaire*, ou simplement *système*, toute famille d'équations de la forme

$$\begin{aligned} a_{11}x_1 &+ a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 &+ a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \dots\dots& \\ a_{m1}x_1 &+ a_{m2}x_2 + \dots + a_{mn}x_n = b_m, \end{aligned} \tag{3.1}$$

où $a_{11}, a_{12}, \dots, a_{mn}$ sont des nombres appelés *coefficients du système*, $b_1, b_2, \dots, b_m$ des nombres appelés *coefficients du second membre* et $x_1, x_2, \dots, x_n$ des nombres inconnus appelés *inconnues* du système. Si les coefficients du second membre sont tous nuls, on dit que le système est *homogène*. On appelle *système homogène associé* au système (3.1) le système que l'on obtient de (3.1) en substituant des zéros aux coefficients du second membre.

3.1.3 Ecriture abrégée

On écrit souvent le système (3.1) sous la forme abrégée suivante:

$$\sum_{j=1}^{n} a_{ij}x_j = b_i, \quad i = 1, 2, \dots, m. \tag{3.2}$$

Une écriture encore plus condensée sera introduite dans l'exemple (5) de 4.1.7.

3.1.4 Solutions d'un système

On appelle *solution* du système (3.1) tout n-uplet de nombres $(x_1^0, x_2^0, ..., x_n^0)$ tel que

$$\sum_{j=1}^{n} a_{ij}x_j^0 = b_i, \quad i = 1, 2, ..., m.$$

Résoudre un système signifie trouver l'ensemble des solutions de ce système. Deux systèmes à n inconnues sont dits *équivalents* si toute solution de l'un est solution de l'autre, autrement dit, s'ils admettent le même ensemble de solutions. On dit parfois que les équations d'un système sont *compatibles* ou *incompatibles*, suivant que ce système admet au moins une solution ou n'en admet aucune.

3.1.5 Interprétation géométrique

Supposons que les premiers membres des équations du système (3.1) soient non nuls. D'après 1.11.8, chacune de ces équations représente alors un hyperplan d'un espace affine de dimension n. Par conséquent, l'ensemble des solutions du système, regardé comme ensemble de n-uplets de coordonnées, représente une intersection finie d'hyperplans. Selon 1.10.7, une telle intersection est un sous-espace affine ou l'ensemble vide. Du point de vue géométrique, résoudre un système admettant au moins une solution revient à chercher les équations paramétriques (1.19) de ce sous-espace affine.

3.1.6 Exemples

(1) Le système

$$\begin{aligned} x_1 + x_2 &= 1 \\ x_1 + x_2 &= 0 \end{aligned}$$

n'admet aucune solution, car les deux équations se contredisent l'une l'autre. Ces équations représentent deux droites parallèles.

(2) Nous nous proposons de résoudre le système

$$\begin{aligned} 3x_1 - 3x_2 - 2x_3 &= -1 \\ 4x_1 - x_2 + 3x_3 &= -1 \\ -2x_1 + 4x_2 + 3x_3 &= 3 \end{aligned} \tag{3.3}$$

par la méthode d'élimination des inconnues.

A l'aide de la première équation, exprimons x_1 en fonction de x_2 et x_3:

$$x_1 = \tfrac{1}{3}(-1 + 3x_2 + 2x_3). \tag{3.4}$$

Substituons ensuite l'expression ainsi obtenue à x_1 dans les deux autres équations de (3.3):

$$\begin{aligned} \tfrac{4}{3}(-1 + 3x_2 + 2x_3) - x_2 + 3x_3 &= -1 \\ -\tfrac{2}{3}(-1 + 3x_2 + 2x_3) + 4x_2 + 3x_3 &= 3. \end{aligned}$$

En multipliant chaque membre par 3 et en groupant les termes, nous obtenons

$$\begin{aligned} 9x_2 + 17x_3 &= 1 \\ 6x_2 + 5x_3 &= 7. \end{aligned} \tag{3.5}$$

A l'aide de la première de ces deux équations, exprimons à présent x_2 en fonction de x_3 et substituons l'expression obtenue à x_2 dans la deuxième équation. Après simplification, il reste

$$x_3 = -1 \tag{3.6}$$

Par substitution de cette valeur à x_3 dans l'une des deux équations (3.5), nous déterminons la valeur de x_2, soit $x_2 = 2$. Nous calculons finalement la valeur de x_1 par (3.4), ce qui nous donne $x_1 = 1$. Le système (3.3) admet donc l'unique solution $x_1 = 1$, $x_2 = 2$, $x_3 = -1$.

Les deux équations (3.5) s'obtiennent plus rapidement en soustrayant la première équation de (3.3) multipliée par 4 de la deuxième multipliée par 3 et en additionnant la même équation multipliée par 2 à la troisième multipliée par 3. De même, nous obtenons l'équation (3.6) en soustrayant la première équation de (3.5) multipliée par 2 de la deuxième multipliée par 3, puis en divisant les deux membres par -19. Ces opérations réduisent le système (3.3) au système équivalent

$$\begin{aligned} 3x_1 - 3x_2 - 2x_3 &= -1 \\ 9x_2 + 17x_3 &= 1 \\ x_3 &= -1 \end{aligned} \tag{3.7}$$

que l'on résout par substitution, comme il a été indiqué ci-dessus. Par la suite, c'est cette variante de la méthode d'élimination des inconnues qui retiendra notre attention.

(3) Nous allons maintenant résoudre le système

$$\begin{aligned} x_1 - x_2 + 2x_3 &= 0 \\ 3x_1 + 2x_2 - 3x_3 &= 2. \end{aligned} \tag{3.8}$$

Eliminons l'inconnue x_1 de la deuxième équation en soustrayant la première équation multipliée par 3 de la deuxième. Cette opération réduit le système (3.8) au système équivalent

$$\begin{aligned} x_1 - x_2 + 2x_3 &= 0 \\ 5x_2 - 9x_3 &= 2. \end{aligned} \tag{3.9}$$

Pour résoudre ce système, nous assignons à x_3 une valeur arbitraire α, exprimons x_2 en fonction de α à l'aide de la deuxième équation, puis x_1 à l'aide de la première équation. Il en résulte que les solutions du système (3.8) sont les triplets de nombres de la forme

$$\begin{aligned} x_1 &= \tfrac{2}{5} - \tfrac{1}{5}\alpha \\ x_2 &= \tfrac{2}{5} + \tfrac{9}{5}\alpha \\ x_3 &= \alpha, \end{aligned} \tag{3.10}$$

où α parcourt $\mathbb{R}$. L'ensemble des solutions représente donc une droite.

3.2 EXISTENCE ET UNICITÉ DES SOLUTIONS

3.2.1 Matrices

On appelle *matrice* à m lignes et n colonnes, ou matrice de type $m \times n$, tout tableau rectangulaire de nombres

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2j} & \cdots & a_{2n} \\ & & \cdots & \cdots & & \\ a_{i1} & a_{i2} & \cdots & a_{ij} & \cdots & a_{in} \\ & & \cdots & \cdots & & \\ a_{m1} & a_{m2} & \cdots & a_{mj} & \cdots & a_{mn} \end{pmatrix}. \tag{3.11}$$

On désigne souvent une matrice de type $m \times n$ plus brièvement par $(a_{ij})_{\substack{1 \leqslant i \leqslant m \\ 1 \leqslant j \leqslant n}}$ ou simplement par (a_{ij}).

Le nombre a_{ij} est appelé *terme d'indices i, j*. L'indice i est appelé *indice de ligne* et l'indice j *indice de colonne*.

Lorsque $m = n$, on dit que (a_{ij}) est une *matrice carrée* d'ordre n. Dans ce cas, les termes a_{11}, a_{22}, ..., a_{nn} sont appelés *termes diagonaux*.

On appelle une matrice à une seule ligne *matrice-ligne* et une matrice à une seule colonne *matrice-colonne*. Il est clair qu'une matrice-colonne n'est rien d'autre qu'un vecteur-colonne. Par la suite, les lignes d'une matrice seront assimilées à des matrices-lignes et les colonnes à des matrices-colonnes.

L'intérêt de la notion de matrice apparaîtra tout au long de ce livre, à partir de cette section, mais la raison d'être immédiate de cette notion est simplement de permettre à certaines familles finies de nombres d'être conçues sous la forme d'un tableau rectangulaire.

3.2.2 Notation

Nous assignerons aux matrices des symboles propres, à savoir les lettres latines majuscules imprimées en caractère gras: **A**, **B**, ... Toutefois, les matrices-colonnes seront également désignées par les symboles réservés aux vecteurs: **a**, **b**, ... ; nous les appellerons d'ailleurs indifféremment matrices-colonnes ou vecteurs-colonnes.

3.2.3 Matrices nulles et matrices-unités

On appelle *matrice nulle*, et on note **O**, toute matrice dont chaque terme est nul. Les matrices-colonnes nulles sont également désignées par le symbole vectoriel **0**.

On appelle *matrice-unité* d'ordre n, et on note $\mathbf{I}_n$ ou simplement **I**, la matrice carrée d'ordre n

$$\begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ & \dots & \dots & \\ 0 & 0 & \dots & 1 \end{bmatrix}. \tag{3.12}$$

Nous verrons au chapitre suivant que **O** joue le rôle d'élément neutre de l'addition matricielle et **I** d'élément neutre de la multiplication matricielle.

3.2.4 Rang d'une matrice

Soit $\mathbf{a}_1$, $\mathbf{a}_2$, ..., $\mathbf{a}_n$ les colonnes d'une matrice **A**. On appelle *rang* de **A**, et on note rg**A**, le rang de la famille $(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n)$ (cf. 1.8.2).

Par exemple, le rang de **O** est 0 et le rang de $\mathbf{I}_n$ est n; les rangs des matrices

$$\begin{bmatrix} 2 & 2 & 5 & 2 \\ 0 & 3 & -1 & 2 \\ 0 & 0 & 4 & -1 \end{bmatrix}, \quad \begin{bmatrix} 2 & 1 & 2 & 5 \\ -1 & 3 & 1 & -2 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 3 \\ -2 & 0 & -6 \\ 1 & 0 & 3 \end{bmatrix}$$

sont respectivement 3, 2 et 1.

3.2.5 Matrices associées à un système linéaire

On appelle *matrice associée au système* (3.1) la matrice (3.11), c'est-à-dire la matrice **A** dont les termes sont les coefficients du système. On appelle *matrice du second membre du système* (3.1), ou simplement *second membre du système* (3.1), la matrice-colonne $\mathbf{b} = (b_i)$ dont les termes sont les coefficients du second membre de ce système. On appelle *matrice augmentée associée au système* (3.1) la matrice obtenue de **A** en y ajoutant **b** comme $(n + 1)$-ième colonne.

Nous venons de définir les différentes matrices associées à un système, mais il est également utile de pouvoir parler du système homogène associé à une matrice donnée ou du système associé à une matrice augmentée donnée. Cette terminologie s'explique d'elle-même et nous ne nous attarderons donc pas à la préciser davantage.

3.2.6 Ecriture vectorielle d'un système linéaire

Considérons un système de matrice associée $\mathbf{A}$ et de second membre $\mathbf{b}$. Désignons par $\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_n$ les colonnes de $\mathbf{A}$. Le système s'écrit alors de manière équivalente sous la forme d'une équation vectorielle linéaire:

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + ... + x_n\mathbf{a}_n = \mathbf{b}. \tag{3.13}$$

En effet, les deux membres de cette équation sont les vecteurs-colonnes dont les i-ièmes termes sont les deux membres de la i-ième équation du système.

D'après la proposition 1.8.4, l'équation (3.13) admet au moins une solution $(x_1^0, x_2^0, ..., x_n^0)$ si et seulement si le rang de la famille $(\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_n)$ est égal au rang de la famille augmentée $(\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_n, \mathbf{b})$. Selon la proposition 1.5.6, cette solution est unique si et seulement si le rang de la famille $(\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_n)$ est n.

A l'aide de la définition 3.2.4, nos conclusions se résument de la manière suivante:

3.2.7 Proposition. Existence et unicité des solutions d'un système linéaire

Pour qu'un système linéaire de matrice associée $\mathbf{A}$ *et de second membre* $\mathbf{b}$ *admette au moins une solution, il faut et il suffit que le rang de* $\mathbf{A}$ *soit égal au rang de la matrice augmentée* $(\mathbf{A} \mid \mathbf{b})$. *Si cette condition est remplie, le système admet une seule solution si et seulement si le rang de* $\mathbf{A}$ *est égal au nombre d'inconnues, autrement dit, les colonnes de* $\mathbf{A}$ *sont linéairement indépendantes.*

3.3 MATRICES ÉCHELONNÉES

3.3.1 Exemple

La méthode employée pour réduire les systèmes (3.3) et (3.8) aux systèmes équivalents (3.7) et (3.9) est, en fait, un procédé d'annulation de certains termes de la matrice associée au système par des opérations sur les lignes de la matrice augmentée. Voici, indiquée à côté de cette matrice, la suite des opérations que nous avons effectuées pour aboutir au système (3.7):

$$\begin{array}{cl} \begin{bmatrix} 3 & -3 & -2 & -1 \\ 4 & -1 & 3 & -1 \\ -2 & 4 & 3 & 3 \end{bmatrix} & \begin{array}{l} \\ 3\mathbf{L}_2 - 4\mathbf{L}_1 \\ 3\mathbf{L}_3 + 2\mathbf{L}_1 \end{array} \end{array} \tag{3.14}$$

$$\begin{array}{cl} \begin{bmatrix} 3 & -3 & -2 & -1 \\ 0 & 9 & 17 & 1 \\ 0 & 6 & 5 & 7 \end{bmatrix} & \begin{array}{l} \\ \\ -\frac{1}{19}(3\mathbf{L}_3 - 2\mathbf{L}_2) \end{array} \end{array}$$

$$\begin{bmatrix} 3 & -3 & -2 & -1 \\ 0 & 9 & 17 & 1 \\ 0 & 0 & 1 & -1 \end{bmatrix}. \tag{3.15}$$

Le système associé à la matrice augmentée (3.15) est le système (3.7). Suivant l'expression que nous allons introduire dans 3.3.5, la matrice (3.14) a été réduite à la forme échelonnée (3.15) par des opérations élémentaires sur les lignes. Il est peut-être superflu de préciser que le symbole $\mathbf{L}_i$ utilisé ci-dessus désigne la i-ième ligne de la matrice contiguë et que la formule suivant la i-ième ligne indique les opérations à effectuer pour obtenir la i-ième ligne de la matrice successive.

3.3.2 Matrices échelonnées

On dit qu'une matrice est *échelonnée* si ses lignes satisfont aux deux conditions suivantes:

(a) Toute ligne nulle n'est suivie que de lignes nulles.

(b) L'indice de colonne du premier terme non nul de toute ligne non nulle est supérieur à l'indice de colonne du premier terme non nul de la ligne qui la précède.

Une matrice échelonnée non nulle est donc de la forme

$$\begin{bmatrix} a_{1j_1} & \cdots & & & & \cdots & a_{1n} \\ & & a_{2j_2} & \cdots & & \cdots & a_{2n} \\ & & & \cdots\cdots\cdots & & & \\ & 0 & & & a_{rj_r} & \cdots & a_{rn} \\ & & & & & & \end{bmatrix}, \tag{3.16}$$

où $j_1 < j_2 < \ldots < j_r$ et $a_{1j_1}, a_{2j_2}, \ldots, a_{rj_r}$ sont des termes non nuls. Bien entendu, les lignes nulles terminales peuvent manquer.

Les matrices nulles et les matrices-unités $\mathbf{I}_n$ sont des matrices échelonnées. Deux autres exemples de matrices échelonnées sont les suivants:

$$\begin{bmatrix} 2 & 0 & -2 & 1 & 2 & -1 \\ 0 & 0 & 0 & 4 & 3 & 0 \\ 0 & 0 & 0 & 0 & 4 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \qquad \begin{bmatrix} 0 & 2 & -3 & 1 & 0 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

3.3.3 Rang d'une matrice échelonnée

Les colonnes d'indice $j_1 j_2, ..., j_r$ de la matrice (3.16) sont clairement linéairement indépendantes. Envisagées comme des vecteurs-colonnes de $\mathbb{R}^r$, elles forment donc une base de cet espace vectoriel. En considérant les autres colonnes également comme des vecteurs-colonnes de $\mathbb{R}^r$, nous en déduisons qu'elles sont combinaison linéaire de celles d'indice $j_1, j_2, ..., j_r$ et donc que le rang de la matrice (3.16) est r.

On notera que r est le nombre de lignes non nulles de la matrice (3.16) et également le rang de la famille des lignes de cette matrice, puisque les lignes non nulles sont manifestement linéairement indépendantes.

3.3.4 Opérations élémentaires sur les lignes d'une matrice

On appelle *opération élémentaire sur les lignes d'une matrice* toute opération de l'un des trois types suivants:

Type 1: échanger deux lignes.
Type 2: additionner à une ligne un multiple d'une autre ligne.
Type 3: multiplier une ligne par un nombre non nul.

3.3.5 Réduction à la forme échelonnée

Toute matrice peut être *transformée en une matrice échelonnée* par une suite finie d'opérations de type 1 et 2. On dit également que toute matrice peut être *réduite à la forme échelonnée* par une telle suite d'opérations. Soit en effet $\mathbf{A} = (a_{ij})$ une matrice de type $m \times n$ non nulle. Si $m = 1$, $\mathbf{A}$ est échelonnée. Si $m > 1$, désignons par j_1 le plus petit indice de colonne non nulle. En échangeant deux lignes, si nécessaire, nous transformons $\mathbf{A}$ en une matrice $\mathbf{A}^{(1)} = (a_{ij}^{(1)})$ dont le terme $a_{1j_1}^{(1)}$ est non nul. En additionnant ensuite à chaque ligne d'indice $i \geqslant 2$ la première ligne multipliée par $-a_{ij_1}^{(1)}/a_{1j_1}^{(1)}$, nous voyons que $\mathbf{A}^{(1)}$ se réduit soit à une matrice de la forme

$$\begin{bmatrix} a_{1j_1}^{(1)} \;\cdots & & & \cdots\; a_{1n}^{(1)} \\ & a'_{2j_2} \;\cdots & & \cdots\; a'_{2n} \\ 0 & & \cdots\cdots\cdots\cdots & \\ & a'_{mj_2} \;\cdots & & \cdots\; a'_{mn} \end{bmatrix},$$

où les termes $a'_{2j_2}, a'_{3j_2}, ..., a'_{mj_2}$ ne sont pas tous nuls, soit à une matrice dont les lignes d'indice $i \geqslant 2$ sont nulles. Dans le deuxième cas, la matrice est échelonnée. Dans le premier cas, le même procédé de réduction appliqué à la matrice

$$\begin{bmatrix} & a'_{2j_2} & \cdots & \cdots & a'_{2n} \\ 0 & & \cdots\cdots\cdots & & \\ & a'_{mj_2} & \cdots & \cdots & a'_{mn} \end{bmatrix}$$

fournit une matrice $\mathbf{A}^{(2)}$, puis une matrice $\mathbf{A}^{(3)}$ et, ainsi de suite, jusqu'à $\mathbf{A}^{(m)}$ ou jusqu'à ce que la seule ligne non nulle d'une matrice $\mathbf{A}^{(r)}$ soit la première. A ce point, la matrice transformée est de la forme

$$\begin{bmatrix} a^{(1)}_{1j_1} & \cdots & & & & \cdots & a^{(1)}_{1n} \\ & & a^{(2)}_{2j_2} & \cdots & & \cdots & a^{(2)}_{2n} \\ & & & \cdots\cdots\cdots & & & \\ & 0 & & & a^{(r)}_{rj_r} & \cdots & a^{(r)}_{rn} \\ & & & & & & \end{bmatrix}. \tag{3.17}$$

3.3.6 Pivots

Les termes $a^{(1)}_{1j_1}$, $a^{(2)}_{2j_2}$, ..., $a^{(r)}_{rj_r}$ de la matrice (3.17) sont appelés *pivots*. Dans les opérations de réduction, les pivots servent à annuler certains termes des colonnes dont ils font partie. Par extension, on appelle *pivot d'une matrice échelonnée* non nulle tout premier terme non nul d'une ligne non nulle. Les pivots de la matrice (3.16) sont donc a_{1j_1}, a_{2j_2}, ..., a_{rj_r}. D'après 3.3.3, le rang d'une matrice échelonnée est égal au nombre de pivots qu'elle comprend.

3.3.7 Réduction à la forme échelonnée simplifiée

Le procédé employé pour parvenir à la matrice échelonnée (3.17) peut être poursuivi. Plus précisément, chaque pivot peut être utilisé pour annuler tous les autres termes de la colonne dont il fait partie. Par exemple, nous annulons le terme $a^{(1)}_{1j_2}$ en additionnant à la première ligne la deuxième multipliée par $-a^{(1)}_{1j_2}/a^{(2)}_{2j_2}$. Après cette suite d'opérations, nous divisons encore chaque ligne non nulle par son pivot. Nous obtenons ainsi une matrice échelonnée dont les pivots sont des 1 et les colonnes dont ils font partie des vecteurs-colonnes de la base canonique de $\mathbb{R}^m$. Une telle matrice est appelée *matrice échelonnée simplifiée*. Ainsi, toute matrice non nulle peut être *transformée en une matrice échelonnée simplifiée* par une suite finie d'opérations élémentaires sur les lignes. On dit également que toute matrice non nulle peut être *réduite à la forme échelonnée simplifiée* par une telle suite d'opérations.

Voici deux exemples de matrices échelonnées simplifiées:

$$\begin{bmatrix} 0 & 1 & 2 & 0 & 3 & 0 & -2 \\ 0 & 0 & 0 & 1 & -2 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \qquad \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -3 \\ 0 & 0 & 1 & 1 \end{bmatrix}.$$

3.3.8 Exemple de réduction

(1) Nous allons réduire la matrice suivante à la forme échelonnée:

$$\begin{bmatrix} 5 & 10 & 11 & -1 & 1 & -41 & 0 \\ -3 & -6 & -7 & 1 & -1 & 27 & 0 \\ 6 & 12 & 14 & -2 & 2 & -52 & 6 \\ 2 & 4 & 5 & -1 & 1 & -19 & 3 \end{bmatrix} \begin{matrix} L_1 - 2L_4 \\ L_2 + L_4 \\ L_3 - 3L_4 \\ \end{matrix} \tag{3.18}$$

$$\begin{bmatrix} 1 & 2 & 1 & 1 & -1 & -3 & -6 \\ -1 & -2 & -2 & 0 & 0 & 8 & 3 \\ 0 & 0 & -1 & 1 & -1 & 5 & -3 \\ 2 & 4 & 5 & -1 & 1 & -19 & 3 \end{bmatrix} \begin{matrix} \\ L_2 + L_1 \\ \\ L_4 - 2L_1 \end{matrix}$$

$$\begin{bmatrix} 1 & 2 & 1 & 1 & -1 & -3 & -6 \\ 0 & 0 & -1 & 1 & -1 & 5 & -3 \\ 0 & 0 & -1 & 1 & -1 & 5 & -3 \\ 0 & 0 & 3 & -3 & 3 & -13 & 15 \end{bmatrix} \begin{matrix} \\ \\ L_3 - L_2 \\ L_4 + 3L_2 \end{matrix} \quad \begin{matrix} \text{puis échanger} \\ L_3 \text{ et } L_4 \end{matrix}$$

$$\begin{bmatrix} 1 & 2 & 1 & 1 & -1 & -3 & -6 \\ 0 & 0 & -1 & 1 & -1 & 5 & -3 \\ 0 & 0 & 0 & 0 & 0 & 2 & 6 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}. \tag{3.19}$$

(2) Pour réduire la matrice (3.18) à la forme échelonnée simplifiée, nous multiplions d'abord la deuxième ligne de (3.19) par -1 et divisons la troisième par 2; ensuite nous procédons ainsi:

$$\begin{bmatrix} 1 & 2 & 1 & 1 & -1 & -3 & -6 \\ 0 & 0 & 1 & -1 & 1 & -5 & 3 \\ 0 & 0 & 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{matrix} L_1 - L_2 - 2L_3 \\ L_2 + 5L_3 \\ \\ \end{matrix}$$

$$\begin{bmatrix} 1 & 2 & 0 & 2 & -2 & 0 & -15 \\ 0 & 0 & 1 & -1 & 1 & 0 & 18 \\ 0 & 0 & 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}. \tag{3.20}$$

3.3.9 Remarques

(1) D'après 3.3.5, la réduction d'une matrice à la forme échelonnée peut se faire sans l'emploi d'opérations de type 3. En pratique, toutefois, il est souvent plus commode d'utiliser également ce type d'opération, notamment pour éviter l'introduction de fractions, lorsque les termes de la matrice sont des entiers (cf. exemple 3.3.1).

(2) Le lecteur aura noté que la suite d'opérations réduisant la matrice (3.18) à la forme échelonnée (3.19) n'est pas celle qui est exposée dans 3.3.5. Des modifications du mécanisme de réduction analogues à celles que nous avons opérées sont souvent avantageuses et varient de cas en cas. Notre but évident était d'éviter l'apparition de fractions.

(3) Lorsqu'une ligne a été modifiée, il faut se garder d'utiliser encore sa forme non modifiée. Cela peut entraîner de graves erreurs, comme le montre l'exemple simple suivant:

$$\begin{bmatrix} 1 & 3 \\ 2 & -1 \end{bmatrix} \begin{matrix} \mathbf{L}_1 + \mathbf{L}_2 \\ \mathbf{L}_2 + \mathbf{L}_1 \end{matrix} \qquad \begin{bmatrix} 3 & 2 \\ 3 & 2 \end{bmatrix} \begin{matrix} \mathbf{L}_1 - \mathbf{L}_2 \\ \mathbf{L}_2 - \mathbf{L}_1 \end{matrix} \qquad \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

3.3.10 Conservation du rang

Les opérations élémentaires sur les lignes d'une matrice ne modifient pas son rang. En effet, si $\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_n$ sont les colonnes d'une matrice $\mathbf{A}$ et $\mathbf{a}'_1, \mathbf{a}'_2, ..., \mathbf{a}'_n$ celles de la matrice $\mathbf{A}'$ déduite de $\mathbf{A}$ par une opération de type 1, 2 ou 3, alors toute relation de la forme

$$\alpha_1 \mathbf{a}_{j1} + \alpha_2 \mathbf{a}_{j2} + ... + \alpha_k \mathbf{a}_{jk} = \mathbf{0}$$

implique la relation

$$\alpha_1 \mathbf{a}'_{j1} + \alpha_2 \mathbf{a}'_{j2} + ... + \alpha_k \mathbf{a}'_{jk} = \mathbf{0}$$

et inversement, ce qui prouve qu'à toute famille libre de colonnes de $\mathbf{A}$ correspond une famille libre de colonnes de $\mathbf{A}'$ et inversement, donc que $\mathrm{rg}\mathbf{A} = \mathrm{rg}\mathbf{A}'$.

Il est d'autre part évident que les opérations élémentaires sur les lignes d'une matrice ne modifient pas le rang de la famille des lignes de cette matrice. Or, nous avons observé dans 3.3.3 que le rang de la famille des lignes d'une matrice échelonnée est égal au rang de la famille des colonnes, c'est-à-dire au rang de cette matrice. Compte tenu de 3.3.5, nous en concluons que *le rang de n'importe quelle matrice est également le rang de la famille des lignes de cette matrice.*

Comme corollaire de cette conclusion, il apparaît que

$$\mathrm{rg}\mathbf{A} \leqslant \min(m, n), \tag{3.21}$$

où $\mathbf{A}$ est une matrice de type $m \times n$.

3.3.11 Réduction d'une matrice carrée d'ordre n et de rang n

Le rang étant conservé par les opérations élémentaires sur les lignes, la forme échelonnée simplifiée de toute matrice carrée d'ordre n et de rang n est la matrice-unité $\mathbf{I}_n$.

3.3.12 Calcul du rang d'une matrice

Le rang d'une matrice peut être déterminé par réduction à la forme échelonnée. Très souvent, cependant, ce rang peut être trouvé avant que la forme échelonnée ne soit atteinte. Voici un exemple:

$$\begin{pmatrix} 1 & -1 & 2 & -1 \\ 3 & 1 & 4 & 3 \\ -5 & -3 & -6 & -7 \\ -2 & 2 & -4 & 2 \\ -1 & 1 & 1 & 1 \end{pmatrix} \begin{matrix} \\ \mathbf{L}_2 - 3\mathbf{L}_1 \\ \mathbf{L}_3 + 5\mathbf{L}_1 \\ \mathbf{L}_4 + 2\mathbf{L}_1 \\ \mathbf{L}_5 + \mathbf{L}_1 \end{matrix} \qquad \begin{pmatrix} 1 & -1 & 2 & -1 \\ 0 & 4 & -2 & 6 \\ 0 & -8 & 4 & -12 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 \end{pmatrix}.$$

Le rang de la famille des lignes et donc de la matrice est visiblement 3.

3.4 MÉTHODE DE RÉSOLUTION DE GAUSS

3.4.1 Opérations élémentaires sur les équations d'un système linéaire

Effectuées sur les lignes de la matrice augmentée associée à un système, les opérations de type 1, 2 et 3 sont l'expression d'opérations sur les équations de ce système. Ces opérations transforment le système en un système équivalent. En effet, échanger deux équations ou multiplier une équation par un nombre non nul ne change en rien l'ensemble des solutions. D'autre part, toute solution d'un couple d'équations de la forme

$$\begin{aligned} a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n &= b_i \\ a_{l1}x_1 + a_{l2}x_2 + \dots + a_{ln}x_n &= b_l \end{aligned}$$

est solution du couple d'équations

$$\begin{aligned} (a_{i1} + \alpha a_{l1})x_1 + (a_{i2} + \alpha a_{l2})x_2 + \dots + (a_{in} + \alpha a_{ln})x_n &= b_i + \alpha b_l \\ a_{l1}x_1 + a_{l2}x_2 + \dots + a_{ln}x_n &= b_l \end{aligned}$$

et inversement, ce qui montre que l'addition à une équation d'un multiple d'une autre équation laisse l'ensemble des solutions inchangé.

3.4.2 Résolution d'un système linéaire admettant exactement une solution

Considérons un système de m équations à n inconnues, de matrice associée $\mathbf{A} = (a_{ij})$ et de second membre $\mathbf{b} = (b_i)$. Supposons que ce système admette une solution unique. Nous allons déterminer cette solution par le procédé de réduction. D'après la proposition 3.2.7, $n = \text{rg}\mathbf{A} = \text{rg}(\mathbf{A} \mid \mathbf{b})$. On notera, au passage, que n est, dans ce cas, inférieur ou égal à m, par (3.21). D'après 3.3.5, la matrice augmentée $(\mathbf{A} \mid \mathbf{b})$ peut être réduite à la forme échelonnée par des opérations élémentaires sur les lignes. Puisqu'en vertu de 3.3.10 les rangs de $\mathbf{A}$ et de $(\mathbf{A} \mid \mathbf{b})$ ne sont pas modifiés par les opérations de réduction, la matrice échelonnée aura la forme

$$\begin{pmatrix} a'_{11} & \cdots & & \cdots & b'_1 \\ & a'_{22} & \cdots & \cdots & b'_2 \\ & & \cdots & & \\ & \mathbf{0} & & & \\ & & & a'_{nn} & b'_n \\ & & & & \end{pmatrix},$$

où $a'_{11}, a'_{22}, \ldots, a'_{nn}$ sont les pivots. Le système associé à cette matrice augmentée s'écrit

$$\begin{aligned} a'_{11}x_1 + a'_{12}x_2 + \ldots + a'_{1n}x_n &= b'_1 \\ a'_{22}x_2 + \ldots + a'_{2n}x_n &= b'_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots & \\ a'_{nn}x_n &= b'_n. \end{aligned} \tag{3.22}$$

D'après 3.4.1, ce système est équivalent au système initial. Les inconnues $x_1, x_2, \ldots, x_n$ peuvent être déterminées dans l'ordre inverse, en résolvant les équations successivement de la dernière à la première, par substitution.

A titre d'exemple, nous allons résoudre le système

$$\begin{aligned} x_1 + x_2 + 2x_3 - 3x_4 &= 1 \\ 2x_1 + 3x_2 + 6x_3 - 5x_4 &= 2 \\ 3x_1 - x_2 + 2x_3 - 7x_4 &= 5 \\ x_1 + 2x_2 + 3x_3 - x_4 &= -1. \end{aligned}$$

Réduisons d'abord la matrice augmentée à la forme échelonnée:

$$\begin{pmatrix} 1 & 1 & 2 & -3 & 1 \\ 2 & 3 & 6 & -5 & 2 \\ 3 & -1 & 2 & -7 & 5 \\ 1 & 2 & 3 & -1 & -1 \end{pmatrix} \begin{matrix} \\ \mathbf{L}_2 - 2\mathbf{L}_1 \\ \mathbf{L}_3 - 3\mathbf{L}_1 \\ \mathbf{L}_4 - \mathbf{L}_1 \end{matrix} \qquad \begin{pmatrix} 1 & 1 & 2 & -3 & 1 \\ 0 & 1 & 2 & 1 & 0 \\ 0 & -4 & -4 & 2 & 2 \\ 0 & 1 & 1 & 2 & -2 \end{pmatrix} \begin{matrix} \\ \\ \frac{1}{2}(\mathbf{L}_3 + 4\mathbf{L}_2) \\ \mathbf{L}_4 - \mathbf{L}_2 \end{matrix}$$

$$\begin{pmatrix} 1 & 1 & 2 & -3 & 1 \\ 0 & 1 & 2 & 1 & 0 \\ 0 & 0 & 2 & 3 & 1 \\ 0 & 0 & -1 & 1 & -2 \end{pmatrix} \begin{matrix} \\ \\ \\ 2\mathbf{L}_4 + \mathbf{L}_3 \end{matrix} \qquad \begin{pmatrix} 1 & 1 & 2 & -3 & 1 \\ 0 & 1 & 2 & 1 & 0 \\ 0 & 0 & 2 & 3 & 1 \\ 0 & 0 & 0 & 5 & -3 \end{pmatrix}.$$

Le système associé à cette matrice augmentée s'écrit

$$\begin{aligned} x_1 + x_2 + 2x_3 - 3x_4 &= 1 \\ x_2 + 2x_3 + x_4 &= 0 \\ 2x_3 + 3x_4 &= 1 \\ 5x_4 &= -3. \end{aligned}$$

La dernière équation détermine x_4, puis la troisième x_3 et ainsi de suite. L'unique solution du système est

$$x_1 = -\frac{7}{5},\ x_2 = -\frac{11}{5},\ x_3 = \frac{7}{5},\ x_4 = -\frac{3}{5}.$$

3.4.3 Résolution d'un système linéaire dans le cas général

Considérons à nouveau un système de m équations à n inconnues, de matrice associée $\mathbf{A} = (a_{ij})$ et de second membre $\mathbf{b} = (b_i)$. Si la matrice $\mathbf{A}$ est nulle, ce système n'admet de solutions qu'à la condition que $\mathbf{b}$ aussi soit nul. Sous cette condition, l'ensemble des solutions est manifestement $\mathbb{R}^n$. Si la matrice $\mathbf{A}$ est non nulle, nous transformons la matrice augmentée $(\mathbf{A} \vdots \mathbf{b})$ en une matrice échelonnée, selon ce qui a été établi dans 3.3.5. Le résultat de cette transformation est soit une matrice de la forme

$$\begin{pmatrix} a'_{1j_1} & \cdots & & & \cdots & a'_{1n} & b'_1 \\ & & a'_{2j_2} & \cdots & \cdots & a'_{2n} & b'_2 \\ & & & \cdots\cdots\cdots & & & \\ & 0 & & & a'_{rj_r} \ \cdots & a'_{rn} & b'_r \\ & & & & & & \end{pmatrix}, \tag{3.23}$$

où $a'_{1j_1}, a'_{2j_2}, \ldots, a'_{rj_r}$ sont les pivots, soit une matrice de cette même forme, mais avec un pivot supplémentaire b'_{r+1} dans la dernière colonne. Dans le deuxième cas, le système n'admet aucune solution, car la dernière colonne n'est pas combinaison linéaire des autres colonnes. Dans le premier cas, le rang de la matrice associée au système est égal au rang de la matrice augmentée, donc le système admet au moins une solution, d'après la proposition 3.2.7. Pour trouver l'ensemble des solutions, nous écrivons d'abord le système réduit équivalent au système initial:

$$\begin{aligned} a'_{1j_1}x_{j_1} + \ldots \qquad\qquad \ldots + a'_{1n}x_n &= b'_1 \\ a'_{2j_2}x_{j_2} + \ldots \qquad \ldots + a'_{2n}x_n &= b'_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots & \\ a'_{rj_r}x_{j_r} + \ldots + a'_{rn}x_n &= b'_r. \end{aligned} \tag{3.24}$$

Nous attribuons ensuite des valeurs arbitraires $\alpha_1, \alpha_2, \ldots, \alpha_{n-r}$ aux $n - r$ inconnues d'indice j différent de $j_1, j_2, \ldots, j_r$ (s'il n'y a pas de telles inconnues, le système se réduit au système (3.22)). En groupant alors les termes contenant les valeurs α_j

dans le second membre, nous obtenons un système de r équations à r inconnues (appelées *inconnues principales*) que nous résolvons comme il est indiqué dans 3.4.2. La solution s'exprimera en fonction des valeurs α_j, ces valeurs jouant le rôle de paramètres.

Résolvons, par exemple, le système

$$\begin{aligned} 5x_1 + 10x_2 + 11x_3 - x_4 + x_5 - 41x_6 &= 0 \\ -3x_1 - 6x_2 - 7x_3 + x_4 - x_5 + 27x_6 &= 0 \\ 6x_1 + 12x_2 + 14x_3 - 2x_4 + 2x_5 - 52x_6 &= 6 \\ 2x_1 + 4x_2 + 5x_3 - x_4 + x_5 - 19x_6 &= 3. \end{aligned} \tag{3.25}$$

La matrice augmentée associée à ce système est la matrice (3.18). Par des opérations élémentaires sur les lignes, cette matrice a été réduite à la forme échelonnée (3.19). Le système associé à la matrice sous cette forme s'écrit

$$\begin{aligned} x_1 + 2x_2 + x_3 + x_4 - x_5 - 3x_6 &= -6 \\ -x_3 + x_4 - x_5 + 5x_6 &= -3 \\ 2x_6 &= 6. \end{aligned} \tag{3.26}$$

Posons

$$x_2 = \alpha_1, \ x_4 = \alpha_2, \ x_5 = \alpha_3, \tag{3.27}$$

où α_1, α_2 et α_3 sont des nombres arbitraires. Le système

$$\begin{aligned} x_1 + x_3 - 3x_6 &= -6 - 2\alpha_1 - \alpha_2 + \alpha_3 \\ -x_3 + 5x_6 &= -3 - \alpha_2 + \alpha_3 \\ 2x_6 &= 6 \end{aligned}$$

peut alors être résolu, par substitution, de la dernière à la première équation. La solution générale est

$$\begin{aligned} x_1 &= -15 - 2\alpha_1 - 2\alpha_2 + 2\alpha_3 \\ x_2 &= \alpha_1 \\ x_3 &= 18 + \alpha_2 - \alpha_3 \\ x_4 &= \alpha_2 \\ x_5 &= \alpha_3 \\ x_6 &= 3. \end{aligned} \tag{3.28}$$

3.4.4 Utilisation de la forme échelonnée simplifiée

Lorsque le système à résoudre admet plus d'une solution, il est souvent plus commode de poursuivre la réduction jusqu'à la forme échelonnée simplifiée. On peut alors résoudre le système réduit, comme dans 3.4.3, en attribuant des valeurs arbitraires aux inconnues dont l'indice n'est pas l'indice de colonne d'un pivot et en exprimant les autres inconnues en fonction des valeurs attribuées.

Par exemple, le système (3.25) est équivalent au système associé à la matrice augmentée (3.20) (obtenue par réduction de la matrice (3.18) à la forme échelonnée simplifiée):

$$\begin{aligned} x_1 + 2x_2 \quad + 2x_4 - 2x_5 \quad &= -15 \\ x_3 - x_4 + x_5 \quad &= 18 \\ x_6 &= 3. \end{aligned}$$

On résout ce système en posant $x_2 = \alpha_1$, $x_4 = \alpha_2$, $x_5 = \alpha_3$ et en écrivant, sans plus aucun calcul, x_1, x_3, x_6 en fonction de α_1, α_2, α_3. Le résultat est la solution (3.28).

3.4.5 Résolution simultanée de plusieurs systèmes linéaires

Pour résoudre plusieurs systèmes associés à une même matrice $\mathbf{A} = (a_{ij})$ et dont les seconds membres sont $\mathbf{b} = (b_i)$, $\mathbf{c} = (c_i)$, $\mathbf{d} = (d_i)$, ..., il y a avantage à effectuer les opérations de réduction une seule fois sur la matrice $\mathbf{A}$ augmentée des colonnes $\mathbf{b}$, $\mathbf{c}$, $\mathbf{d}$, ...

Par exemple, pour résoudre les deux systèmes

$$\begin{aligned} x_1 + x_2 + 2x_3 &= 1 \\ 3x_1 + 2x_2 + 3x_3 &= 4 \\ 5x_1 + 4x_2 + 7x_3 &= 6, \end{aligned} \qquad \begin{aligned} x_1 + x_2 + 2x_3 &= 3 \\ 3x_1 + 2x_2 + 3x_3 &= -3 \\ 5x_1 + 4x_2 + 7x_3 &= 3, \end{aligned}$$

on réduit la matrice associée, augmentée des deux seconds membres, à la forme échelonnée simplifiée:

$$\begin{pmatrix} 1 & 1 & 2 & 1 & 3 \\ 3 & 2 & 3 & 4 & -3 \\ 5 & 4 & 7 & 6 & 3 \end{pmatrix} \begin{matrix} \\ -\mathbf{L}_2 + 3\mathbf{L}_1 \\ \mathbf{L}_3 - 5\mathbf{L}_1 \end{matrix}$$

$$\begin{pmatrix} 1 & 1 & 2 & 1 & 3 \\ 0 & 1 & 3 & -1 & 12 \\ 0 & -1 & -3 & 1 & -12 \end{pmatrix} \begin{matrix} \mathbf{L}_1 - \mathbf{L}_2 \\ \\ \mathbf{L}_3 + \mathbf{L}_2 \end{matrix}$$

$$\begin{pmatrix} 1 & 0 & -1 & 2 & -9 \\ 0 & 1 & 3 & -1 & 12 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Les deux systèmes réduits s'écrivent alors

$$\begin{aligned} x_1 \quad - x_3 &= 2 \\ x_2 + 3x_3 &= -1, \end{aligned} \qquad \begin{aligned} x_1 \quad - x_3 &= -9 \\ x_2 + 3x_3 &= 12. \end{aligned}$$

On obtient les solutions en attribuant une valeur arbitraire α à l'inconnue x_3 et en exprimant les deux autres inconnues en fonction de α:

$$\begin{aligned} x_1 &= 2 + \alpha \\ x_2 &= -1 - 3\alpha \\ x_3 &= \alpha, \end{aligned} \qquad \begin{aligned} x_1 &= -9 + \alpha \\ x_2 &= 12 - 3\alpha \\ x_3 &= \alpha. \end{aligned}$$

3.4.6 Choix du pivot

Lorsqu'on résout un système par réduction à la forme échelonnée en calculant avec un nombre limité de chiffres significatifs, on amoindrit les erreurs d'arrondi en choisissant, parmi les coefficients pouvant jouer le rôle de pivot, le plus grand en valeur absolue. Le choix d'un pivot petit en valeur absolue diminue la précision des calculs. Les erreurs proviennent, en effet, de l'addition à certains coefficients de multiples d'autres coefficients divisés par un pivot. Ces erreurs seront d'autant plus petites que le pivot est grand en valeur absolue.

3.4.7 Nombre d'opérations

Le nombre d'opérations nécessaires pour résoudre un système de n équations à n inconnues par réduction de la matrice augmentée associée à la forme échelonnée est, pour n grand, de l'ordre de n^3. La preuve de cette assertion est laissée comme exercice.

3.5 STRUCTURE ET DIMENSION DE L'ENSEMBLE DES SOLUTIONS

3.5.1 Introduction

Pour étudier l'ensemble des solutions d'un système, il est utile de considérer toute solution comme un vecteur-colonne de $\mathbb{R}^n$, où n est le nombre d'inconnues de ce système. L'ensemble des solutions est alors un sous-ensemble de $\mathbb{R}^n$ dont nous allons examiner la structure.

3.5.2 Systèmes linéaires homogènes

Considérons un système homogène de m équations à n inconnues et de matrice associée $\mathbf{A} = (a_{ij})$. Il est clair que $\mathbf{0}$ est une solution (appelée *solution nulle* ou *triviale*). En outre, toute combinaison linéaire $\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 = (\alpha_1 x_j^1 + \alpha_2 x_j^2)$ de solutions $\mathbf{x}_1 = (x_j^1)$ et $\mathbf{x}_2 = (x_j^2)$ est également une solution, car

$$\sum_{j=1}^{n} a_{ij}(\alpha_1 x_j^1 + \alpha_2 x_j^2) = \alpha_1 \sum_{j=1}^{n} a_{ij} x_j^1 + \alpha_2 \sum_{j=1}^{n} a_{ij} x_j^2 = 0, \quad i = 1, 2, ..., m.$$

Cela montre que l'ensemble des solutions est un sous-espace vectoriel de $\mathbb{R}^n$. Si $\mathbf{A}$ est la matrice nulle, il est évident que ce sous-espace est $\mathbb{R}^n$ lui-même. Ecartons ce cas et réduisons $\mathbf{A}$ à la forme échelonnée simplifiée. Supposons, pour alléger l'exposition, que la matrice réduite soit de la forme particulière

$$\begin{pmatrix} 1 & a'_{12} & 0 & a'_{14} & 0 & a'_{16} & a'_{17} \\ 0 & 0 & 1 & a'_{24} & 0 & a'_{26} & a'_{27} \\ 0 & 0 & 0 & 0 & 1 & a'_{36} & a'_{37} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \tag{3.29}$$

En procédant comme il est indiqué dans 3.4.4, c'est-à-dire en attribuant aux inconnues x_2, x_4, x_6, x_7 des valeurs arbitraires α_1, α_2, α_3, α_4, nous voyons que la solution générale du système s'exprime par

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \end{pmatrix} = \alpha_1 \begin{pmatrix} -a'_{12} \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \alpha_2 \begin{pmatrix} -a'_{14} \\ 0 \\ -a'_{24} \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \alpha_3 \begin{pmatrix} -a'_{16} \\ 0 \\ -a'_{26} \\ 0 \\ -a'_{36} \\ 1 \\ 0 \end{pmatrix} + \alpha_4 \begin{pmatrix} -a'_{17} \\ 0 \\ -a'_{27} \\ 0 \\ -a'_{37} \\ 0 \\ 1 \end{pmatrix}. \tag{3.30}$$

Désignons les quatre vecteurs-colonnes du second membre de (3.30) par $\mathbf{x}_1$, $\mathbf{x}_2$, $\mathbf{x}_3$ et $\mathbf{x}_4$. Ces vecteurs sont des solutions du système ($\mathbf{x}_i$ s'obtient en posant $\alpha_i = 1$ et $\alpha_j = 0$ pour $j \neq i$) et, de plus, ils sont linéairement indépendants. En outre, d'après (3.30), toute solution du système est combinaison linéaire de ces quatre vecteurs. Ils forment donc une base du sous-espace vectoriel des solutions, ce sous-espace étant par conséquent de dimension 4. On remarquera que 4 est égal au nombre d'inconnues moins le rang de $\mathbf{A}$. Le raisonnement dans le cas général est le même et entraîne la proposition 3.5.3.

3.5.3 Proposition

Les solutions d'un système linéaire homogène à n inconnues forment un sous-espace vectoriel de $\mathbb{R}^n$ de dimension $n - r$, où r est le rang de la matrice associée au système.

Le corollaire suivant dérive également de la proposition 3.2.7.

3.5.4 Corollaire

Pour qu'un système linéaire homogène à n inconnues n'admette que la solution nulle, il faut et il suffit que $r = n$, où r est le rang de la matrice associée au système.

En raison de (3.21), un système homogène dont le nombre d'équations est inférieur au nombre d'inconnues admet donc toujours des solutions non nulles.

3.5.5 Systèmes linéaires dont le second membre n'est pas nul

Considérons un système de m équations à n inconnues, de matrice associée $\mathbf{A} = (a_{ij})$ et de second membre $\mathbf{b} = (b_i)$. Supposons que ce système admette au moins une solution $\mathbf{x}_0 = (x_j^0)$. Si $\mathbf{y} = (y_j)$ est aussi une solution, alors

$$\sum_{j=1}^{n} a_{ij}(y_j - x_j^0) = \sum_{j=1}^{n} a_{ij}y_j - \sum_{j=1}^{n} a_{ij}x_j^0 = b_i - b_i = 0, \ \ i = 1, 2, ..., m,$$

ce qui montre que $\mathbf{y} - \mathbf{x}_0 = (y_j - x_j^0)$ est une solution du système homogène associé et donc que toute solution du système s'obtient en additionnant à $\mathbf{x}_0$ une solution du système homogène associé. Compte tenu de la proposition 3.5.3, la solution générale s'écrit donc sous la forme

$$\mathbf{x} = \mathbf{x}_0 + \alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 + ... + \alpha_{n-r}\mathbf{x}_{n-r}, \tag{3.31}$$

où $\alpha_1, \alpha_2, ..., \alpha_{n-r}$ sont des nombres arbitraires et $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_{n-r}$ forment une base du sous-espace vectoriel des solutions du système homogène associé (si $r = n$, la solution $\mathbf{x}_0$ est unique). Il en résulte que l'ensemble des solutions du système est un sous-espace affine de $\mathbb{R}^n$, plus exactement le sous-espace affine

$$\mathbf{x}_0 + S, \tag{3.32}$$

où S désigne le sous-espace vectoriel des solutions du système homogène associé (cf. exemples (3) et (4) de 1.9.7).

Les résultats obtenus sont résumés dans les paragraphes 3.5.6 à 3.5.9.

3.5.6 Proposition

Toute solution d'un système linéaire admettant au moins une solution $\mathbf{x}_0$ *s'obtient en additionnant à* $\mathbf{x}_0$ *une solution du système homogène associé.*

3.5.7 Proposition

L'ensemble des solutions d'un système linéaire à n *inconnues admettant au moins une solution est un sous-espace affine de* $\mathbb{R}^n$ *de dimension* $n - r$*, où* r *est le rang de la matrice associée au système.*

D'après cette proposition, un système linéaire dont le nombre d'équations m est inférieur au nombre d'inconnues n n'admet en aucun cas une seule solution, car $n - r > m - r \geqslant 0$.

3.5.8 Corollaire

Pour qu'un système linéaire admette une solution unique, il faut et il suffit qu'il admette au moins une solution et que le système homogène associé n'ait que la solution nulle.

Si le nombre d'équations m, le nombre d'inconnues n et le rang r sont égaux, le rang de la matrice augmentée associée au système ne peut être supérieur à r. Dans ce cas, la proposition suivante peut être considérée comme corollaire de la proposition 3.2.7.

3.5.9 Proposition

Un système linéaire de n équations à n inconnues admet une solution unique si et seulement si le rang de la matrice associée au système est n.

3.6 EXERCICES

3.6.1 A quelle condition doit satisfaire le coefficient α, pour que le système

$$\begin{aligned} x_2 + x_3 &= b_1 \\ x_1 \quad + x_3 &= b_2 \\ x_1 + x_2 + \alpha x_3 &= b_3 \end{aligned}$$

admette au moins une solution quel que soit le choix de b_1, b_2 et b_3? Cette condition étant remplie, ce système a-t-il plus d'une solution?

3.6.2 Dans chacun des cas suivants, dire, sans le résoudre, si le système n'a aucune solution, une seule solution, plus d'une solution:

(a) $$\begin{aligned} 2x_1 - 2x_2 + 4x_3 &= 1 \\ -3x_1 + 3x_2 - 6x_3 &= 2 \\ x_1 - x_2 + 2x_3 &= 1. \end{aligned}$$

(b) $$\begin{aligned} 5x_1 - 3x_2 - x_3 + 4x_4 &= 6 \\ -x_1 + 2x_2 + x_3 - x_4 &= -2 \\ 2x_2 - 2x_3 + 3x_4 &= 3. \end{aligned}$$

(c) $$\begin{aligned} 2x_1 + x_2 + 5x_3 &= 3 \\ x_1 - 2x_2 - 2x_3 &= -1 \\ -3x_1 + 2x_2 + x_3 &= -1 \\ 3x_1 - 3x_2 - 2x_3 &= 0. \end{aligned}$$

3.6.3 Déterminer le polynôme p de degré 3 qui satisfait à l'équation différentielle $\ddot{p}(t) + \dot{p}(t) - p(t) + t - t^3 \equiv 0$, où le point désigne la dérivation par rapport à t.

3.6.4 Réduire les matrices suivantes à la forme échelonnée:

$$\begin{pmatrix} 3 & 2 & 5 & -1 & 2 \\ 4 & -1 & 2 & 3 & -3 \\ -2 & 3 & -1 & 1 & 4 \\ 5 & -2 & 2 & 1 & 2 \end{pmatrix}, \qquad \begin{pmatrix} 2 & -1 & -1 & 4 & -1 & 2 \\ -3 & 2 & 1 & -3 & 2 & -3 \\ 4 & 3 & -4 & 2 & 0 & 5 \\ 5 & 1 & 0 & -3 & 4 & -2 \end{pmatrix}.$$

3.6.5 Décrire l'ensemble des triplets (a, b, c) rendant le rang de la matrice

$$\begin{pmatrix} 1 & 1 & a & b \\ -1 & 2 & b & c \\ 2 & -1 & c & a \end{pmatrix}$$

égal à 2.

3.6.6 Soit $a_1, a_2, \ldots, a_n$ des nombres distincts et k un entier positif inférieur à n. Montrer que le rang de la matrice

$$\begin{bmatrix} 1 & a_1 & a_1^2 & \ldots & a_1^k \\ 1 & a_2 & a_2^2 & \ldots & a_2^k \\ & & \ldots\ldots\ldots & & \\ 1 & a_n & a_n^2 & \ldots & a_n^k \end{bmatrix}$$

est égal à $k+1$.

3.6.7 A quelle condition doivent satisfaire b_1, b_2 et b_3 pour que le système

$$\begin{aligned} x_1 + 2x_2 - 3x_3 &= b_1 \\ 2x_1 + 6x_2 - 11x_3 &= b_2 \\ x_1 - 2x_2 + 7x_3 &= b_3 \end{aligned}$$

admette au moins une solution?

3.6.8 Dans un espace affine de dimension 4 muni d'un repère, on considère les deux plans d'équations paramétriques respectives

$$\begin{aligned} x_1 &= 1 + \alpha_1 + 2\alpha_2 \\ x_2 &= -1 - \alpha_1 \\ x_3 &= 2 + 3\alpha_1 - \alpha_2 \\ x_4 &= 1 + \alpha_1 + \alpha_2, \end{aligned} \qquad \begin{aligned} x_1 &= 2 + \alpha_1' + 2\alpha_2' \\ x_2 &= -2 + 2\alpha_2' \\ x_3 &= 5 + \alpha_1' - 5\alpha_2' \\ x_4 &= 2 + 2\alpha_1' + 3\alpha_2'. \end{aligned}$$

Trouver l'intersection de ces deux plans.

3.6.9 Réduire la matrice augmentée associée au système

$$\begin{aligned} 3x_1 - 3x_2 - x_3 + 2x_4 - 9x_5 &= 13 \\ x_1 - x_2 + 2x_3 - x_4 - 6x_5 &= -6 \\ x_1 - x_2 + x_3 + x_4 - 6x_5 &= 1 \\ -x_1 + x_2 - x_3 - 2x_4 + 7x_5 &= -3 \end{aligned}$$

à la forme échelonnée, puis en tirer la solution générale du système.

3.6.10 Résoudre le système

$$\begin{aligned} ax_1 + bx_2 + cx_3 &= 1 \\ a^2x_1 + b^2x_2 + c^2x_3 &= 1 \\ a^3x_1 + b^3x_2 + c^3x_3 &= 1, \end{aligned}$$

où a, b et c sont des nombres distincts et non nuls.

3.6.11 Par réduction simultanée, résoudre les deux systèmes associés à la matrice

$$\begin{bmatrix} 1 & 5 & 2 \\ -1 & 2 & 1 \\ 5 & 4 & 1 \end{bmatrix},$$

dont les seconds membres respectifs sont $\begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}$ et $\begin{bmatrix} 2 \\ 3 \\ -5 \end{bmatrix}$.

3.6.12 Résoudre le système

$$\begin{aligned}
x_1 &= \phantom{ax_1 + {}} bx_2 && + 1 \\
x_2 &= ax_1 + bx_3 && + 1 \\
x_3 &= \phantom{ax_1 + {}} ax_2 + bx_4 && + 1 \\
&\cdots\cdots\cdots\cdots\cdots\cdots \\
x_{n-1} &= \qquad\qquad ax_{n-2} + bx_n && + 1 \\
x_n &= \qquad\qquad\qquad ax_{n-1} && + 1,
\end{aligned}$$

où a et b sont des nombres distincts non nuls dont la somme est 1. (Poser $x_0 = x_{n+1} = 0$, introduire de nouvelles inconnues $y_j = x_j - x_{j-1}$, $j = 1, 2, ..., n+1$, et résoudre récursivement.)

CHAPITRE 4

Algèbre matricielle

4.1 OPÉRATIONS SUR LES MATRICES

4.1.1 Introduction

Définies dans le chapitre précédent en vue de faciliter l'étude des systèmes linéaires, les matrices deviendront, dès maintenant, des objets pouvant se lier les uns aux autres, en accord avec des règles que nous établirons.

4.1.2 Addition de matrices

Soit $\mathbf{A} = (a_{ij})$ et $\mathbf{B} = (b_{ij})$ des matrices de type $m \times n$. On appelle *somme* de $\mathbf{A}$ et $\mathbf{B}$, et on note $\mathbf{A} + \mathbf{B}$, la matrice $\mathbf{C} = (c_{ij})$ de type $m \times n$ dont les termes sont définis par la relation

$$c_{ij} = a_{ij} + b_{ij}.$$

L'opération qui associe à tout couple de matrices du même type leur somme est appelée *addition matricielle*. Exemple:

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{pmatrix} = \begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} & a_{13}+b_{13} \\ a_{21}+b_{21} & a_{22}+b_{22} & a_{23}+b_{23} \end{pmatrix}.$$

De toute évidence, l'opération d'addition matricielle est associative et commutative:

$$(\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C}), \quad \mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}.$$

L'élément neutre est la matrice nulle et la matrice opposée de $\mathbf{A} = (a_{ij})$ est la matrice $-\mathbf{A} = (-a_{ij})$.

Si $\mathbf{A}$ et $\mathbf{B}$ sont des matrices du même type, la matrice $\mathbf{A} + (-\mathbf{B})$ est notée plus simplement $\mathbf{A} - \mathbf{B}$ et appelée *différence* de $\mathbf{A}$ et $\mathbf{B}$. L'opération qui associe à tout couple de matrices du même type leur différence est appelée *soustraction matricielle*.

L'addition matricielle s'étend de manière évidente au cas d'une famille finie quelconque de matrices du même type.

4.1.3 Multiplication d'une matrice par un scalaire

On appelle *produit du nombre* α *par la matrice* $\mathbf{A}$ la matrice, notée $\alpha\mathbf{A}$, dont les termes sont ceux de $\mathbf{A}$ multipliés par α. L'opération consistant à effectuer le produit d'un nombre par une matrice est appelée *multiplication par un scalaire.* Exemple:

$$\alpha \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} = \begin{bmatrix} \alpha a_{11} & \alpha a_{12} & \alpha a_{13} \\ \alpha a_{21} & \alpha a_{22} & \alpha a_{23} \end{bmatrix}.$$

Les propriétés suivantes découlent immédiatement des définitions:

(a) $(\alpha + \beta)\mathbf{A} = \alpha\mathbf{A} + \beta\mathbf{A}$.
(b) $\alpha(\mathbf{A} + \mathbf{B}) = \alpha\mathbf{A} + \alpha\mathbf{B}$.
(c) $\alpha(\beta\mathbf{A}) = (\alpha\beta)\mathbf{A}$.
(d) $(-1)\mathbf{A} = -\mathbf{A}$, $0\mathbf{A} = \mathbf{O}$, $\alpha\mathbf{O} = \mathbf{O}$.

4.1.4 Espaces vectoriels de matrices

Muni des opérations d'addition et de multiplication par un scalaire, l'ensemble des matrices de type $m \times n$ devient un espace vectoriel de dimension mn. Les matrices dont un des termes est égal à 1 et les autres sont nuls forment une base de cet espace appelée *base canonique.*

4.1.5 Multiplication de matrices

Soit $\mathbf{A} = (a_{ij})$ une matrice de type $m \times n$ et $\mathbf{B} = (b_{jk})$ une matrice de type $n \times p$. On appelle *produit* de $\mathbf{A}$ et $\mathbf{B}$, et on note $\mathbf{AB}$, la matrice $\mathbf{C} = (c_{ik})$ de type $m \times p$ dont les termes sont définis par la relation

$$c_{ik} = a_{i1}b_{1k} + a_{i2}b_{2k} + \dots + a_{in}b_{nk} = \sum_{j=1}^{n} a_{ij}b_{jk}. \tag{4.1}$$

L'opération qui associe à tout couple de matrices leur produit (lorsque celui-ci est défini) est appelée *multiplication matricielle.* Exemples:

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{bmatrix} = \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} & a_{11}b_{13} + a_{12}b_{23} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} & a_{21}b_{13} + a_{22}b_{23} \end{bmatrix},$$

$$\begin{bmatrix} a_1 & a_2 & \dots & a_n \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \cdot \\ \cdot \\ \cdot \\ b_n \end{bmatrix} = \begin{bmatrix} a_1b_1 + a_2b_2 + \dots + a_nb_n \end{bmatrix},$$

$$\begin{bmatrix} a_1 \\ a_2 \\ \cdot \\ \cdot \\ \cdot \\ a_n \end{bmatrix} \begin{bmatrix} b_1 & b_2 & \ldots & b_n \end{bmatrix} = \begin{bmatrix} a_1b_1 & a_1b_2 & \ldots & a_1b_n \\ a_2b_1 & a_2b_2 & \ldots & a_2b_n \\ \ldots & \ldots & \ldots & \ldots \\ a_nb_1 & a_nb_2 & \ldots & a_nb_n \end{bmatrix}.$$

On retiendra que le produit **AB** n'est défini que si le nombre de colonnes de **A** est égal au nombre de lignes de **B**. Schématiquement, cela peut s'exprimer ainsi:

$$ {}_{m}\!\updownarrow \overset{\overset{p}{\leftrightarrow}}{\mathbf{C}} = {}_{m}\!\updownarrow \overset{\overset{n}{\leftrightarrow}}{\mathbf{A}} \;\; {}_{n}\!\updownarrow \overset{\overset{p}{\leftrightarrow}}{\mathbf{B}}, $$

ou encore

$$m \times p = (m \times n)\cdot(n \times p).$$

En particulier, la multiplication est toujours possible entre matrices carrées du même ordre.

Le terme d'indices i, k du produit **AB** est, selon (4.1), le produit scalaire de la i-ième ligne de **A** par la k-ième colonne de **B**, toutes deux considérées comme des vecteurs de $\mathbb{R}^n$.

La multiplication matricielle n'est pas une opération commutative. En fait, le produit **BA** n'est pas nécessairement défini lorsque **AB** l'est. D'autre part, les produits **AB** et **BA** peuvent être définis et ne pas être du même type. Mais même lorsqu'ils sont du même type, ils sont généralement différents, comme le montre l'exemple simple suivant:

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \qquad \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}.$$

En revanche, la multiplication matricielle est une opération associative et distributive. Plus précisément, chacune des égalités suivantes est vraie, à la seule condition que l'un de ses deux membres soit défini (l'autre l'est alors également):

(a) $\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}$ (associativité).

(b) $\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$
(c) $(\mathbf{A} + \mathbf{B})\mathbf{C} = \mathbf{AC} + \mathbf{BC}$ (distributivité).

Il est en outre évident qu'à la même condition,

(d) $(\alpha\mathbf{A})(\beta\mathbf{B}) = (\alpha\beta)\mathbf{AB}$.

A titre d'exemple, démontrons (a). Soit $\mathbf{A} = (a_{ij})$, $\mathbf{B} = (b_{jk})$ et $\mathbf{C} = (c_{kl})$ des matrices respectivement de type $m \times n$, $n \times p$ et $p \times q$. Le terme d'indices i, l de $\mathbf{A}(\mathbf{BC})$ est

$$\sum_{j=1}^{n} a_{ij}\left(\sum_{k=1}^{p} b_{jk}c_{kl}\right)$$

et le terme de mêmes indices de $(\mathbf{AB})\mathbf{C}$ est

$$\sum_{k=1}^{p}\left(\sum_{j=1}^{n} a_{ij}b_{jk}\right)c_{kl}.$$

S'agissant de sommes finies, l'ordre des sommations sur j et sur k peut être inversé, ce qui montre l'égalité des deux expressions.

4.1.6 Produits de plusieurs facteurs et puissances

On définit un produit matriciel de plusieurs facteurs par une succession de produits de deux facteurs. Grâce à l'associativité, un tel produit peut s'écrire sans parenthèses. Par exemple, $\mathbf{ABCD}$ désigne la succession de produits $((\mathbf{AB})\mathbf{C})\mathbf{D}$.

Si $\mathbf{A}$ est une matrice carrée, le produit $\mathbf{AA} \ldots \mathbf{A}$ (k facteurs) est désigné par $\mathbf{A}^k$ et appelé *puissance k-ième* de $\mathbf{A}$. Par convention, $\mathbf{A}^0 = \mathbf{I}$. Il est évident que

$$\mathbf{A}^k\mathbf{A}^l = \mathbf{A}^{k+l} \quad \text{et} \quad (\mathbf{A}^k)^l = \mathbf{A}^{kl}, \tag{4.2}$$

où k et l sont des entiers non négatifs.

4.1.7 Exemples et remarques

(1) Voici deux exemples de multiplication par la matrice-unité $\mathbf{I}_3$:

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix},$$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$

D'une manière générale, on constate aisément que la matrice-unité joue le rôle d'élément neutre de la multiplication matricielle, autrement dit que

$$\mathbf{AI}_n = \mathbf{A} \quad \text{et} \quad \mathbf{I}_m\mathbf{A} = \mathbf{A},$$

pour toute matrice $\mathbf{A}$ de type $m \times n$. Grâce à la propriété (d) de 4.1.5, il s'ensuit que

$$\mathbf{A}(\alpha\mathbf{I}_n) = \alpha\mathbf{A} \quad \text{et} \quad (\alpha\mathbf{I}_m)\mathbf{A} = \alpha\mathbf{A}.$$

(2) On dit que deux matrices carrées du même ordre $\mathbf{A}$ et $\mathbf{B}$ *commutent*, ou que l'une des deux *commute avec* l'autre, si $\mathbf{AB} = \mathbf{BA}$. Nous venons de voir que les matrices $\alpha\mathbf{I}_n$ commutent avec toutes les matrices carrées d'ordre n. Ces matrices sont d'ailleurs les seules qui jouissent de cette propriété (cf. exercice 4.7.5). A noter que si $\mathbf{A}$ et $\mathbf{B}$ commutent,

$$(\mathbf{A} + \mathbf{B})^2 = \mathbf{A}^2 + \mathbf{AB} + \mathbf{BA} + \mathbf{B}^2 = \mathbf{A}^2 + 2\mathbf{AB} + \mathbf{B}^2.$$

Plus généralement, sous la même hypothèse, il est facile de constater, par récurrence sur k, que

$$(\mathbf{A} + \mathbf{B})^k = \sum_{l=0}^{k} \binom{k}{l} \mathbf{A}^l \mathbf{B}^{k-l} \quad \textit{(formule du binôme)}, \tag{4.3}$$

où k désigne un entier positif.

(3) Il est évident que $\mathbf{AB} = \mathbf{O}$ si $\mathbf{A} = \mathbf{O}$ ou $\mathbf{B} = \mathbf{O}$. Mais le produit $\mathbf{AB}$ peut être nul sans que $\mathbf{A}$ ou $\mathbf{B}$ soient des matrices nulles, comme le montre l'exemple simple suivant:

$$\begin{pmatrix} 0 & a \\ 0 & b \end{pmatrix} \begin{pmatrix} c & d \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Il est donc aussi possible que

$$\mathbf{AB} = \mathbf{AC},$$

sans que $\mathbf{A} = \mathbf{O}$ ou que $\mathbf{B} = \mathbf{C}$.

(4) En posant $a = d = 1$ et $b = c = 0$ dans l'exemple ci-dessus, nous voyons que le carré d'une matrice non nulle peut être nul. A l'aide de la multiplication par blocs, nous verrons que la puissance n-ième de toute matrice carrée de la forme

$$\mathbf{A}_n = \begin{pmatrix} 0 & a_{12} & a_{13} & \dots & a_{1n} \\ 0 & 0 & a_{23} & \dots & a_{2n} \\ & & \cdots\cdots & & \\ 0 & 0 & 0 & \dots & a_{n-1,n} \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix}. \tag{4.4}$$

est la matrice nulle.

D'une manière générale, on appelle matrice *nilpotente* toute matrice carrée $\mathbf{A}$ telle que $\mathbf{A}^k = \mathbf{O}$ pour un entier positif k.

(5) L'opération de multiplication matricielle permet d'exprimer tout système linéaire sous la forme d'une équation matricielle. En effet, résoudre un système de matrice associée $\mathbf{A} = (a_{ij})$ et de second membre $\mathbf{b} = (b_i)$ revient à chercher les matrices-colonnes $\mathbf{x} = (x_j)$ telles que

$$\mathbf{Ax} = \mathbf{b},$$

ou encore,

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ & \cdots\cdots & & \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \cdot \\ \cdot \\ \cdot \\ b_m \end{pmatrix}.$$

(6) Soit $\mathbf{A}$ une matrice de type $m \times n$. Si $\mathbf{Ax} = \mathbf{0}$ pour tout vecteur-colonne $\mathbf{x}$ de $\mathbb{R}^n$, alors $\mathbf{A}$ est la matrice nulle. Cela découle immédiatement de la proposition 3.5.3, mais se démontre également en observant que $\mathbf{Ax}$ est la j-ième colonne de $\mathbf{A}$ si $\mathbf{x}$ est le j-ième vecteur-colonne de la base canonique de $\mathbb{R}^n$.

L'assertion suivante en est un corollaire:

Soit $\mathbf{A}$ et $\mathbf{B}$ des matrices de type $m \times n$. Si $\mathbf{Ax} = \mathbf{Bx}$ pour tout vecteur-colonne $\mathbf{x}$ de $\mathbb{R}^n$, alors $\mathbf{A} = \mathbf{B}$.

(7) On appelle *trace* d'une matrice carrée $\mathbf{A}$, et on note $\operatorname{tr}\mathbf{A}$, la somme des termes diagonaux de $\mathbf{A}$.

Si $\mathbf{A} = (a_{ij})$ et $\mathbf{B} = (b_{ij})$ sont deux matrices carrées d'ordre n,

$$\operatorname{tr}(\mathbf{AB}) = \operatorname{tr}(\mathbf{BA}). \tag{4.5}$$

En effet,

$$\operatorname{tr}(\mathbf{AB}) = \sum_{i=1}^{n}\left(\sum_{j=1}^{n} a_{ij}b_{ji}\right) = \sum_{j=1}^{n}\left(\sum_{i=1}^{n} b_{ji}a_{ij}\right) = \operatorname{tr}(\mathbf{BA}).$$

4.1.8 Multiplication matricielle par blocs

On dit qu'une matrice $\mathbf{A}$ est partagée en sous-matrices si elle est de la forme

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} & \cdots & \mathbf{A}_{1r} \\ \mathbf{A}_{21} & \mathbf{A}_{22} & \cdots & \mathbf{A}_{2r} \\ & \cdots & \cdots & \\ \mathbf{A}_{q1} & \mathbf{A}_{q2} & \cdots & \mathbf{A}_{qr} \end{pmatrix}, \tag{4.6}$$

où $\mathbf{A}_{ij}$ est, pour $i = 1, 2, \ldots, q$ et $j = 1, 2, \ldots, r$, une matrice dont le nombre de lignes est indépendant de j et le nombre de colonnes indépendant de i.

On peut multiplier des matrices partagées en sous-matrices, en opérant comme si chaque sous-matrice était un terme ordinaire d'une matrice. La seule condition que doivent remplir les partitions est que les produits matriciels entrant en jeu soient définis. On appelle ce mode de multiplication *multiplication matricielle par blocs*.

Voici un exemple:

$$\begin{array}{cc} & \begin{array}{cc} n & n' \end{array} \\ \begin{array}{c} m \\ m' \end{array} & \left(\begin{array}{c:c} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \hdashline \mathbf{A}_{21} & \mathbf{A}_{22} \end{array}\right) \end{array}
\begin{array}{cc} & \begin{array}{ccc} p & p' & p'' \end{array} \\ \begin{array}{c} n \\ n' \end{array} & \left(\begin{array}{c:c:c} \mathbf{B}_{11} & \mathbf{B}_{12} & \mathbf{B}_{13} \\ \hdashline \mathbf{B}_{21} & \mathbf{B}_{22} & \mathbf{B}_{23} \end{array}\right) \end{array}
=
\begin{array}{cc} & \begin{array}{ccc} p & p' & p'' \end{array} \\ \begin{array}{c} m \\ m' \end{array} & \left(\begin{array}{c:c:c} \mathbf{A}_{11}\mathbf{B}_{11} + \mathbf{A}_{12}\mathbf{B}_{21} & \mathbf{A}_{11}\mathbf{B}_{12} + \mathbf{A}_{12}\mathbf{B}_{22} & \mathbf{A}_{11}\mathbf{B}_{13} + \mathbf{A}_{12}\mathbf{B}_{23} \\ \hdashline \mathbf{A}_{21}\mathbf{B}_{11} + \mathbf{A}_{22}\mathbf{B}_{21} & \mathbf{A}_{21}\mathbf{B}_{12} + \mathbf{A}_{22}\mathbf{B}_{22} & \mathbf{A}_{21}\mathbf{B}_{13} + \mathbf{A}_{22}\mathbf{B}_{23} \end{array}\right) \end{array}.$$

D'une manière générale, si **A** est la matrice (4.6) et

$$\mathbf{B} = \begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} & \dots & \mathbf{B}_{1s} \\ \mathbf{B}_{21} & \mathbf{B}_{22} & \dots & \mathbf{B}_{2s} \\ & & \dots\dots & \\ \mathbf{B}_{r1} & \mathbf{B}_{r2} & \dots & \mathbf{B}_{rs} \end{bmatrix},$$

sous réserve que les produits $\mathbf{A}_{ij}\mathbf{B}_{jk}$ soient définis pour tout choix des indices i, j, k,

$$\mathbf{AB} = \begin{bmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} & \dots & \mathbf{C}_{1s} \\ \mathbf{C}_{21} & \mathbf{C}_{22} & \dots & \mathbf{C}_{2s} \\ & & \dots\dots & \\ \mathbf{C}_{q1} & \mathbf{C}_{q2} & \dots & \mathbf{C}_{qs} \end{bmatrix},$$

où

$$\mathbf{C}_{ik} = \mathbf{A}_{i1}\mathbf{B}_{1k} + \mathbf{A}_{i2}\mathbf{B}_{2k} + \dots + \mathbf{A}_{ir}\mathbf{B}_{rk}. \tag{4.7}$$

En vue d'applications ultérieures, nous allons utiliser la multiplication par blocs pour calculer le produit **AB**, sous l'hypothèse que **B** est partagée en vecteurs-colonnes. Il vient:

$$\mathbf{A}\left[\begin{array}{c:c:c:c} \mathbf{b}_1 & \mathbf{b}_2 & \dots & \mathbf{b}_n \end{array}\right] = \left[\begin{array}{c:c:c:c} \mathbf{Ab}_1 & \mathbf{Ab}_2 & \dots & \mathbf{Ab}_n \end{array}\right]. \tag{4.8}$$

Comme deuxième application, nous nous proposons de démontrer, par récurrence sur n, que la puissance n-ième de la matrice (4.4) est la matrice nulle. Pour $n = 1$ il n'y a rien à démontrer. Supposons que $\mathbf{A}_{n-1}^{n-1} = \mathbf{O}$ et partageons $\mathbf{A}_n$ de la manière suivante:

$$\mathbf{A}_n = \left[\begin{array}{cccc:c} & & & & a_{1n} \\ & & & & \cdot \\ & \mathbf{A}_{n-1} & & & \cdot \\ & & & & \cdot \\ & & & & a_{n-1,n} \\ \hdashline 0 & 0 & \dots & 0 & 0 \end{array}\right].$$

En multipliant par blocs, nous voyons aisément que

$$\mathbf{A}_n^2 = \left[\begin{array}{cccc:c} & & & & \times \\ & & & & \cdot \\ & \mathbf{A}_{n-1}^2 & & & \cdot \\ & & & & \cdot \\ & & & & \times \\ \hdashline 0 & 0 & \dots & 0 & 0 \end{array}\right], \quad \dots, \quad \mathbf{A}_n^{n-1} = \left[\begin{array}{cccc:c} & & & & \times \\ & & & & \cdot \\ & \mathbf{A}_{n-1}^{n-1} & & & \cdot \\ & & & & \cdot \\ & & & & \times \\ \hdashline 0 & 0 & \dots & 0 & 0 \end{array}\right],$$

où les croix occupent la place de termes numériques qu'il n'est pas nécessaire de spécifier. Mais alors

$$\mathbf{A}_n^n = \mathbf{A}_n^{n-1}\mathbf{A}_n = \left[\begin{array}{ccc:c} & & & \times \\ & & & \cdot \\ & \mathbf{O} & & \cdot \\ & & & \cdot \\ & & & \times \\ \hdashline 0\ 0 & \dots & 0 & 0 \end{array}\right] \left[\begin{array}{ccc:c} & & & a_{1n} \\ & & & \cdot \\ & \mathbf{A}_{n-1} & & \cdot \\ & & & \cdot \\ & & & a_{n-1,n} \\ \hdashline 0\ 0 & \dots & 0 & 0 \end{array}\right] = \mathbf{O},$$

ce qui achève la démonstration.

4.1.9 Transposition d'une matrice

On appelle *transposée* d'une matrice $\mathbf{A} = (a_{ij})$, et on note ${}^t\mathbf{A} = (a'_{ij})$ la matrice dont les lignes sont les colonnes de $\mathbf{A}$ (et, par conséquent, les colonnes sont les lignes de $\mathbf{A}$). Entre les termes de ${}^t\mathbf{A}$ et ceux de $\mathbf{A}$ il y a donc la relation

$$a'_{ij} = a_{ji}. \tag{4.9}$$

En outre, ${}^t\mathbf{A}$ est de type $n \times m$ si $\mathbf{A}$ est de type $m \times n$.

Voici un exemple de transposition:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}, \qquad {}^t\mathbf{A} = \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \\ a_{13} & a_{23} \end{bmatrix}.$$

On notera que l'opération de transposition laisse toute matrice de type 1×1 inchangée, transforme les matrices-lignes en matrices-colonnes et les matrices-colonnes en matrices-lignes.

Chacune des égalités suivantes est vraie chaque fois que l'un de ses deux membres est défini (l'autre l'est alors également):

(a) ${}^t(\mathbf{A} + \mathbf{B}) = {}^t\mathbf{A} + {}^t\mathbf{B}$.
(b) ${}^t(\alpha\mathbf{A}) = \alpha\,{}^t\mathbf{A}$.
(c) ${}^t({}^t\mathbf{A}) = \mathbf{A}$.
(d) ${}^t(\mathbf{AB}) = {}^t\mathbf{B}\,{}^t\mathbf{A}$.

La seule qui ne soit pas évidente est la dernière. En voici la preuve. Soit $\mathbf{A} = (a_{ij})$ une matrice de type $m \times n$ et $\mathbf{B} = (b_{jk})$ une matrice de type $n \times p$. Le terme d'indices k, i de ${}^t(\mathbf{AB})$ est le terme d'indices i, k de $\mathbf{AB}$, c'est-à-dire $\sum_{j=1}^{n} a_{ij}b_{jk}$. Le terme d'indices k, i de ${}^t\mathbf{B}\,{}^t\mathbf{A}$ est $\sum_{j=1}^{n} b'_{kj}a'_{ji}$. L'égalité a donc bien lieu en vertu de (4.9).

L'extension de (d) à une famille finie quelconque de matrices est immédiate:

(e) ${}^t(\mathbf{A}_1\mathbf{A}_2 \dots \mathbf{A}_k) = {}^t\mathbf{A}_k\,{}^t\mathbf{A}_{k-1} \dots {}^t\mathbf{A}_1$.

En particulier, lorsque $\mathbf{A}$ est une matrice carrée (et k un entier positif),

(f) ${}^t(\mathbf{A}^k) = ({}^t\mathbf{A})^k$.

4.1.10 Rang d'une matrice transposée

Pour toute matrice $\mathbf{A}$

$$\operatorname{rg}{}^t\mathbf{A} = \operatorname{rg}\mathbf{A}, \tag{4.10}$$

car la famille des lignes et celle des colonnes d'une même matrice ont le même rang (cf. 3.3.10).

4.2 MATRICES INVERSIBLES

4.2.1 Introduction

On sait que tout nombre non nul α admet un inverse unique, c'est-à-dire un nombre, noté α^{-1}, tel que $\alpha\alpha^{-1} = 1$. En est-il de même pour les matrices? La réponse à cette question est négative. Par exemple, le produit de la matrice

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

par une matrice quelconque

$$\mathbf{B} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

n'est jamais $\mathbf{I}_2$, car

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b \\ 0 & 0 \end{pmatrix}.$$

Dans cette section, nous nous proposons d'étudier les matrices qui admettent un inverse et d'exposer une méthode de calcul de cet inverse.

4.2.2 Matrices inversibles

On dit qu'une matrice carrée $\mathbf{A}$ est *inversible* ou *régulière* s'il existe une matrice $\mathbf{B}$ (nécessairement carrée et du même ordre que $\mathbf{A}$) telle que

$$\mathbf{AB} = \mathbf{BA} = \mathbf{I}. \tag{4.11}$$

On notera que si $\mathbf{AB} = \mathbf{I}$ et $\mathbf{B}'\mathbf{A} = \mathbf{I}$, alors $\mathbf{B} = \mathbf{B}'$, car $\mathbf{B}' = \mathbf{B}'(\mathbf{AB}) = (\mathbf{B}'\mathbf{A})\mathbf{B} = \mathbf{B}$. Lorsque $\mathbf{A}$ est inversible, la matrice $\mathbf{B}$ de la relation (4.11) est donc unique. Elle est appelée *matrice inverse* de $\mathbf{A}$ et notée $\mathbf{A}^{-1}$.

4.2.3 Proposition

Soit $\mathbf{A}$ *une matrice carrée. S'il existe une matrice* $\mathbf{B}$ *(nécessairement carrée et du même ordre que* $\mathbf{A}$*) telle que* $\mathbf{AB} = \mathbf{I}$ *ou* $\mathbf{BA} = \mathbf{I}$*, alors* $\mathbf{B}$ *vérifie* (4.11), *autrement dit,* $\mathbf{A}$ *est inversible et* $\mathbf{B} = \mathbf{A}^{-1}$.

DÉMONSTRATION

Il suffit de montrer que $\mathbf{AB} = \mathbf{I}$ entraîne $\mathbf{BA} = \mathbf{I}$. Soit n l'ordre de $\mathbf{A}$ et de $\mathbf{B}$. Le système homogène de matrice associée $\mathbf{B}$ n'admet que la solution triviale, car si $\mathbf{x}$ est une solution,

$$\mathbf{x} = \mathbf{ABx} = \mathbf{A(Bx)} = \mathbf{A0} = \mathbf{0}.$$

Il s'ensuit, d'après le corollaire 3.5.4, que le rang de $\mathbf{B}$ est n. Soit maintenant $\mathbf{x}$ un vecteur-colonne quelconque de $\mathbb{R}^n$. En vertu de la proposition 3.5.9, le système de matrice associée $\mathbf{B}$ et de second membre $\mathbf{x}$ admet une solution (unique) $\mathbf{y}$. Par conséquent,

$$\mathbf{BAx} = \mathbf{BA(By)} = \mathbf{B(AB)y} = \mathbf{By} = \mathbf{x}.$$

Comme $\mathbf{x}$ est arbitraire, la remarque (6) de 4.1.7 nous permet de conclure que $\mathbf{BA} = \mathbf{I}$.

4.2.4 Proposition. Rang d'une matrice inversible

Une matrice carrée $\mathbf{A}$ *d'ordre* n *est inversible si et seulement si son rang est* n.

DÉMONSTRATION

Supposons que $\mathbf{A}$ soit inversible. Alors $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$, donc le rang de $\mathbf{A}$ est n, d'après ce qui a été établi dans la première partie de la démonstration précédente. Réciproquement, supposons que le rang de $\mathbf{A}$ soit n. Le procédé exposé dans le paragraphe suivant démontre l'existence d'une matrice $\mathbf{B}$ telle que $\mathbf{AB} = \mathbf{I}$. D'après la proposition 4.2.3, $\mathbf{A}$ est donc inversible.

4.2.5 Calcul de la matrice inverse

Soit $\mathbf{A}$ une matrice carrée d'ordre n et de rang n. Nous nous proposons de résoudre l'équation matricielle

$$\mathbf{AX} = \mathbf{I}, \tag{4.12}$$

où $\mathbf{X}$ est une matrice carrée inconnue. En vertu de (4.8), cette équation s'exprime, de manière équivalente, par les n équations matricielles

$$\mathbf{Ax}_1 = \mathbf{e}_1,\ \mathbf{Ax}_2 = \mathbf{e}_2,\ \ldots,\ \mathbf{Ax}_n = \mathbf{e}_n, \tag{4.13}$$

où $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_n$ sont les colonnes de $\mathbf{X}$ et $\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n$ celles de $\mathbf{I}$. Les équations (4.13) sont l'expression matricielle de n systèmes de matrice associée commune $\mathbf{A}$ et de seconds membres $\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n$. Suivant 3.4.5, ces systèmes peuvent être résolus simultanément en réduisant la matrice augmentée n fois

$$\left[\begin{array}{c|c|c|c} \mathbf{A} & \mathbf{e}_1 & \cdots & \mathbf{e}_n \end{array}\right] = \left[\begin{array}{c|c} \mathbf{A} & \mathbf{I} \end{array}\right]$$

à la forme échelonnée simplifiée. Compte tenu de 3.3.11, le résultat de cette réduction est une matrice de la forme

$$\left[\begin{array}{c|cccc} & b_{11} & b_{12} & \cdots & b_{1n} \\ & b_{21} & b_{22} & \cdots & b_{2n} \\ \mathbf{I} & & \cdots & \cdots & \\ & b_{n1} & b_{n2} & \cdots & b_{nn} \end{array}\right]. \tag{4.14}$$

Posons $\mathbf{B} = (b_{ij})$ et désignons par $\mathbf{b}_1, \mathbf{b}_2, ..., \mathbf{b}_n$ les colonnes de $\mathbf{B}$. Les solutions des n systèmes réduits associés à la matrice augmentée (4.14) sont

$$\mathbf{x}_1 = \mathbf{b}_1, \ \mathbf{x}_2 = \mathbf{b}_2, \ ..., \ \mathbf{x}_n = \mathbf{b}_n.$$

Par conséquent, $\mathbf{B}$ est une solution de l'équation (4.12). La proposition 4.2.3 nous permet alors de conclure que $\mathbf{A}$ est inversible et que $\mathbf{B} = \mathbf{A}^{-1}$.

Une méthode de calcul de la matrice inverse faisant appel aux déterminants sera présentée dans 5.3.6.

4.2.6 Proposition. Propriétés de l'opération d'inversion

Soit $\mathbf{A}$ *et* $\mathbf{B}$ *des matrices carrées du même ordre.*
(a) *Si* $\mathbf{A}$ *est inversible,* $\mathbf{A}^{-1}$ *et* ${}^t\mathbf{A}$ *le sont également et* $(\mathbf{A}^{-1})^{-1} = \mathbf{A}$, $({}^t\mathbf{A})^{-1} = {}^t(\mathbf{A}^{-1})$.
(b) *Si* $\mathbf{A}$ *et* $\mathbf{B}$ *sont inversibles, le produit* $\mathbf{AB}$ *l'est également et* $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$.
(c) *Si le produit* $\mathbf{AB}$ *est inversible,* $\mathbf{A}$ *et* $\mathbf{B}$ *le sont également.*

DÉMONSTRATION

Assertion (a). Par définition de l'inverse, $\mathbf{A}$ est la matrice inverse de $\mathbf{A}^{-1}$. D'autre part, en vertu de (d) de 4.1.9, ${}^t\mathbf{A}\,{}^t(\mathbf{A}^{-1}) = {}^t(\mathbf{A}^{-1}\mathbf{A}) = {}^t\mathbf{I} = \mathbf{I}$, donc, par la proposition 4.2.3, ${}^t(\mathbf{A}^{-1})$ est la matrice inverse de ${}^t\mathbf{A}$.

Assertion (b). La conclusion procède encore de la proposition 4.2.3, car $(\mathbf{AB})(\mathbf{B}^{-1}\mathbf{A}^{-1}) = \mathbf{A}(\mathbf{BB}^{-1})\mathbf{A}^{-1} = \mathbf{I}$.

Assertion (c). On peut écrire $\mathbf{A}(\mathbf{B}(\mathbf{AB})^{-1}) = (\mathbf{AB})(\mathbf{AB})^{-1} = \mathbf{I}$, ainsi que $((\mathbf{AB})^{-1}\mathbf{A})\mathbf{B} = (\mathbf{AB})^{-1}(\mathbf{AB}) = \mathbf{I}$, d'où la conclusion s'ensuit, grâce à la proposition 4.2.3.

4.2.7 Corollaire. Inversion d'un produit de plusieurs facteurs et d'une puissance

Si $\mathbf{A}_1$, $\mathbf{A}_2$, ..., $\mathbf{A}_k$ *sont des matrices inversibles du même ordre, le produit* $\mathbf{A}_1\mathbf{A}_2...\mathbf{A}_k$ *est une matrice inversible et* $(\mathbf{A}_1\mathbf{A}_2...\mathbf{A}_k)^{-1} = \mathbf{A}_k^{-1}\mathbf{A}_{k-1}^{-1}...\mathbf{A}_1^{-1}$. *En particulier, si* $\mathbf{A}$ *est une matrice inversible, la matrice* $\mathbf{A}^k$ *l'est également pour tout entier positif* k *et* $(\mathbf{A}^k)^{-1} = (\mathbf{A}^{-1})^k$.

4.2.8 Remarques

(1) On désigne la matrice $(\mathbf{A}^{-1})^k$ plus simplement par $\mathbf{A}^{-k}$. La notion de puissance est ainsi étendue aux exposants entiers négatifs. A noter que lorsque $\mathbf{A}$ est inversible, les relations (4.2) sont vraies pour tous les entiers k et l; en outre, grâce à (a) de la proposition 4.2.6, la relation (f) de 4.1.9 s'étend à tous les entiers k.

(2) La somme de deux matrices inversibles n'est en général pas inversible. Par exemple, si $\mathbf{A}$ est inversible, $-\mathbf{A}$ l'est également, mais $\mathbf{A} + (-\mathbf{A}) = \mathbf{O}$ ne l'est pas.

4.2.9 Exemples

(1) La matrice-unité $\mathbf{I}$ est inversible et $\mathbf{I}^{-1} = \mathbf{I}$.

(2) On constate facilement que

$$\begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -4 \end{bmatrix}^{-1} = \begin{bmatrix} 3^{-1} & 0 & 0 \\ 0 & 2^{-1} & 0 \\ 0 & 0 & -4^{-1} \end{bmatrix}.$$

Cet exemple sera généralisé dans la section suivante.

(3) Nous allons calculer la matrice inverse de

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & -2 \\ 1 & 3 & -1 \\ 2 & 1 & -5 \end{bmatrix}$$

par réduction de $(\mathbf{A} \,\vdots\, \mathbf{I})$ à la forme échelonnée simplifiée. Voici la suite des transformations:

$$\begin{bmatrix} 1 & 2 & -2 & 1 & 0 & 0 \\ 1 & 3 & -1 & 0 & 1 & 0 \\ 2 & 1 & -5 & 0 & 0 & 1 \end{bmatrix} \begin{matrix} \\ \mathbf{L}_2 - \mathbf{L}_1 \\ \mathbf{L}_3 - 2\mathbf{L}_1 \end{matrix} \qquad \begin{bmatrix} 1 & 2 & -2 & 1 & 0 & 0 \\ 0 & 1 & 1 & -1 & 1 & 0 \\ 0 & -3 & -1 & -2 & 0 & 1 \end{bmatrix} \begin{matrix} \mathbf{L}_1 - 2\mathbf{L}_2 \\ \\ \mathbf{L}_3 + 3\mathbf{L}_2 \end{matrix}$$

$$\begin{bmatrix} 1 & 0 & -4 & 3 & -2 & 0 \\ 0 & 1 & 1 & -1 & 1 & 0 \\ 0 & 0 & 2 & -5 & 3 & 1 \end{bmatrix} \begin{matrix} \mathbf{L}_1 + 2\mathbf{L}_3 \\ 2\mathbf{L}_2 - \mathbf{L}_3 \\ \\ \end{matrix} \qquad \begin{bmatrix} 1 & 0 & 0 & -7 & 4 & 2 \\ 0 & 2 & 0 & 3 & -1 & -1 \\ 0 & 0 & 2 & -5 & 3 & 1 \end{bmatrix} \begin{matrix} \\ \frac{1}{2}\mathbf{L}_2 \\ \frac{1}{2}\mathbf{L}_3 \end{matrix}.$$

La matrice inverse est donc

$$\mathbf{A}^{-1} = \tfrac{1}{2}\begin{bmatrix} -14 & 8 & 4 \\ 3 & -1 & -1 \\ -5 & 3 & 1 \end{bmatrix}.$$

(4) Nous nous proposons de calculer la matrice inverse de

$$\mathbf{A} = \begin{bmatrix} a & -b & -c & -d \\ b & a & d & -c \\ c & -d & a & b \\ d & c & -b & a \end{bmatrix},$$

où a, b, c et d sont des nombres non tous nuls. Les colonnes de $\mathbf{A}$ sont orthogonales deux à deux et de même norme. Cela entraîne aussitôt que

$$^t\mathbf{A}\mathbf{A} = (a^2 + b^2 + c^2 + d^2)\mathbf{I}_4.$$

Il en résulte que

$$\mathbf{A}^{-1} = \frac{1}{a^2 + b^2 + c^2 + d^2}\,{}^t\mathbf{A}.$$

(5) Soit $\mathbf{A}$ une matrice inversible d'ordre m, $\mathbf{B}$ une matrice inversible d'ordre n et $\mathbf{C}$ une matrice de type $m \times n$. Considérons l'équation matricielle

$$\begin{bmatrix} \mathbf{A} & \mathbf{C} \\ \mathbf{O} & \mathbf{B} \end{bmatrix}\begin{bmatrix} \mathbf{X} & \mathbf{Z} \\ \mathbf{O} & \mathbf{Y} \end{bmatrix} = \begin{bmatrix} \mathbf{I}_m & \mathbf{O} \\ \mathbf{O} & \mathbf{I}_n \end{bmatrix},$$

où $\mathbf{X}$, $\mathbf{Y}$ et $\mathbf{Z}$ sont des matrices inconnues. En multipliant par blocs et en égalant les sous-matrices du produit ainsi obtenu à celles du second membre, nous voyons que

$$\mathbf{AX} = \mathbf{I}_m, \quad \mathbf{BY} = \mathbf{I}_n, \quad \mathbf{AZ} + \mathbf{CY} = \mathbf{O}.$$

Les solutions de ces équations sont

$$\mathbf{X} = \mathbf{A}^{-1}, \quad \mathbf{Y} = \mathbf{B}^{-1}, \quad \mathbf{Z} = -\mathbf{A}^{-1}\mathbf{C}\mathbf{B}^{-1}.$$

Ainsi,

$$\begin{bmatrix} \mathbf{A} & \mathbf{C} \\ \mathbf{O} & \mathbf{B} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{A}^{-1} & -\mathbf{A}^{-1}\mathbf{C}\mathbf{B}^{-1} \\ \mathbf{O} & \mathbf{B}^{-1} \end{bmatrix}. \tag{4.15}$$

4.3 MATRICES CARRÉES PARTICULIÈRES

4.3.1 Matrices de type 1×1

Effectuées sur les matrices de type 1×1 (c'est-à-dire les matrices à un seul terme), les opérations définies dans 4.1.2, 4.1.3, 4.1.5 et 4.2.2 se confondent avec les opérations ordinaires sur les nombres. C'est pourquoi nous identifierons les matrices de ce type aux nombres.

Grâce à cette identification nous pourrons, par exemple, écrire le produit scalaire canonique dans $\mathbb{R}^n$ et la norme correspondante sous la forme

$$(\mathbf{x} \mid \mathbf{y}) = {}^t\mathbf{x}\mathbf{y}, \quad \| \mathbf{x} \| = \sqrt{{}^t\mathbf{x}\mathbf{x}}.$$

4.3.2 Matrices scalaires

On appelle *matrice scalaire* toute matrice de la forme

$$a\mathbf{I} = \begin{bmatrix} a & 0 & \cdots & 0 \\ 0 & a & \cdots & 0 \\ & & \cdots\cdots & \\ 0 & 0 & \cdots & a \end{bmatrix}.$$

Nous avons déjà constaté que la multiplication d'une matrice scalaire $a\mathbf{I}$ par une matrice $\mathbf{A}$ est équivalente à la multiplication du nombre a par $\mathbf{A}$:

$$(a\mathbf{I})\mathbf{A} = a\mathbf{A}, \quad \mathbf{A}(a\mathbf{I}) = a\mathbf{A}.$$

4.3.3 Matrices diagonales

On appelle *matrice diagonale*, et on note $\mathrm{diag}(a_1, a_2, ..., a_n)$, toute matrice de la forme

$$\begin{bmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ & & \cdots\cdots & \\ 0 & 0 & \cdots & a_n \end{bmatrix}.$$

Un calcul facile montre que

$$\mathrm{diag}(a_1, a_2, ..., a_n)\mathrm{diag}(b_1, b_2, ..., b_n) = \mathrm{diag}(a_1b_1, a_2b_2, ..., a_nb_n). \tag{4.16}$$

D'après la proposition 4.2.4, une matrice diagonale est inversible si et seulement si ses termes diagonaux sont tous non nuls. Si cette condition est remplie, il découle de (4.16) que

$$(\mathrm{diag}(a_1, a_2, ..., a_n))^{-1} = \mathrm{diag}(a_1^{-1}, a_2^{-1}, ..., a_n^{-1}).$$

La formule (4.16) s'étend, de manière évidente, au cas d'une famille finie quelconque de matrices diagonales. En particulier,

$$(\mathrm{diag}(a_1, a_2, ..., a_n))^k = \mathrm{diag}(a_1^k, a_2^k, ..., a_n^k), \tag{4.17}$$

où k est un entier quelconque ou un entier non négatif suivant que $\mathrm{diag}(a_1, a_2, ..., a_n)$ est inversible ou non.

Plus généralement, l'effet de la multiplication d'une matrice (non nécessairement carrée) par une matrice diagonale est le suivant:

$$\mathrm{diag}(a_1, a_2, ..., a_n)\begin{pmatrix} \mathbf{L}_1 \\ \hdashline \cdot \\ \cdot \\ \cdot \\ \hdashline \mathbf{L}_n \end{pmatrix} = \begin{pmatrix} a_1\mathbf{L}_1 \\ \hdashline \cdot \\ \cdot \\ \cdot \\ \hdashline a_n\mathbf{L}_n \end{pmatrix},$$

$$\left(\begin{array}{c:c:c} \mathbf{b}_1 & \cdots & \mathbf{b}_n \end{array}\right)\mathrm{diag}(a_1, a_2, ..., a_n) = \left(\begin{array}{c:c:c} a_1\mathbf{b}_1 & \cdots & a_n\mathbf{b}_n \end{array}\right).$$

4.3.4 Matrices triangulaires

On appelle *matrice triangulaire supérieure* toute matrice de la forme

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ & & \cdots\cdots & \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}.$$

On appelle *matrice triangulaire inférieure* la transposée d'une matrice triangulaire supérieure.

Pour qu'une matrice triangulaire supérieure (inférieure) $\mathbf{A}$ soit inversible il faut et il suffit que ses termes diagonaux $a_{11}, a_{22}, ..., a_{nn}$ soient tous non nuls. Cette condition étant remplie, $\mathbf{A}^{-1}$ est encore une matrice triangulaire supérieure (inférieure); en outre, les termes diagonaux de $\mathbf{A}^{-1}$ sont $a_{11}^{-1}, a_{22}^{-1}, ..., a_{nn}^{-1}$. Pour démontrer ces assertions, il suffit d'observer l'effet des opérations élémentaires sur la matrice $(\mathbf{A} \,\vdots\, \mathbf{I})$, lorsqu'on la transforme en une matrice échelonnée simplifiée. On peut également procéder par récurrence sur n en s'appuyant sur la formule (4.15).

4.3.5 Matrices symétriques

On appelle *matrice symétrique* toute matrice égale à sa transposée. Entre les termes a_{ij} d'une matrice symétrique il y a donc la relation

$$a_{ij} = a_{ji}. \tag{4.18}$$

Toute matrice diagonale est évidemment symétrique. Voici un exemple de matrice symétrique non diagonale:

$$\begin{bmatrix} -1 & 2 & 5 \\ 2 & 0 & -1 \\ 5 & -1 & 3 \end{bmatrix}.$$

L'inverse d'une matrice symétrique inversible est encore une matrice symétrique, car, d'après (a) de la proposition 4.2.6,

$$^t(\mathbf{A}^{-1}) = (^t\mathbf{A})^{-1} = \mathbf{A}^{-1}.$$

Si **A** et **B** sont des matrices symétriques du même ordre, le produit **AB** n'est symétrique que si **A** et **B** commutent. En effet, d'après la propriété (d) de 4.1.9,

$$^t(\mathbf{AB}) = {^t\mathbf{B}}\,{^t\mathbf{A}} = \mathbf{BA}.$$

4.3.6 Remarque

L'addition et la multiplication par un scalaire de matrices de l'une quelconque des classes présentées dans cette section est une matrice de la même classe. En d'autres termes, les matrices d'ordre n de chaque classe forment un sous-espace vectoriel de l'espace vectoriel des matrices carrées d'ordre n.

4.4 RETOUR AUX OPÉRATIONS ÉLÉMENTAIRES

4.4.1 Matrices élémentaires

On appelle *matrice élémentaire* toute matrice pouvant se déduire de la matrice-unité par l'une des opérations élémentaires définies dans 3.3.4. Toute matrice carrée d'ordre 1 est donc une matrice élémentaire et toute matrice élémentaire d'ordre supérieur à 1 est de l'un des trois types suivants:

Type 1: $\mathbf{I}_{il}$ ($i \neq l$) se déduit de $\mathbf{I}$ par échange de la i-ième et la l-ième ligne; par exemple, si $i < l$,

$$\mathbf{I}_{il} = \begin{array}{ll} \begin{bmatrix} 1 & \cdot\,\cdot & 0 & \cdot\,\cdot & 0 & \cdot\,\cdot & 0 \\ \cdot & & \cdot & & \cdot & & \cdot \\ 0 & \cdot\,\cdot & 0 & \cdot\,\cdot & 1 & \cdot\,\cdot & 0 \\ \cdot & & \cdot & & \cdot & & \cdot \\ 0 & \cdot\,\cdot & 1 & \cdot\,\cdot & 0 & \cdot\,\cdot & 0 \\ \cdot & & \cdot & & \cdot & & \cdot \\ 0 & \cdot\,\cdot & 0 & \cdot\,\cdot & 0 & \cdot\,\cdot & 1 \end{bmatrix} & \begin{array}{l} \\ \\ \leftarrow i\text{-ième} \\ \\ \leftarrow l\text{-ième} \\ \\ \\ \end{array} \\ \begin{array}{ccc} \uparrow & & \uparrow \\ i\text{-ième} & & l\text{-ième} \end{array} & \end{array} .$$

Type 2: $\mathbf{I}_{il}^{a}$ $(i \neq l)$ se déduit de $\mathbf{I}$ par addition à la i-ième ligne du produit de la l-ième ligne par le nombre a; par exemple, si $i < l$,

$$\mathbf{I}_{il}^{a} = \begin{array}{ll} \begin{bmatrix} 1 & \cdot\,\cdot & 0 & \cdot\,\cdot & 0 & \cdot\,\cdot & 0 \\ \cdot & & \cdot & & \cdot & & \cdot \\ 0 & \cdot\,\cdot & 1 & \cdot\,\cdot & a & \cdot\,\cdot & 0 \\ \cdot & & \cdot & & \cdot & & \cdot \\ 0 & \cdot\,\cdot & 0 & \cdot\,\cdot & 1 & \cdot\,\cdot & 0 \\ \cdot & & \cdot & & \cdot & & \cdot \\ 0 & \cdot\,\cdot & 0 & \cdot\,\cdot & 0 & \cdot\,\cdot & 1 \end{bmatrix} & \begin{array}{l} \\ \\ \leftarrow i\text{-ième} \\ \\ \leftarrow l\text{-ième} \\ \\ \\ \end{array} \\ \begin{array}{ccc} \uparrow & & \uparrow \\ i\text{-ième} & & l\text{-ième} \end{array} & \end{array} .$$

Type 3: $\mathbf{I}_{i}^{a}$ se déduit de $\mathbf{I}$ par multiplication de la i-ième ligne par le nombre a; autrement dit,

$$\mathbf{I}_{i}^{a} = \begin{array}{ll} \begin{bmatrix} 1 & 0 & \cdots & & & \cdots & 0 \\ 0 & 1 & \cdots & & & \cdots & 0 \\ & & \ddots & & & & \\ & & & 1 & & & \\ \cdots & \cdots & \cdots & & a & \cdots & \cdots \\ & & & & & 1 & \\ & & & & & & \ddots \\ 0 & 0 & \cdots & & & \cdots & 1 \end{bmatrix} & \begin{array}{l} \\ \\ \\ \\ \leftarrow i\text{-ième} \\ \\ \\ \\ \end{array} \end{array} .$$

4.4.2 Effet de la multiplication par une matrice élémentaire

L'intérêt des matrices élémentaires vient du fait qu'elles permettent d'exprimer les opérations de réduction d'une matrice par des multiplications matricielles. Supposons que $\mathbf{I}_{il}$, $\mathbf{I}_{il}^{a}$ et $\mathbf{I}_{i}^{a}$ soient d'ordre n. Alors, pour toute matrice $\mathbf{A}$ à n lignes:

(a) $\mathbf{I}_{il}\mathbf{A}$ se déduit de $\mathbf{A}$ par échange de la i-ième et la l-ième ligne.
(b) $\mathbf{I}_{il}^{a}\mathbf{A}$ se déduit de $\mathbf{A}$ par addition à la i-ième ligne du produit de la l-ième ligne par a.
(c) $\mathbf{I}_{i}^{a}\mathbf{A}$ se déduit de $\mathbf{A}$ par multiplication de la i-ième ligne par a.

De même, pour toute matrice $\mathbf{A}$ à n colonnes, $\mathbf{AI}_{il}$, $\mathbf{AI}_{li}^{a}$ et $\mathbf{AI}_{i}^{a}$ se déduisent de $\mathbf{A}$ par les opérations correspondantes sur les colonnes.

4.4.3 Transposée et inverse d'une matrice élémentaire

La transposée et l'inverse d'une matrice élémentaire sont encore des matrices élémentaires. Plus précisément:

(a) ${}^t\mathbf{I}_{il} = \mathbf{I}_{il}$, $\mathbf{I}_{il}^{-1} = \mathbf{I}_{il}$.
(b) ${}^t(\mathbf{I}_{il}^a) = \mathbf{I}_{li}^a$, $(\mathbf{I}_{il}^a)^{-1} = \mathbf{I}_{il}^{-a}$.
(c) ${}^t(\mathbf{I}_i^a) = \mathbf{I}_i^a$, $(\mathbf{I}_i^a)^{-1} = \mathbf{I}_i^{1/a}$ $(a \neq 0)$.

4.4.4 Matrices inversibles comme produits de matrices élémentaires

Selon 3.3.5, 3.3.7 et 4.4.2, toute matrice peut être transformée en une matrice échelonnée ou en une matrice échelonnée simplifiée au moyen d'une suite finie de multiplications par une matrice élémentaire. En particulier, pour toute matrice inversible $\mathbf{A}$, il existe, vu 3.3.11, des matrices élémentaires $\mathbf{M}_1$, $\mathbf{M}_2$, ..., $\mathbf{M}_k$ telles que

$$\mathbf{M}_1\mathbf{M}_2 \dots \mathbf{M}_k\mathbf{A} = \mathbf{I}.$$

En multipliant les deux membres par $(\mathbf{M}_1\mathbf{M}_2 \dots \mathbf{M}_k)^{-1}$, nous en concluons, grâce au corollaire 4.2.7, que

$$\mathbf{A} = (\mathbf{M}_1\mathbf{M}_2 \dots \mathbf{M}_k)^{-1} = \mathbf{M}_k^{-1}\mathbf{M}_{k-1}^{-1} \dots \mathbf{M}_1^{-1},$$

ce qui atteste que toute matrice inversible est un produit de matrices élémentaires.

4.5 FONCTIONS MATRICIELLES

4.5.1 Polynômes matriciels

Si $\mathbf{A}$ est une matrice carrée et $f(x) = \alpha_0 + \alpha_1 x + \dots + \alpha_k x^k$ un polynôme, nous poserons

$$f(\mathbf{A}) = \alpha_0\mathbf{I} + \alpha_1\mathbf{A} + \dots + \alpha_k\mathbf{A}^k. \tag{4.19}$$

La fonction qui associe à toute matrice carrée $\mathbf{A}$ la matrice $f(\mathbf{A})$ est appelée *polynôme matriciel.*

4.5.2 Fonctions matricielles

Nous pouvons définir des fonctions matricielles plus générales que les polynômes à l'aide de séries matricielles. Soit $\mathbf{A}$ une matrice carrée d'ordre n. Nous dirons qu'une série

$$\sum_{k=0}^{\infty} \alpha_k \mathbf{A}^k \tag{4.20}$$

converge (*absolument*) si chacun des n^2 termes de la matrice

$$\sum_{k=0}^{l} \alpha_k \mathbf{A}^k$$

converge (absolument) quand l tend vers l'infini. Cette condition étant remplie, nous appellerons *somme* de la série (4.20) la matrice formée des termes-limites. Cette somme est encore désignée par l'expression (4.20).

A noter que si la matrice $\mathbf{A}$ est nilpotente, la série (4.20) se réduit à une somme finie, car $\mathbf{A}^k = \mathbf{O}$ pour tout entier k assez grand.

Supposons qu'une fonction réelle f soit définie par une série entière dans un voisinage de 0:

$$f(x) = \sum_{k=0}^{\infty} \alpha_k x^k, \quad |x| < r. \tag{4.21}$$

La série (4.20) converge alors absolument pour toute matrice carrée $\mathbf{A}$ d'ordre n telle que $\| \mathbf{A} \| < r$ (cf. exercices 4.7.19 et 4.7.18), où

$$\| \mathbf{A} \| = \left(\sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}^2 \right)^{1/2}. \tag{4.22}$$

Pour une telle matrice $\mathbf{A}$ nous poserons

$$f(\mathbf{A}) = \sum_{k=0}^{\infty} \alpha_k \mathbf{A}^k. \tag{4.23}$$

Dans le cas où $\mathbf{A}$ est diagonale, cette définition de $f(\mathbf{A})$ se réduit, grâce à (4.17), à

$$f(\operatorname{diag}(a_1, a_2, \ldots, a_n)) = \operatorname{diag}(f(a_1), f(a_2), \ldots, f(a_n)). \tag{4.24}$$

4.5.3 Dérivation de fonctions matricielles

On dit qu'une matrice dépendant d'un paramètre réel t est *dérivable* si tous ses termes sont dérivables. Soit $\mathbf{A}(t) = (a_{ij}(t))$ une matrice dérivable. On appelle *dérivée* de $\mathbf{A}(t)$ la matrice $\dot{\mathbf{A}}(t) = (\dot{a}_{ij}(t))$, où $\dot{a}_{ij}(t)$ désigne la dérivée de $a_{ij}(t)$.

Les seules matrices dérivables qui nous intéressent se déduisent de (4.23) par substitution de $t\mathbf{A}$ à $\mathbf{A}$:

$$f(t\mathbf{A}) = \sum_{k=0}^{\infty} \alpha_k t^k \mathbf{A}^k.$$

D'après ce qui a été dit dans 4.5.2, cette série converge absolument pour tout t tel que

$$|t| < \frac{r}{\| \mathbf{A} \|},$$

où r est la borne apparaissant dans (4.21) (par convention $r/0 = \infty$). En outre, $f(t\mathbf{A})$ est dérivable et

$$(f(t\mathbf{A}))^{\boldsymbol{\cdot}} = \mathbf{A}\dot{f}(t\mathbf{A}), \tag{4.25}$$

où le point dans le premier membre indique la dérivation par rapport à t de la matrice entre parenthèses et $\dot{f}$ dans le second membre désigne la dérivée de f. Dans le cas où f est un polynôme, cette relation résulte des calculs suivants:

$$\begin{aligned}(f(t\mathbf{A}))^{\cdot} &= (\alpha_0\mathbf{I} + \alpha_1 t\mathbf{A} + \alpha_2 t^2\mathbf{A}^2 + \dots + \alpha_k t^k\mathbf{A}^k)^{\cdot}\\ &= \alpha_1\mathbf{A} + 2\alpha_2 t\mathbf{A}^2 + \dots + k\alpha_k t^{k-1}\mathbf{A}^k\\ &= \mathbf{A}(\alpha_1\mathbf{I} + 2\alpha_2 t\mathbf{A} + \dots + k\alpha_k t^{k-1}\mathbf{A}^{k-1}) = \mathbf{A}\dot{f}(t\mathbf{A}).\end{aligned}$$

Le cas général s'ensuit par passage à la limite.

4.5.4 Exponentielle d'une matrice

L'*exponentielle* d'une matrice carrée $\mathbf{A}$ s'obtient à partir de (4.21) en posant $\alpha_k = 1/k!$ (dans ce cas, $r = \infty$):

$$\exp(\mathbf{A}) = \sum_{k=0}^{\infty} \frac{1}{k!}\mathbf{A}^k. \tag{4.26}$$

Il est évident que $\exp(\mathbf{O}) = \mathbf{I}$ et $\exp(a\mathbf{I}) = \mathrm{e}^a\mathbf{I}$. En outre, si $\mathbf{A}$ et $\mathbf{B}$ commutent

$$\exp(\mathbf{A} + \mathbf{B}) = \exp(\mathbf{A})\exp(\mathbf{B}). \tag{4.27}$$

En effet, compte tenu de (4.3),

$$\begin{aligned}\exp(\mathbf{A})\exp(\mathbf{B}) &= \left(\sum_{k=0}^{\infty}\frac{1}{k!}\mathbf{A}^k\right)\left(\sum_{l=0}^{\infty}\frac{1}{l!}\mathbf{B}^l\right) = \sum_{l=0}^{\infty}\left(\sum_{k=0}^{l}\frac{1}{k!(l-k)!}\mathbf{A}^k\mathbf{B}^{l-k}\right)\\ &= \sum_{l=0}^{\infty}\frac{1}{l!}(\mathbf{A} + \mathbf{B})^l = \exp(\mathbf{A} + \mathbf{B}).\end{aligned}$$

En particulier,

$$\exp(\mathbf{A})\exp(-\mathbf{A}) = \exp(\mathbf{A} - \mathbf{A}) = \exp(\mathbf{O}) = \mathbf{I},$$

d'où il résulte que $\exp(\mathbf{A})$ est inversible et

$$(\exp(\mathbf{A}))^{-1} = \exp(-\mathbf{A}). \tag{4.28}$$

Appliquée à la fonction exponentielle, la relation (4.25) devient

$$(\exp(t\mathbf{A}))^{\cdot} = \mathbf{A}\exp(t\mathbf{A}). \tag{4.29}$$

4.6 MATRICES DE TRANSITION

4.6.1 Chaînes de Markov finies

Imaginons qu'une particule occupe une parmi n positions possibles, désignées par 1, 2, ..., n et appelées états. Supposons qu'aux instants 0, 1, 2, ... la particule change d'état de manière aléatoire. Ce changement est gouverné par des probabilités appelées *probabilités de transition*. La suite des états successifs occupés par la particule au cours du temps constitue ce que l'on appelle un processus stochastique. Si la probabilité de transition d'un état i à un état j ne dépend que de i et de j, et non pas d'autres états ou de l'instant auquel le changement d'état a lieu, on dit que le processus stochastique est une *chaîne de Markov homogène* ou simplement une *chaîne de Markov*.

4.6.2 Matrices de transition

Etant donné une chaîne de Markov, désignons par a_{ij} la probabilité de transition de l'état i à l'état j en une unité de temps. On appelle la matrice $\mathbf{A} = (a_{ij})$ *matrice de transition* de la chaîne. Cette matrice satisfait aux conditions suivantes:

$$0 \leqslant a_{ij} \leqslant 1, \quad i, j = 1, 2, ..., n,$$
$$a_{i1} + a_{i2} + ... + a_{in} = 1, \quad i = 1, 2, ..., n.$$

La première condition exprime simplement le fait qu'une probabilité se mesure par un nombre compris entre 0 et 1; la deuxième que la chaîne passe, avec certitude, de i à un des états 1, 2, ..., n.

4.6.3 Loi initiale

Supposons que l'état initial de la chaîne ne soit pas fixe, mais gouverné, lui aussi, par des probabilités: $p_j^{(0)}$ est la probabilité qu'à l'instant 0 la chaîne occupe l'état j. Ces probabilités satisfont aux conditions suivantes:

$$0 \leqslant p_j^{(0)} \leqslant 1, \quad j = 1, 2, ..., n,$$
$$p_1^{(0)} + p_2^{(0)} + ... + p_n^{(0)} = 1.$$

On appelle la matrice-ligne

$$\mathbf{L}^{(0)} = \begin{bmatrix} p_1^{(0)} & p_2^{(0)} & \cdots & p_n^{(0)} \end{bmatrix}$$

loi initiale de la chaîne.

4.6.4 Loi à l'instant k

La loi initiale et la matrice de transition d'une chaîne de Markov déterminent la probabilité de tout événement lié à cette chaîne. Par exemple, désignons par $p_j^{(k)}$ la probabilité qu'à l'instant k la chaîne occupe l'état j. On appelle la matrice-ligne

$$\mathbf{L}^{(k)} = \begin{bmatrix} p_1^{(k)} & p_2^{(k)} & \cdots & p_n^{(k)} \end{bmatrix}$$

loi de la chaîne à l'instant k. Il est intuitivement clair que

$$p_j^{(1)} = p_1^{(0)} a_{1j} + p_2^{(0)} a_{2j} + ... + p_n^{(0)} a_{nj}$$

et, plus généralement, pour tout instant positif k, que

$$p_j^{(k)} = p_1^{(k-1)}a_{1j} + p_2^{(k-1)}a_{2j} + \ldots + p_n^{(k-1)}a_{nj}.$$

En termes matriciels cette relation s'écrit

$$\mathbf{L}^{(k)} = \mathbf{L}^{(k-1)}\mathbf{A}.$$

Par récurrence, il s'ensuit que

$$\mathbf{L}^{(k)} = \mathbf{L}^{(0)}\mathbf{A}^k. \tag{4.30}$$

Nous voyons ainsi que la loi initiale et la matrice de transition permettent de calculer la loi de la chaîne à n'importe quel instant.

4.6.5 Exemple

Un magasin vend deux marques m_1 et m_2 d'un certain produit et un client habituel de la maison achète m_1 avec probabilité 0,7 si son dernier achat était m_1 et avec probabilité 0,4 si son dernier achat était m_2. Nous allons calculer les probabilités d'achat de m_1 et de m_2 lors du sixième achat (l'achat initial est fait à l'instant 0). Nous avons à faire à une chaîne de Markov à deux états, l'état 1 désignant la marque m_1 et l'état 2 la marque m_2. La matrice de transition de cette chaîne est

$$\mathbf{A} = \begin{bmatrix} 0{,}7 & 0{,}3 \\ 0{,}4 & 0{,}6 \end{bmatrix}.$$

Supposons, par exemple, que la loi initiale soit

$$\mathbf{L}^{(0)} = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} \end{bmatrix}.$$

La loi à l'instant 5 est alors, d'après (4.30),

$$\mathbf{L}^{(5)} = \mathbf{L}^{(0)}\mathbf{A}^5 \cong \begin{bmatrix} \frac{1}{3} & \frac{2}{3} \end{bmatrix} \begin{bmatrix} 0{,}57247 & 0{,}42753 \\ 0{,}57004 & 0{,}42996 \end{bmatrix} \cong \begin{bmatrix} 0{,}57084 & 0{,}42914 \end{bmatrix}.$$

Lors du sixième achat, les probabilités d'achat de m_1 et de m_2 sont donc respectivement 0,57084 et 0,42914.

On peut se demander si l'influence du choix initial tend à disparaître après un grand nombre d'achats. Nous répondrons à cette question plus loin, lorsque nous disposerons des moyens nécessaires pour étudier le comportement asymptotique des puissances de certaines matrices (cf. 8.4.5).

4.7 EXERCICES

4.7.1 Soit

$$\mathbf{A} = \begin{bmatrix} 2 & -1 & 3 \\ 3 & 2 & -1 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} -1 & 1 & 4 \\ 1 & 2 & -1 \\ -1 & 1 & 2 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}.$$

Calculer, lorsque la multiplication est possible, les produits **AA**, **BB**, **AB**, **AC**, **BC**, ${}^t\mathbf{C}\mathbf{A}$, ${}^t\mathbf{C}\mathbf{B}$, ${}^t\mathbf{C}\mathbf{C}$, $\mathbf{C}{}^t\mathbf{C}$.

4.7.2 Soit $\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_n$ les colonnes d'une matrice **A** de type $m \times n$ et $\mathbf{B}_1, \mathbf{B}_2, ..., \mathbf{B}_n$ les lignes d'une matrice **B** de type $n \times p$. Montrer que

$$\mathbf{AB} = \mathbf{a}_1\mathbf{B}_1 + \mathbf{a}_2\mathbf{B}_2 + ... + \mathbf{a}_n\mathbf{B}_n.$$

4.7.3 Par réduction simultanée, trouver l'ensemble des solutions de l'équation matricielle $\mathbf{AX} = \mathbf{B}$, où

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix} \quad \text{et} \quad \mathbf{B} = \begin{bmatrix} 2 & 3 \\ 3 & 5 \\ 5 & 2 \end{bmatrix}.$$

4.7.4 Ecrire la matrice

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}$$

sous la forme $\mathbf{I} + \mathbf{B}$ et calculer $\mathbf{A}^k$, pour tout nombre entier positif k, à l'aide de la formule du binôme.

4.7.5 Montrer que les seules matrices qui commutent avec la matrice

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & & & & \\ & 0 & 1 & & 0 & \\ & & 0 & 1 & & \\ & & & \ddots & \ddots & \\ & 0 & & & \ddots & 1 \\ & & & & & 0 \end{bmatrix} \quad \text{sont celles de la forme} \quad \begin{bmatrix} a_1 & a_2 & a_3 & \cdots & a_n \\ & a_1 & a_2 & \cdots & a_{n-1} \\ & & a_1 & \cdots & a_{n-2} \\ & 0 & & \ddots & \vdots \\ & & & & a_1 \end{bmatrix}.$$

En déduire que les matrices qui commutent avec **A** et ${}^t\mathbf{A}$ sont de la forme $a\mathbf{I}$. Relever, au passage, que si deux matrices commutent, aucune des deux, en général, ne commute avec la transposée de l'autre.

4.7.6 Par récurrence sur k, démontrer que la formule du binôme (4.3) s'applique à tout couple $(\mathbf{A}, \mathbf{B})$ de matrices carrées qui commutent.

4.7.7 Soit $\mathbf{A}$ une matrice de type $m \times n$ et de rang n.
(a) Montrer que le rang de ${}^t\mathbf{AA}$ est n. (Multiplier les deux membres de l'équation ${}^t\mathbf{AAx} = \mathbf{0}$ par ${}^t\mathbf{x}$, en déduire que $\mathbf{Ax} = \mathbf{0}$ et appliquer le corollaire 3.5.4.)
(b) En déduire que la seule solution de toute équation de la forme $\mathbf{AX} = \mathbf{O}$ est $\mathbf{X} = \mathbf{O}$.

4.7.8 *Méthode des moindres carrés.* Soit $P_1(a_1, b_1)$, $P_2(a_2, b_2)$, ..., $P_n(a_n, b_n)$ n points de $\mathbb{R}^2$ tels que $a_i \neq a_j$ si $i \neq j$. Soit k un entier positif inférieur à n. Pour tout polynôme p de degré inférieur ou égal à k, on pose

$$\delta^2(p) = (b_1 - p(a_1))^2 + (b_2 - p(a_2))^2 + \dots + (b_n - p(a_n))^2.$$

Le but de cet exercice est de trouver le polynôme $p(t) \equiv \alpha_0 + \alpha_1 t + \dots + \alpha_k t^k$ qui minimise $\delta(p)$. On désigne par $\mathbf{p}$ le vecteur-colonne formé des coefficients inconnus $\alpha_0, \alpha_1, \dots, \alpha_k$ et on pose

$$\mathbf{A} = \begin{pmatrix} 1 & a_1 & a_1^2 & \dots & a_1^k \\ 1 & a_2 & a_2^2 & \dots & a_2^k \\ & & \dots\dots & & \\ 1 & a_n & a_n^2 & \dots & a_n^k \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \cdot \\ \cdot \\ \cdot \\ b_n \end{pmatrix}.$$

Les assertions suivantes sont à démontrer:
(a) $\delta(p) = \| \mathbf{b} - \mathbf{Ap} \|$.
(b) $\delta(p)$ est minimal lorsque $\mathbf{Ap}$ est la projection orthogonale de $\mathbf{b}$ sur le sous-espace vectoriel formé des vecteurs-colonnes $\mathbf{Ax}$, $\mathbf{x}$ parcourant $\mathbb{R}^{k+1}$.
(c) $\mathbf{Ap}$ est cette projection lorsque $\mathbf{p}$ est l'unique solution de l'équation

$${}^t\mathbf{AAp} = {}^t\mathbf{Ab}.$$

(Poser que $\mathbf{b} - \mathbf{Ap}$ est orthogonal à $\mathbf{Ax}$ pour tout $\mathbf{x}$ de $\mathbb{R}^{k+1}$; utiliser les exercices 3.6.6 et 4.7.7 pour prouver l'unicité de la solution.)

Application numérique: étant donné $P_1(0, 1)$, $P_2(1, 2)$, $P_3(2, 4)$, $P_4(3, 8)$, trouver le polynôme p de degré inférieur ou égal à 2 qui minimise $\delta(p)$.

4.7.9 Montrer que si $\mathbf{A}$ est une matrice inversible, $\alpha\mathbf{A}$ l'est également pour tout nombre α non nul et $(\alpha\mathbf{A})^{-1} = \frac{1}{\alpha}\mathbf{A}^{-1}$.

4.7.10 Montrer que si deux matrices inversibles $\mathbf{A}$ et $\mathbf{B}$ commutent, alors $\mathbf{A}^{-1}$ et $\mathbf{B}^{-1}$ commutent également.

4.7.11 Montrer que si une matrice inversible $\mathbf{A}$ commute avec une matrice $\mathbf{B}$, alors $\mathbf{A}^{-1}$ commute avec $\mathbf{B}$.

4.7.12 Soit $\mathbf{A}$ une matrice carrée. Soit $\mathbf{B}$ et $\mathbf{C}$ des matrices inversibles du même ordre que $\mathbf{A}$ telles que ${}^t(\mathbf{BA})\mathbf{B}^{-1}(\mathbf{CB}^{-1})^{-1} = \mathbf{I}$. Montrer que $\mathbf{A}$ est inversible et exprimer $\mathbf{A}^{-1}$ en fonction de $\mathbf{B}$ et $\mathbf{C}$. (Appliquer la proposition 4.2.3.)

4.7.13 Soit $\mathbf{A}$ une matrice carrée telle que $\mathbf{A}^3 - \mathbf{A}^2 + 2\mathbf{A} - 3\mathbf{I} = \mathbf{O}$. Montrer que $\mathbf{A}$ est inversible et exprimer $\mathbf{A}^{-1}$ sous la forme d'une combinaison linéaire de puissances entières non négatives de $\mathbf{A}$. (Appliquer la proposition 4.2.3.)

4.7.14 Soit $\mathbf{A}$ une matrice inversible d'ordre m, $\mathbf{B}$ une matrice inversible d'ordre n et $\mathbf{C}$ une matrice de type $m \times n$. Montrer que

$$\begin{bmatrix} \mathbf{C} & \mathbf{A} \\ \mathbf{B} & \mathbf{O} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{O} & \mathbf{B}^{-1} \\ \mathbf{A}^{-1} & -\mathbf{A}^{-1}\mathbf{C}\mathbf{B}^{-1} \end{bmatrix}.$$

A l'aide de cette formule, calculer la matrice inverse de

$$\begin{bmatrix} -1 & 1 & 1 & 0 \\ 1 & -1 & 0 & 1 \\ -1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix}.$$

4.7.15 Calculer la matrice inverse de

$$\begin{bmatrix} 1 & -1 & 2 & 0 \\ 2 & 0 & 2 & 3 \\ 2 & -5 & 1 & 0 \\ 1 & -2 & -2 & 2 \end{bmatrix}.$$

4.7.16 On appelle *matrice antisymétrique* toute matrice égale à sa transposée multipliée par -1.

(a) Montrer que pour toute matrice carrée $\mathbf{A}$, $\mathbf{A} + {}^t\mathbf{A}$ est une matrice symétrique et $\mathbf{A} - {}^t\mathbf{A}$ une matrice antisymétrique.

(b) Montrer que l'ensemble des matrices symétriques d'ordre n et celui des matrices antisymétriques d'ordre n sont des sous-espaces complémentaires dans l'espace vectoriel des matrices carrées d'ordre n.

(c) Trouver la dimension de ces deux sous-espaces et vérifier la relation (1.12).

4.7.17 Soit $\mathbf{D}$ une matrice diagonale d'ordre n dont les termes diagonaux sont distincts. Démontrer que toute matrice diagonale d'ordre n s'écrit de manière unique sous la forme d'une combinaison linéaire des matrices $\mathbf{I}, \mathbf{D}, \mathbf{D}^2, \ldots, \mathbf{D}^{n-1}$. (Utiliser l'exercice 3.6.6.)

4.7.18 Dans l'espace vectoriel des matrices carrées d'ordre n, on considère l'opération $(\cdot \mid \cdot)$ définie par la formule

$$(\mathbf{A} \mid \mathbf{B}) = \operatorname{tr}({}^t\mathbf{A}\mathbf{B}).$$

(a) Montrer que cette opération est un produit scalaire.

(b) Montrer que la norme associée à ce produit scalaire est celle qui est définie dans (4.22).

(c) Montrer que

$$\| \mathbf{AB} \| \leqslant \| \mathbf{A} \| \| \mathbf{B} \|.$$

(d) Montrer que si $\mathbf{x}$ est un vecteur-colonne de $\mathbb{R}^n$,

$$\| \mathbf{Ax} \| \leqslant \| \mathbf{A} \| \| \mathbf{x} \|.$$

4.7.19 Montrer que la série (4.20) converge absolument si $\| \mathbf{A} \| < r$, où r est la borne qui apparaît sous (4.21).

4.7.20 Soit $\mathbf{A}$ une matrice carrée telle que $\| \mathbf{A} \| < 1$. Montrer que

$$(\mathbf{I} - \mathbf{A})^{-1} = \sum_{k=0}^{\infty} \mathbf{A}^k.$$

(Utiliser l'exercice précédent.)

4.7.21 Montrer que si

$$\mathbf{A} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad \text{alors} \quad \exp(t\mathbf{A}) = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}.$$

(Utiliser les développements en série de Taylor des fonctions cos et sin.)

CHAPITRE 5

Déterminants

5.1 DÉFINITION ET PROPRIÉTÉS DES DÉTERMINANTS

5.1.1 Introduction

La notion de déterminant est apparue en relation avec la résolution de certains systèmes linéaires. Elle est également liée à l'idée de volume, comme le montrent les conclusions de 2.7.14 et 2.7.15. Dans cette section, nous nous proposons d'exposer et de commenter les propriétés principales dérivant de la définition de cette nouvelle notion, ainsi que d'étendre le concept de volume à un parallélépipède n-dimensionnel. Dans la section 5.3, nous étudierons les développements et appliquerons les résultats obtenus à la résolution de certains systèmes linéaires, ainsi qu'au calcul de l'inverse d'une matrice.

5.1.2 Notation

Le *déterminant d'une matrice* carrée $\mathbf{A} = (a_{ij})$ est un nombre associé à $\mathbf{A}$ que nous nous apprêtons à définir et noterons

$$|\mathbf{A}| \quad \text{ou} \quad \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & \cdots\cdots & & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix},$$

ou aussi

$$\det(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n),$$

lorsque la dépendance de $\mathbf{A}$ est exprimée au moyen de ses colonnes $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$. En accord avec ce dernier symbole, le déterminant est également appelé *déterminant des vecteurs-colonnes* $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$.

Par la suite, il sera souvent question de colonnes ou de lignes d'un déterminant. Il est entendu que ces colonnes ou ces lignes sont celles de la matrice à laquelle le déterminant est associé.

Si $\mathbf{A}$ est une matrice carrée d'ordre n supérieur à 1, nous désignerons par $\mathbf{A}_{ij}$ la matrice carrée d'ordre $n-1$ obtenue en supprimant la i-ième ligne et la j-ième colonne de $\mathbf{A}$.

5.1.3 Déterminant d'ordre n

On appelle *déterminant d'ordre n* la fonction qui associe à toute matrice carrée d'ordre n son déterminant. Nous définissons cette fonction par récurrence sur n:

- le déterminant d'une matrice carrée $\mathbf{A} = (a)$ d'ordre 1 est a;
- supposons que le déterminant d'ordre $n-1$ ait été défini pour un entier n supérieur à 1; le déterminant d'une matrice carrée $\mathbf{A} = (a_{ij})$ d'ordre n est alors le nombre

$$\begin{aligned} |\mathbf{A}| &= a_{11}|\mathbf{A}_{11}| - a_{21}|\mathbf{A}_{21}| + \ldots + (-1)^{n+1}a_{n1}|\mathbf{A}_{n1}| \\ &= \sum_{i=1}^{n}(-1)^{i+1}a_{i1}|\mathbf{A}_{i1}|. \end{aligned} \tag{5.1}$$

On notera que pour $n = 2$ et $n = 3$, notre définition coïncide avec celle introduite dans 2.7.2:

$$n = 2:\ |\mathbf{A}| = a_{11}|(a_{22})| - a_{21}|(a_{12})| = a_{11}a_{22} - a_{21}a_{12};$$

$$\begin{aligned} n = 3:\ |\mathbf{A}| &= a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{21}\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{31}\begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} \\ &= a_{11}a_{22}a_{33} + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{23} \\ &\quad - a_{11}a_{32}a_{23} - a_{21}a_{12}a_{33} - a_{31}a_{22}a_{13}. \end{aligned}$$

5.1.4 Exemples

(1) Exemple de calcul d'un déterminant:

$$\begin{vmatrix} 0 & 1 & 1 & 2 \\ 3 & 2 & 0 & 1 \\ 0 & 1 & 2 & 4 \\ 2 & 3 & 2 & 0 \end{vmatrix} = -3\begin{vmatrix} 1 & 1 & 2 \\ 1 & 2 & 4 \\ 3 & 2 & 0 \end{vmatrix} - 2\begin{vmatrix} 1 & 1 & 2 \\ 2 & 0 & 1 \\ 1 & 2 & 4 \end{vmatrix}$$

$$= -3(1\begin{vmatrix} 2 & 4 \\ 2 & 0 \end{vmatrix} - 1\begin{vmatrix} 1 & 2 \\ 2 & 0 \end{vmatrix} + 3\begin{vmatrix} 1 & 2 \\ 2 & 4 \end{vmatrix}) - 2(1\begin{vmatrix} 0 & 1 \\ 2 & 4 \end{vmatrix} - 2\begin{vmatrix} 1 & 2 \\ 2 & 4 \end{vmatrix} + 1\begin{vmatrix} 1 & 2 \\ 0 & 1 \end{vmatrix})$$

$$= -3(-8 + 4 + 0) - 2(-2 - 0 + 1) = 14.$$

(2) *Déterminant d'une matrice diagonale.* Par application répétée de (5.1), nous obtenons:

$$\begin{vmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ & \cdots\cdots & & \\ 0 & 0 & \cdots & a_n \end{vmatrix} = a_1 \begin{vmatrix} a_2 & 0 & \cdots & 0 \\ 0 & a_3 & \cdots & 0 \\ & \cdots\cdots & & \\ 0 & 0 & \cdots & a_n \end{vmatrix} = \ldots = a_1 a_2 \ldots a_n. \qquad (5.2)$$

En particulier,

$$| a\mathbf{I}_n | = a^n. \qquad (5.3)$$

(3) *Déterminant d'une matrice triangulaire.* Par application répétée de (5.1), nous obtenons:

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ & \cdots\cdots & & \\ 0 & 0 & \cdots & a_{nn} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & a_{33} & \cdots & a_{3n} \\ & \cdots\cdots & & \\ 0 & 0 & \cdots & a_{nn} \end{vmatrix} = \ldots = a_{11} a_{22} \ldots a_{nn}. \qquad (5.4)$$

Nous verrons dans la proposition 5.1.8 que le déterminant d'une matrice carrée $\mathbf{A}$ est égal au déterminant de sa transposée ${}^t\mathbf{A}$. Nous en déduisons que le déterminant d'une matrice triangulaire inférieure est aussi égal au produit des termes diagonaux.

5.1.5 Théorème. Propriétés fondamentales du déterminant d'ordre n

Le déterminant d'ordre n jouit des propriétés suivantes:

(a) $\det(\mathbf{a}_1, \ldots, \mathbf{a}_j, \ldots, \mathbf{a}_k, \ldots, \mathbf{a}_n) = -\det(\mathbf{a}_1, \ldots, \mathbf{a}_k, \ldots, \mathbf{a}_j, \ldots, \mathbf{a}_n)$ $(n > 1, j < k)$.
(b) $\det(\mathbf{a}_1, \ldots, \alpha\mathbf{a}_j, \ldots, \mathbf{a}_n) = \alpha\det(\mathbf{a}_1, \ldots, \mathbf{a}_j, \ldots, \mathbf{a}_n)$.
(c) $\det(\mathbf{a}_1, \ldots, \mathbf{a}_j + \mathbf{a}'_j, \ldots, \mathbf{a}_n) = \det(\mathbf{a}_1, \ldots, \mathbf{a}_j, \ldots, \mathbf{a}_n) + \det(\mathbf{a}_1, \ldots, \mathbf{a}'_j, \ldots, \mathbf{a}_n)$.
(d) $\det(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n) = 1$, *où* $(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ *est la base canonique de* $\mathbb{R}^n$.

En d'autres termes, le déterminant est multiplié par -1 si deux de ses colonnes sont échangées, il est multiplié par α si une de ses colonnes est multipliée par α, il est additif en chacune de ses colonnes et attribue la valeur 1 à la matrice-unité.

L'union des propriétés (a), (b) et (c) s'énonce en disant que le déterminant d'ordre n est une *forme n-linéaire alternée* dans $\mathbb{R}^n$ (cf. 6.4.3).

Le déterminant d'ordre n n'est pas additif (sauf si $n = 1$); autrement dit, $| \mathbf{A} + \mathbf{B} |$ n'est en général pas égal à $| \mathbf{A} | + | \mathbf{B} |$. Par exemple,

$$\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} \neq \begin{vmatrix} 1 & 0 \\ 0 & 0 \end{vmatrix} + \begin{vmatrix} 0 & 0 \\ 0 & 1 \end{vmatrix}.$$

Nous verrons dans 5.2.4 que les propriétés (a)–(d) caractérisent le déterminant d'ordre n. Ces propriétés en entraînent d'autres, que nous énoncerons dans les deux propositions suivantes et démontrerons, avec le théorème 5.1.5, dans la section 5.2.

On remarquera d'ores et déjà que le déterminant est nul si deux de ses colonnes sont égales ou si l'une d'entre elles est nulle. En effet, un échange de deux colonnes égales laisse le déterminant inchangé et, d'après (a), change en même temps ce déterminant en son opposé, ce qui démontre la première assertion; quant à la seconde, il suffit de poser $\alpha = 0$ dans (b).

On notera en outre què

$$|\alpha \mathbf{A}| = \alpha^n |\mathbf{A}| \tag{5.5}$$

pour toute matrice carrée $\mathbf{A}$ d'ordre n. En effet, cela résulte de (b) appliquée à chaque colonne de $\mathbf{A}$.

5.1.6 Proposition. Propriétés du déterminant d'ordre n

Le déterminant d'ordre n jouit des propriétés suivantes:

(a) $\det(\mathbf{a}_1, ..., \mathbf{a}_j + \alpha\mathbf{a}_k, ..., \mathbf{a}_n) = \det(\mathbf{a}_1, ..., \mathbf{a}_j, ..., \mathbf{a}_n)$ $(n > 1, j \neq k)$, *ou, plus généralement*, $\det(\mathbf{a}_1, ..., \mathbf{a}_j + \sum_{k:k\neq j} \alpha_k\mathbf{a}_k, ..., \mathbf{a}_n) = \det(\mathbf{a}_1, ..., \mathbf{a}_j, ..., \mathbf{a}_n)$.

(b) $\det(\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_n) \neq 0$ *si et seulement si les vecteurs-colonnes* $\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_n$ *sont linéairement indépendants.*

En d'autres termes, le déterminant reste inchangé lorsqu'on additionne à l'une de ses colonnes un multiple d'une autre colonne ou une combinaison linéaire des autres colonnes; en outre, vu la proposition 4.2.4, le déterminant d'une matrice est non nul si et seulement si cette matrice est inversible (ou nul si et seulement si cette matrice n'est pas inversible).

5.1.7 Déterminant d'une matrice élémentaire

Le déterminant d'une matrice élémentaire se calcule par application répétée de (5.1) ($|\mathbf{I}_i^a|$ et $|\mathbf{I}_{il}^a|$ avec $i < l$ sont des cas particuliers de (5.4)). Il peut également être déduit de (d) du théorème 5.1.5 et (a) ou (b) du même théorème, ou (a) de la proposition 5.1.6. Le résultat est le suivant:

$$|\mathbf{I}_{il}| = -1, \quad |\mathbf{I}_{il}^a| = 1, \quad |\mathbf{I}_i^a| = a. \tag{5.6}$$

Vu l'effet de la multiplication par une matrice élémentaire (cf. 4.4.2), il s'ensuit, grâce à (a) et (b) du théorème 5.1.5 et à (a) de la proposition 5.1.6, que

$$|\mathbf{AM}| = |\mathbf{A}|\,|\mathbf{M}| \tag{5.7}$$

pour toute matrice carrée **A** et toute matrice élémentaire **M** du même ordre que **A**.

Des relations (a), (b) et (c) de 4.4.3 et du fait que $|\mathbf{I}_{ii}^{a}| = 1$, il résulte en outre que

$$|{}^t\mathbf{M}| = |\mathbf{M}| \tag{5.8}$$

pour toute matrice élémentaire **M**.

Les propriétés (5.7) et (5.8) serviront à démontrer leur généralisation que voici:

5.1.8 Proposition. Autres propriétés du déterminant d'ordre *n*

Si **A** *et* **B** *sont des matrices carrées d'ordre n, alors*

(a) $|\mathbf{AB}| = |\mathbf{A}|\,|\mathbf{B}|$;
(b) $|{}^t\mathbf{A}| = |\mathbf{A}|$.

En d'autres termes, le déterminant du produit de deux matrices carrées du même ordre est égal au produit de leurs déterminants et le déterminant de la transposée d'une matrice carrée est égal au déterminant de cette matrice.

Il résulte aussitôt de (a) que

$$|\mathbf{A}^k| = |\mathbf{A}|^k \tag{5.9}$$

pour tout entier positif k.

5.1.9 Déterminant comme fonction des lignes

Le déterminant peut être conçu comme fonction de la famille des lignes d'une matrice. Grâce à (b) de la proposition 5.1.8, cette fonction jouit des propriétés énoncées dans le théorème 5.1.5 et dans la proposition 5.1.6, à savoir: elle est une forme n-linéaire alternée, elle attribue la valeur 1 à la famille des lignes de la matrice-unité, sa valeur ne change pas lorsqu'on additionne à une ligne un multiple d'une autre ligne et cette valeur est non nulle si et seulement si les lignes sont linéairement indépendantes.

5.1.10 Déterminant de l'inverse d'une matrice

Soit **A** une matrice inversible. D'après (d) du théorème 5.1.5 et (a) de la proposition 5.1.8,

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

d'où nous déduisons que

$$|\mathbf{A}^{-1}| = |\mathbf{A}|^{-1}. \tag{5.10}$$

Il en découle aussitôt que dans le cas où **A** est inversible, la relation (5.9) est vraie pour tout entier k.

5.1.11 Proposition. Rang d'une matrice par les déterminants

Le rang d'une matrice **A** *(non nécessairement carrée) est le plus grand entier k tel que l'on puisse extraire de* **A** *une sous-matrice carrée d'ordre k et de déterminant non nul.*

DÉMONSTRATION

Soit r ce plus grand entier et m le nombre de lignes de $\mathbf{A} = (a_{ij})$. Désignons par $j_1, j_2, ..., j_r$ les indices de colonne d'une sous-matrice carrée de **A** de déterminant non nul. D'après (b) de la proposition 5.1.6, les colonnes et donc les lignes de cette sous-matrice sont linéairement indépendantes. Il s'ensuit que le rang de la matrice

$$\begin{pmatrix} a_{1j_1} & a_{1j_2} & \cdots & a_{1j_r} \\ a_{2j_1} & a_{2j_2} & \cdots & a_{2j_r} \\ & \cdots\cdots & & \\ a_{mj_1} & a_{mj_2} & \cdots & a_{mj_r} \end{pmatrix}$$

est r, car r de ses lignes sont linéairement indépendantes. Par conséquent, $\mathrm{rg}\mathbf{A} \geqslant r$. D'autre part, r' étant le rang de **A**, nous pouvons extraire de **A** r' colonnes linéairement indépendantes formant une sous-matrice $\mathbf{A}'$ de rang r'. En prenant alors r' lignes linéairement indépendantes de $\mathbf{A}'$, nous constituons une sous-matrice carrée de **A** d'ordre r' et de rang r'. D'après (b) de la proposition 5.1.6, le déterminant de cette sous-matrice est non nul. Par conséquent, $r' = \mathrm{rg}\mathbf{A} \leqslant r$. Nous en concluons que $\mathrm{rg}\mathbf{A} = r$, ce qu'il fallait démontrer.

5.1.12 Volume d'un parallélépipède

Soit $\mathscr{E}$ un espace affine euclidien de dimension finie non nulle n. On appelle *parallélépipède* construit sur les points $P_0, P_1, ..., P_n$ de $\mathscr{E}$ le sous-ensemble de $\mathscr{E}$

$$\{P: \overrightarrow{P_0P} = \sum_{j=1}^{n} \alpha_j \overrightarrow{P_0P_j},\ 0 \leqslant \alpha_j \leqslant 1, j = 1, 2, ..., n\}. \tag{5.11}$$

Munissons $\mathscr{E}$ d'un repère orthonormal et désignons par $\mathbf{a}_j$ le vecteur-colonne de $\mathbb{R}^n$ formé des composantes du vecteur $\overrightarrow{P_0P_j}$. On appelle *volume* du parallélépipède (5.11), et on note $\mathrm{vol}(P_0, P_1, ..., P_n)$, la valeur absolue du déterminant des vecteurs-colonnes $\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_n$:

$$\mathrm{vol}(P_0, P_1, ..., P_n) = |\det(\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_n)|. \tag{5.12}$$

Dans le cas où $n = 1$, un parallélépipède est un intervalle et son volume est la longueur.

Dans le cas où $n = 2$, un parallélépipède est appelé *parallélogramme* et son volume *aire*. D'après (5.12)

$$\mathrm{aire}(P_0, P_1, P_2) = |\det(\mathbf{a}_1, \mathbf{a}_2)| = \left|\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}\right| = |a_{11}a_{22} - a_{21}a_{12}|.$$

Montrons que le dernier membre est bien l'aire ordinaire du parallélogramme construit sur P_0, P_1, P_2. Le carré de celle-ci est égal à

$$\begin{aligned} &\|\overrightarrow{P_0P_1}\|^2\,\|\overrightarrow{P_0P_2}\|^2\sin^2(P_1\hat{P}_0P_2) = \|\mathbf{a}_1\|^2\,\|\mathbf{a}_2\|^2 - \|\mathbf{a}_1\|^2\,\|\mathbf{a}_2\|^2\cos^2(P_1\hat{P}_0P_2) \\ &= \|\mathbf{a}_1\|^2\,\|\mathbf{a}_2\|^2 - (\mathbf{a}_1 \mid \mathbf{a}_2)^2 = (a_{11}a_{22} - a_{21}a_{12})^2, \end{aligned}$$

ce qui démontre l'assertion si P_1 et P_2 sont distincts de P_0; autrement, elle est évidente.

Compte tenu de 2.7.14 et de 2.7.15, nous voyons également que

$$\mathrm{vol}(P_0, P_1, P_2, P_3) = |\det(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3)| = \left|\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}\right|$$

représente le volume ordinaire du parallélépipède construit sur P_0, P_1, P_2, P_3.

Nous montrerons dans 7.1.12 que la définition de volume (5.12) est indépendante du choix du repère orthonormal. De ce fait, le volume d'un cube-unité, c'est-à-dire d'un parallélépipède tel que les vecteurs $\overrightarrow{P_0P_1}$, $\overrightarrow{P_0P_2}$, ..., $\overrightarrow{P_0P_n}$ sont orthogonaux deux à deux et unitaires, est égal à 1, car dans le repère $(P_0; \overrightarrow{P_0P_1}, \overrightarrow{P_0P_2}, ..., \overrightarrow{P_0P_n})$ ce volume s'écrit

$$\mathrm{vol}(P_0, P_1, ..., P_n) = |\det(\mathbf{e}_1, \mathbf{e}_2, .., \mathbf{e}_n)|$$

et vaut donc 1, par la propriété (d) du théorème 5.1.5.

A titre d'exercice, le lecteur pourra écrire les propriétés du volume qui découlent de celles du déterminant et illustrer ces propriétés, à l'aide d'une figure, dans le cas de l'aire et du volume ordinaires.

5.2 DÉMONSTRATIONS DES PROPRIÉTÉS DES DÉTERMINANTS

5.2.1 Démonstration du théorème 5.1.5

Propriété (a). Cette propriété est évidente si $n = 2$. Supposons qu'elle soit vérifiée par le déterminant d'ordre $n - 1$, n étant supérieur à 2. En vertu de (5.1), elle est alors vérifiée par le déterminant d'ordre n dans le cas où $1 < j < k$. Le cas où $j = 1$ et $k > 2$ se réduit à celui où $j = 1$ et $k = 2$ par un triple échange: d'abord de $\mathbf{a}_2$ et $\mathbf{a}_k$, ensuite de $\mathbf{a}_1$ et $\mathbf{a}_k$ et finalement de $\mathbf{a}_1$ et $\mathbf{a}_2$. Le premier et le troisième échange transforment chacun la valeur du déterminant en sa valeur opposée, par ce que nous venons de constater. Le deuxième échange en fait autant, en vertu de ce que nous allons démontrer. Pour tout couple (i, l) tel que $i \neq l$, désignons par $\mathbf{A}_{il,12}$ la matrice carrée d'ordre $n - 2$ déduite de $\mathbf{A}$ par suppression de la i-ième et la l-ième ligne, ainsi que des deux premières colonnes. D'après (5.1),

$$|\mathbf{A}_{i1}| = \sum_{l=1}^{i-1} (-1)^{l+1} a_{l2} |\mathbf{A}_{il,12}| + \sum_{l=i+1}^{n} (-1)^{l} a_{l2} |\mathbf{A}_{il,12}|,$$

où il est entendu que le second membre se réduit à la deuxième somme si $i = 1$ et à la première si $i = n$. Substituons ce second membre à $|\mathbf{A}_{i1}|$ dans (5.1). Il vient:

$$\begin{aligned} |\mathbf{A}| &= \sum_{i=2}^{n} \sum_{l=1}^{i-1} (-1)^{i+l} a_{i1} a_{l2} |\mathbf{A}_{il,12}| - \sum_{i=1}^{n-1} \sum_{l=i+1}^{n} (-1)^{i+l} a_{i1} a_{l2} |\mathbf{A}_{il,12}| \\ &= \sum_{(i,l)\in\Delta_1} (-1)^{i+l} a_{i1} a_{l2} |\mathbf{A}_{il,12}| - \sum_{(i,l)\in\Delta_2} (-1)^{i+l} a_{i1} a_{l2} |\mathbf{A}_{il,12}|, \end{aligned} \tag{5.13}$$

où $\Delta_1 = \{(i, l)\colon\ 2 \leqslant i \leqslant n,\ 1 \leqslant l \leqslant i - 1\}$ et $\Delta_2 = \{(i, l)\colon\ 1 \leqslant i \leqslant l - 1,\ 2 \leqslant l \leqslant n\}$. Si nous échangeons la première et la deuxième colonne de $\mathbf{A}$, les matrices $\mathbf{A}_{il,12}$ ne subissent aucune modification. En revanche, $a_{i1}a_{l2}$ change en $a_{l1}a_{i2}$, pour tout couple (i, l). Nous voyons ainsi que la deuxième somme du dernier membre de (5.13) devient la première et, inversement, la première devient la deuxième. L'échange de la première et la deuxième colonne a donc pour seul effet de multiplier le déterminant de $\mathbf{A}$ par -1, ce qu'il fallait démontrer.

Propriétés (b) *et* (c). Vu la propriété (a), nous pouvons admettre que $j = 1$. Les conclusions découlent alors directement de (5.1).

Propriété (d). Il suffit de poser $a = 1$ dans (5.3).

5.2.2 Démonstration de la proposition 5.1.6

Propriété (a). En vertu de (a) du théorème 5.1.5, nous pouvons admettre que $j = 1$. D'après (b) et (c) du même théorème,

$$\det(\mathbf{a}_1 + \alpha\mathbf{a}_k, \mathbf{a}_2, \ldots, \mathbf{a}_n) = \det(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n) + \alpha\det(\mathbf{a}_k, \mathbf{a}_2, \ldots, \mathbf{a}_n).$$

Puisque k est différent de $j = 1$ par hypothèse, le vecteur-colonne $\mathbf{a}_k$ apparaît deux fois dans le dernier déterminant. Celui-ci est donc nul, d'après ce qui a été remarqué dans 5.1.5. La formulation générale s'ensuit par itération.

Propriété (b). Supposons que les vecteurs-colonnes $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ soient linéairement dépendants. En vertu de la proposition 1.5.5, un de ces vecteurs, par exemple $\mathbf{a}_j$, est combinaison linéaire des autres:

$$\mathbf{a}_j = \sum_{k:k\neq j} \alpha_k \mathbf{a}_k.$$

Il s'ensuit, grâce à la propriété (a) déjà établie, que

$$\det(\mathbf{a}_1, ..., \mathbf{a}_j, ..., \mathbf{a}_n) = \det(\mathbf{a}_1, ..., \mathbf{a}_j - \sum_{k:k \neq j} \alpha_k \mathbf{a}_k, ..., \mathbf{a}_n)$$
$$= \det(\mathbf{a}_1, ..., \mathbf{0}, ..., \mathbf{a}_n) = 0.$$

Inversement, supposons que les vecteurs-colonnes $\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_n$ soient linéairement indépendants. Selon la proposition 4.2.4, la matrice **A** de colonnes $\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_n$ est alors inversible. D'après (d) du théorème 5.1.5 et (a) de la proposition 5.1.8 (démontrée ci-dessous),

$$1 = |\mathbf{I}| = |\mathbf{A}\mathbf{A}^{-1}| = |\mathbf{A}||\mathbf{A}^{-1}|,$$

ce qui entraîne que $|\mathbf{A}| = \det(\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_n)$ est non nul.

5.2.3 Démonstration de la proposition 5.1.8

Propriété (a). Si **B** n'est pas inversible, le produit **AB** ne l'est pas non plus, selon (c) de la proposition 4.2.6. Or, nous venons d'établir que le déterminant d'une matrice non inversible (c'est-à-dire dont les colonnes sont linéairement dépendantes) est nul. L'égalité $|\mathbf{AB}| = |\mathbf{A}||\mathbf{B}|$ est donc vraie dans ce cas. Si **B** est inversible, il existe, d'après 4.4.4, des matrices élémentaires $\mathbf{M}_1, \mathbf{M}_2, ..., \mathbf{M}_k$ telles que $\mathbf{B} = \mathbf{M}_1\mathbf{M}_2 ... \mathbf{M}_k$. Compte tenu de (5.7), nous pouvons donc écrire

$$|\mathbf{AB}| = |\mathbf{A}\mathbf{M}_1\mathbf{M}_2 ... \mathbf{M}_k| = |\mathbf{A}\mathbf{M}_1\mathbf{M}_2 ... \mathbf{M}_{k-1}||\mathbf{M}_k|$$
$$= |\mathbf{A}||\mathbf{M}_1||\mathbf{M}_2| ... |\mathbf{M}_k| = |\mathbf{A}||\mathbf{M}_1\mathbf{M}_2 ... \mathbf{M}_k| = |\mathbf{A}||\mathbf{B}|,$$

ce qui établit la propriété (a).

Propriété (b). Si **A** n'est pas inversible, ${}^t\mathbf{A}$ ne l'est pas non plus, d'après (a) de 4.2.6, donc $|\mathbf{A}| = |{}^t\mathbf{A}| = 0$, d'après (b) de la proposition 5.1.6. Si **A** est inversible, $\mathbf{A} = \mathbf{M}_1\mathbf{M}_2 ... \mathbf{M}_k$, où $\mathbf{M}_1, \mathbf{M}_2, ..., \mathbf{M}_k$ sont des matrices élémentaires. Dans ce cas, grâce à la relation (e) de 4.1.9, à la propriété (a) démontrée ci-dessus et à (5.8), nous pouvons écrire

$$|{}^t\mathbf{A}| = |{}^t\mathbf{M}_k{}^t\mathbf{M}_{k-1} ... {}^t\mathbf{M}_1| = |{}^t\mathbf{M}_k||{}^t\mathbf{M}_{k-1}| ... |{}^t\mathbf{M}_1|$$
$$= |\mathbf{M}_1||\mathbf{M}_2| ... |\mathbf{M}_k| = |\mathbf{M}_1\mathbf{M}_2 ... \mathbf{M}_k| = |\mathbf{A}|,$$

ce qui établit la propriété (b).

5.2.4 Caractérisation du déterminant d'ordre *n*

Le déterminant d'ordre n est la seule fonction réelle définie dans l'ensemble des matrices carrées d'ordre n vérifiant les propriétés (a)–(d) énoncées dans le théorème 5.1.5. En effet, nous avons démontré qu'une telle fonction, tout comme le déterminant d'ordre n, attribue à **A** la valeur 0 si **A** est une matrice non inversible et la valeur $|\mathbf{M}_1||\mathbf{M}_2| ... |\mathbf{M}_k|$ si **A** est un produit de matrices élémentaires $\mathbf{M}_1$, $\mathbf{M}_2, ..., \mathbf{M}_k$. Toute matrice carrée étant de l'une de ces deux sortes, l'assertion est démontrée.

5.3 DÉVELOPPEMENTS, FORMULE DE CRAMER

5.3.1 Développement d'un déterminant suivant une colonne

Soit $\mathbf{A} = (a_{ij})$ une matrice carrée d'ordre $n > 1$ et $\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_n$ ses colonnes. Fixons $j > 1$ et désignons par $\mathbf{A}' = (a'_{ij})$ la matrice obtenue à partir de $\mathbf{A}$ en échangeant successivement $\mathbf{a}_j$ et $\mathbf{a}_{j-1}$, $\mathbf{a}_j$ et $\mathbf{a}_{j-2}$, ..., $\mathbf{a}_j$ et $\mathbf{a}_1$. Les colonnes de $\mathbf{A}'$ (écrites dans l'ordre) étant $\mathbf{a}_j, \mathbf{a}_1, ..., \mathbf{a}_{j-1}, \mathbf{a}_{j+1}, ..., \mathbf{a}_n$, il est clair que, pour tout indice de ligne i,

$$a'_{i1} = a_{ij} \quad \text{et} \quad \mathbf{A}'_{i1} = \mathbf{A}_{ij}.$$

D'après (5.1), le déterminant de $\mathbf{A}'$ s'écrit donc sous la forme

$$|\mathbf{A}'| = \sum_{i=1}^{n} (-1)^{i+1} a_{ij} |\mathbf{A}_{ij}|. \tag{5.14}$$

D'autre part, puisque $\mathbf{A}'$ est déduite de $\mathbf{A}$ par $j - 1$ échanges de colonnes, la propriété (a) du théorème 5.1.5 entraîne

$$|\mathbf{A}'| = (-1)^{j-1} |\mathbf{A}|. \tag{5.15}$$

En comparant (5.14) à (5.15), nous constatons que

$$|\mathbf{A}| = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} |\mathbf{A}_{ij}|, \quad j = 1, 2, ..., n. \tag{5.16}$$

On appelle le second membre de (5.16) *développement du déterminant de* $\mathbf{A}$ *suivant la j-ième colonne*. On remarquera que (5.1) est un cas particulier de (5.16), à savoir le cas où $j = 1$.

5.3.2 Développement d'un déterminant suivant une ligne

Soit $\mathbf{A} = (a_{ij})$ une matrice carrée d'ordre $n > 1$. D'après (5.16), le développement de la matrice ${}^t\mathbf{A} = (a'_{ji})$ suivant la i-ième colonne s'écrit

$$|{}^t\mathbf{A}| = \sum_{j=1}^{n} (-1)^{i+j} a'_{ji} |({}^t\mathbf{A})_{ji}|. \tag{5.17}$$

Comme

$$a'_{ji} = a_{ij} \quad \text{et} \quad ({}^t\mathbf{A})_{ji} = {}^t(\mathbf{A}_{ij}),$$

il s'ensuit, vu la propriété (b) de la proposition 5.1.8, que

$$|\mathbf{A}| = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} |\mathbf{A}_{ij}|, \quad i = 1, 2, ..., n. \tag{5.18}$$

On appelle le second membre de (5.18) *développement du déterminant de* $\mathbf{A}$ *suivant la i-ième ligne*.

5.3.3 Mineurs, cofacteurs

On appelle $|\mathbf{A}_{ij}|$ *mineur* d'indices i, j et $(-1)^{i+j}|\mathbf{A}_{ij}|$ *cofacteur* de a_{ij}. Le signe $(-1)^{i+j}$ varie selon le tableau suivant:

$$\begin{bmatrix} + & - & + & . & . & . \\ - & + & - & . & . & . \\ + & - & + & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & + & - \\ . & . & . & . & - & + \end{bmatrix}.$$

Ainsi, d'après (5.16) et (5.18), le déterminant d'une matrice carrée est égal à la somme des termes d'une colonne ou d'une ligne multipliés par leurs cofacteurs respectifs.

5.3.4 Exemple d'application

Les développements s'emploient avantageusement pour calculer le déterminant d'une matrice dont plusieurs termes sont nuls. Dans l'exemple suivant, la flèche indique la ligne ou la colonne suivant laquelle le déterminant est développé:

$$\begin{array}{c} \begin{vmatrix} 1 & 0 & -3 & 0 & 7 & 2 \\ 11 & -3 & 2 & 1 & 10 & -2 \\ 0 & 0 & 1 & 0 & 6 & 0 \\ 5 & 1 & 4 & 0 & -9 & 6 \\ 0 & 0 & 0 & 0 & 2 & 0 \\ -3 & 0 & -10 & 0 & 4 & 1 \end{vmatrix} \\ \begin{array}{cccccc} & & & \uparrow & & \end{array} \end{array} = \begin{array}{c} \begin{vmatrix} 1 & 0 & -3 & 7 & 2 \\ 0 & 0 & 1 & 6 & 0 \\ 5 & 1 & 4 & -9 & 6 \\ 0 & 0 & 0 & 2 & 0 \\ -3 & 0 & -10 & 4 & 1 \end{vmatrix} \\ \begin{array}{ccccc} & \uparrow & & & \end{array} \end{array}$$

$$= -\begin{vmatrix} 1 & -3 & 7 & 2 \\ 0 & 1 & 6 & 0 \\ 0 & 0 & 2 & 0 \\ -3 & -10 & 4 & 1 \end{vmatrix} \begin{array}{c} \\ \\ \leftarrow \\ \\ \end{array} = -2\begin{vmatrix} 1 & -3 & 2 \\ 0 & 1 & 0 \\ -3 & -10 & 1 \end{vmatrix} \begin{array}{c} \\ \leftarrow \\ \\ \end{array}$$

$$= -2\begin{vmatrix} 1 & 2 \\ -3 & 1 \end{vmatrix} = -2 \cdot 7 = -14.$$

5.3.5 Formule de Cramer

Considérons un système de n équations à n inconnues, de matrice associée $\mathbf{A} = (a_{ij})$ et de second membre $\mathbf{b} = (b_i)$. Supposons que le déterminant de $\mathbf{A}$ soit non nul. D'après la proposition 3.5.9, ce système admet une solution unique $\mathbf{x}_0 = (x_j^0)$. Nous allons exposer une méthode de calcul de $\mathbf{x}_0$ s'appuyant sur les propriétés des déterminants. Désignons par $\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_n$ les colonnes de $\mathbf{A}$. Par hypothèse,

$$x_1^0\mathbf{a}_1 + x_2^0\mathbf{a}_2 + ... + x_n^0\mathbf{a}_n = \mathbf{b}.$$

Substituons le premier membre de cette égalité à $\mathbf{b}$ dans le déterminant

$$\det(\mathbf{a}_1, ..., \mathbf{a}_{j-1}, \mathbf{b}, \mathbf{a}_{j+1}, ..., \mathbf{a}_n).$$

D'après (a) de la proposition 5.1.6, ce déterminant est égal à

$$\det(\mathbf{a}_1, ..., \mathbf{a}_{j-1}, x_j^0\mathbf{a}_j, \mathbf{a}_{j+1}, ..., \mathbf{a}_n) = x_j^0\det(\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_n).$$

Il s'ensuit que

$$x_j^0 = \frac{1}{|\mathbf{A}|}\begin{vmatrix} a_{11} & \cdots & b_1 & \cdots & a_{1n} \\ a_{21} & \cdots & b_2 & \cdots & a_{2n} \\ & & \cdots\cdots & & \\ a_{n1} & \cdots & b_n & \cdots & a_{nn} \end{vmatrix}, \quad j = 1, 2, ..., n. \tag{5.19}$$
$$\begin{array}{c}\uparrow \\ j\text{-ième}\end{array}$$

Cette relation est connue sous le nom de *formule de Cramer* pour la résolution d'un système linéaire.

Résolvons, par exemple, le système linéaire

$$\begin{aligned} 3x_1 + 2x_2 + 6x_3 &= 5 \\ 2x_1 + 3x_2 + 3x_3 &= 2 \\ 2x_1 - 2x_2 - 4x_3 &= 1. \end{aligned}$$

Selon (5.19),

$$x_1 = \frac{\begin{vmatrix} 5 & 2 & 6 \\ 2 & 3 & 3 \\ 1 & -2 & -4 \end{vmatrix}}{\begin{vmatrix} 3 & 2 & 6 \\ 2 & 3 & 3 \\ 2 & -2 & -4 \end{vmatrix}} = 1, \quad x_2 = \frac{\begin{vmatrix} 3 & 5 & 6 \\ 2 & 2 & 3 \\ 2 & 1 & -4 \end{vmatrix}}{\begin{vmatrix} 3 & 2 & 6 \\ 2 & 3 & 3 \\ 2 & -2 & -4 \end{vmatrix}} = -\tfrac{1}{2}, \quad x_3 = \frac{\begin{vmatrix} 3 & 2 & 5 \\ 2 & 3 & 2 \\ 2 & -2 & 1 \end{vmatrix}}{\begin{vmatrix} 3 & 2 & 6 \\ 2 & 3 & 3 \\ 2 & -2 & -4 \end{vmatrix}} = \tfrac{1}{2}.$$

5.3.6 Calcul de la matrice inverse par la méthode des cofacteurs

Soit $\mathbf{A} = (a_{ij})$ une matrice carrée d'ordre n et de déterminant non nul. D'après (b) de la proposition 5.1.6 et la proposition 4.2.4, $\mathbf{A}$ est inversible. Les colonnes de la matrice inverse $\mathbf{A}^{-1}$ sont les solutions des n équations matricielles (4.13):

$$\mathbf{A}\begin{pmatrix} x_{1j} \\ x_{2j} \\ \cdot \\ \cdot \\ \cdot \\ x_{nj} \end{pmatrix} = \begin{pmatrix} 0 \\ \cdot \\ \cdot \\ \cdot \\ 1 \\ \cdot \\ 0 \end{pmatrix} \begin{matrix} \\ \\ \\ , \\ \leftarrow j\text{-ième} \\ \\ \\ \end{matrix} \qquad j = 1, 2, \ldots, n.$$

Ces équations sont l'expression matricielle de n systèmes linéaires que nous résolvons à l'aide de la formule de Cramer (5.19). La solution unique du j-ième système s'écrit

$$x_{ij} = \frac{1}{|\mathbf{A}|}\begin{vmatrix} a_{11} & \ldots & 0 & \ldots & a_{1n} \\ & & \cdots\cdots & & \\ a_{j1} & \ldots & 1 & \ldots & a_{jn} \\ & & \cdots\cdots & & \\ a_{n1} & \ldots & 0 & \ldots & a_{nn} \end{vmatrix}, \quad i = 1, 2, \ldots, n,$$
$$\begin{matrix} \uparrow \\ i\text{-ième} \end{matrix}$$

ou encore, après développement du dernier déterminant suivant la i-ième colonne,

$$x_{ij} = \frac{(-1)^{i+j}|\mathbf{A}_{ji}|}{|\mathbf{A}|} = \frac{\operatorname{cof} a_{ji}}{|\mathbf{A}|}, \quad i = 1, 2, \ldots, n, \tag{5.20}$$

où $\operatorname{cof} a_{ji}$ désigne le cofacteur de a_{ji}. Posons $\tilde{\mathbf{A}} = (\tilde{a}_{ij})$, où $\tilde{a}_{ij} = \operatorname{cof} a_{ji}$. La matrice inverse de $\mathbf{A}$ s'écrit alors sous la forme

$$\mathbf{A}^{-1} = \frac{1}{|\mathbf{A}|}\tilde{\mathbf{A}}. \tag{5.21}$$

En d'autres termes, $\mathbf{A}^{-1}$ est égale à la matrice des cofacteurs transposée divisée par le déterminant de $\mathbf{A}$.

On notera le cas particulier où $n = 2$:

$$\mathbf{A}^{-1} = \frac{1}{a_{11}a_{22} - a_{21}a_{12}}\begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}. \tag{5.22}$$

Calculons, par exemple, la matrice inverse de

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

En développant le déterminant de $\mathbf{A}$ suivant la troisième ligne, nous voyons aisément que $|\mathbf{A}| = -2$. D'autre part,

$$\operatorname{cof} a_{11} = \begin{vmatrix} 1 & 1 \\ 1 & 0 \end{vmatrix} = -1, \quad \operatorname{cof} a_{12} = -\begin{vmatrix} 0 & 1 \\ -1 & 0 \end{vmatrix} = -1,$$

$$\operatorname{cof} a_{13} = \begin{vmatrix} 0 & 1 \\ -1 & 1 \end{vmatrix} = 1, \quad \operatorname{cof} a_{21} = -\begin{vmatrix} 1 & 2 \\ 1 & 0 \end{vmatrix} = 2, \ldots,$$

$$\tilde{\mathbf{A}} = {}^t\begin{pmatrix} -1 & -1 & 1 \\ 2 & 2 & -4 \\ -1 & -3 & 3 \end{pmatrix}.$$

D'après (5.21), nous concluons que

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

5.4 EXEMPLES ET REMARQUES DIVERSES

5.4.1 Quelques exemples de calcul d'un déterminant

Le calcul d'un déterminant est d'ordinaire précédé d'un travail de simplification fait principalement à l'aide de (a) de la proposition 5.1.6 ou de la propriété correspondante relative aux lignes.

(1) Pour calculer le déterminant

$$\begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & 1 & 1 \end{vmatrix},$$

on peut commencer par additionner la première colonne à la troisième et à la quatrième et appliquer (b) du théorème 5.1.5 à ces deux colonnes:

$$\begin{vmatrix} 1 & 1 & 2 & 2 \\ 1 & 1 & 0 & 0 \\ 1 & -1 & 2 & 0 \\ 1 & -1 & 2 & 2 \end{vmatrix} = 4\begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & -1 & 1 & 0 \\ 1 & -1 & 1 & 1 \end{vmatrix}.$$

Ensuite, il convient de soustraire de la première ligne deux fois la deuxième et une fois la quatrième:

$$4\begin{vmatrix} -2 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & -1 & 1 & 0 \\ 1 & -1 & 1 & 1 \end{vmatrix} = 4\cdot(-2) = -8.$$

(2) Pour calculer le déterminant

$$\begin{vmatrix} 1 & a & a^2 \\ 1 & b & b^2 \\ 1 & c & c^2 \end{vmatrix},$$

on peut soustraire la deuxième ligne de la première et la troisième de la deuxième:

$$\begin{vmatrix} 0 & a-b & a^2-b^2 \\ 0 & b-c & b^2-c^2 \\ 1 & c & c^2 \end{vmatrix} = (a-b)(b-c)\begin{vmatrix} 1 & a+b \\ 1 & b+c \end{vmatrix} = (a-b)(b-c)(c-a).$$

(3) On notera, à titre d'exemple, que

$$\begin{vmatrix} 2 & -3 & 1 \\ 2 & 3 & -2 \\ 4 & 0 & -1 \end{vmatrix} = 0, \text{ car la troisième ligne est la somme des deux premières.}$$

(4) Le déterminant

$$\begin{vmatrix} 1+a & 1 & 1 & 1 \\ 1 & 1+b & 1 & 1 \\ 1 & 1 & 1+c & 1 \\ 1 & 1 & 1 & 1+d \end{vmatrix}$$

peut être calculé en appliquant (c) du théorème 5.1.5 et (a) de la proposition 5.1.6. A cet effet, posons

$$\mathbf{v} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \quad \mathbf{a} = \begin{pmatrix} a \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 0 \\ b \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{c} = \begin{pmatrix} 0 \\ 0 \\ c \\ 0 \end{pmatrix}, \quad \mathbf{d} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ d \end{pmatrix}.$$

Alors le déterminant s'écrit:

$$\begin{aligned} & \det(\mathbf{v}+\mathbf{a}, \mathbf{v}+\mathbf{b}, \mathbf{v}+\mathbf{c}, \mathbf{v}+\mathbf{d}) \\ = {} & \det(\mathbf{v}, \mathbf{v}+\mathbf{b}, \mathbf{v}+\mathbf{c}, \mathbf{v}+\mathbf{d}) + \det(\mathbf{a}, \mathbf{v}+\mathbf{b}, \mathbf{v}+\mathbf{c}, \mathbf{v}+\mathbf{d}) \\ = {} & \det(\mathbf{v}, \mathbf{b}, \mathbf{c}, \mathbf{d}) + \det(\mathbf{a}, \mathbf{v}, \mathbf{v}+\mathbf{c}, \mathbf{v}+\mathbf{d}) + \det(\mathbf{a}, \mathbf{b}, \mathbf{v}+\mathbf{c}, \mathbf{v}+\mathbf{d}) \\ = {} & bcd + \det(\mathbf{a}, \mathbf{v}, \mathbf{c}, \mathbf{d}) + \det(\mathbf{a}, \mathbf{b}, \mathbf{v}, \mathbf{v}+\mathbf{d}) + \det(\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{v}+\mathbf{d}) \\ = {} & bcd + acd + \det(\mathbf{a}, \mathbf{b}, \mathbf{v}, \mathbf{d}) + \det(\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{v}) + \det(\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}) \\ = {} & bcd + acd + abd + abc + abcd. \end{aligned}$$

(5) Nous vous proposons de calculer le déterminant de la matrice carrée d'ordre n

$$\mathbf{A}_n = \begin{bmatrix} 2 & 1 & 0 & 0 & . & . & . & 0 \\ 1 & 2 & 1 & 0 & . & . & . & 0 \\ 0 & 1 & 2 & 1 & . & . & . & 0 \\ & & & \cdots & \cdots & & & \\ 0 & . & . & . & . & 1 & 2 & 1 \\ 0 & . & . & . & . & 0 & 1 & 2 \end{bmatrix}.$$

Supposons $n > 2$ et développons $|\mathbf{A}_n|$ suivant la première ligne:

$$|\mathbf{A}_n| = 2|\mathbf{A}_{n-1}| - \begin{vmatrix} 1 & 1 \quad 0 \ . \ . \ . \ 0 \\ 0 & \\ . & \\ . & \mathbf{A}_{n-2} \\ . & \\ 0 & \end{vmatrix} = 2|\mathbf{A}_{n-1}| - |\mathbf{A}_{n-2}|.$$

Vu que $|\mathbf{A}_1| = 2$ et $|\mathbf{A}_2| = 3$, cette formule nous permet de calculer $|\mathbf{A}_n|$ par récurrence. Le résultat est $|\mathbf{A}_n| = n + 1$.

5.4.2 Déterminant d'une matrice partagée en sous-matrices

Commençons par observer que, pour toute matrice carrée $\mathbf{A}$ d'ordre m et toute matrice $\mathbf{C}$ de type $m \times n$

$$\begin{vmatrix} \mathbf{I}_n & \mathbf{O} \\ \mathbf{C} & \mathbf{A} \end{vmatrix} = \begin{vmatrix} \mathbf{A} & \mathbf{C} \\ \mathbf{O} & \mathbf{I}_n \end{vmatrix} = |\mathbf{A}|. \tag{5.23}$$

Pour s'en rendre compte, il suffit de développer le premier (deuxième) déterminant n fois suivant la première (dernière) ligne.

Soit maintenant $\mathbf{A}$ et $\mathbf{B}$ des matrices carrées respectivement d'ordre m et n et $\mathbf{C}$ une matrice de type $m \times n$. En vertu de (4.7),

$$\begin{bmatrix} \mathbf{A} & \mathbf{C} \\ \mathbf{O} & \mathbf{B} \end{bmatrix} = \begin{bmatrix} \mathbf{I}_m & \mathbf{O} \\ \mathbf{O} & \mathbf{B} \end{bmatrix} \begin{bmatrix} \mathbf{A} & \mathbf{C} \\ \mathbf{O} & \mathbf{I}_n \end{bmatrix}.$$

Il s'ensuit, grâce à (a) de la proposition 5.1.8 et à (5.23), que

$$\begin{vmatrix} \mathbf{A} & \mathbf{C} \\ \mathbf{O} & \mathbf{B} \end{vmatrix} = |\mathbf{A}||\mathbf{B}|. \tag{5.24}$$

On notera cependant que si **A**, **B**, **C** et **D** sont des matrices carrées du même ordre, en général

$$\begin{vmatrix} \mathbf{A} & \mathbf{C} \\ \mathbf{D} & \mathbf{B} \end{vmatrix} \neq |\mathbf{A}||\mathbf{B}| - |\mathbf{D}||\mathbf{C}|.$$

Pour s'en rendre compte, il suffit de prendre $\mathbf{A} = \mathbf{B} = a\mathbf{I}_2$ et $\mathbf{C} = \mathbf{D} = c\mathbf{I}_2$, $a \neq \pm c$, $c \neq 0$ (cf. exercice 5.5.11).

5.4.3 Retour au volume d'un parallélépipède

Considérons de nouveau le parallélépipède défini par (5.11). Munissons l'espace affine euclidien $\mathscr{E}$ d'un repère orthonormal et désignons par x_{j1}, x_{j2}, ..., x_{jn} les coordonnées de P_j, où $j = 0, 1, ..., n$. Le volume $\mathrm{vol}(P_0, P_1, ..., P_n)$ de ce parallélépipède est alors la valeur absolue du déterminant

$$\begin{vmatrix} 1 & x_{01} & x_{02} & \cdots & x_{0n} \\ 1 & x_{11} & x_{12} & \cdots & x_{1n} \\ & & \cdots\cdots & & \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nn} \end{vmatrix}. \tag{5.25}$$

Pour établir cela, il suffit de soustraire la première ligne de toutes les autres et d'appliquer (5.1). Compte tenu de (b) de la proposition 5.1.8, le résultat de ces opérations est le déterminant $\det(\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_n)$ dont la valeur absolue est, selon la définition (5.12), le volume du parallélépipède.

5.4.4 Equation cartésienne d'un hyperplan

Soit $\mathscr{E}$ un espace affine de dimension finie non nulle n et P_1, P_2, ..., P_n n points de $\mathscr{E}$ engendrant un hyperplan. Supposons que $\mathscr{E}$ soit muni d'un repère et désignons par x_{j1}, x_{j2}, ..., x_{jn} les coordonnées de P_j, où $j = 1, 2, ..., n$. L'équation cartésienne de cet hyperplan peut alors s'écrire sous la forme

$$\begin{vmatrix} 1 & x_1 & x_2 & \cdots & x_n \\ 1 & x_{11} & x_{12} & \cdots & x_{1n} \\ & & \cdots\cdots & & \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nn} \end{vmatrix} = 0. \tag{5.26}$$

En effet, en développant le déterminant suivant la première ligne, nous obtenons une équation de la forme (1.21). En outre, chaque point P_j appartient à l'hyperplan défini par cette équation, car en substituant les coordonnées x_{j1}, x_{j2}, ..., x_{jn} de P_j aux inconnues x_1, x_2, ..., x_n, nous voyons que le déterminant a deux lignes égales et donc qu'il est nul.

5.5 EXERCICES

5.5.1 Montrer que

$$\begin{vmatrix} 1 & 1 & 3 & 0 & 2 \\ 3 & 1 & 0 & 1 & 2 \\ 0 & 1 & 3 & 0 & 2 \\ 4 & -2 & 3 & 1 & 0 \\ 5 & 1 & 0 & 0 & 6 \end{vmatrix} = 60.$$

5.5.2 Calculer le déterminant de la matrice

$$\mathbf{A} = \begin{pmatrix} a & 1 & -1 & 1 \\ -1 & a & -1 & -1 \\ 1 & 1 & a & -1 \\ -1 & 1 & 1 & a \end{pmatrix}$$

(a) d'une manière directe,
(b) en effectuant, au préalable, le produit $\mathbf{A}^t\mathbf{A}$.

5.5.3 A l'aide du théorème 5.1.5 et de la proposition 5.1.6, montrer que

$$\begin{vmatrix} \alpha a_1 + b_1 & \alpha b_1 + c_1 & \alpha c_1 + a_1 \\ \alpha a_2 + b_2 & \alpha b_2 + c_2 & \alpha c_2 + a_2 \\ \alpha a_3 + b_3 & \alpha b_3 + c_3 & \alpha c_3 + a_3 \end{vmatrix} = (\alpha^3 + 1) \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}.$$

5.5.4 Quel est le degré du polynôme p défini par

$$p(t) \equiv \begin{vmatrix} a_{11}+t & a_{12}+t & \dots & a_{1n}+t \\ a_{21}+t & a_{22}+t & \dots & a_{2n}+t \\ & \cdots\cdots & & \\ a_{n1}+t & a_{n2}+t & \dots & a_{nn}+t \end{vmatrix} ?$$

5.5.5 Soit $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k)$ une famille de vecteurs d'un espace vectoriel euclidien. Démontrer que cette famille est liée si et seulement si son *déterminant de Gram*

$$\begin{vmatrix} (\mathbf{x}_1 \mid \mathbf{x}_1) & \dots & (\mathbf{x}_1 \mid \mathbf{x}_k) \\ & \cdots\cdots & \\ (\mathbf{x}_k \mid \mathbf{x}_1) & \dots & (\mathbf{x}_k \mid \mathbf{x}_k) \end{vmatrix}$$

est nul. (Utiliser la proposition 1.5.5 et les propriétés du déterminant.)

5.5.6 Montrer que pour toute famille de nombres $(a_1, a_2, ..., a_n)$ $(n > 1)$,

$$\begin{vmatrix} 1 & a_1 & a_1^2 & \dots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \dots & a_2^{n-1} \\ & & \dots\dots & & \\ 1 & a_n & a_n^2 & \dots & a_n^{n-1} \end{vmatrix} = \prod_{i>j} (a_i - a_j).$$

Ce déterminant est appelé *déterminant de Vandermonde.*

5.5.7 Résoudre le système de l'exercice 3.6.10 au moyen de la formule de Cramer.

5.5.8 Calculer la matrice inverse de

$$\begin{bmatrix} 1 & 2 & -2 \\ 1 & 3 & -1 \\ 2 & 1 & -5 \end{bmatrix}$$

par la méthode des cofacteurs.

5.5.9 Soit $\mathbf{A} = (a_{ij})$ une matrice carrée d'ordre $n > 1$.

(a) Montrer que

$\tilde{\mathbf{A}}\mathbf{A} = |\mathbf{A}|\,\mathbf{I}_n.$

(Appliquer (5.16) à la matrice déduite de $\mathbf{A}$ en remplaçant sa j-ième colonne par sa k-ième.)

(b) Dans l'hypothèse que $|\mathbf{A}| \neq 0$, en déduire la formule d'inversion (5.21), ainsi que la relation

$|\tilde{\mathbf{A}}| = |\mathbf{A}|^{n-1}.$

(Utiliser (a) de la proposition 5.1.8 et (5.3).)

5.5.10 Démontrer, par récurrence sur le nombre n de lignes, que

$$\begin{vmatrix} 2\cos\theta & 1 & & & & \\ 1 & 2\cos\theta & 1 & & 0 & \\ & 1 & 2\cos\theta & 1 & & \\ & & \cdot & \cdot & \cdot & \\ & 0 & & \cdot & \cdot & \cdot \\ & & & \cdot & \cdot & 1 \\ & & & & 1 & 2\cos\theta \end{vmatrix} = \begin{cases} \dfrac{\sin((n+1)\theta)}{\sin\theta} & \text{si } \theta \neq k\pi, \\ n+1 & \text{si } \theta = 2k\pi, \\ (-1)^n(n+1) & \text{si } \theta = (2k+1)\pi. \end{cases}$$

5.5.11 Soit **A** une matrice inversible d'ordre m, **B** une matrice carrée d'ordre n, **C** une matrice de type $m \times n$ et **D** une matrice de type $n \times m$.

(a) Montrer que

$$\begin{vmatrix} \mathbf{A} & \mathbf{C} \\ \mathbf{D} & \mathbf{B} \end{vmatrix} = |\mathbf{A}|\,|\mathbf{B} - \mathbf{D}\mathbf{A}^{-1}\mathbf{C}|.$$

(Vérifier que

$$\begin{bmatrix} \mathbf{A} & \mathbf{C} \\ \mathbf{D} & \mathbf{B} \end{bmatrix} = \begin{bmatrix} \mathbf{I}_m & \mathbf{O} \\ \mathbf{D}\mathbf{A}^{-1} & \mathbf{I}_n \end{bmatrix} \begin{bmatrix} \mathbf{A} & \mathbf{O} \\ \mathbf{O} & \mathbf{B} - \mathbf{D}\mathbf{A}^{-1}\mathbf{C} \end{bmatrix} \begin{bmatrix} \mathbf{I}_m & \mathbf{A}^{-1}\mathbf{C} \\ \mathbf{O} & \mathbf{I}_n \end{bmatrix}$$

et appliquer (5.24).)

(b) En déduire que si **A** et **B** sont du même ordre et **A** et **D** commutent

$$\begin{vmatrix} \mathbf{A} & \mathbf{C} \\ \mathbf{D} & \mathbf{B} \end{vmatrix} = |\mathbf{AB} - \mathbf{DC}|.$$

5.5.12 Soit $P_1(a_1, a_2)$, $P_2(b_1, b_2)$ et $P_3(c_1, c_2)$ trois points non alignés d'un plan affine euclidien, muni d'un repère orthonormal. Montrer que

$$\begin{vmatrix} 1 & x_1^2+x_2^2 & x_1 & x_2 \\ 1 & a_1^2+a_2^2 & a_1 & a_2 \\ 1 & b_1^2+b_2^2 & b_1 & b_2 \\ 1 & c_1^2+c_2^2 & c_1 & c_2 \end{vmatrix} = 0$$

est l'équation du cercle passant par P_1, P_2 et P_3.

5.5.13 Dans un espace affine euclidien $\mathscr{E}$ de dimension finie n supérieure à 2 et muni d'un repère orthonormal, on considère un hyperplan $\mathscr{S}$ dont le j-ième vecteur directeur a pour composantes les termes de la j-ième colonne de la matrice

$$\mathbf{V} = \begin{bmatrix} v_{11} & v_{12} & \cdots & v_{1,n-1} \\ v_{21} & v_{22} & \cdots & v_{2,n-1} \\ & & \cdots\cdots & \\ v_{n1} & v_{n2} & \cdots & v_{n,n-1} \end{bmatrix}.$$

(a) Soit $\mathbf{V}_i$ ($i = 1, 2, \ldots, n$) la matrice obtenue en supprimant la i-ième ligne de **V**. Montrer que le vecteur **n** de composantes

$$v_i = (-1)^{i+n}\,|\mathbf{V}_i|, \quad i = 1, 2, \ldots, n,$$

est normal à $\mathscr{S}$. (Appliquer (a) de l'exercice 5.5.9 à la matrice $\mathbf{A} = (\mathbf{V} \vdots \mathbf{a})$, où **a** est un vecteur-colonne quelconque de $\mathbb{R}^n$.)

(b) Interpréter ce résultat dans le cas où $n = 3$.

CHAPITRE 6

Applications linéaires et applications affines

6.1 GÉNÉRALITÉS

6.1.1 Applications

On appelle *application* d'un ensemble E dans un ensemble F, ou *fonction* définie dans E et à valeurs dans F, une relation entre éléments de E et éléments de F qui associe à chaque élément de E un et un seul élément de F. On dit que E est l'*ensemble de départ* et F l'*ensemble d'arrivée* de l'application. Nous désignerons les applications par les lettres grecques φ, ψ, χ, Φ, Ψ. Les applications à valeur dans $\mathbb{R}$ (autrement dit, les fonctions réelles) seront également notées par des lettres latines minuscules telles que f ou g.

Tout au long de cette section, E, F, G et H désigneront des ensembles.

Soit φ une application de E dans F. Pour tout élément x de E, $\varphi(x)$ désigne l'élément de F associé à x par l'application φ. On dit que $\varphi(x)$ est l'*image* de x par φ, ou la *valeur* de φ en x. Lorsqu'on sait écrire explicitement l'image $\varphi(x)$ de chaque élément x, l'application φ sera désignée également par $x \to \varphi(x)$.

Soit ψ, de même que φ, une application de E dans F. On dit que φ et ψ sont égales, et on écrit $\varphi = \psi$, si $\varphi(x) = \psi(x)$ pour tout élément x de E.

6.1.2 Exemples

(1) L'application φ de E dans E qui associe à tout élément x de E l'élément x lui-même est appelée *application identique*, ou *identité*, et est notée id_E.

(2) Une application φ de E dans F est dite *constante* si à tout élément x de E est associé un même élément c de F.

(3) φ est l'application de $\mathbb{R}^2$ dans $\mathbb{R}^2$

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \longrightarrow \begin{pmatrix} 2x_1 + 3x_2 \\ -x_1 + 2x_2 \end{pmatrix}.$$

(4) φ est l'application de $\mathbb{R}^2$ dans $\mathbb{R}^3$

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \longrightarrow \begin{pmatrix} x_1 + x_2^3 \\ x_1 - 2 \\ x_1 - x_2^3 \end{pmatrix}.$$

(5) φ est l'application de $C_{[0,1]}$ dans $\mathbb{R}$

$$\mathbf{f} \longrightarrow \int_0^1 \mathbf{f}(t)\mathrm{d}t.$$

(6) Si $\varphi_1, \varphi_2, \ldots, \varphi_n$ sont des applications de E dans $\mathbb{R}$,

$$\varphi(x) = \begin{bmatrix} \varphi_1(x) \\ \varphi_2(x) \\ \cdot \\ \cdot \\ \cdot \\ \varphi_n(x) \end{bmatrix}$$

définit une application de E dans $\mathbb{R}^n$.

6.1.3 Eléments invariants, sous-ensembles invariants, sous-ensembles stables

Soit φ une application de E dans E. On dit qu'un élément x de E est *invariant* ou *fixe* si $\varphi(x) = x$. On dit qu'un sous-ensemble S de E est *invariant* si tout élément de S est invariant. On dit qu'un sous-ensemble S de E est *stable* si $\varphi(x)$ appartient à S pour tout élément x de S.

6.1.4 Images directes

Soit φ une application de E dans F. On appelle *image d'un sous-ensemble S* de E, et on note $\varphi(S)$, l'ensemble $\{y \in F\colon y = \varphi(x),\ x \in S\}$. En gros, on peut donc dire que $\varphi(S)$ est l'ensemble des images des éléments de S. Le symbole $\varphi(S)$ peut paraître abusif, mais il est utilisé couramment. Le lecteur prendra garde de ne pas confondre $\varphi(S)$ (sous-ensemble de F) et $\varphi(x)$ (élément de F). On remarquera que lorsque $F = E$, l'inclusion $\varphi(S) \subset S$ caractérise les sous-ensembles stables S de E.

L'image de E s'appelle également *image de l'application* φ et se note $\mathrm{Im}\,\varphi$. Par exemple, l'image de l'application φ de l'exemple (4) de 6.1.2 est un plan de l'espace affine $\mathbb{R}^3$.

On appelle *image d'une famille* $(x_1, x_2, \ldots)$ d'éléments de E la famille $(\varphi(x_1), \varphi(x_2), \ldots)$ d'éléments de F.

6.1.5 Images réciproques

Soit φ une application de E dans F. On appelle *image réciproque d'un sous-ensemble T* de F, et on note $\varphi^{-1}(T)$, l'ensemble $\{x \in E\colon \varphi(x) \in T\}$. En d'autres termes, $\varphi^{-1}(T)$ est l'ensemble des éléments de E dont l'image appartient à T.

6.1.6 Applications injectives, surjectives, bijectives

On dit qu'une application φ de E dans F est *injective* si $\varphi(x_1) \neq \varphi(x_2)$ chaque fois que $x_1 \neq x_2$, autrement dit, si des éléments distincts ont des images distinctes, ou encore, si la relation $\varphi(x_1) = \varphi(x_2)$ entraîne $x_1 = x_2$. On dit que φ est *surjective* si $\mathrm{Im}\,\varphi = F$, autrement dit, si tout élément de F est l'image d'au moins un élément de E. On dit que φ est *bijective* si φ est à la fois injective et surjective. En général, une application n'est ni injective ni surjective. L'application de l'exemple (3) de 6.1.2 est bijective, celle de l'exemple (4) est injective mais non surjective, celle de l'exemple (5) est surjective mais non injective.

6.1.7 Inversion d'une application bijective

Si φ est une application bijective de E dans F, l'application qui associe à chaque élément y de F l'unique élément x de E dont il est l'image s'appelle *application inverse* de φ et se note φ^{-1}. On remarquera que ce symbole n'a de sens que si φ est bijective, tandis que l'image réciproque $\varphi^{-1}(T)$ est définie pour toute application φ. Il est évident que φ^{-1} est une application bijective de F dans E et que $(\varphi^{-1})^{-1} = \varphi$.

6.1.8 Composition d'applications

Soit φ une application de E dans F et ψ une application de F dans G. L'application qui associe à chaque élément x de E l'élément $\psi(\varphi(x))$ de G s'appelle *application composée* de φ et de ψ et se note $\psi\circ\varphi$. Ainsi, $(\psi\circ\varphi)(x) = \psi(\varphi(x))$ pour tout élément x de E.

Soit, de plus, χ une application de G dans H. On vérifie aussitôt que l'opération de composition est associative:

$$\chi\circ(\psi\circ\varphi) = (\chi\circ\psi)\circ\varphi.$$

Cela nous permet d'omettre les parenthèses et d'écrire plus simplement

$$\chi\circ\psi\circ\varphi.$$

Dans le cas particulier où φ est une application de E dans E, on note φ^k l'application composée $\varphi\circ\varphi\circ\ldots\circ\varphi$ (k fois).

La vérification des remarques suivantes est laissée comme exercice:

(1) Si φ et ψ sont des applications de E dans E, l'application $\psi\circ\varphi$ est en général différente de l'application $\varphi\circ\psi$.

(2) Si φ est une application bijective de E dans F, alors $\varphi^{-1}\circ\varphi = \mathrm{id}_E$ et $\varphi\circ\varphi^{-1} = \mathrm{id}_F$.

(3) Une application φ de E dans F est bijective si et seulement s'il existe une application ψ de F dans E telle que $\psi\circ\varphi = \mathrm{id}_E$ et $\varphi\circ\psi = \mathrm{id}_F$. Dans ce cas, $\psi = \varphi^{-1}$.

(4) Si φ est une application bijective de E dans F et ψ une application bijective de F dans G, alors $\psi \circ \varphi$ est une application bijective de E dans G et $(\psi \circ \varphi)^{-1} = \varphi^{-1} \circ \psi^{-1}$.

6.1.9 Addition et multiplication par un scalaire

Soit φ et ψ des applications d'un ensemble E dans un espace vectoriel F. On appelle *somme* de φ et ψ, et on note $\varphi + \psi$, l'application qui associe à chaque élément x de E le vecteur $\varphi(x) + \psi(x)$ de F. On appelle *produit du nombre* λ *par* φ, et on note $\lambda\varphi$, l'application qui associe à chaque élément x de E le vecteur $\lambda\varphi(x)$ de F. Ces deux opérations sont appelées respectivement *addition* et *multiplication par un scalaire*. Elles munissent l'ensemble des applications de E dans F d'une structure d'espace vectoriel.

6.2 APPLICATIONS LINÉAIRES

6.2.1 Introduction

La linéarité de certaines opérations définies au cours des précédents chapitres (le produit scalaire, par exemple) s'est révélée riche de possibilités. Nous allons à présent aborder l'étude systématique de cette propriété et montrer le rôle important qu'elle joue aussi bien en algèbre qu'en géométrie.

6.2.2 Applications linéaires

On dit qu'une application φ d'un espace vectoriel E dans un espace vectoriel F est *linéaire* si, pour tout couple $(\mathbf{x}_1, \mathbf{x}_2)$ de vecteurs de E et tout couple (α_1, α_2) de nombres,

$$\varphi(\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2) = \alpha_1\varphi(\mathbf{x}_1) + \alpha_2\varphi(\mathbf{x}_2). \tag{6.1}$$

Cette condition s'exprime de manière équivalente ainsi: pour tout couple $(\mathbf{x}_1, \mathbf{x}_2)$ de vecteurs de E et tout nombre α,

$$\varphi(\mathbf{x}_1 + \mathbf{x}_2) = \varphi(\mathbf{x}_1) + \varphi(\mathbf{x}_2) \quad \text{et} \quad \varphi(\alpha\mathbf{x}_1) = \alpha\varphi(\mathbf{x}_1). \tag{6.2}$$

En d'autres termes, une application φ est linéaire si elle conserve les opérations d'addition vectorielle et de multiplication par un scalaire.

6.2.3 Remarques

(1) En posant $\alpha_1 = \alpha_2 = 0$ dans (6.1), nous voyons que l'image du vecteur nul de E est le vecteur nul de F.

(2) Toute application linéaire φ de E dans F jouit de la propriété suivante: pour toute famille $(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k)$ de vecteurs de E et toute famille $(\alpha_1, \alpha_2, ..., \alpha_k)$ de nombres,

$$\varphi(\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 + ... + \alpha_k\mathbf{x}_k) = \alpha_1\varphi(\mathbf{x}_1) + \alpha_2\varphi(\mathbf{x}_2) + ... + \alpha_k\varphi(\mathbf{x}_k). \quad (6.3)$$

La vérification se fait par récurrence sur k, en partant de (6.1).

(3) Soit $(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k)$ une famille génératrice de E. Toute application linéaire φ de E dans F est déterminée par l'image $(\varphi(\mathbf{x}_1), \varphi(\mathbf{x}_2), ..., \varphi(\mathbf{x}_k))$ de $(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k)$. En effet, tout vecteur $\mathbf{x}$ de E s'écrit sous la forme

$$\mathbf{x} = \alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 + ... + \alpha_k\mathbf{x}_k,$$

d'où il résulte, grâce à (6.3), que

$$\varphi(\mathbf{x}) = \alpha_1\varphi(\mathbf{x}_1) + \alpha_2\varphi(\mathbf{x}_2) + ... + \alpha_k\varphi(\mathbf{x}_k).$$

Cela montre que les vecteurs $\varphi(\mathbf{x}_1), \varphi(\mathbf{x}_2), ..., \varphi(\mathbf{x}_k)$ déterminent l'image de tout vecteur $\mathbf{x}$ de E.

On peut dire également que l'image par φ d'une famille génératrice de E est une famille génératrice de $\mathrm{Im}\,\varphi$.

(4) Soit φ une application linéaire de E dans F. L'image par φ de toute famille liée $(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k)$ de vecteurs de E est une famille liée de vecteurs de F. Cela résulte de (6.3) et la remarque (1).

6.2.4 Proposition

Soit E un espace vectoriel de dimension finie non nulle n et F un espace vectoriel de dimension finie ou infinie. Soit en outre $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ une base de E. Pour toute famille $(\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_n)$ de vecteurs de F, il existe une application linéaire φ de E dans F et une seule telle que $(\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_n)$ est l'image de $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ par φ.

DÉMONSTRATION

Considérons l'application φ de E dans F définie par la formule

$$\varphi(\mathbf{x}) = x_1\mathbf{y}_1 + x_2\mathbf{y}_2 + ... + x_n\mathbf{y}_n,$$

où $x_1, x_2, ..., x_n$ sont les composantes de $\mathbf{x}$ dans la base $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$. Vu la proposition 1.6.5, φ est linéaire. En outre, $\varphi(\mathbf{e}_i) = \mathbf{y}_i$ pour $i = 1, 2, ..., n$.

Afin d'établir l'unicité, désignons par ψ une application linéaire de E dans F telle que $\psi(\mathbf{e}_i) = \mathbf{y}_i$ pour $i = 1, 2, ..., n$. Soit $\mathbf{x}$ un vecteur quelconque de E et x_1,

$x_2, ..., x_n$ ses composantes dans la base $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$. D'après (6.3),

$$\begin{aligned}\psi(\mathbf{x}) &= x_1\psi(\mathbf{e}_1) + x_2\psi(\mathbf{e}_2) + ... + x_n\psi(\mathbf{e}_n)\\ &= x_1\mathbf{y}_1 + x_2\mathbf{y}_2 + ... + x_n\mathbf{y}_n = \varphi(\mathbf{x}),\end{aligned}$$

ce qui démontre que $\psi = \varphi$.

6.2.5 Exemples d'applications linéaires

(1) D'après la proposition 6.2.4, la donnée d'une application linéaire φ de $\mathbb{R}^n$ dans $\mathbb{R}$ équivaut à la donnée d'une famille de nombres $(a_1, a_2, ... a_n)$ représentant l'image de la base canonique de $\mathbb{R}^n$. L'image d'un vecteur quelconque de $\mathbb{R}^n$ s'écrit alors sous la forme

$$\varphi(\begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix}) = a_1x_1 + a_2x_2 + ... + a_nx_n. \tag{6.4}$$

Les seules applications linéaires de $\mathbb{R}$ dans $\mathbb{R}$ sont donc celles définies par $\varphi(x) = ax$, où a est un nombre quelconque fixé.

(2) L'application de $\mathbb{R}^2$ dans $\mathbb{R}^3$

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \longrightarrow \begin{bmatrix} 2 & 1 \\ 3 & 0 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \tag{6.5}$$

est linéaire. Plus généralement, si $\mathbf{A}$ est une matrice de type $m \times n$, l'application de $\mathbb{R}^n$ dans $\mathbb{R}^m$

$$\mathbf{x} \longrightarrow \mathbf{A}\mathbf{x} \tag{6.6}$$

est linéaire.

(3) Soit E un espace vectoriel euclidien et $\mathbf{v}$ un vecteur de E arbitrairement choisi. L'application de E dans $\mathbb{R}$

$$\mathbf{x} \longrightarrow (\mathbf{x} \mid \mathbf{v}) \tag{6.7}$$

est linéaire, en vertu de (b) de 2.2.2.

(4) Pour tout nombre λ, l'application d'un espace vectoriel E dans lui-même

$$\mathbf{x} \longrightarrow \lambda\mathbf{x}$$

est linéaire. On l'appelle *homothétie de rapport* λ. L'homothétie de rapport 1 est l'application identique id_E, celle de rapport 0 est appelée *application nulle*, celle de rapport -1 *symétrie centrale*.

(5) Soit S un plan vectoriel d'un espace vectoriel E de dimension 3. Soit $(\mathbf{v}_1, \mathbf{v}_2)$ une base de S, $\mathbf{v}_3$ un vecteur n'appartenant pas à S et D la droite vectorielle engendrée par $\mathbf{v}_3$. Par les propositions 1.8.4 et 1.7.9, $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ est une base de E, donc tout vecteur $\mathbf{x}$ de E s'exprime de manière unique sous la forme

$$\mathbf{x} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + x_3\mathbf{v}_3.$$

Pour tout nombre λ, l'application de E dans E

$$\mathbf{x} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + x_3\mathbf{v}_3 \longrightarrow x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \lambda x_3\mathbf{v}_3 \tag{6.8}$$

est linéaire. On l'appelle *dilatation relativement à S, de direction D et de rapport λ.* La dilatation de rapport 0 est appelée *projection sur S parallèlement à D* et la dilatation de rapport -1 *symétrie par rapport à S parallèlement à D* (dans la figure 6.1, φ désigne la projection et ψ une dilatation).

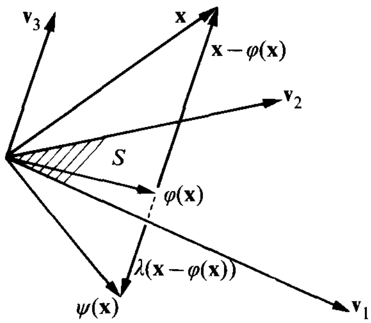


Fig. 6.1

(6) L'exemple (5) se généralise de la manière suivante. Supposons qu'un espace vectoriel E soit somme directe de deux sous-espaces vectoriels S et T différents de $\{\mathbf{0}\}$. Nous avons vu dans 1.8.7 que tout vecteur $\mathbf{x}$ de E s'écrit de manière unique sous la forme

$$\mathbf{x} = \mathbf{s} + \mathbf{t}, \tag{6.9}$$

où $\mathbf{s}$ est un vecteur de S et $\mathbf{t}$ un vecteur de T. Pour tout nombre λ, l'application

$$\mathbf{x} = \mathbf{s} + \mathbf{t} \longrightarrow \mathbf{s} + \lambda\mathbf{t} \tag{6.10}$$

est linéaire. On l'appelle *dilatation relativement à S, de direction T et de rapport λ.* La dilatation de rapport 0 est appelée *projection sur S parallèlement à T* et celle de rapport -1 *symétrie par rapport à S parallèlement à T.*

Si E est euclidien et $T = S^{\perp}$, on parle préférablement de *dilatation orthogonale à S*, de *projection orthogonale sur S* (cf. 2.6.3) et de *symétrie orthogonale par rapport à S.*

On notera que $\varphi^2 = \varphi$ si φ est une projection et $\varphi^2 = \mathrm{id}_E$ (c'est-à-dire $\varphi^{-1} = \varphi$) si φ est une symétrie (cf. exercices 6.8.3 et 6.8.4).

Soit φ une dilatation relativement à S, de direction T et de rapport λ différent de 1. Soit en outre $\mathbf{x}$ un vecteur invariant. A l'aide de la décomposition (6.9), l'égalité $\varphi(\mathbf{x}) = \mathbf{x}$ s'écrit sous la forme

$$\mathbf{s} + \lambda\mathbf{t} = \mathbf{s} + \mathbf{t}$$

et entraîne

$$\lambda\mathbf{t} = \mathbf{t}.$$

Comme λ est différent de 1, il s'ensuit que $\mathbf{t} = \mathbf{0}$, autrement dit que $\mathbf{x}$ est un vecteur de S. Réciproquement, tout vecteur de S est manifestement invariant, ce qui nous permet de conclure que S est constitué de tous les vecteurs invariants. On appelle *S sous-espace invariant de la dilatation φ.*

Il est par ailleurs évident que T est un sous-espace stable. De plus, $\varphi(\mathbf{x}) - \mathbf{x}$ est un vecteur non nul de T, pour tout vecteur $\mathbf{x}$ n'appartenant pas à S, ce qui justifie l'expression «de direction T». A noter que, lorsque T est une droite vectorielle, une seule différence $\varphi(\mathbf{x}) - \mathbf{x}$ suffit à déterminer T.

Les termes qui précisent la notion de dilatation ne se justifient plus lorsque le rapport λ est égal à 1, c'est-à-dire lorsque la dilatation est l'application identique de E. Cela ne créera toutefois aucune confusion.

(7) Supposons que l'espace vectoriel E de l'exemple précédent soit euclidien et de dimension finie supérieure à 1 et que S soit un hyperplan vectoriel de E de vecteur normal $\mathbf{n}$. L'image d'un vecteur par une projection ou, plus généralement, par une dilatation peut alors être écrite explicitement à l'aide du produit scalaire, comme nous allons le voir.

La projection de $\mathbf{x}$ sur S parallèlement à une droite D engendrée par un vecteur $\mathbf{v}$ n'appartenant pas à S est l'unique vecteur de la forme $\mathbf{x} + \alpha\mathbf{v}$ tel que $(\mathbf{x} + \alpha\mathbf{v} \mid \mathbf{n}) = 0$. Il en découle que

$$\alpha = -\frac{(\mathbf{x} \mid \mathbf{n})}{(\mathbf{v} \mid \mathbf{n})},$$

donc que l'image de $\mathbf{x}$ par la projection φ sur S parallèlement à D est

$$\varphi(\mathbf{x}) = \mathbf{x} - \frac{(\mathbf{x} \mid \mathbf{n})}{(\mathbf{v} \mid \mathbf{n})}\mathbf{v}. \tag{6.11}$$

Notons maintenant que la décomposition (6.9) s'écrit, de manière équivalente, sous la forme

$$\mathbf{x} = \varphi(\mathbf{x}) + (\mathbf{x} - \varphi(\mathbf{x})).$$

L'image de $\mathbf{x}$ par la dilatation ψ relativement à S, de direction D et de rapport λ est donc

$$\psi(\mathbf{x}) = \varphi(\mathbf{x}) + \lambda(\mathbf{x} - \varphi(\mathbf{x})) = \lambda\mathbf{x} + (1 - \lambda)\,\varphi(\mathbf{x}).$$

En substituant à $\varphi(\mathbf{x})$ son expression (6.11), nous en concluons que

$$\psi(\mathbf{x}) = \lambda\mathbf{x} + (1 - \lambda)(\mathbf{x} - \frac{(\mathbf{x} \mid \mathbf{n})}{(\mathbf{v} \mid \mathbf{n})}\mathbf{v}) = \mathbf{x} - (1 - \lambda)\frac{(\mathbf{x} \mid \mathbf{n})}{(\mathbf{v} \mid \mathbf{n})}\mathbf{v}. \tag{6.12}$$

(8) Les exemples (3) et (5) de 6.1.2 sont des applications linéaires.

(9) Une application constante est linéaire si et seulement si la constante est le vecteur nul.

(10) Un autre exemple tiré de l'analyse (l'exemple (5) de 6.1.2 en est un) est l'application linéaire de $C^1_{(a,\,b)}$ dans $C_{(a,\,b)}$ qui associe à toute fonction continûment dérivable $\mathbf{f}$ sa dérivée $\dot{\mathbf{f}}$.

6.3 NOYAUX ET IMAGES

6.3.1 Proposition. Images directes et réciproques de sous-espaces vectoriels

Soit φ une application linéaire d'un espace vectoriel E dans un espace vectoriel F.

(a) *L'image $\varphi(S)$ de tout sous-espace vectoriel S de E est un sous-espace vectoriel de F.*

(b) *L'image réciproque $\varphi^{-1}(T)$ de tout sous-espace vectoriel T de F est un sous-espace vectoriel de E.*

DÉMONSTRATION

Assertion (a). Soit $(\mathbf{y}_1, \mathbf{y}_2)$ un couple de vecteurs de $\varphi(S)$ et (α_1, α_2) un couple de nombres. Il existe des vecteurs $\mathbf{x}_1$ et $\mathbf{x}_2$ de S tels que $\varphi(\mathbf{x}_1) = \mathbf{y}_1$ et $\varphi(\mathbf{x}_2) = \mathbf{y}_2$. En outre, par la linéarité de φ,

$$\varphi(\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2) = \alpha_1\varphi(\mathbf{x}_1) + \alpha_2\varphi(\mathbf{x}_2) = \alpha_1\mathbf{y}_1 + \alpha_2\mathbf{y}_2,$$

ce qui prouve que $\alpha_1\mathbf{y}_1 + \alpha_2\mathbf{y}_2$ est un vecteur de $\varphi(S)$, puisque $\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2$ est un vecteur de S. Comme $\varphi(S)$ n'est pas vide (le vecteur nul de F y appartient), la conclusion suit de la proposition 1.4.6.

Assertion (b). Soit $(\mathbf{x}_1, \mathbf{x}_2)$ un couple de vecteurs de $\varphi^{-1}(T)$ et (α_1, α_2) un couple de nombres. Par la linéarité de φ,

$$\varphi(\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2) = \alpha_1\varphi(\mathbf{x}_1) + \alpha_2\varphi(\mathbf{x}_2),$$

ce qui nous montre que $\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2$ est un vecteur de $\varphi^{-1}(T)$, puisque $\alpha_1\varphi(\mathbf{x}_1) + \alpha_2\varphi(\mathbf{x}_2)$ est un vecteur de T. Comme $\varphi^{-1}(T)$ n'est pas vide (le vecteur nul de E y appartient), la conclusion suit de la proposition 1.4.6.

6.3.2 Rang d'une application linéaire

Soit φ une application linéaire d'un espace vectoriel E dans un espace vectoriel F. D'après la proposition 6.3.1, $\mathrm{Im}\varphi$ est un sous-espace vectoriel de F. Lorsque ce sous-espace est de dimension finie, on appelle *rang de* φ, et on note $\mathrm{rg}\varphi$, la dimension de $\mathrm{Im}\varphi$. Vu la remarque (3) de 6.2.3, si $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k)$ est une famille génératrice de E, le rang de φ est donc égal au rang de la famille $(\varphi(\mathbf{x}_1), \varphi(\mathbf{x}_2), \ldots, \varphi(\mathbf{x}_k))$.

On notera que $\mathrm{rg}\varphi = 0$ si et seulement si φ est l'application nulle. On notera également que le rang de φ est défini si E ou F sont de dimension finie.

Par exemple, le rang de l'application (6.4) est 0 si tous les a_i sont nuls, 1 si au moins un des a_i est non nul; le rang de l'application (6.5) est 2; le rang de l'application (6.6) est égal à la dimension du sous-espace vectoriel de $\mathbb{R}^m$ engendré par les colonnes de $\mathbf{A}$, c'est-à-dire au rang de $\mathbf{A}$; le rang de l'application (6.8) est 2 si $\lambda = 0$, 3 si $\lambda \neq 0$; celui de l'application (6.10) est $\dim S$ si $\lambda = 0$, $\dim E$ si $\lambda \neq 0$ (à condition que S, respectivement E, soient de dimension finie).

6.3.3 Noyau d'une application linéaire

Soit φ une application linéaire d'un espace vectoriel E dans un espace vectoriel F. On appelle *noyau de* φ, et on note $\mathrm{Ker}\varphi$, l'image réciproque de $\{\mathbf{0}\}$, autrement dit, l'ensemble des vecteurs de E dont l'image est le vecteur nul de F. D'après la proposition 6.3.1, $\mathrm{Ker}\varphi$ est un sous-espace vectoriel de E.

Par exemple, le noyau de l'application (6.4) est $\mathbb{R}^n$ ou l'hyperplan vectoriel de $\mathbb{R}^n$ d'équation $a_1x_1 + a_2x_2 + \ldots + a_nx_n = 0$, suivant que tous les a_i sont nuls ou non; le noyau de l'application (6.6) est le sous-espace vectoriel de $\mathbb{R}^n$ formé des solutions de l'équation (système homogène) $\mathbf{Ax} = \mathbf{0}$; le noyau de l'application (6.10) est T si $\lambda = 0$, $\{\mathbf{0}\}$ si $\lambda \neq 0$.

6.3.4 Proposition. Caractérisation des applications linéaires injectives

Soit φ une application linéaire d'un espace vectoriel E dans un espace vectoriel F. Pour que φ soit injective, il faut et il suffit que l'une des conditions équivalentes suivantes soit satisfaite:

(a) $\mathrm{Ker}\varphi = \{\mathbf{0}\}$.

(b) *L'image de toute famille libre* $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k)$ *de vecteurs de E est une famille libre de vecteurs de F.*

DÉMONSTRATION

Si φ est injective, $\varphi(\mathbf{x}) \neq \varphi(\mathbf{0}) = \mathbf{0}$ pour tout vecteur $\mathbf{x} \neq \mathbf{0}$, donc $\mathrm{Ker}\varphi = \{\mathbf{0}\}$. Réciproquement, si $\mathrm{Ker}\varphi = \{\mathbf{0}\}$ et $\mathbf{x}_1$, $\mathbf{x}_2$ sont des vecteurs distincts de E, alors $\mathbf{x}_1 - \mathbf{x}_2 \neq \mathbf{0}$, donc $\varphi(\mathbf{x}_1 - \mathbf{x}_2) \neq \mathbf{0}$, ce qui entraîne $\varphi(\mathbf{x}_1) \neq \varphi(\mathbf{x}_2)$ et démontre l'injectivité de φ.

Supposons que $\mathrm{Ker}\varphi$ se réduise à $\{\mathbf{0}\}$ et considérons une famille libre $(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k)$ de vecteurs de E. Par la linéarité de φ, la relation

$$\alpha_1\varphi(\mathbf{x}_1) + \alpha_2\varphi(\mathbf{x}_2) + ... + \alpha_k\varphi(\mathbf{x}_k) = \mathbf{0}$$

s'écrit aussi sous la forme

$$\varphi(\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 + ... + \alpha_k\mathbf{x}_k) = \mathbf{0}$$

et entraîne

$$\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 + ... + \alpha_k\mathbf{x}_k = \mathbf{0},$$

donc $\alpha_1 = \alpha_2 = ... = \alpha_k = 0$, puisque $(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k)$ est une famille libre, ce qui prouve que $(\varphi(\mathbf{x}_1), \varphi(\mathbf{x}_2), ..., \varphi(\mathbf{x}_k))$ est une famille libre. Réciproquement, si $\mathrm{Ker}\varphi$ comprend un vecteur non nul $\mathbf{x}$, la famille $(\mathbf{x})$ est libre et son image $(\mathbf{0})$ est liée. L'équivalence des conditions (a) et (b) est ainsi démontrée.

6.3.5 Proposition

Soit φ une application linéaire d'un espace vectoriel E dans un espace vectoriel F. Si E est de dimension finie, alors

$$\dim E = \dim \mathrm{Ker}\varphi + \mathrm{rg}\varphi. \tag{6.13}$$

DÉMONSTRATION

Ecartons le cas banal où la dimension n de E est nulle et considérons les trois cas suivants:

Cas où $\dim\mathrm{Ker}\varphi = 0$. D'après (b) de la proposition 6.3.4, l'image d'une base quelconque de E est une famille libre. En outre, par la remarque (3) de 6.2.3, cette famille engendre $\mathrm{Im}\varphi$. Il s'ensuit que $\mathrm{rg}\varphi = n$, ce qui montre que la relation (6.13) est satisfaite.

Cas où $\dim\mathrm{Ker}\varphi = n$. D'après la proposition 1.8.5, $\mathrm{Ker}\varphi = E$, donc $\mathrm{Im}\varphi = \{\mathbf{0}\}$, ce qui montre que la relation (6.13) est satisfaite.

Cas où $0 < \dim\mathrm{Ker}\varphi < n$. Soit $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_k)$ une base de $\mathrm{Ker}\varphi$. D'après le théorème 1.7.3, cette base se prolonge en une base $(\mathbf{e}_1, ..., \mathbf{e}_k, \mathbf{e}_{k+1}, ..., \mathbf{e}_n)$ de E. Vu ce qui a été relevé dans 6.3.2, $\mathrm{rg}\varphi$ est égal au rang de la famille $(\varphi(\mathbf{e}_1), ..., \varphi(\mathbf{e}_k), \varphi(\mathbf{e}_{k+1}), ..., \varphi(\mathbf{e}_n))$, donc au rang de la famille $(\varphi(\mathbf{e}_{k+1}), ..., \varphi(\mathbf{e}_n))$, puisque $\varphi(\mathbf{e}_1) = ... = \varphi(\mathbf{e}_k) = \mathbf{0}$, du fait que $\mathbf{e}_1, ..., \mathbf{e}_k$ sont des vecteurs de $\mathrm{Ker}\varphi$. Montrons que $(\varphi(\mathbf{e}_{k+1}), ..., \varphi(\mathbf{e}_n))$ est une famille libre. Par la linéarité de φ, la relation

$$\alpha_{k+1}\varphi(\mathbf{e}_{k+1}) + ... + \alpha_n\varphi(\mathbf{e}_n) = \mathbf{0}$$

s'écrit aussi sous la forme

$$\varphi(\alpha_{k+1}\mathbf{e}_{k+1} + ... + \alpha_n\mathbf{e}_n) = \mathbf{0}$$

et entraîne que $\alpha_{k+1}\mathbf{e}_{k+1} + \ldots + \alpha_n\mathbf{e}_n$ est un vecteur de $\operatorname{Ker}\varphi$. Or, cela n'est possible que si $\alpha_{k+1} = \ldots = \alpha_n = 0$, ce qui montre que $(\varphi(\mathbf{e}_{k+1}), \ldots, \varphi(\mathbf{e}_n))$ est une famille libre. Il s'ensuit que $\operatorname{rg}\varphi = n - k = \dim E - \dim \operatorname{Ker}\varphi$.

6.3.6 Corollaire

Soit E et F des espaces vectoriels de dimension finie. Si $\dim E = \dim F$, *les assertions suivantes sont équivalentes:*

(a) *φ est une application linéaire injective de E dans F.*

(b) *φ est une application linéaire surjective de E dans F.*

(c) *φ est une application linéaire bijective de E dans F.*

En d'autres termes, φ est une application linéaire bijective de E dans F si et seulement si son noyau se réduit au vecteur nul de E, ou encore, si et seulement si son rang est égal à $\dim E$ *(ou* $\dim F$*).*

DÉMONSTRATION

Si φ est injective, $\operatorname{Ker}\varphi = \{\mathbf{0}\}$, par la proposition 6.3.4. D'après (6.13), $\dim E = \operatorname{rg}\varphi$, ou encore $\dim F = \dim \operatorname{Im}\varphi$. Vu la proposition 1.8.5, cela entraîne que $F = \operatorname{Im}\varphi$, autrement dit que φ est surjective.

Si φ est surjective, $F = \operatorname{Im}\varphi$, donc $\dim F = \dim \operatorname{Im}\varphi = \operatorname{rg}\varphi$ et, par conséquent, $\dim E = \operatorname{rg}\varphi$. D'après (6.13), $\dim \operatorname{Ker}\varphi = 0$, donc φ est injective.

6.3.7 Isomorphies d'espaces vectoriels

On appelle *isomorphisme* d'un espace vectoriel E dans un espace vectoriel F toute application linéaire bijective de E dans F. On dit que deux espaces vectoriels sont *isomorphes*, ou que l'un est *isomorphe à* l'autre, s'il existe un isomorphisme de l'un dans l'autre.

Des propositions 6.3.4 et 6.3.5, il résulte que deux espaces vectoriels isomorphes sont de même dimension. Nous avons vu dans 1.7.10 que tout espace vectoriel de dimension finie non nulle n est isomorphe à $\mathbb{R}^n$.

6.4 OPÉRATIONS SUR LES APPLICATIONS LINÉAIRES

6.4.1 Espace vectoriel des applications linéaires de E dans F

Si φ et ψ sont des applications linéaires d'un espace vectoriel E dans un espace vectoriel F, $\varphi + \psi$ et $\lambda\varphi$ (pour tout nombre λ) sont également des applications linéaires. Montrons, par exemple, que $\varphi + \psi$ est linéaire. Par la définition

6.1.9 et la linéarité de φ et de ψ, pour tout couple $(\mathbf{x}_1, \mathbf{x}_2)$ de vecteurs de E et tout couple (α_1, α_2) de nombres,

$$\begin{aligned}(\varphi + \psi)(\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2) &= \varphi(\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2) + \psi(\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2)\\ &= \alpha_1\varphi(\mathbf{x}_1) + \alpha_2\varphi(\mathbf{x}_2) + \alpha_1\psi(\mathbf{x}_1) + \alpha_2\psi(\mathbf{x}_2)\\ &= \alpha_1(\varphi + \psi)(\mathbf{x}_1) + \alpha_2(\varphi + \psi)(\mathbf{x}_2),\end{aligned}$$

ce qui montre que $\varphi + \psi$ est linéaire.

Désignons par $L(E, F)$ l'ensemble des applications linéaires de E dans F. Cet ensemble n'est pas vide, car il comprend au moins l'application nulle. En outre, en même temps que φ et ψ, il comprend toutes leurs combinaisons linéaires $\lambda\varphi + \mu\psi$. D'après la proposition 1.4.6, $L(E, F)$ est donc un sous-espace vectoriel de l'espace vectoriel de toutes les applications de E dans F.

6.4.2 Proposition. Linéarité de l'application inverse et de l'application composée

Soit E, F et G des espaces vectoriels.

(a) *Si φ est une application linéaire bijective de E dans F, l'application inverse φ^{-1} est linéaire. En d'autres termes, l'inverse de tout isomorphisme est un isomorphisme.*

(b) *Si φ est une application linéaire de E dans F et ψ une application linéaire de F dans G, l'application composée $\psi\circ\varphi$ est linéaire.*

DÉMONSTRATION

Assertion (a). Soit $(\mathbf{y}_1, \mathbf{y}_2)$ un couple de vecteurs de F et (α_1, α_2) un couple de nombres. Posons $\mathbf{x}_1 = \varphi^{-1}(\mathbf{y}_1)$ et $\mathbf{x}_2 = \varphi^{-1}(\mathbf{y}_2)$. Alors $\mathbf{y}_1 = \varphi(\mathbf{x}_1)$ et $\mathbf{y}_2 = \varphi(\mathbf{x}_2)$; en outre, par la linéarité de φ,

$$\alpha_1\mathbf{y}_1 + \alpha_2\mathbf{y}_2 = \alpha_1\varphi(\mathbf{x}_1) + \alpha_2\varphi(\mathbf{x}_2) = \varphi(\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2),$$

ce qui entraîne

$$\varphi^{-1}(\alpha_1\mathbf{y}_1 + \alpha_2\mathbf{y}_2) = \alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 = \alpha_1\varphi^{-1}(\mathbf{y}_1) + \alpha_2\varphi^{-1}(\mathbf{y}_2)$$

et établit ainsi la linéarité de l'application φ^{-1}.

Assertion (b). Soit $(\mathbf{x}_1, \mathbf{x}_2)$ un couple de vecteurs de E et (α_1, α_2) un couple de nombres. D'après la définition 6.1.8 et la linéarité de φ et de ψ,

$$\begin{aligned}&(\psi\circ\varphi)(\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2) = \psi(\varphi(\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2)) = \psi(\alpha_1\varphi(\mathbf{x}_1) + \alpha_2\varphi(\mathbf{x}_2))\\ &= \alpha_1\psi(\varphi(\mathbf{x}_1)) + \alpha_2\psi(\varphi(\mathbf{x}_2)) = \alpha_1(\psi\circ\varphi)(\mathbf{x}_1) + \alpha_2(\psi\circ\varphi)(\mathbf{x}_2),\end{aligned}$$

ce qui démontre la linéarité de l'application $\psi\circ\varphi$.

6.4.3 Formes linéaires, espace dual

Soit E un espace vectoriel. On appelle *forme linéaire* dans E toute application linéaire de E dans $\mathbb{R}$.

L'application de l'exemple (5) de 6.1.2, ainsi que les applications (6.4) et (6.7) sont des formes linéaires.

On appelle *espace dual* de E, et on note E^*, l'espace vectoriel des formes linéaires dans E.

6.4.4 Bases duales

Soit E un espace vectoriel de dimension finie non nulle n. Soit en outre $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ une base de E. Vu la proposition 6.2.4, nous définissons n formes linéaires $\mathbf{e}_1^*, \mathbf{e}_2^*, ..., \mathbf{e}_n^*$ dans E en posant

$$\mathbf{e}_i^*(\mathbf{e}_j) = \begin{cases} 1 & \text{si } i = j, \\ 0 & \text{si } i \neq j. \end{cases} \tag{6.14}$$

Pour tout indice i et tout vecteur $\mathbf{x}$ de E, $\mathbf{e}_i^*(\mathbf{x}) = x_i$, où x_i désigne la i-ième composante de $\mathbf{x}$ dans la base $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$. En effet, par la linéarité de $\mathbf{e}_i^*$

$$\begin{aligned} \mathbf{e}_i^*(\mathbf{x}) &= \mathbf{e}_i^*(x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + ... + x_n\mathbf{e}_n) \\ &= x_1\mathbf{e}_i^*(\mathbf{e}_1) + x_2\mathbf{e}_i^*(\mathbf{e}_2) + ... + x_n\mathbf{e}_i^*(\mathbf{e}_n) = x_i. \end{aligned}$$

Nous allons maintenant démontrer que $(\mathbf{e}_1^*, \mathbf{e}_2^*, ..., \mathbf{e}_n^*)$ est une base de E^*. Supposons que

$$\alpha_1\mathbf{e}_1^* + \alpha_2\mathbf{e}_2^* + ... + \alpha_n\mathbf{e}_n^* = \mathbf{0}^*,$$

où $\mathbf{0}^*$ désigne la forme linéaire nulle. Alors, pour $j = 1, 2, ..., n$,

$$\begin{aligned} 0 = \mathbf{0}^*(\mathbf{e}_j) &= (\alpha_1\mathbf{e}_1^* + \alpha_2\mathbf{e}_2^* + ... + \alpha_n\mathbf{e}_n^*)(\mathbf{e}_j) \\ &= \alpha_1\mathbf{e}_1^*(\mathbf{e}_j) + \alpha_2\mathbf{e}_2^*(\mathbf{e}_j) + ... + \alpha_n\mathbf{e}_n^*(\mathbf{e}_j) = \alpha_j, \end{aligned}$$

ce qui prouve que $(\mathbf{e}_1^*, \mathbf{e}_2^*, ..., \mathbf{e}_n^*)$ est une famille libre. D'autre part, soit $\mathbf{x}^*$ une forme linéaire quelconque dans E. Posons

$$\alpha_1 = \mathbf{x}^*(\mathbf{e}_1), \ \alpha_2 = \mathbf{x}^*(\mathbf{e}_2), ..., \alpha_n = \mathbf{x}^*(\mathbf{e}_n).$$

D'après (6.14), pour $j = 1, 2, ..., n$,

$$\mathbf{x}^*(\mathbf{e}_j) = (\alpha_1\mathbf{e}_1^* + \alpha_2\mathbf{e}_2^* + ... + \alpha_n\mathbf{e}_n^*)(\mathbf{e}_j),$$

ce qui entraîne, grâce à la proposition 6.2.4, que

$$\mathbf{x}^* = \alpha_1\mathbf{e}_1^* + \alpha_2\mathbf{e}_2^* + ... + \alpha_n\mathbf{e}_n^*.$$

En d'autres termes, $(\mathbf{e}_1^*, \mathbf{e}_2^*, ..., \mathbf{e}_n^*)$ est une famille génératrice de E^*.

On appelle la base $(\mathbf{e}_1^*, \mathbf{e}_2^*, ..., \mathbf{e}_n^*)$ de E^* *base duale* de $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$.

6.4.5 Proposition. Hyperplans vectoriels comme noyaux de formes linéaires

Un sous-espace vectoriel d'un espace vectoriel E de dimension finie non nulle est un hyperplan vectoriel si et seulement s'il est le noyau d'une forme linéaire non nulle dans E.

DÉMONSTRATION

Désignons par n la dimension de E. Soit $\mathbf{x}^*$ une forme linéaire non nulle dans E. D'après (6.13), la dimension de $\operatorname{Ker}\mathbf{x}^*$ est $n - 1$, puisque celle de $\operatorname{Im}\mathbf{x}^*$ est 1. En d'autres termes, $\operatorname{Ker}\mathbf{x}^*$ est un hyperplan vectoriel de E. Réciproquement, soit S un hyperplan vectoriel de E. Si $n = 1$, S se réduit au seul vecteur nul de E, donc S est le noyau de n'importe quelle forme linéaire non nulle dans E. Si $n > 1$, désignons par $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_{n-1})$ une base de S. En vertu du théorème 1.7.3, cette base se prolonge en une base $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ de E. La forme linéaire $\mathbf{e}_n^*$ définie par (6.14) (avec $i = n$) est non nulle et $\operatorname{Ker}\mathbf{e}_n^* = S$.

6.5 REPRÉSENTATION MATRICIELLE D'UNE APPLICATION LINÉAIRE

6.5.1 Introduction

Lorsque les espaces de départ et d'arrivée sont munis chacun d'une base, à toute application linéaire est associée une matrice permettant de calculer les composantes de l'image de chaque vecteur.

Dans cette section, E, F et G désigneront des espaces vectoriels de dimensions finies non nulles respectives n, m et p.

6.5.2 Matrice d'une application linéaire

Soit φ une application linéaire de E dans F. Soit en outre $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ une base de E et $(\mathbf{f}_1, \mathbf{f}_2, ..., \mathbf{f}_m)$ une base de F. Pour $j = 1, 2, ..., n$, désignons par $a_{1j}, a_{2j}, ..., a_{mj}$ les composantes de $\varphi(\mathbf{e}_j)$ dans la base $(\mathbf{f}_1, \mathbf{f}_2, ..., \mathbf{f}_m)$. On appelle la matrice de type $m \times n$

$$\begin{pmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2j} & \cdots & a_{2n} \\ & & \cdots\cdots & & \cdot \\ a_{m1} & \cdots & a_{mj} & \cdots & a_{mn} \end{pmatrix} \tag{6.15}$$

matrice de φ dans les bases $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ *et* $(\mathbf{f}_1, \mathbf{f}_2, ..., \mathbf{f}_m)$. Cette matrice sera notée $\mathbf{A}_\varphi$, ou simplement $\mathbf{A}$, lorsque aucune confusion n'est à craindre. Manifestement, la matrice de φ dépend du choix des bases de E et de F, quoique son symbole ne fasse pas état de cette dépendance.

D'après la proposition 6.2.4, l'application linéaire φ est déterminée par l'image $(\varphi(\mathbf{e}_1), \varphi(\mathbf{e}_2), ..., \varphi(\mathbf{e}_n))$ de la base $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$. D'autre part, chaque vecteur

$\varphi(\mathbf{e}_j)$ est déterminé par ses composantes, c'est-à-dire par la j-ième colonne de $\mathbf{A}_\varphi$. L'application linéaire φ est donc déterminée par la matrice $\mathbf{A}_\varphi$.

On remarquera que

$$\text{rg}\varphi = \text{rg}\mathbf{A}_\varphi. \tag{6.16}$$

En effet, le rang de φ est, d'après 6.3.2, égal au rang de $(\varphi(\mathbf{e}_1), \varphi(\mathbf{e}_2), ..., \varphi(\mathbf{e}_n))$, donc au rang de la famille des colonnes de $\mathbf{A}_\varphi$, qui n'est autre que le rang de $\mathbf{A}_\varphi$, selon la définition 3.2.4.

Grâce au corollaire 6.3.6 et à la proposition 4.2.4, il découle de (6.16) que φ est bijective si et seulement si $m = n$ et $\mathbf{A}_\varphi$ est une matrice inversible.

Lorsque $F = E$, la base $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ joue le rôle de base de l'espace de départ et de base de l'espace d'arrivée, à moins que le contraire ne soit explicitement mentionné. Dans ce cas, $\mathbf{A}_\varphi$ est la matrice carrée dont les termes de la j-ième colonne sont les composantes de $\varphi(\mathbf{e}_j)$ dans la base $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$. On l'appelle *matrice de φ dans la base* $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$.

6.5.3 Application linéaire associée à une matrice

En présence d'une base $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ de E et d'une base $(\mathbf{f}_1, \mathbf{f}_2, ..., \mathbf{f}_m)$ de F, toute matrice $\mathbf{A} = (a_{ij})$ de type $m \times n$ définit, grâce à la proposition 6.2.4, une application linéaire φ de E dans F: l'image de $\mathbf{e}_j$ par φ est le vecteur de F dont les composantes dans la base $(\mathbf{f}_1, \mathbf{f}_2, ..., \mathbf{f}_m)$ sont les termes de la j-ième colonne de $\mathbf{A}$. Il est clair que la matrice de φ est $\mathbf{A}$. On dit que φ est l'*application linéaire associée à* $\mathbf{A}$. Lorsque $E = \mathbb{R}^n$, $F = \mathbb{R}^m$ et les bases de ces deux espaces sont les bases canoniques, on dit que φ est l'*application linéaire associée canoniquement à* $\mathbf{A}$.

6.5.4 Exemples

(1) La matrice de l'application (6.4) dans les bases canoniques de $\mathbb{R}^n$ et $\mathbb{R}$ est

$$\begin{bmatrix} a_1 & a_2 & . & . & . & . & a_n \end{bmatrix}.$$

(2) La matrice de l'application (6.5) dans les bases canoniques de $\mathbb{R}^2$ et $\mathbb{R}^3$ est

$$\begin{bmatrix} 2 & 1 \\ 3 & 0 \\ -1 & 2 \end{bmatrix}.$$

(3) La matrice de l'application linéaire de $\mathbb{R}^2$ dans $\mathbb{R}^4$

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \longrightarrow \begin{bmatrix} 3x_1 \\ -x_1 \\ 0 \\ x_2 \end{bmatrix}$$

dans les bases canoniques de $\mathbb{R}^2$ et $\mathbb{R}^4$ est

$$\begin{bmatrix} 3 & 0 \\ -1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

(4) La matrice de l'application (6.6) dans les bases canoniques de $\mathbb{R}^n$ et $\mathbb{R}^m$ est $\mathbf{A}$.

(5) Lorsque l'espace vectoriel E est de dimension finie non nulle n, la matrice d'une homothétie de rapport λ dans une base quelconque de E est $\lambda\mathbf{I}_n$. En particulier, la matrice de l'application identique est la matrice-unité $\mathbf{I}_n$ et celle de l'application nulle est la matrice nulle $\mathbf{O}$.

(6) La matrice de l'application (6.8) dans la base $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ est

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \lambda \end{bmatrix}.$$

En particulier, la matrice de la projection et celle de la symétrie sont respectivement

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{et} \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

(7) L'espace vectoriel E de l'exemple (6) de 6.2.5 étant supposé de dimension finie n, si $(\mathbf{v}_1, ..., \mathbf{v}_k)$ est une base de S et $(\mathbf{v}_{k+1}, ..., \mathbf{v}_n)$ une base de T, la matrice de l'application (6.10) dans la base $(\mathbf{v}_1, ..., \mathbf{v}_k, \mathbf{v}_{k+1}, ..., \mathbf{v}_n)$ de E est

$$k\Big\updownarrow\begin{bmatrix} 1 & & & & & \\ & \ddots & & & 0 & \\ & & 1 & & & \\ & & & \lambda & & \\ & 0 & & & \ddots & \\ & & & & & \lambda \end{bmatrix}. \tag{6.17}$$

6.5.5 Isomorphie de $L(E, F)$ et de l'espace vectoriel des matrices de type $m \times n$

Les espaces vectoriels E et F étant munis chacun d'une base, si φ et ψ sont des applications linéaires de E dans F et λ est un nombre, on voit immédiatement que

$$\mathbf{A}_{\varphi+\psi} = \mathbf{A}_\varphi + \mathbf{A}_\psi \quad \text{et} \quad \mathbf{A}_{\lambda\varphi} = \lambda\mathbf{A}_\varphi. \tag{6.18}$$

Cela signifie que l'application $\varphi \to \mathbf{A}_\varphi$ de l'espace vectoriel $L(E, F)$ dans l'espace vectoriel des matrices de type $m \times n$ est linéaire. Comme elle est également bijective (par ce qui a été vu dans 6.5.2 et 6.5.3), ces deux espaces sont isomorphes. En particulier, la dimension de $L(E, F)$ est mn.

6.5.6 Notation

Le vecteur-colonne des composantes d'un vecteur $\mathbf{x}$ dans une base donnée sera dorénavant noté par $\mathbf{x}$ indexé par la lettre désignant les vecteurs de cette base. Par exemple,

$$\mathbf{x}_e \text{ désigne le vecteur-colonne } \begin{pmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{pmatrix}, \tag{6.19}$$

où $x_1, x_2, ..., x_n$ sont les composantes de $\mathbf{x}$ dans la base $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$.

6.5.7 Calcul des composantes de l'image d'un vecteur

Les données étant celles de 6.5.2, considérons un vecteur quelconque $\mathbf{x}$ de E et posons $\mathbf{y} = \varphi(\mathbf{x})$. De la décomposition

$$\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + ... + x_n\mathbf{e}_n$$

nous déduisons, grâce à la linéarité de φ,

$$\mathbf{y} = x_1\varphi(\mathbf{e}_1) + x_2\varphi(\mathbf{e}_2) + ... + x_n\varphi(\mathbf{e}_n).$$

En égalant les composantes respectives des deux membres de cette relation, nous obtenons, vu la définition de la matrice (6.15),

$$\begin{pmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ \cdot \\ y_m \end{pmatrix} = x_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \cdot \\ \cdot \\ \cdot \\ a_{m1} \end{pmatrix} + x_2 \begin{pmatrix} a_{12} \\ a_{22} \\ \cdot \\ \cdot \\ \cdot \\ a_{m2} \end{pmatrix} + ... + x_n \begin{pmatrix} a_{1n} \\ a_{2n} \\ \cdot \\ \cdot \\ \cdot \\ a_{mn} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & \cdots\cdots & & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{pmatrix},$$

autrement dit

$$\mathbf{y}_f = \mathbf{A}_\varphi \mathbf{x}_e. \tag{6.20}$$

On dit que (6.20) est l'*équation matricielle de l'application linéaire φ*.

On notera que toute relation de la forme

$$\mathbf{y}_f = \mathbf{A}\mathbf{x}_e \tag{6.21}$$

définit une application linéaire φ de E dans F dont la matrice n'est autre que $\mathbf{A}$. On peut dire également que (6.21) est l'équation matricielle de l'application linéaire φ associée à la matrice $\mathbf{A}$ selon 6.5.3.

6.5.8 Exemple

Nous nous proposons de calculer la matrice de la projection définie par (6.11), dans une base orthonormale $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ de E arbitrairement choisie. A cet effet, observons d'abord que le produit scalaire de deux vecteurs $\mathbf{u}$ et $\mathbf{v}$ peut s'écrire, compte tenu de (2.28) et de 4.3.1, sous la forme

$$(\mathbf{u} \mid \mathbf{v}) = {}^t\mathbf{u}_e\mathbf{v}_e = {}^t\mathbf{v}_e\mathbf{u}_e. \tag{6.22}$$

A l'aide de cette expression, la relation vectorielle

$$\mathbf{y} = \mathbf{x} - \frac{(\mathbf{x} \mid \mathbf{n})}{(\mathbf{v} \mid \mathbf{n})}\mathbf{v}$$

s'exprime, de manière équivalente, sous la forme

$$\mathbf{y}_e = \mathbf{x}_e - \frac{{}^t\mathbf{x}_e\mathbf{n}_e}{{}^t\mathbf{v}_e\mathbf{n}_e}\mathbf{v}_e.$$

Or, le produit du nombre ${}^t\mathbf{x}_e\mathbf{n}_e = {}^t\mathbf{n}_e\mathbf{x}_e$ par le vecteur-colonne $\mathbf{v}_e$ peut s'écrire sous la forme matricielle

$$\mathbf{v}_e({}^t\mathbf{n}_e\mathbf{x}_e) = (\mathbf{v}_e{}^t\mathbf{n}_e)\mathbf{x}_e.$$

Par conséquent,

$$\mathbf{y}_e = \mathbf{x}_e - \frac{1}{{}^t\mathbf{v}_e\mathbf{n}_e}(\mathbf{v}_e{}^t\mathbf{n}_e)\mathbf{x}_e = (\mathbf{I} - \frac{1}{{}^t\mathbf{v}_e\mathbf{n}_e}\mathbf{v}_e{}^t\mathbf{n}_e)\mathbf{x}_e,$$

d'où nous concluons que la matrice de la projection définie par (6.11) est

$$\mathbf{I} - \frac{1}{{}^t\mathbf{v}_e\mathbf{n}_e}\mathbf{v}_e{}^t\mathbf{n}_e. \tag{6.23}$$

Il résulte de même que la matrice de la dilatation définie par (6.12) est

$$\mathbf{I} - \frac{1-\lambda}{{}^t\mathbf{v}_e\mathbf{n}_e}\mathbf{v}_e{}^t\mathbf{n}_e. \tag{6.24}$$

Exemple numérique pour la projection et la symétrie:

$$\mathbf{n}_e = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{v}_e = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}, \quad {}^t\mathbf{v}_e\mathbf{n}_e = 2, \quad \mathbf{v}_e{}^t\mathbf{n}_e = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \\ -1 & -1 & -1 \end{bmatrix},$$

$$\mathbf{I} - \frac{1}{{}^t\mathbf{v}_e\mathbf{n}_e}\mathbf{v}_e{}^t\mathbf{n}_e = \frac{1}{2}\begin{bmatrix} 1 & -1 & -1 \\ -2 & 0 & -2 \\ 1 & 1 & 3 \end{bmatrix}, \quad \mathbf{I} - \frac{2}{{}^t\mathbf{v}_e\mathbf{n}_e}\mathbf{v}_e{}^t\mathbf{n}_e = \begin{bmatrix} 0 & -1 & -1 \\ -2 & -1 & -2 \\ 1 & 1 & 2 \end{bmatrix}.$$

6.5.9 Interprétation matricielle de l'inverse d'une application linéaire

Soit φ une application linéaire bijective de E dans F, $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ une base de E et $(\mathbf{f}_1, \mathbf{f}_2, ..., \mathbf{f}_n)$ une base de F. L'équation matricielle de l'application φ^{-1} est

$$\mathbf{x}_e = \mathbf{A}_{\varphi^{-1}}\mathbf{y}_f. \tag{6.25}$$

D'autre part, $\mathbf{A}_\varphi$ étant une matrice inversible, de (6.20) il résulte que

$$\mathbf{x}_e = \mathbf{A}_\varphi^{-1}\mathbf{y}_f. \tag{6.26}$$

En comparant (6.26) à (6.25), nous concluons, grâce à la remarque (6) de 4.1.7, que

$$\mathbf{A}_{\varphi^{-1}} = \mathbf{A}_\varphi^{-1}. \tag{6.27}$$

En d'autres termes, la matrice de l'application inverse de φ est l'inverse de la matrice de φ. Cette conclusion résulte également de la relation (6.28) ci-dessous.

6.5.10 Interprétation matricielle de la composée d'applications linéaires

Soit φ une application linéaire de E dans F et ψ une application linéaire de F dans G. Soit en outre $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$, $(\mathbf{f}_1, \mathbf{f}_2, ..., \mathbf{f}_m)$ et $(\mathbf{g}_1, \mathbf{g}_2, ..., \mathbf{g}_p)$ des bases respectives de E, F et G. En vertu de (6.20),

$$\mathbf{y}_f = \mathbf{A}_\varphi\mathbf{x}_e \quad \text{et} \quad \mathbf{z}_g = \mathbf{A}_\psi\mathbf{y}_f.$$

Par substitution de $\mathbf{A}_\varphi\mathbf{x}_e$ à $\mathbf{y}_f$ dans la deuxième relation, il résulte que

$$\mathbf{z}_g = \mathbf{A}_\psi\mathbf{A}_\varphi\mathbf{x}_e.$$

En comparant cette relation à l'équation matricielle de l'application $\psi\circ\varphi$, nous concluons, encore par la remarque (6) de 4.1.7, que

$$\mathbf{A}_{\psi\circ\varphi} = \mathbf{A}_\psi\mathbf{A}_\varphi. \tag{6.28}$$

En d'autres termes, la matrice de l'application composée de φ et de ψ est le produit des matrices de ψ et de φ. Dans le cas particulier où φ est une application linéaire de E dans E, il s'avère ainsi que

$$\mathbf{A}_{\varphi^k} = \mathbf{A}_\varphi^k \quad (k \text{ entier positif}). \tag{6.29}$$

Exemple d'application. Nous avons déjà remarqué que toute projection (symétrie) φ satisfait à la relation $\varphi^2 = \varphi$ ($\varphi^2 = \mathrm{id}_E$). Vu (6.28), la matrice d'une projection (symétrie) vérifie donc la relation $\mathbf{A}^2 = \mathbf{A}$ ($\mathbf{A}^2 = \mathbf{I}$).

6.6 CHANGEMENTS DE BASE

6.6.1 Introduction

La résolution de certains problèmes est facilitée lorsqu'on passe de la base initiale à une base auxiliaire mieux appropriée aux données du problème. Dans cette section, nous établirons la liaison entre composantes d'un même vecteur dans des bases différentes, ainsi qu'entre matrices représentatives d'une même application linéaire dans des bases différentes. Nous verrons dans le chapitre 8 par quels critères le choix d'une nouvelle base pourra être fait.

Les lettres E et F désigneront encore des espaces vectoriels de dimensions finies non nulles respectives n et m.

6.6.2 Matrices de passage

Soit $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ et $(\mathbf{e}'_1, \mathbf{e}'_2, ..., \mathbf{e}'_n)$ des bases de E. On appelle *matrice de passage* de $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ à $(\mathbf{e}'_1, \mathbf{e}'_2, ..., \mathbf{e}'_n)$, et on note $\mathbf{P}_{\mathrm{ee}'}$, la matrice carrée d'ordre n dont les termes de la j-ième colonne sont les composantes de $\mathbf{e}'_j$ dans la base $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$. Autrement dit,

$$\mathbf{P}_{\mathrm{ee}'} = \begin{pmatrix} p_{11} & \cdots & p_{1j} & \cdots & p_{1n} \\ p_{21} & \cdots & p_{2j} & \cdots & p_{2n} \\ & & \cdots\cdots & & \\ p_{n1} & \cdots & p_{nj} & \cdots & p_{nn} \end{pmatrix}, \tag{6.30}$$

où $p_{1j}, p_{2j}, .., p_{nj}$ sont définis par la décomposition

$$\mathbf{e}'_j = p_{1j}\mathbf{e}_1 + p_{2j}\mathbf{e}_2 + ... + p_{nj}\mathbf{e}_n.$$

On remarquera que le rang de $\mathbf{P}_{\mathrm{ee}'}$ est n, ce qui équivaut à dire que $\mathbf{P}_{\mathrm{ee}'}$ est une matrice inversible.

6.6.3 Matrices de passage vues comme matrices d'applications linéaires

Voici deux interprétations de la notion de matrice de passage:

(1) Soit φ l'application linéaire de E dans E définie par $\varphi(\mathbf{e}_1) = \mathbf{e}'_1$, $\varphi(\mathbf{e}_2) = \mathbf{e}'_2$, ..., $\varphi(\mathbf{e}_n) = \mathbf{e}'_n$. La matrice de φ dans la base $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ est $\mathbf{P}_{\mathrm{ee}'}$.

(2) Soit φ l'application identique $\mathbf{x} \to \mathbf{x}$ de E, muni de la base $(\mathbf{e}'_1, \mathbf{e}'_2, ..., \mathbf{e}'_n)$, dans E muni de la base $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$. La matrice de φ dans les bases $(\mathbf{e}'_1, \mathbf{e}'_2, ..., \mathbf{e}'_n)$ et $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ est $\mathbf{P}_{\mathrm{ee}'}$.

6.6.4 Transformation des composantes par suite d'un changement de base

Les composantes d'un vecteur $\mathbf{x}$ de E dans les bases $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ et $(\mathbf{e}'_1, \mathbf{e}'_2, ..., \mathbf{e}'_n)$ sont liées par la relation

$$\mathbf{x}_{\mathrm{e}} = \mathbf{P}_{\mathrm{ee}'}\mathbf{x}_{\mathrm{e}'}. \tag{6.31}$$

En effet, cette relation n'est autre que (6.20) appliquée au cas où φ est l'application identique $\mathbf{x} \to \mathbf{x}$ de E, muni de la base $(\mathbf{e}'_1, \mathbf{e}'_2, ..., \mathbf{e}'_n)$, dans E muni de la base $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$.

6.6.5 Remarque

Si $\mathbf{P}$ est une matrice carrée d'ordre n telle que

$$\mathbf{x}_{\mathrm{e}} = \mathbf{P}\mathbf{x}_{\mathrm{e}'} \tag{6.32}$$

pour tout vecteur $\mathbf{x}$ de E, alors $\mathbf{P} = \mathbf{P}_{\mathrm{ee}'}$. Cela résulte de la comparaison de (6.32) à (6.31), compte tenu de la remarque (6) de 4.1.7.

6.6.6 Changements de base inverses

En multipliant les deux membres de (6.31) par $\mathbf{P}_{\mathrm{ee}'}^{-1}$, nous obtenons

$$\mathbf{x}_{\mathrm{e}'} = \mathbf{P}_{\mathrm{ee}'}^{-1}\mathbf{x}_{\mathrm{e}},$$

d'où nous déduisons, grâce à la remarque 6.6.5, que

$$\mathbf{P}_{\mathrm{ee}'}^{-1} = \mathbf{P}_{\mathrm{e}'\mathrm{e}}. \tag{6.33}$$

6.6.7 Changements de base successifs

Soit $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$, $(\mathbf{e}'_1, \mathbf{e}'_2, ..., \mathbf{e}'_n)$ et $(\mathbf{e}''_1, \mathbf{e}''_2, ... \mathbf{e}''_n)$ des bases de E. D'après (6.31),

$$\mathbf{x}_{\mathrm{e}} = \mathbf{P}_{\mathrm{ee}'}\mathbf{x}_{\mathrm{e}'} \quad \text{et} \quad \mathbf{x}_{\mathrm{e}'} = \mathbf{P}_{\mathrm{e}'\mathrm{e}''}\mathbf{x}_{\mathrm{e}''}.$$

En substituant $\mathbf{P}_{\mathrm{e}'\mathrm{e}''}\mathbf{x}_{\mathrm{e}''}$ à $\mathbf{x}_{\mathrm{e}'}$ dans la première relation, nous obtenons

$$\mathbf{x}_{\mathrm{e}} = \mathbf{P}_{\mathrm{ee}'}\mathbf{P}_{\mathrm{e}'\mathrm{e}''}\mathbf{x}_{\mathrm{e}''},$$

d'où nous déduisons, encore par la remarque 6.6.5, que

$$\mathbf{P}_{\mathrm{ee}'}\mathbf{P}_{\mathrm{e}'\mathrm{e}''} = \mathbf{P}_{\mathrm{ee}''}. \tag{6.34}$$

6.6.8 Déterminants de passage, orientation d'un espace vectoriel

On appelle *déterminant de passage* d'une base $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ à une base $(\mathbf{e}'_1, \mathbf{e}'_2, ..., \mathbf{e}'_n)$ le déterminant de la matrice de passage $\mathbf{P}_{\mathrm{ee}'}$.

L'ensemble des bases de E peut être partagé en deux classes de la manière suivante: $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ étant une base arbitrairement choisie, on range dans la première classe toutes les bases $(\mathbf{e}'_1, \mathbf{e}'_2, ..., \mathbf{e}'_n)$ telles que le déterminant de passage $|\mathbf{P}_{ee'}|$ est positif et dans la seconde toutes les bases $(\mathbf{e}'_1, \mathbf{e}'_2, ..., \mathbf{e}'_n)$ telles que le déterminant de passage $|\mathbf{P}_{ee'}|$ est négatif. De (6.33), (6.34), (5.10) et (a) de la proposition 5.1.8, il résulte aisément que le déterminant de passage entre bases d'une même classe est positif, tandis qu'entre bases de classes différentes ce même déterminant est négatif.

On oriente l'espace vectoriel E en choisissant une des deux classes et en disant que les bases de la classe choisie sont *directes* ou *positivement orientées*. Les bases de l'autre classe sont alors dites *indirectes* ou *négativement orientées*. On dit également que la classe choisie définit l'*orientation* de E. D'habitude, on indique ce choix par la donnée d'une base.

6.6.9 Orientation d'un espace affine

On dit qu'un espace affine $\mathscr{E}$ de direction E est orienté si E est orienté.

On dit qu'un repère d'un espace affine orienté est direct ou indirect, suivant que la base de ce repère est directe ou indirecte.

6.6.10 Transformation de la matrice d'une application linéaire par suite d'un changement de bases

Soit φ une application linéaire de E dans F. Soit en outre $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ et $(\mathbf{e}'_1, \mathbf{e}'_2, ..., \mathbf{e}'_n)$ des bases de E, $(\mathbf{f}_1, \mathbf{f}_2, ..., \mathbf{f}_m)$ et $(\mathbf{f}'_1, \mathbf{f}'_2, ..., \mathbf{f}'_m)$ des bases de F. Désignons par $\mathbf{A}$ la matrice de φ dans les bases $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ et $(\mathbf{f}_1, \mathbf{f}_2, ..., \mathbf{f}_m)$ et par $\mathbf{A}'$ la matrice de φ dans les bases $(\mathbf{e}'_1, \mathbf{e}'_2, ..., \mathbf{e}'_n)$ et $(\mathbf{f}'_1, \mathbf{f}'_2, ..., \mathbf{f}'_m)$. Soit $\mathbf{x}$ un vecteur quelconque de E. D'après (6.20),

$$(\varphi(\mathbf{x}))_f = \mathbf{A}\mathbf{x}_e \quad \text{et} \quad (\varphi(\mathbf{x}))_{f'} = \mathbf{A}'\mathbf{x}_{e'}.$$

D'autre part, en vertu de (6.31),

$$(\varphi(\mathbf{x}))_f = \mathbf{P}_{ff'}(\varphi(\mathbf{x}))_{f'}.$$

Il s'ensuit que

$$\mathbf{A}\mathbf{x}_e = \mathbf{P}_{ff'}\mathbf{A}'\mathbf{x}_{e'}.$$

De là, par substitution de $\mathbf{P}_{ee'}\mathbf{x}_{e'}$ à $\mathbf{x}_e$, nous concluons, grâce à la remarque (6) de 4.1.7, que

$$\mathbf{A}\mathbf{P}_{ee'} = \mathbf{P}_{ff'}\mathbf{A}',$$

ce qui s'écrit également sous les formes

$$\mathbf{A}' = \mathbf{P}_{ff'}^{-1}\mathbf{A}\mathbf{P}_{ee'} \quad \text{ou} \quad \mathbf{A} = \mathbf{P}_{ff'}\mathbf{A}'\mathbf{P}_{ee'}^{-1}. \tag{6.35}$$

Dans le cas particulier où $F = E$, ces relations deviennent respectivement

$$\mathbf{A}' = \mathbf{P}_{ee'}^{-1}\mathbf{A}\mathbf{P}_{ee'} \quad \text{et} \quad \mathbf{A} = \mathbf{P}_{ee'}\mathbf{A}'\mathbf{P}_{ee'}^{-1}. \tag{6.36}$$

6.6.11 Exemple

L'espace vectoriel E étant de dimension 3 et muni d'une base $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$, considérons le plan vectoriel S engendré par $\mathbf{v}_1(0, 2, 1)$ et $\mathbf{v}_2(1, 0, 3)$ et la droite vectorielle D engendrée par $\mathbf{v}_3(0, -1, 1)$. Nous nous proposons de trouver la matrice $\mathbf{A}$ de la dilatation relativement à S, de direction D et de rapport -3. Posons $\mathbf{e}'_1 = \mathbf{v}_1$, $\mathbf{e}'_2 = \mathbf{v}_2$, $\mathbf{e}'_3 = \mathbf{v}_3$. La matrice de la dilatation dans la base $(\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3)$ est

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

D'autre part,

$$\mathbf{P}_{ee'} = \begin{pmatrix} 0 & 1 & 0 \\ 2 & 0 & -1 \\ 1 & 3 & 1 \end{pmatrix}, \quad \mathbf{P}_{ee'}^{-1} = \frac{1}{3}\begin{pmatrix} -3 & 1 & 1 \\ 3 & 0 & 0 \\ -6 & -1 & 2 \end{pmatrix}.$$

En vertu de (6.36), la matrice de la dilatation dans la base $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ est donc

$$\mathbf{A} = \mathbf{P}_{ee'}\mathbf{A}'\mathbf{P}_{ee'}^{-1} = \frac{1}{3}\begin{pmatrix} 3 & 0 & 0 \\ -24 & -1 & 8 \\ 24 & 4 & -5 \end{pmatrix}.$$

Ce résultat peut être obtenu plus rapidement à l'aide de la formule (6.24). Pour pouvoir appliquer cette formule, on considère E comme étant muni du produit scalaire rendant la base $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ orthonormale (cf. 2.5.7).

6.6.12 Matrices semblables

Soit $\mathbf{A}$ et $\mathbf{B}$ des matrices carrées du même ordre. On dit que $\mathbf{A}$ et $\mathbf{B}$ sont *semblables*, ou que $\mathbf{A}$ est *semblable à* $\mathbf{B}$, s'il existe une matrice inversible $\mathbf{P}$ telle que

$$\mathbf{B} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}. \tag{6.37}$$

D'après (6.36), deux matrices d'une même application linéaire de E dans E sont semblables. La réciproque de cette assertion est également vraie. En effet, si $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ est une base quelconque de E, φ l'application linéaire associée à $\mathbf{A}$ selon 6.5.3, $\mathbf{e}'_j$ le vecteur de E dont les composantes dans la base $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ sont les termes de la j-ième colonne de $\mathbf{P}$, alors le premier membre de (6.37) est la matrice de φ dans la base $(\mathbf{e}'_1, \mathbf{e}'_2, ..., \mathbf{e}'_n)$.

Deux matrices semblables ont le même rang, à savoir le rang d'une application linéaire quelconque qui leur est simultanément associée. Elles ont la même trace, puisqu'en vertu de (4.5),

$$\mathrm{tr}\,\mathbf{B} = \mathrm{tr}(\mathbf{P}^{-1}\mathbf{A}\mathbf{P}) = \mathrm{tr}(\mathbf{P}\mathbf{P}^{-1}\mathbf{A}) = \mathrm{tr}\,\mathbf{A}.$$

Elles ont aussi le même déterminant, car, d'après (a) de la proposition 5.1.8 et (5.10),

$$|\mathbf{B}| = |\mathbf{P}^{-1}\mathbf{A}\mathbf{P}| = |\mathbf{P}|^{-1}|\mathbf{A}||\mathbf{P}| = |\mathbf{A}|.$$

Cette dernière conclusion nous suggère la définition suivante.

6.6.13 Déterminant d'une application linéaire

Soit φ une application linéaire de E dans E. On appelle *déterminant* de φ, et on note $\det\varphi$, le déterminant de la matrice de φ dans une base quelconque de E.

Par exemple, si φ est une homothétie de rapport λ (une dilatation de rapport λ relativement à un sous-espace vectoriel de dimension positive k),

$$\det\varphi = \lambda^n \quad (\det\varphi = \lambda^{n-k}). \tag{6.38}$$

En effet, les matrices de ces deux applications sont respectivement $\lambda\mathbf{I}_n$ et la matrice (6.17). En particulier, si φ est une symétrie par rapport à un hyperplan vectoriel,

$$\det\varphi = -1. \tag{6.39}$$

On notera qu'une application linéaire de E dans E est bijective si et seulement si son déterminant est non nul. En effet, pour qu'une telle application soit inversible, il faut et il suffit que sa matrice soit inversible, d'après ce qui a été établi dans 6.5.2.

La première des deux relations suivantes découle de (6.27) et (5.10) (φ étant supposée bijective) et la seconde de (6.28) et (a) de la proposition 5.1.8:

$$\det\varphi^{-1} = (\det\varphi)^{-1}, \quad \det(\psi\circ\varphi) = \det\psi\det\varphi. \tag{6.40}$$

6.6.14 Valeurs semblables de fonctions matricielles

La relation (6.37) s'étend, par récurrence, aux puissances positives de $\mathbf{A}$ et de $\mathbf{B}$. En effet, si $\mathbf{B}^{k-1} = \mathbf{P}^{-1}\mathbf{A}^{k-1}\mathbf{P}$, alors $\mathbf{B}^k = \mathbf{B}^{k-1}\mathbf{B} = (\mathbf{P}^{-1}\mathbf{A}^{k-1}\mathbf{P})(\mathbf{P}^{-1}\mathbf{A}\mathbf{P})$, ce qui entraîne

$$\mathbf{B}^k = \mathbf{P}^{-1}\mathbf{A}^k\mathbf{P}. \tag{6.41}$$

Si $\mathbf{A}$ ou $\mathbf{B}$ sont inversibles, cette relation est également vraie pour les puissances entières négatives, car les deux membres de (6.41) peuvent, dans ce cas, être inversés.

En passant aux combinaisons linéaires de puissances, nous voyons aussi que (6.37) entraîne

$$f(\mathbf{B}) = \mathbf{P}^{-1}f(\mathbf{A})\mathbf{P}, \tag{6.42}$$

où f est un polynôme ou, plus généralement, une fonction définie par une série entière appropriée (cf. 4.5.2).

6.7 APPLICATIONS AFFINES

6.7.1 Introduction

Les applications affines sont des applications linéaires entre espaces affines convenablement vectorialisés. La plupart de leurs propriétés se déduisent immédiatement des propriétés des applications linéaires étudiées dans les sections précédentes de ce chapitre. Dans cette section, nous aborderons l'étude de la notion d'application affine et présenterons quelques exemples qui en illustrent l'importance. Nous énoncerons, en outre, un certain nombre de résultats généraux.

Tout au long de la section, $\mathscr{E}$ désignera un espace affine de direction E et $\mathscr{F}$ un espace affine de direction F.

6.7.2 Applications affines

On dit qu'une application Φ de $\mathscr{E}$ dans $\mathscr{F}$ est *affine* s'il existe un point O de $\mathscr{E}$ et une application linéaire φ de E dans F tels que

$$\Phi(P) = \Phi(O) + \varphi(\overrightarrow{OP}) \quad (\text{ou } \overrightarrow{\Phi(O)\Phi(P)} = \varphi(\overrightarrow{OP})) \tag{6.43}$$

pour tout point P de $\mathscr{E}$.

Contrairement à ce que laisse entendre cette définition, le point O ne joue aucun rôle distinctif. Nous allons voir, en effet, que tout point P_0 de $\mathscr{E}$ remplit les mêmes fonctions que O. Supposons que la relation (6.43) soit vraie pour tout point P de $\mathscr{E}$. Alors

$$\begin{aligned} \overrightarrow{\Phi(P_0)\Phi(P)} &= \overrightarrow{\Phi(O)\Phi(P)} - \overrightarrow{\Phi(O)\Phi(P_0)} \\ &= \varphi(\overrightarrow{OP}) - \varphi(\overrightarrow{OP_0}) = \varphi(\overrightarrow{P_0P}), \end{aligned}$$

autrement dit,

$$\Phi(P) = \Phi(P_0) + \varphi(\overrightarrow{P_0P}) \tag{6.44}$$

pour tout point P de $\mathscr{E}$.

En revanche, l'application linéaire φ est déterminée par l'application affine Φ. En effet, (6.43) détermine $\varphi(\overrightarrow{OP})$ pour tout point P de $\mathscr{E}$, donc l'image $\varphi(\mathbf{x})$ de tout vecteur $\mathbf{x}$ de E. On dit que φ est l'*application linéaire associée à* Φ.

Nous nous référerons à (6.44) en disant que Φ est *exprimée relativement au point* P_0.

On remarquera qu'en vertu de (6.44), un point quelconque P_0 de $\mathscr{E}$, son image $\Phi(P_0)$ et l'application linéaire φ déterminent l'image $\Phi(P)$ de tout point P de $\mathscr{E}$ et donc l'application affine Φ.

6.7.3 Applications affines vues comme applications linéaires

Par l'intermédiaire de la notion d'espace vectorialisé d'un espace affine (cf. 1.9.6), les applications affines peuvent être considérées comme des applications linéaires. La relation (6.43) (version entre parenthèses) est, en effet, l'exacte expression du fait que Φ est une application linéaire du vectorialisé $\mathscr{E}_O$ dans le vectorialisé $\mathscr{F}_{\Phi(O)}$ (fig. 6.2).

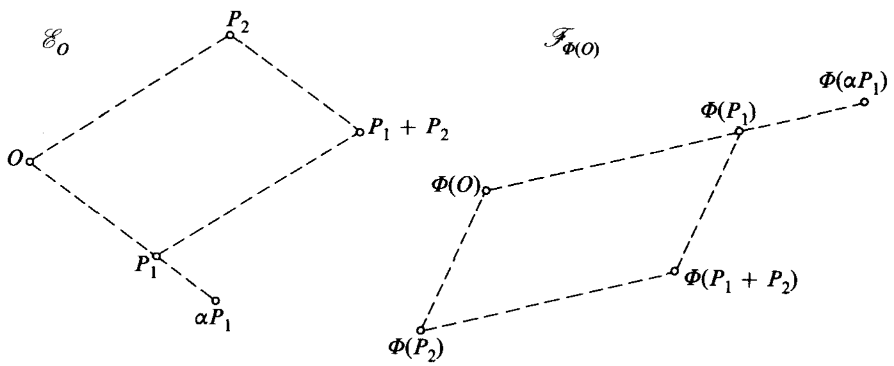


Fig. 6.2

6.7.4 Applications affines entre espaces vectoriels

Relativement au vecteur nul de E, une application affine d'un espace vectoriel E dans un espace vectoriel F (considérés comme des espaces affines) s'exprime par la relation

$$\Phi(\mathbf{x}) = \Phi(\mathbf{0}) + \varphi(\mathbf{x}), \tag{6.45}$$

où $\mathbf{x}$ est un vecteur quelconque de E et φ est l'application linéaire associée à Φ. Nous voyons ainsi que Φ est une application linéaire si et seulement si $\Phi(\mathbf{0})$ est le vecteur nul de F.

6.7.5 Applications constantes

Toute application constante de $\mathscr{E}$ dans $\mathscr{F}$ est une application affine. En effet, une application Φ de cette sorte vérifie la relation (6.44) avec $\varphi = 0$, où 0 désigne l'application nulle de E dans F. Réciproquement, si l'application linéaire associée à une application affine est nulle, cette application affine est constante.

6.7.6 Translations

Soit $\mathbf{v}$ un vecteur de E. On appelle *translation de vecteur* $\mathbf{v}$ l'application $\Phi_{\mathbf{v}}$ de $\mathscr{E}$ dans $\mathscr{E}$ définie par

$$\Phi_{\mathbf{v}}(P) = P + \mathbf{v}. \tag{6.46}$$

Toute translation $\Phi_{\mathbf{v}}$ est une application affine. En effet, P_0 étant un point arbitrairement choisi, il découle de (6.46) que $\mathbf{v} = \overrightarrow{P_0\Phi_{\mathbf{v}}(P_0)}$, donc que

$$\Phi_{\mathbf{v}}(P) = P + \overrightarrow{P_0\Phi_{\mathbf{v}}(P_0)} = \Phi_{\mathbf{v}}(P_0) + \overrightarrow{P_0P}, \tag{6.47}$$

la dernière égalité étant justifiée par la règle (4) de 1.9.5. Cela montre que $\Phi_{\mathbf{v}}$ vérifie la relation (6.44) avec $\varphi = \mathrm{id}_E$.

Réciproquement, toute application affine Φ de $\mathscr{E}$ dans $\mathscr{E}$ dont l'application linéaire associée est id_E est une translation. En effet, l'expression de Φ relativement à un point P_0 quelconque devient, grâce à la règle mentionnée ci-dessus,

$$\Phi(P) = \Phi(P_0) + \overrightarrow{P_0P} = P + \overrightarrow{P_0\Phi(P_0)},$$

ce qui entraîne, par comparaison avec (6.47) ou (6.46), que $\Phi = \Phi_{\mathbf{v}}$, où $\mathbf{v} = \overrightarrow{P_0\Phi(P_0)}$.

6.7.7 Points invariants

Soit Φ une application affine de $\mathscr{E}$ dans $\mathscr{E}$, φ l'application linéaire associée à Φ et P_0 un point invariant. Relativement à P_0, Φ s'exprime par la relation

$$\Phi(P) = P_0 + \varphi(\overrightarrow{P_0P}) \quad (\text{ou } \overrightarrow{P_0\Phi(P)} = \varphi(\overrightarrow{P_0P})), \tag{6.48}$$

qui n'est autre que (6.44) appliquée au cas où $\Phi(P_0) = P_0$. Les autres points invariants sont donc les points P tels que

$$\overrightarrow{P_0P} = \varphi(\overrightarrow{P_0P}).$$

En d'autres termes, l'ensemble des points invariants est le sous-espace affine de $\mathscr{E}$

$$\mathscr{S} = P_0 + S,$$

où S est le sous-espace vectoriel de E formé des vecteurs invariants par φ.

6.7.8 Homothéties

Soit P_0 un point de $\mathscr{E}$ et λ un nombre. On appelle *homothétie de centre* P_0 *et de rapport* λ l'application Φ de $\mathscr{E}$ dans $\mathscr{E}$ définie par

$$\Phi(P) = P_0 + \lambda\overrightarrow{P_0P}. \tag{6.49}$$

L'homothétie de rapport 1 est l'identité et celle de rapport -1 est appelée *symétrie centrale de centre* P_0.

Le centre P_0 étant un point invariant, en comparant (6.49) à (6.48), nous concluons qu'une homothétie de rapport λ est une application affine dont l'application linéaire associée est $\lambda \mathrm{id}_E$ (c'est-à-dire une homothétie vectorielle de rapport λ).

Réciproquement, nous allons montrer que si l'application linéaire associée à une application affine Φ de $\mathscr{E}$ dans $\mathscr{E}$ est $\lambda \mathrm{id}_E$, Φ est une homothétie de rapport λ ou une translation, suivant que $\lambda \neq 1$ ou $\lambda = 1$. La conclusion dans l'hypothèse où $\lambda = 1$ ayant été prouvée dans 6.7.6, il nous reste à considérer le cas où $\lambda \neq 1$. Il suffit d'établir l'existence d'un point invariant P_0, car en exprimant Φ relativement à ce point, nous voyons que Φ vérifie la relation (6.49). Exprimons Φ relativement à un point O arbitrairement choisi. L'équation $\Phi(P) = P$ s'écrit alors sous la forme

$$P = \Phi(O) + \lambda\overrightarrow{OP},$$

ou encore

$$\overrightarrow{OP} = \overrightarrow{O\Phi(O)} + \lambda\overrightarrow{OP}.$$

L'unique solution de cette équation est le point P_0 tel que

$$\overrightarrow{OP_0} = \frac{1}{1-\lambda}\overrightarrow{O\Phi(O)},$$

c'est-à-dire le point

$$P_0 = O + \frac{1}{1-\lambda}\overrightarrow{O\Phi(O)}. \tag{6.50}$$

6.7.9 Dilatations affines

On dit qu'une application affine de $\mathscr{E}$ dans $\mathscr{E}$ est une *dilatation affine* (*projection affine*, *symétrie affine*) si elle admet au moins un point invariant et si l'application linéaire associée est une dilatation (projection, symétrie) de l'espace directeur E.

Plus explicitement, supposons que E soit somme directe de deux sous-espaces vectoriels S et T différents de $\{\mathbf{0}\}$ et désignons par φ la dilatation (vectorielle) relativement à S, de direction T et de rapport λ. Considérons une dilatation affine

Φ admettant φ comme application linéaire associée. Relativement à un point invariant P_0, Φ s'exprime par la relation

$$\Phi(P) = P_0 + \varphi(\overrightarrow{P_0P}) = P_0 + \mathbf{s} + \lambda\mathbf{t}, \tag{6.51}$$

où $\mathbf{s} + \mathbf{t}$ est la décomposition (6.9) du vecteur $\overrightarrow{P_0P}$ suivant S et T. Ecartons le cas où $\lambda = 1$, c'est-à-dire où Φ est l'application identique de $\mathscr{E}$. D'après ce qui a été montré dans (6) de 6.2.5, S est alors le sous-espace invariant de la dilatation (vectorielle) φ. Il s'ensuit, grâce aux résultats établis dans 6.7.7, que l'ensemble des points invariants est le sous-espace affine $\mathscr{S} = P_0 + S$. On l'appelle *sous-espace invariant de la dilatation affine* Φ. Posons maintenant

$$\operatorname{proj} P = P_0 + \mathbf{s}. \tag{6.52}$$

On notera que $\operatorname{proj} P$ n'est autre que l'image de P par la projection affine obtenue en posant $\lambda = 0$ dans (6.51). A l'aide de (6.52), nous pouvons écrire

$$P = P_0 + \overrightarrow{P_0P} = P_0 + \mathbf{s} + \mathbf{t} = \operatorname{proj} P + \mathbf{t},$$

d'où nous déduisons que

$$\mathbf{t} = \overrightarrow{(\operatorname{proj} P)P}. \tag{6.53}$$

Grâce à (6.52) et (6.53), la relation (6.51) peut donc être récrite sous la forme

$$\Phi(P) = \operatorname{proj} P + \lambda\overrightarrow{(\operatorname{proj} P)P}. \tag{6.54}$$

Les conclusions de la discussion qui précède suggèrent la terminologie plus explicite que voici. On appelle la dilatation affine Φ *dilatation relativement à* $\mathscr{S}$, *de direction* T *et de rapport* λ. La dilatation de rapport 0 est appelée *projection sur* $\mathscr{S}$ *parallèlement à* T et celle de rapport -1 *symétrie par rapport à* $\mathscr{S}$ *parallèlement à* T (fig. 6.3).

Si $\mathscr{E}$ est euclidien et $T = S^{\perp}$, on parle préférablement de *dilatation orthogonale à* $\mathscr{S}$, de *projection orthogonale sur* $\mathscr{S}$ et de *symétrie orthogonale par rapport à* $\mathscr{S}$.

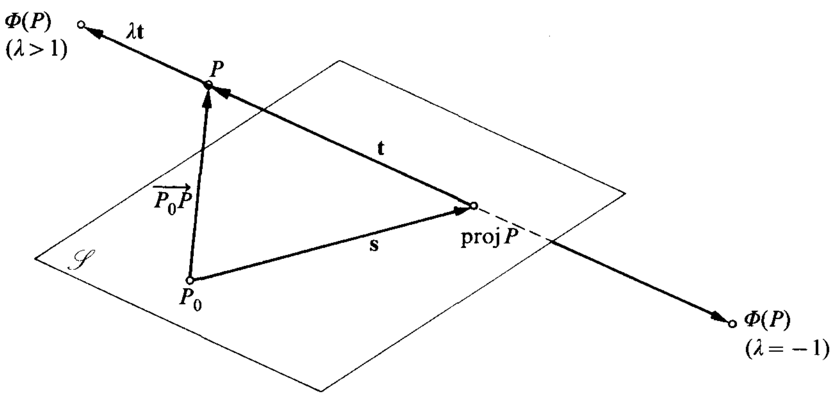


Fig. 6.3

Si $\mathscr{E}$ est de dimension finie non nulle et $\mathscr{S}$ est un hyperplan, une dilatation relativement à $\mathscr{S}$ est également appelée *affinité*. En accord avec ce terme, on dit que $\mathscr{S}$ est l'*hyperplan d'affinité*, T la *direction d'affinité* et λ le *rapport d'affinité*. On notera que dans ce cas

$$\lambda = \det\varphi, \tag{6.55}$$

où φ est la dilatation (vectorielle) associée à l'affinité (poser $k = n - 1$ dans (6.38)).

6.7.10 Equation matricielle d'une application affine

Supposons que les espaces $\mathscr{E}$ et $\mathscr{F}$ soient de dimensions finies non nulles respectives n et m et choisissons un repère $(O_{\mathscr{E}}; \mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ de $\mathscr{E}$ et un repère $(O_{\mathscr{F}}; \mathbf{f}_1, \mathbf{f}_2, ..., \mathbf{f}_m)$ de $\mathscr{F}$. Soit Φ une application affine de $\mathscr{E}$ dans $\mathscr{F}$ et φ l'application linéaire associée à Φ. Désignons par $\mathbf{x}$ et $\mathbf{y}$ les vecteurs-colonnes des coordonnées respectives d'un point générique P de $\mathscr{E}$ et de son image $\Phi(P)$ et par $\mathbf{b}$ le vecteur-colonne des coordonnées de $\Phi(O_{\mathscr{E}})$. Soit $\mathbf{A}$ la matrice de φ dans les bases $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ et $(\mathbf{f}_1, \mathbf{f}_2, ..., \mathbf{f}_m)$ (dans la base $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ si $\mathscr{F} = \mathscr{E}$). La relation (6.43) (avec $O_{\mathscr{E}}$ à la place de O) s'écrit, de manière équivalente, sous la forme

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{b}. \tag{6.56}$$

On dit que (6.56) est l'*équation matricielle de l'application affine* Φ dans les repères $(O_{\mathscr{E}}; \mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ et $(O_{\mathscr{F}}; \mathbf{f}_1, \mathbf{f}_2, ..., \mathbf{f}_m)$ (dans le repère $(O_{\mathscr{E}}; \mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ si $\mathscr{F} = \mathscr{E}$).

On notera que si $\Phi(O_{\mathscr{E}}) = O_{\mathscr{F}}$ (en particulier, si $\mathscr{F} = \mathscr{E}$ et $O_{\mathscr{E}}$ est invariant par Φ), l'équation (6.56) se simplifie et devient

$$\mathbf{y} = \mathbf{A}\mathbf{x}. \tag{6.57}$$

Grâce à (6.56), la recherche des points invariants d'une application affine de $\mathscr{E}$ dans $\mathscr{E}$ se réduit à la recherche des solutions d'une équation de la forme

$$\mathbf{x} = \mathbf{A}\mathbf{x} + \mathbf{b},$$

qui s'écrit également

$$(\mathbf{I} - \mathbf{A})\mathbf{x} = \mathbf{b}. \tag{6.58}$$

6.7.11 Changements de repère

La recherche de l'équation matricielle d'une application affine se fait souvent par l'intermédiaire d'un repère auxiliaire. Supposons, par exemple, qu'une application affine Φ de $\mathscr{E}$ dans $\mathscr{E}$ admette un point invariant P_0 et que son équation matricielle dans un repère auxiliaire $(P_0; \mathbf{e}'_1, \mathbf{e}'_2, ..., \mathbf{e}'_n)$ soit

$$\mathbf{y}' = \mathbf{A}'\mathbf{x}'.$$

L'équation matricielle de Φ dans le repère initial $(O; \mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ est alors

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{b},$$

où

$$\mathbf{A} = \mathbf{P}_{ee'}\mathbf{A}'\mathbf{P}_{ee'}^{-1} \tag{6.59}$$

et $\mathbf{b}$ est déterminé par la condition exprimant que P_0 est invariant, c'est-à-dire par l'équation

$$\mathbf{x}_0 = \mathbf{A}\mathbf{x}_0 + \mathbf{b},$$

qui entraîne

$$\mathbf{b} = (\mathbf{I} - \mathbf{A})\mathbf{x}_0, \tag{6.60}$$

$\mathbf{x}_0$ désignant le vecteur-colonne des coordonnées de P_0 dans le repère $(O; \mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$.

On remarquera que les vecteurs-colonnes $\mathbf{x}$ et $\mathbf{x}'$ (ainsi que $\mathbf{y}$ et $\mathbf{y}'$) sont liés par la relation

$$\mathbf{x} = \mathbf{P}_{ee'}\mathbf{x}' + \mathbf{x}_0. \tag{6.61}$$

6.7.12 Exemples

(1) Soit S et D le plan et la droite vectoriels introduits dans l'exemple 6.6.11. Soit en outre $\mathscr{S}$ le plan passant par $P_0(-1, 3, -2)$ et de direction S. Nous nous proposons de trouver l'équation matricielle de l'affinité relativement à $\mathscr{S}$, de direction D et de rapport -3.

La matrice de la dilatation vectorielle associée à cette affinité a été calculée dans l'exemple 6.6.11:

$$\mathbf{A} = \frac{1}{3}\begin{pmatrix} 3 & 0 & 0 \\ -24 & -1 & 8 \\ 24 & 4 & -5 \end{pmatrix}.$$

Il ne reste donc plus qu'à calculer le vecteur-colonne $\mathbf{b}$ au moyen de (6.60):

$$\mathbf{b} = (\mathbf{I} - \mathbf{A})\mathbf{x}_0 = \frac{1}{3}\begin{pmatrix} 0 & 0 & 0 \\ 24 & 4 & -8 \\ -24 & -4 & 8 \end{pmatrix}\begin{pmatrix} -1 \\ 3 \\ -2 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 0 \\ 4 \\ -4 \end{pmatrix}.$$

L'équation matricielle cherchée est ainsi

$$\mathbf{y} = \frac{1}{3}\begin{pmatrix} 3 & 0 & 0 \\ -24 & -1 & 8 \\ 24 & 4 & -5 \end{pmatrix}\mathbf{x} + \frac{1}{3}\begin{pmatrix} 0 \\ 4 \\ -4 \end{pmatrix}.$$

(2) Nous nous proposons de déterminer le plan, la direction et le rapport d'affinité de l'affinité d'équation matricielle

$$\mathbf{y} = \frac{1}{3}\begin{pmatrix} 5 & 1 & -1 \\ -4 & 1 & 2 \\ 8 & 4 & -1 \end{pmatrix}\mathbf{x} + \begin{pmatrix} 1 \\ -2 \\ 4 \end{pmatrix}. \tag{6.62}$$

La direction d'affinité est la droite vectorielle engendrée par le vecteur $\overrightarrow{O\Phi(O)}$, O étant l'origine du repère. Le vecteur-colonne des composantes de ce vecteur (c'est-à-dire des coordonnées de $\Phi(O)$) s'obtient en posant $\mathbf{x} = \mathbf{0}$ dans (6.62), ce qui donne

$$\begin{pmatrix} 1 \\ -2 \\ 4 \end{pmatrix}.$$

Le plan d'affinité est déterminé par l'équation matricielle (6.58):

$$\frac{1}{3}(3\mathbf{I} - \begin{pmatrix} 5 & 1 & -1 \\ -4 & 1 & 2 \\ 8 & 4 & -1 \end{pmatrix})\mathbf{x} = \begin{pmatrix} 1 \\ -2 \\ 4 \end{pmatrix}.$$

Cette équation est équivalente au système linéaire

$$\begin{aligned} -2x_1 - x_2 + x_3 &= 3 \\ 4x_1 + 2x_2 - 2x_3 &= -6 \\ -8x_1 - 4x_2 + 4x_3 &= 12. \end{aligned}$$

La deuxième équation et la troisième étant multiples de la première, l'équation cartésienne du plan d'affinité est

$$-2x_1 - x_2 + x_3 = 3.$$

D'après (6.55), le rapport d'affinité est le déterminant de la matrice carrée de l'équation (6.62), soit $-\frac{1}{3}$.

6.7.13 Compléments

Les compléments suivants sont énoncés sans démonstration. Ils peuvent être déduits des propositions démontrées dans les sections précédentes de ce chapitre.

(1) L'espace affine $\mathscr{E}$ étant supposé de dimension finie non nulle n, soit P_0, P_1, ..., P_n les points d'un repère de $\mathscr{E}$ (cf. 1.11.3) et Q_0, Q_1, ..., Q_n des points quelconques de $\mathscr{F}$. Il existe une application affine unique Φ de $\mathscr{E}$ dans $\mathscr{F}$ telle que $\Phi(P_0) = Q_0$, $\Phi(P_1) = Q_1$, ..., $\Phi(P_n) = Q_n$. De toute évidence, cette conclusion est également vraie dans le cas où $n = 0$, P_0 étant dans ce cas l'unique point de $\mathscr{E}$.

(2) Soit Φ une application affine de $\mathscr{E}$ dans $\mathscr{F}$, φ l'application linéaire associée à Φ, $\mathscr{S}$ un sous-espace affine de $\mathscr{E}$ de direction S et $\mathscr{T}$ un sous-espace

affine de $\mathscr{F}$ de direction T. L'image $\Phi(\mathscr{S})$ de $\mathscr{S}$ est un sous-espace affine de $\mathscr{F}$ de direction $\varphi(S)$. L'image réciproque $\Phi^{-1}(\mathscr{T})$ de $\mathscr{T}$ est soit l'ensemble vide, soit un sous-espace affine de $\mathscr{E}$ de direction $\varphi^{-1}(T)$.

En particulier, toute application affine conserve le parallélisme entre sous-espaces affines.

(3) Une application affine de $\mathscr{E}$ dans $\mathscr{F}$ est injective (surjective, bijective) si et seulement si l'application linéaire associée est injective (surjective, bijective).

(4) L'image d'un sous-espace affine de dimension k par une application affine injective Φ est un sous-espace affine de dimension k. En particulier, l'image d'une droite est une droite et celle d'un plan est un plan. L'image du sous-espace affine de représentation paramétrique

$$P = P_0 + \alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \dots + \alpha_n \mathbf{v}_n$$

est le sous-espace affine de représentation paramétrique

$$Q = Q_0 + \alpha_1 \varphi(\mathbf{v}_1) + \alpha_2 \varphi(\mathbf{v}_2) + \dots + \alpha_n \varphi(\mathbf{v}_n),$$

où $Q = \Phi(P)$, $Q_0 = \Phi(P_0)$ et φ est l'application linéaire associée à Φ.

(5) Soit Φ une application affine de $\mathscr{E}$ dans $\mathscr{F}$, Ψ une application affine de $\mathscr{F}$ dans $\mathscr{G}$, φ et ψ les applications linéaires associées respectivement à Φ et à Ψ. La composée $\Psi \circ \Phi$ et, lorsque Φ est bijective, l'inverse Φ^{-1} sont des applications affines admettant respectivement $\psi \circ \varphi$ et φ^{-1} comme applications linéaires associées.

6.7.14 Image d'un parallélépipède

Supposons que $\mathscr{E}$ soit de dimension finie non nulle n. Soit Φ une application affine de $\mathscr{E}$ dans $\mathscr{E}$. L'image par Φ du parallélépipède construit sur les points P_0, P_1, ..., P_n est le parallélépipède construit sur les points $\Phi(P_0)$, $\Phi(P_1)$, ..., $\Phi(P_n)$.

Lorsque $\mathscr{E}$ est euclidien et muni d'un repère orthonormal, le volume du parallélépipède construit sur P_0, P_1, ..., P_n est déterminé par la formule (cf. (5.12))

$$\operatorname{vol}(P_0, P_1, \dots, P_n) = |\det(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n)|,$$

où $\mathbf{a}_j$ désigne le vecteur-colonne des composantes du vecteur $\overrightarrow{P_0P_j}$ dans la base du repère. Le volume du parallélépipède construit sur $\Phi(P_0)$, $\Phi(P_1)$, ..., $\Phi(P_n)$ est donc déterminé par la formule

$$\operatorname{vol}(\Phi(P_0), \Phi(P_1), \dots, \Phi(P_n)) = |\det(\mathbf{A}\mathbf{a}_1, \mathbf{A}\mathbf{a}_2, \dots, \mathbf{A}\mathbf{a}_n)|, \tag{6.63}$$

car le vecteur-colonne des composantes du vecteur $\overrightarrow{\Phi(P_0)\Phi(P_j)} = \varphi(\overrightarrow{P_0P_j})$ est $\mathbf{A}\mathbf{a}_j$, où $\mathbf{A}$ est la matrice de l'application linéaire φ associée à Φ. Or, en vertu de (4.8), de la propriété (a) de la proposition 5.1.8 et de 6.6.13,

$$|\det(\mathbf{A}\mathbf{a}_1, \mathbf{A}\mathbf{a}_2, \dots, \mathbf{A}\mathbf{a}_n)| = |\det\varphi|\,|\det(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n)|. \tag{6.64}$$

Par conséquent,

$$\mathrm{vol}(\Phi(P_0), \Phi(P_1), ..., \Phi(P_n)) = |\det\varphi|\,\mathrm{vol}(P_0, P_1, ..., P_n). \tag{6.65}$$

En résumant, on peut donc dire que toute application affine de $\mathscr{E}$ dans $\mathscr{E}$ multiplie les volumes par la valeur absolue du déterminant de l'application linéaire associée.

Par exemple, une homothétie et une dilatation de rapport λ multiplient les volumes respectivement par $|\lambda|^n$ et $|\lambda|^{n-k}$, k désignant la dimension du sous-espace invariant de la dilatation.

6.8 EXERCICES

6.8.1 Dans chacun des cas suivants, dire si l'application φ de l'espace vectoriel E dans l'espace vectoriel F est linéaire.

(a) $\varphi\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix} = x_1\mathbf{y}_1 + x_2\mathbf{y}_2 + x_3\mathbf{y}_3$ ($E = \mathbb{R}^3$, F quelconque).

(b) $\varphi\begin{pmatrix}x_1\\x_2\end{pmatrix} = x_1x_2$ ($E = \mathbb{R}^2$, $F = \mathbb{R}$).

(c) $\varphi(\mathbf{x}) = \|\mathbf{x}\|$ (E euclidien, $F = \mathbb{R}$).

(d) $\varphi(\mathbf{x}) = \sum_{i=1}^{k} (\mathbf{x} \mid \mathbf{a}_i)\mathbf{b}_i$ (E euclidien, F quelconque).

(e) $\varphi(\mathbf{x}) = \det(\mathbf{x}, \mathbf{a}_2, ..., \mathbf{a}_n)$ ($E = \mathbb{R}^n$, $F = \mathbb{R}$).

(f) $\varphi(\mathbf{f}) = \int_a^b \mathbf{f}(t)\mathrm{d}t$ ($E = C_{[a,b]}$, $F = \mathbb{R}$).

(g) $\varphi(\mathbf{f}) = \alpha_0 + \alpha_1\mathbf{f} + ... + \alpha_k\mathbf{f}^{(k)}$ ($E = C^k_{(a,b)}$, $F = C_{(a,b)}$).

(h) $\varphi\begin{pmatrix}x_1\\x_2\end{pmatrix} = \begin{pmatrix}2x_1+1\\x_2\end{pmatrix}$ ($E = F = \mathbb{R}^2$).

6.8.2 Soit φ une application linéaire d'un espace vectoriel E dans lui-même telle que $\mathrm{Ker}\varphi = \mathrm{Im}\varphi$. Démontrer:
(a) $\varphi^2 = 0$ (0 désigne l'application nulle).
(b) Si E est de dimension finie, $\dim E$ est un nombre pair.

Fournir en outre un exemple d'une application linéaire φ telle que $\mathrm{Ker}\varphi = \mathrm{Im}\varphi$, ainsi qu'un exemple d'une application linéaire ψ telle que $\psi^2 = 0$, mais $\mathrm{Ker}\psi \neq \mathrm{Im}\psi$. (Définir les images des vecteurs d'une base de manière appropriée.)

6.8.3 Soit φ une application linéaire d'un espace vectoriel E dans lui-même telle que $\varphi^2 = \varphi$. Soit $\psi = \mathrm{id}_E - \varphi$. Démontrer les assertions suivantes:

(a) $\varphi \circ \psi = \psi \circ \varphi = 0$ (0 désigne l'application nulle).

(b) $\psi^2 = \psi$.

(c) $S = \mathrm{Im}\,\varphi$ et $T = \mathrm{Im}\,\psi$ sont des sous-espaces complémentaires dans E.

(d) Si φ n'est ni l'application nulle ni l'application identique, S et T sont différents de $\{\mathbf{0}\}$, φ est la projection sur S parallèlement à T et ψ la projection sur T parallèlement à S.

6.8.4 Soit φ une application linéaire d'un espace vectoriel E dans lui-même telle que $\varphi^2 = \mathrm{id}_E$. Soit $\psi = \frac{1}{2}(\mathrm{id}_E + \varphi)$.

(a) Montrer que $\psi^2 = \psi$.

(b) Dans l'hypothèse où φ n'est ni l'application identique ni la symétrie centrale, déduire de l'exercice précédent que ψ est une projection et en conclure que φ est une symétrie.

6.8.5 Soit φ une application linéaire d'un espace vectoriel de dimension finie E dans lui-même. Montrer que les assertions suivantes sont équivalentes:

(a) $\mathrm{Ker}\,\varphi$ et $\mathrm{Im}\,\varphi$ sont des sous-espaces complémentaires dans E.

(b) $\mathrm{Ker}\,\varphi \cap \mathrm{Im}\,\varphi = \{\mathbf{0}\}$.

(c) $\mathrm{Ker}(\varphi^2) = \mathrm{Ker}\,\varphi$.

(d) $\mathrm{Im}(\varphi^2) = \mathrm{Im}\,\varphi$.

(e) $\mathrm{rg}(\varphi^2) = \mathrm{rg}\,\varphi$.

6.8.6 Soit E, F, G, φ et ψ comme dans (b) de la proposition 6.4.2.

(a) Montrer que $\mathrm{Ker}(\psi \circ \varphi) \supset \mathrm{Ker}\,\varphi$ et, par un exemple simple, que $\mathrm{Ker}(\psi \circ \varphi)$ peut se réduire à $\{\mathbf{0}\}$, sans que $\mathrm{Ker}\,\psi$ soit égal à $\{\mathbf{0}\}$ (autrement dit, que $\psi \circ \varphi$ peut être injective, sans que ψ le soit).

(b) Montrer que $\mathrm{Im}(\psi \circ \varphi) = \psi(\mathrm{Im}\,\varphi) \subset \mathrm{Im}\,\psi$.

(c) En déduire que si $\mathrm{Im}\,\varphi$ et $\mathrm{Im}\,\psi$ sont de dimension finie,

$$\mathrm{rg}(\psi \circ \varphi) \leqslant \min(\mathrm{rg}\,\varphi, \mathrm{rg}\,\psi),$$

avec égalité si ψ est injective ou φ surjective.

6.8.7 Soit $\mathbf{A}$ une matrice de type $m \times n$ et $\mathbf{B}$ une matrice de type $n \times p$.

(a) A l'aide des applications linéaires associées canoniquement à $\mathbf{A}$ et à $\mathbf{B}$, déduire de l'exercice précédent que

$$\mathrm{rg}(\mathbf{AB}) \leqslant \min(\mathrm{rg}\,\mathbf{A}, \mathrm{rg}\,\mathbf{B}),$$

avec égalité si $\mathrm{rg}\,\mathbf{A} = n$ ou $\mathrm{rg}\,\mathbf{B} = n$ (en particulier, si $\mathbf{A}$ ou $\mathbf{B}$ sont carrées et inversibles).

(b) Au moyen d'un exemple, montrer qu'en général les deux membres ne sont pas égaux.

6.8.8 Soit E un espace vectoriel de dimension 3, muni d'une base. Dans chacun des cas suivants, dire quelle est la nature de l'application linéaire de E dans E dont la matrice est:

$$(a)\begin{bmatrix}1&0&0\\0&1&0\\0&0&-1\end{bmatrix}, (b)\begin{bmatrix}1&0&0\\0&-1&0\\0&0&-1\end{bmatrix}, (c)\begin{bmatrix}1&0&0\\0&0&0\\0&0&1\end{bmatrix}, (d)\begin{bmatrix}0&1&0\\1&0&0\\0&0&1\end{bmatrix}, (e)\begin{bmatrix}0&0&2\\0&2&0\\2&0&0\end{bmatrix}.$$

6.8.9 Soit E un espace vectoriel de dimension 3, muni d'une base. Soit en outre S le plan vectoriel d'équation $x_1 - 3x_2 + 2x_3 = 0$ et D la droite vectorielle engendrée par $\mathbf{v}(2, 1, 1)$.
(a) Trouver la matrice de la projection sur S parallèlement à D.
(b) Trouver la matrice de la symétrie par rapport à S parallèlement à D.
(c) En déduire, sans calculs, la matrice de la symétrie par rapport à D parallèlement à S.

6.8.10 Soit E un espace vectoriel de dimension 4, muni d'une base $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4)$. Soit S et T les sous-espaces complémentaires dans E engendrés respectivement par les couples $(\mathbf{v}_1(1, 0, 1, 0), \mathbf{v}_2(0, 1, 0, -1))$ et $(\mathbf{v}_3(1, 1, -1, 0), \mathbf{v}_4(0, 1, 1, 1))$. Trouver la matrice dans la base $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4)$ de la dilatation relativement à S, de direction T et de rapport 3.

6.8.11 Soit E et F deux espaces vectoriels munis de bases respectives $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ et $(\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3, \mathbf{f}_4)$. Trouver la matrice, dans ces deux bases, de l'application linéaire φ de E dans F par laquelle l'image de la famille de vecteurs $(\mathbf{x}_1(-11, -4, 6), \mathbf{x}_2(2, 0, -1), \mathbf{x}_3(2, 1, -1))$ est la famille de vecteurs $(\mathbf{y}_1(1, 1, 0, 0), \mathbf{y}_2(0, 1, 1, 0), \mathbf{y}_3(0, 0, 1, 1))$. (Utiliser (6.35) avec $\mathbf{f}' = \mathbf{f}$.)

6.8.12 Soit E un espace vectoriel de dimension 4, muni d'une base. Soit S l'hyperplan vectoriel d'équation $x_1 - 2x_2 + x_3 - x_4 = 0$. Trouver la forme linéaire f définie dans E telle que $\operatorname{Ker} f = S$ et $f(\mathbf{v}) = 3$, où $\mathbf{v}$ est le vecteur de composantes $1, 1, 1, -1$. (On remarquera que pour tout nombre non nul α, $\mathbf{x} \to \alpha(x_1 - 2x_2 + x_3 - x_4)$ est une forme linéaire dont le noyau est S.)

6.8.13 Soit E un espace vectoriel de dimension finie $n > 1$, S un hyperplan vectoriel de E, f une forme linéaire dans E telle que $\operatorname{Ker} f = S$ et $\mathbf{v}$ un vecteur non nul de S. Soit φ l'application linéaire de E dans E définie par $\varphi(\mathbf{x}) = \mathbf{x} + f(\mathbf{x})\mathbf{v}$. On appelle φ *transvection relativement à* S.
(a) Montrer que S est constitué des vecteurs invariants par φ.
(b) Montrer que φ est bijective.
(c) Soit $\mathbf{e}_n$ un vecteur n'appartenant pas à S et $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_{n-1})$ une base de S telle que $\mathbf{e}_{n-1} = \mathbf{v}$. Trouver la matrice de φ dans la base $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$.
(d) Dans les cas où $n = 2$ et $n = 3$, tracer l'image d'un vecteur $\mathbf{x}$ n'appartenant pas à S, ainsi que celle des vecteurs $2\mathbf{x}$ et $3\mathbf{x}$.

6.8.14 Soit E et S comme dans l'exercice précédent. Montrer qu'une application linéaire de E dans E laissant invariant tout vecteur de S est une dilatation ou une transvection.

6.8.15 Soit φ une application linéaire d'un espace vectoriel E dans un espace vectoriel F. On appelle *application duale* de φ, et on note φ^*, l'application linéaire de F^* dans E^* qui associe à tout vecteur $\mathbf{y}^*$ de F^* le vecteur $\varphi^*(\mathbf{y}^*)$ de E^* défini par

$$(\varphi^*(\mathbf{y}))(\mathbf{x}) = \mathbf{y}^*(\varphi(\mathbf{x})), \quad \mathbf{x} \in E.$$

On suppose que E et F soient de dimensions finies non nulles et munis de bases respectives $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ et $(\mathbf{f}_1, \mathbf{f}_2, ..., \mathbf{f}_m)$. Soit $\mathbf{A}$ la matrice de φ dans ces bases. Montrer que la matrice de φ^* dans les bases duales $(\mathbf{f}_1^*, \mathbf{f}_2^*, ..., \mathbf{f}_m^*)$ et $(\mathbf{e}_1^*, \mathbf{e}_2^*, ..., \mathbf{e}_n^*)$ est ${}^t\mathbf{A}$.

6.8.16 Soit φ et ψ des applications linéaires d'un espace vectoriel dans lui-même. On suppose que φ et ψ *commutent*, c'est-à-dire que $\varphi \circ \psi = \psi \circ \varphi$, et on pose $S = \operatorname{Ker} \psi$, $T = \operatorname{Im} \psi$. Montrer que $\varphi(S) \subset S$ et $\varphi(T) \subset T$.

6.8.17 Soit E somme directe de ses sous-espaces vectoriels $S_1, S_2, ..., S_l$. On suppose que chacun de ces sous-espaces admette une base et on réunit ces bases en une base $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ de E. Soit φ une application linéaire de E dans E qui commute avec φ_i pour $i = 1, 2, ..., l$, où φ_i désigne la projection sur S_i parallèlement à $S_1 \oplus ... \oplus \hat{S}_i \oplus ... \oplus S_l$. A l'aide de l'exercice précédent, montrer que la matrice de φ dans la base $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ est *diagonale par blocs*, c'est-à-dire de la forme (4.6), avec $q = r$ et $\mathbf{A}_{ij} = \mathbf{O}$ si $i \neq j$.

6.8.18 Montrer que la seule matrice semblable à $a\mathbf{I}$ est $a\mathbf{I}$ elle-même.

6.8.19 Montrer que la relation «semblable à» jouit des propriétés suivantes (qui en font une relation d'équivalence);

(a) $\mathbf{A}$ est semblable à $\mathbf{A}$.

(b) Si $\mathbf{A}$ est semblable à $\mathbf{B}$, $\mathbf{B}$ est semblable à $\mathbf{A}$.

(c) Si $\mathbf{A}$ est semblable à $\mathbf{B}$ et $\mathbf{B}$ semblable à $\mathbf{C}$, $\mathbf{A}$ est semblable à $\mathbf{C}$.

Exercices sur les applications affines

6.8.20 Trouver l'équation matricielle de l'affinité plane Φ définie par les images $\Phi(\mathscr{D})$ et $\Phi(\mathscr{D}')$ de deux droites $\mathscr{D}$ et $\mathscr{D}'$.

$$\mathscr{D}: \; x_1 = 1, \qquad \Phi(\mathscr{D}): \; x_2 = 2,$$
$$\mathscr{D}': \; 3x_1 + 2x_2 = 5, \qquad \Phi(\mathscr{D}'): \; x_1 = -1.$$

Déterminer également l'axe (hyperplan), la direction et le rapport d'affinité.

6.8.21 Soit $P_0(2, 1, 0)$, $P_1(3, 2, 0)$ et $P_2(3, 3, -1)$ trois points d'un espace affine de dimension 3, muni d'un repère. Soit en outre D la droite vectorielle engendrée par le vecteur $\mathbf{v}(0, 2, -1)$.

(a) Trouver l'équation matricielle de la symétrie Φ parallèlement à D telle que $\Phi(P_0) = P_0$, $\Phi(P_1) = P_1$ et $\Phi(P_2) = P_2$.

(b) Trouver l'équation de l'image par Φ du plan d'équation $x_1 + x_2 = 1$.

6.8.22 Dans un espace affine de dimension 4, muni d'un repère, on considère la projection sur l'hyperplan d'équation $x_1 + x_2 - 2x_3 + x_4 = 2$, parallèlement à la droite vectorielle engendrée par le vecteur $\mathbf{v}(1, -1, 0, 1)$. Trouver l'équation matricielle de cette projection.

6.8.23 On dit qu'une application affine d'un espace affine $\mathscr{E}$ de dimension finie $n > 1$ dans lui-même est une *transvection affine* si elle admet au moins un point invariant et si l'application linéaire associée est une transvection de l'espace directeur de E. Etablir les propriétés des transvections affines.

6.8.24 Les espaces affines $\mathscr{E}$, $\mathscr{F}$ et $\mathscr{G}$ étant munis chacun d'un repère, une application affine Φ de $\mathscr{E}$ dans $\mathscr{F}$ et une application affine Ψ de $\mathscr{F}$ dans $\mathscr{G}$ sont supposées données par leurs équations matricielles respectives. Trouver l'équation matricielle de la composée $\Psi \circ \Phi$ et, dans l'hypothèse où Φ est bijective, celle de l'inverse Φ^{-1}.

6.8.25 Démontrer l'assertion (5) de 6.7.13.

6.8.26 Montrer que toute projection (symétrie) affine Φ vérifie la relation $\Phi^2 = \Phi$ ($\Phi^2 = \mathrm{id}_{\mathscr{E}}$). Réciproquement, montrer que toute application affine d'un espace affine $\mathscr{E}$ dans lui-même telle que $\Phi^2 = \Phi$ ($\Phi^2 = \mathrm{id}_{\mathscr{E}}$) est soit $\mathrm{id}_{\mathscr{E}}$, soit une application constante, soit une projection affine (soit $\mathrm{id}_{\mathscr{E}}$, soit une symétrie centrale, soit une symétrie affine). (Utiliser les exercices 6.8.25, 6.8.3 et 6.8.4.)

6.8.27 Soit Φ l'application affine d'un espace affine euclidien dans lui-même dont l'équation matricielle dans un repère orthonormal donné est

$$\mathbf{y} = \frac{1}{2}\begin{pmatrix} 0 & -1 & 2 & 1 \\ 2 & 1 & 0 & 2 \\ -4 & 2 & 3 & -2 \\ 0 & 1 & 4 & 0 \end{pmatrix}\mathbf{x} + \mathbf{b}.$$

Calculer le volume de l'image par Φ du parallélépipède construit sur les points $P_0(1, 0, -1, 1)$, $P_1(1, 0, 1, 0)$, $P_2(0, 1, -1, 1)$, $P_3(1, 1, 1, 1)$ et $P_4(0, 0, -2, 1)$.

6.8.28 Soit Φ une application affine d'un espace affine $\mathscr{E}$ dans un espace affine $\mathscr{F}$. Montrer que Φ *conserve les barycentres*, c'est-à-dire que $\Phi(G)$ est le barycentre des points $\Phi(P_1)$, $\Phi(P_2)$, ..., $\Phi(P_k)$ affectés des coefficients respectifs α_1, α_2, ..., α_k si G est le barycentre des points P_1, P_2, ..., P_k affectés des mêmes coefficients.

6.8.29 Trouver l'image du sous-espace affine $\mathscr{S}$ par l'application affine d'équation matricielle (6.56) dans les cas suivants:

(a) $\mathscr{S}$ est la droite d'équations $\dfrac{x_1 - 1}{2} = x_2 - 2 = \dfrac{x_3 + 3}{-1} = x_4$,

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 3 & -3 \\ -2 & 1 & 0 & 4 \\ 2 & 0 & 1 & 4 \\ 1 & 1 & 2 & 2 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix}.$$

(b) $\mathscr{S}$ est l'hyperplan d'équation $x_1 - x_2 + 5x_3 - 2x_4 = 2$ et $\mathbf{A}$, $\mathbf{b}$ sont comme dans (a).

(c) $\mathscr{S}$ est le plan d'équation $-x_1 + 5x_2 + 7x_3 = 1$,

$$\mathbf{A} = \begin{bmatrix} 2 & 3 & 1 \\ -1 & 2 & 4 \\ 1 & -1 & -3 \end{bmatrix}, \quad \mathbf{b} = \mathbf{0}.$$

(d) $\mathscr{S}$ est comme dans (c),

$$\mathbf{A} = \begin{bmatrix} -8 & 16 & 8 \\ -3 & 6 & 3 \\ 1 & 2 & -1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}.$$

CHAPITRE 7

Transformations et matrices orthogonales, isométries, similitudes

7.1 TRANSFORMATIONS ET MATRICES ORTHOGONALES

7.1.1 Introduction

Les propriétés métriques établies dans le chapitre 2 découlent de la présence d'un produit scalaire. Dans ce chapitre, nous nous proposons de caractériser et de classifier les applications qui conservent ce produit. Nous appliquerons ensuite les conclusions obtenues à l'étude des isométries et des similitudes.

Tout au long du chapitre, E désignera un espace vectoriel euclidien.

7.1.2 Conservation du produit scalaire et de la norme

Soit φ une application de E dans E. On dit que φ *conserve le produit scalaire* si

$$(\varphi(\mathbf{x}) \mid \varphi(\mathbf{y})) = (\mathbf{x} \mid \mathbf{y}) \tag{7.1}$$

pour tout couple $(\mathbf{x}, \mathbf{y})$ de vecteurs de E. On dit que φ *conserve la norme* si

$$\| \varphi(\mathbf{x}) \| = \| \mathbf{x} \| \tag{7.2}$$

pour tout vecteur $\mathbf{x}$ de E. On dit que φ *conserve l'orthogonalité* si l'image par φ de tout couple de vecteurs orthogonaux de E est un couple de vecteurs orthogonaux.

Une application qui conserve le produit scalaire conserve évidemment la norme et l'orthogonalité. En particulier, l'image d'une famille orthonormale par une telle application est encore une famille orthonormale. Par contre, une application qui conserve la norme et l'orthogonalité ne conserve pas le produit scalaire, en général. Par exemple, si S est un sous-espace vectoriel de E différent de $\{\mathbf{0}\}$ et de E, l'application φ de E dans E définie par

$$\varphi(\mathbf{x}) = \begin{cases} \mathbf{x} & \text{si } \mathbf{x} \text{ appartient à } S, \\ -\mathbf{x} & \text{si } \mathbf{x} \text{ n'appartient pas à } S, \end{cases}$$

conserve la norme et l'orthogonalité, mais ne conserve pas le produit scalaire. On remarquera, cependant, que φ n'est pas linéaire.

On dit qu'une application φ de E dans E *conserve les angles* si l'angle de $\varphi(\mathbf{x})$ et $\varphi(\mathbf{y})$ est défini et égal à l'angle de $\mathbf{x}$ et $\mathbf{y}$ chaque fois que ce dernier est défini. Vu la définition de la notion d'angle, il est évident que la conservation du produit scalaire équivaut à la conservation de la norme et des angles.

7.1.3 Proposition

Soit φ une application de E dans E. Les conditions suivantes sont équivalentes:

(a) *φ est linéaire et conserve la norme.*

(b) $\varphi(\mathbf{0}) = \mathbf{0}$ et $\| \varphi(\mathbf{x}) - \varphi(\mathbf{y}) \| = \| \mathbf{x} - \mathbf{y} \|$ *pour tout couple* $(\mathbf{x}, \mathbf{y})$ *de vecteurs de E.*

(c) *φ conserve le produit scalaire.*

DÉMONSTRATION

(a) entraîne (b). Il suffit d'observer que $\varphi(\mathbf{x}) - \varphi(\mathbf{y}) = \varphi(\mathbf{x} - \mathbf{y})$, en vertu de la linéarité de φ.

(b) entraîne (c). En posant $\mathbf{y} = \mathbf{0}$ dans (b), nous voyons d'abord que φ conserve la norme. D'autre part, d'après (2.14), pour tout couple $(\mathbf{x}, \mathbf{y})$ de vecteurs de E,

$$\| \mathbf{x} - \mathbf{y} \|^2 = \| \mathbf{x} \|^2 + \| \mathbf{y} \|^2 - 2(\mathbf{x} \mid \mathbf{y}),$$
$$\|\varphi(\mathbf{x}) - \varphi(\mathbf{y}) \|^2 = \| \varphi(\mathbf{x}) \|^2 + \| \varphi(\mathbf{y}) \|^2 - 2(\varphi(\mathbf{x}) \mid \varphi(\mathbf{y})).$$

Les deux premiers membres étant égaux par hypothèse, les deux seconds le sont aussi. Comme φ conserve la norme, nous en concluons que les deux produits scalaires sont égaux et donc que φ conserve le produit scalaire.

(c) entraîne (a). La conservation du produit scalaire entraînant celle de la norme, il nous reste à démontrer que φ est linéaire. D'après (2.14), pour tout couple $(\mathbf{x}, \mathbf{y})$ de vecteurs de E et tout couple (α, β) de nombres,

$$\begin{aligned} &\| \varphi(\alpha\mathbf{x} + \beta\mathbf{y}) - \alpha\varphi(\mathbf{x}) - \beta\varphi(\mathbf{y}) \|^2 \\ &= \| \varphi(\alpha\mathbf{x} + \beta\mathbf{y}) \|^2 + \alpha^2\| \varphi(\mathbf{x}) \|^2 + \beta^2\| \varphi(\mathbf{y}) \|^2 - 2\alpha(\varphi(\alpha\mathbf{x} + \beta\mathbf{y}) \mid \varphi(\mathbf{x})) - \\ &\quad 2\beta(\varphi(\alpha\mathbf{x} + \beta\mathbf{y}) \mid \varphi(\mathbf{y})) + 2\alpha\beta(\varphi(\mathbf{x}) \mid \varphi(\mathbf{y})). \end{aligned}$$

Comme φ conserve le produit scalaire et la norme, le second membre peut s'écrire sous la forme

$$\begin{aligned} &\| \alpha\mathbf{x} + \beta\mathbf{y} \|^2 + \alpha^2\| \mathbf{x} \|^2 + \beta^2\| \mathbf{y} \|^2 - \\ &\quad 2\alpha(\alpha\mathbf{x} + \beta\mathbf{y} \mid \mathbf{x}) - 2\beta(\alpha\mathbf{x} + \beta\mathbf{y} \mid \mathbf{y}) + 2\alpha\beta(\mathbf{x} \mid \mathbf{y}) \\ &= \| \alpha\mathbf{x} + \beta\mathbf{y} \|^2 + \alpha^2\| \mathbf{x} \|^2 + \beta^2\| \mathbf{y} \|^2 + 2\alpha\beta(\mathbf{x} \mid \mathbf{y}) - 2\| \alpha\mathbf{x} + \beta\mathbf{y} \|^2 \\ &= 2\| \alpha\mathbf{x} + \beta\mathbf{y} \|^2 - 2\| \alpha\mathbf{x} + \beta\mathbf{y} \|^2 = 0. \end{aligned}$$

Il s'ensuit que le premier membre est également nul, ce qui montre la linéarité de φ.

7.1.4 Transformations orthogonales

On appelle *transformation orthogonale* de E toute application φ de E dans E satisfaisant à l'une des conditions équivalentes (a), (b) ou (c) de la proposition 7.1.3.

Toute transformation orthogonale est injective. Cela résulte de la proposition 6.3.4, car le noyau d'une telle application se réduit à $\{\mathbf{0}\}$, du fait que $\varphi(\mathbf{x}) = \mathbf{0}$ implique $\| \mathbf{x} \| = \| \varphi(\mathbf{x}) \| = 0$ et donc $\mathbf{x} = \mathbf{0}$.

Par le corollaire 6.3.6, il en découle que toute transformation orthogonale d'un espace vectoriel euclidien de dimension finie est bijective. Cette conclusion tombe en défaut lorsque l'espace n'est pas de dimension finie (cf. exercice 7.4.1).

L'inverse φ^{-1} d'une transformation orthogonale bijective φ est encore une transformation orthogonale, car φ^{-1} est linéaire et conserve la norme, puisque

$$\| \varphi^{-1}(\mathbf{x}) \| = \| \varphi(\varphi^{-1}(\mathbf{x})) \| = \| \mathbf{x} \|.$$

De même, la composée de deux transformations orthogonales (non nécessairement distinctes) est encore une transformation orthogonale.

7.1.5 Exemples

(1) L'application id_E et la symétrie centrale $-\mathrm{id}_E$ sont des transformations orthogonales. Lorsque $\dim E = 1$, ce sont les seules transformations orthogonales de E.

(2) Soit E somme directe de S et $S^\perp$, ces deux sous-espaces vectoriels étant supposés différents de $\{\mathbf{0}\}$. La symétrie orthogonale φ par rapport à S est une transformation orthogonale. En effet, φ est linéaire et en outre, d'après (6.10) (avec $\lambda = -1$) et l'orthogonalité de $\mathbf{s}$ et $\mathbf{t}$ dans la décomposition (6.9),

$$\begin{aligned}\| \varphi(\mathbf{x}) \|^2 &= \| \varphi(\mathbf{s} + \mathbf{t}) \|^2 = \| \mathbf{s} - \mathbf{t} \|^2 = \| \mathbf{s} \|^2 + \| \mathbf{t} \|^2 \\ &= \| \mathbf{s} + \mathbf{t} \|^2 = \| \mathbf{x} \|^2,\end{aligned}$$

ce qui montre que φ conserve la norme.

7.1.6 Matrice d'une transformation orthogonale

Supposons que E soit de dimension finie non nulle n. Soit φ une transformation orthogonale de E. Désignons par $\mathbf{A}$ la matrice de φ dans une base orthonormale $(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ de E arbitrairement choisie. Comme $(\varphi(\mathbf{x}))_\mathrm{e} = \mathbf{A}\mathbf{x}_\mathrm{e}$ et $(\varphi(\mathbf{y}))_\mathrm{e} = \mathbf{A}\mathbf{y}_\mathrm{e}$, la condition (7.1) s'exprime, de manière équivalente, sous l'une des deux formes

$${}^t(\mathbf{A}\mathbf{x}_\mathrm{e})\mathbf{A}\mathbf{y}_\mathrm{e} = {}^t\mathbf{x}_\mathrm{e}\mathbf{y}_\mathrm{e}$$

ou

$${}^t\mathbf{x}_\mathrm{e}({}^t\mathbf{A}\mathbf{A} - \mathbf{I})\mathbf{y}_\mathrm{e} = 0. \tag{7.3}$$

En prenant pour $\mathbf{x}_e$ et $\mathbf{y}_e$ respectivement le i-ième et le j-ième vecteur de la base canonique de $\mathbb{R}^n$ (c'est-à-dire les vecteurs-colonnes des composantes de $\mathbf{e}_i$ et de $\mathbf{e}_j$), nous voyons que le premier membre de (7.3) n'est rien d'autre que le terme d'indices i, j de la matrice ${}^t\mathbf{A}\mathbf{A} - \mathbf{I}$. Les indices i et j étant quelconques, cette matrice est donc nulle, autrement dit,

$$
{}^t\mathbf{A}\mathbf{A} = \mathbf{I}. \tag{7.4}
$$

Inversement, si $\mathbf{A}$ est une matrice carrée d'ordre n vérifiant (7.4) et $(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ une base orthonormale de E, l'application linéaire de E dans E associée à $\mathbf{A}$ conserve le produit scalaire et est donc une transformation orthogonale.

Nous sommes ainsi amenés à poser la définition qui va suivre.

7.1.7 Matrices orthogonales

On dit qu'une matrice carrée $\mathbf{A}$ d'ordre n est *orthogonale* si elle vérifie la relation (7.4).

On notera que cette relation est l'exacte expression du fait que le produit scalaire (dans $\mathbb{R}^n$) de la i-ième ligne de ${}^t\mathbf{A}$, c'est-à-dire la i-ième colonne de $\mathbf{A}$, par la j-ième colonne de $\mathbf{A}$ est 1 ou 0, suivant que $i = j$ ou $i \neq j$. En d'autres termes, une matrice carrée est orthogonale si et seulement si ses colonnes forment une famille orthonormale.

Par exemple, les matrices suivantes sont orthogonales, car leurs colonnes sont orthonormales:

$$
\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \frac{1}{\sqrt{13}}\begin{pmatrix} 2 & -3 \\ 3 & 2 \end{pmatrix}, \qquad \begin{pmatrix} \dfrac{1}{\sqrt{11}} & \dfrac{1}{\sqrt{2}} & -\dfrac{3}{\sqrt{22}} \\ -\dfrac{1}{\sqrt{11}} & \dfrac{1}{\sqrt{2}} & \dfrac{3}{\sqrt{22}} \\ \dfrac{3}{\sqrt{11}} & 0 & \dfrac{2}{\sqrt{22}} \end{pmatrix}.
$$

Par contre, la matrice

$$
\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}
$$

n'est pas orthogonale, car ses colonnes (tout en étant orthogonales) ne sont pas des vecteurs unitaires.

Les matrices orthogonales d'ordre 2 peuvent d'ailleurs être décrites très facilement. La première colonne est un vecteur unitaire de $\mathbb{R}^2$

$$
\begin{pmatrix} a \\ b \end{pmatrix}.
$$

Etant orthogonale à la première, la deuxième colonne est de la forme

$$
\alpha\begin{pmatrix} -b \\ a \end{pmatrix}.
$$

Comme ce vecteur doit être également unitaire, $\alpha^2 = 1$, autrement dit, $\alpha = \pm 1$. Les deux seules formes possibles d'une matrice orthogonale d'ordre 2 sont donc:

$$\begin{bmatrix} a & -b \\ b & a \end{bmatrix}, \quad \begin{bmatrix} a & b \\ b & -a \end{bmatrix}, \qquad \text{où } a^2 + b^2 = 1. \tag{7.5}$$

7.1.8 Caractérisation des matrices orthogonales

D'après la proposition 4.2.3, si $\mathbf{A}$ est une matrice orthogonale,

$$\mathbf{A}\,{}^t\mathbf{A} = \mathbf{I}, \tag{7.6}$$

autrement dit, $\mathbf{A}$ est inversible et

$$\mathbf{A}^{-1} = {}^t\mathbf{A}. \tag{7.7}$$

Inversement, toute matrice carrée $\mathbf{A}$ qui vérifie (7.6) ou (7.7) est orthogonale.

La proposition suivante résume les caractères distinctifs des matrices orthogonales:

7.1.9 Proposition

Soit $\mathbf{A}$ *une matrice carrée d'ordre n. Les conditions suivantes sont équivalentes*:

(a) $\mathbf{A}$ *est orthogonale.*

(b) $\mathbf{A}\,{}^t\mathbf{A} = \mathbf{I}$, *autrement dit,* ${}^t\mathbf{A}$ *est orthogonale.*

(c) $\mathbf{A}$ *est inversible et* $\mathbf{A}^{-1} = {}^t\mathbf{A}$.

(d) *Les colonnes de* $\mathbf{A}$ *forment une famille orthonormale.*

(e) *Les lignes de* $\mathbf{A}$ *forment une famille orthonormale.*

(f) *L'espace E étant supposé de dimension n et muni d'une base orthonormale arbitrairement choisie, l'application linéaire de E dans E associée à* $\mathbf{A}$ *est une transformation orthogonale.*

(g) *Même condition que* (f), *mais avec une base donnée.*

7.1.10 Déterminant d'une matrice orthogonale

Le déterminant d'une matrice orthogonale $\mathbf{A}$ vaut 1 ou -1. En effet, d'après (a) et (b) de la proposition 5.1.8,

$$|\mathbf{A}|^2 = |{}^t\mathbf{A}|\,|\mathbf{A}| = |{}^t\mathbf{A}\mathbf{A}| = |\mathbf{I}| = 1,$$

ce qui entraîne

$$|\mathbf{A}| = \pm 1.$$

Le déterminant d'une transformation orthogonale vaut donc aussi 1 ou -1.

On prendra garde à ne pas conclure qu'une matrice est orthogonale du seul fait que son déterminant est égal à 1 ou à -1. Pour se rendre compte de l'erreur qu'on ferait, il suffit de penser que toute matrice inversible peut être légèrement modifiée (par exemple en divisant les termes d'une de ses colonnes par son déterminant) de manière à obtenir une matrice de déterminant 1.

7.1.11 Matrices de passage orthogonales

La matrice de passage d'une base orthonormale à une base orthonormale est orthogonale. Cela résulte immédiatement de la définition 6.6.2, compte tenu de (2.28).

7.1.12 Indépendance du volume du choix du repère orthonormal

Supposons que E soit de dimension finie non nulle n. Soit $\mathscr{E}$ un espace affine de direction E et P_0, P_1, ..., P_n des points de $\mathscr{E}$. Désignons par $\mathbf{a}_j$ et $\mathbf{a}'_j$ ($j = 1, 2, ..., n$) les vecteurs-colonnes des composantes du vecteur $\overrightarrow{P_0P_j}$ dans les bases de deux repères orthonormaux de $\mathscr{E}$. Soit $\mathbf{P}$ la matrice de passage de la base du premier à celle du deuxième repère. Suivant le choix d'un de ces repères, le volume du parallélépipède construit sur P_0, P_1, ..., P_n est, d'après (5.12),

$$|\det(\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_n)| \quad \text{ou} \quad |\det(\mathbf{a}'_1, \mathbf{a}'_2, ..., \mathbf{a}'_n)|.$$

Or, d'après (6.31),

$$\mathbf{a}_j = \mathbf{P}\mathbf{a}'_j.$$

Par conséquent, vu la règle (4.8), la propriété (a) de la proposition 5.1.8, 7.1.11 et 7.1.10

$$\begin{aligned} |\det(\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_n)| &= |\det(\mathbf{P}\mathbf{a}'_1, \mathbf{P}\mathbf{a}'_2, ..., \mathbf{P}\mathbf{a}'_n)| \\ &= ||\mathbf{P}||\cdot|\det(\mathbf{a}'_1, \mathbf{a}'_2, ..., \mathbf{a}'_n)| = |\det(\mathbf{a}'_1, \mathbf{a}'_2, ..., \mathbf{a}'_n)|, \end{aligned}$$

ce qui démontre l'indépendance de la définition (5.12) du choix du repère orthonormal, puisque la valeur absolue de $|\mathbf{P}|$ vaut 1.

7.2 CLASSIFICATION DES TRANSFORMATIONS ORTHOGONALES À DEUX ET À TROIS DIMENSIONS

7.2.1 Angles orientés

L'angle (non orienté) de deux vecteurs non nuls $\mathbf{x}$ et $\mathbf{y}$ a été défini dans 2.4.4: c'est l'unique nombre θ de l'intervalle $[0, \pi]$ tel que

$$\cos\theta = \frac{(\mathbf{x} \mid \mathbf{y})}{\|\mathbf{x}\| \|\mathbf{y}\|}. \tag{7.8}$$

Nous allons maintenant définir la notion d'angle orienté. Supposons que E soit orienté et considérons d'abord le cas où E est de dimension 2.

On appelle *angle orienté* d'un couple $(\mathbf{x}, \mathbf{y})$ de vecteurs non nuls le nombre θ' de l'intervalle $[0, 2\pi)$ défini par

$$\theta' = \begin{cases} \theta & \text{si } (\mathbf{x}, \mathbf{y}) \text{ est une base directe ou si } \mathbf{x} \text{ est multiple de } \mathbf{y}, \\ 2\pi - \theta & \text{si } (\mathbf{x}, \mathbf{y}) \text{ est une base indirecte,} \end{cases} \tag{7.9}$$

où θ désigne l'angle (non orienté) de $\mathbf{x}$ et $\mathbf{y}$ (dans la figure 7.1 l'orientation de E est définie par $(\mathbf{e}_1, \mathbf{e}_2)$).

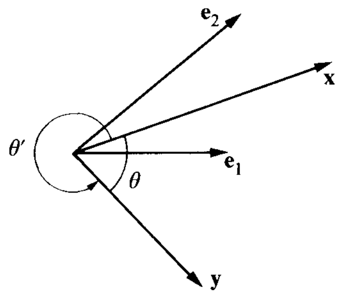


Fig. 7.1

Manifestement, l'angle orienté ainsi défini ne dépend pas des normes de $\mathbf{x}$ et de $\mathbf{y}$. En revanche, si $\mathbf{x}$ n'est pas multiple de $\mathbf{y}$, il dépend de l'ordre d'écriture des vecteurs $\mathbf{x}$ et $\mathbf{y}$, ainsi que de l'orientation de E. Plus précisément, si $\mathbf{x}$ et $\mathbf{y}$ sont échangés, ou si l'orientation de E est changée, l'angle orienté θ' change en $2\pi - \theta'$.

On notera que si θ est l'angle (non orienté) de $\mathbf{x}$ et $\mathbf{y}$ et θ' l'angle orienté de $(\mathbf{x}, \mathbf{y})$, $\cos\theta' = \cos\theta$, tandis que $\sin\theta' = \sin\theta$ ou $\sin\theta' = -\sin\theta$, suivant que $\theta' = \theta$ ou $\theta' = 2\pi - \theta$.

L'orientation de E étant définie par une base $(\mathbf{e}_1, \mathbf{e}_2)$, pour savoir si $\theta' = \theta$ ou $\theta' = 2\pi - \theta$, il suffit de calculer le déterminant $\det(\mathbf{x}_e, \mathbf{y}_e)$. Si $\mathbf{x}$ n'est pas multiple de $\mathbf{y}$, ce déterminant n'est autre que le déterminant de passage de la base $(\mathbf{e}_1, \mathbf{e}_2)$ à la base $(\mathbf{x}, \mathbf{y})$. Il est donc positif ou négatif, suivant que $\theta' = \theta$ ou $\theta' = 2\pi - \theta$. S'il est nul, $\mathbf{x}$ est multiple de $\mathbf{y}$ et dans ce cas $\theta' = \theta$.

Lorsque la base $(\mathbf{e}_1, \mathbf{e}_2)$ est orthonormale, le lien entre $\sin\theta'$ et $\det(\mathbf{x}_e, \mathbf{y}_e)$ est précisé par la relation

$$\sin\theta' = \frac{\det(\mathbf{x}_e, \mathbf{y}_e)}{\| \mathbf{x}_e \| \| \mathbf{y}_e \|}. \tag{7.10}$$

En effet,

$$\cos^2\theta' + \left(\frac{\det(\mathbf{x}_e, \mathbf{y}_e)}{\| \mathbf{x}_e \| \| \mathbf{y}_e \|}\right)^2 = \frac{(x_1y_1 + x_2y_2)^2 + (x_1y_2 - x_2y_1)^2}{(x_1^2 + x_2^2)(y_1^2 + y_2^2)} = 1,$$

ce qui entraîne que les carrés des deux membres de (7.10) sont égaux et donc que les deux membres eux-mêmes sont égaux, puisque l'un est positif si et seulement si l'autre l'est.

Par la suite, nous noterons les angles orientés et non orientés par un seul symbole, principalement par θ. En outre, pour plus de commodité, nous permettrons à tout nombre de représenter un angle orienté, en convenant que deux nombres désigneront le même angle orienté s'ils diffèrent d'un multiple entier de 2π. Par exemple, $2\pi - \theta$ et $-\theta$ désigneront le même angle orienté.

Lorsque la dimension de E est supérieure à 2, la notion d'angle orienté de deux vecteurs linéairement indépendants se définit au moyen du plan vectoriel qu'ils engendrent. Bien entendu, ce plan doit être préalablement orienté. L'angle orienté de deux vecteurs linéairement dépendants non nuls est simplement leur angle (non orienté).

Lorsque E est orienté et de dimension 3, on dit que l'orientation d'un plan vectoriel S de E est définie par un vecteur $\mathbf{e}$ n'appartenant pas à S, ou par la droite vectorielle engendrée et orientée par $\mathbf{e}$, si les bases directes de S sont celles qui se prolongent par $\mathbf{e}$ en une base directe de E.

7.2.2 Transformations orthogonales à deux dimensions

Supposons que E soit orienté, de dimension 2 et muni d'une base orthonormale directe $(\mathbf{e}_1, \mathbf{e}_2)$. Soit φ une transformation orthogonale de E. D'après la proposition 7.1.9, la matrice de φ est orthogonale, donc de l'une des deux formes décrites dans (7.5). Par ailleurs, nous savons qu'il y a correspondance biunivoque entre les angles orientés et les vecteurs unitaires de $\mathbb{R}^2$, cette correspondance étant définie par

$$\theta \leftrightarrow \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}. \tag{7.11}$$

Nous en concluons qu'il existe un angle orienté θ permettant d'écrire la matrice de φ sous l'une des deux formes

$$\mathbf{A}_1 = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \quad \text{ou} \quad \mathbf{A}_2 = \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix}. \tag{7.12}$$

Plus précisément,

$$\begin{aligned} \mathbf{A}_\varphi &= \mathbf{A}_1 \ \text{ si } \det\varphi = 1, \\ \mathbf{A}_\varphi &= \mathbf{A}_2 \ \text{ si } \det\varphi = -1. \end{aligned} \tag{7.13}$$

Si $\mathbf{A}_\varphi = \mathbf{A}_1$, on appelle φ *rotation d'angle orienté* θ (fig. 7.2). Cette dénomination est justifiée par le fait que φ conserve la norme et que l'angle orienté de tout vecteur non nul et son image est θ. Pour prouver cette dernière assertion, désignons par θ' l'angle orienté de $(\mathbf{x}, \varphi(\mathbf{x}))$, où $\mathbf{x}$ est un vecteur non nul quelconque, que nous pouvons supposer unitaire. D'après (7.8), (7.9) et (7.10),

$$\cos\theta' = (\mathbf{x} \mid \varphi(\mathbf{x})) = {}^t\mathbf{x}_e\mathbf{A}_1\mathbf{x}_e = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \cos\theta$$

et

$$\sin\theta' = \det(\mathbf{x}_e, \mathbf{A}_1\mathbf{x}_e) = \begin{vmatrix} x_1 & x_1\cos\theta - x_2\sin\theta \\ x_2 & x_1\sin\theta + x_2\cos\theta \end{vmatrix} = \sin\theta,$$

d'où il s'ensuit que $\theta' = \theta$.

Si $\mathbf{A}_\varphi = \mathbf{A}_2$, φ est une symétrie axiale orthogonale, plus exactement la symétrie orthogonale par rapport à la droite vectorielle engendrée par le vecteur de composantes $\cos\frac{\theta}{2}$, $\sin\frac{\theta}{2}$ (fig. 7.2). En effet, d'après (6.24), la matrice de cette application est

$$\mathbf{I} - 2\mathbf{n}_e{}'\mathbf{n}_e = \mathbf{I} - 2\begin{pmatrix} -\sin\frac{\theta}{2} \\ \cos\frac{\theta}{2} \end{pmatrix}\begin{pmatrix} -\sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{pmatrix} = \begin{pmatrix} 1-2\sin^2\frac{\theta}{2} & 2\sin\frac{\theta}{2}\cos\frac{\theta}{2} \\ 2\sin\frac{\theta}{2}\cos\frac{\theta}{2} & 1-2\cos^2\frac{\theta}{2} \end{pmatrix}$$

$$= \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix} = \mathbf{A}_2.$$

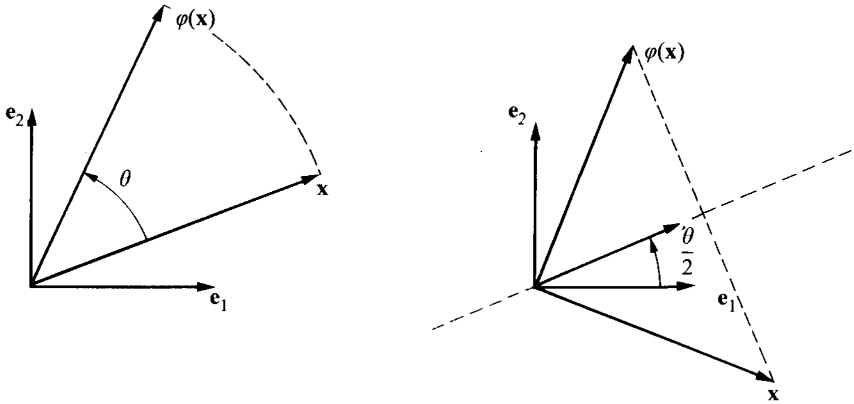


Fig. 7.2

Lorsque E est de dimension 2, nos conclusions se résument donc ainsi: *toute transformation orthogonale de déterminant 1 est une rotation et toute transformation orthogonale de déterminant -1 est une symétrie axiale orthogonale.*

On notera que les rotations conservent les angles orientés, tandis que les symétries axiales transforment ceux-ci en leurs opposés. Cela résulte de (7.10), vu que

$$\det(\mathbf{A}_\varphi\mathbf{x}_e, \mathbf{A}_\varphi\mathbf{y}_e) = \det\varphi\det(\mathbf{x}_e, \mathbf{y}_e) \quad (\text{cf. } (6.64)).$$

7.2.3 Existence de vecteurs invariants

Si E est de dimension finie et impaire n, toute transformation orthogonale φ de E admet un vecteur invariant non nul ou un vecteur non nul dont l'image est son opposé. Cela revient à dire que φ ou $-\varphi$ admettent un vecteur invariant non nul, ou encore qu'il existe un vecteur non nul $\mathbf{x}$ tel que

$$(\varphi - \mathrm{id}_E)(\mathbf{x}) = \mathbf{0} \quad \text{ou} \quad (\varphi + \mathrm{id}_E)(\mathbf{x}) = \mathbf{0}. \tag{7.14}$$

Pour établir cette conclusion, choisissons une base orthonormale quelconque de E et désignons par $\mathbf{A}$ la matrice de φ dans cette base. Les deux équations (7.14) s'expriment alors, de manière équivalente, sous la forme de deux systèmes linéaires homogènes ayant pour matrices associées respectives $\mathbf{A} - \mathbf{I}$ et $\mathbf{A} + \mathbf{I}$. Vu le corollaire 3.5.4 et (b) de la proposition 5.1.6, tout revient donc à démontrer que $|\mathbf{A} - \mathbf{I}| = 0$ ou $|\mathbf{A} + \mathbf{I}| = 0$. Par (b) de la proposition 5.1.8 et (5.5),

$$|\mathbf{A} - {}^t\mathbf{A}| = |{}^t(\mathbf{A} - {}^t\mathbf{A})| = |{}^t\mathbf{A} - \mathbf{A}| = |-(\mathbf{A} - {}^t\mathbf{A})| = (-1)^n |\mathbf{A} - {}^t\mathbf{A}|,$$

ce qui entraîne, n étant impair,

$$|\mathbf{A} - {}^t\mathbf{A}| = 0. \tag{7.15}$$

D'autre part, par (a) de la proposition 5.1.8 et grâce au fait que $\mathbf{A}$ est orthogonale,

$$|\mathbf{A} - \mathbf{I}|\,|\mathbf{A} + \mathbf{I}| = |\mathbf{A}^2 - \mathbf{I}| = |\mathbf{A}(\mathbf{A} - {}^t\mathbf{A})| = |\mathbf{A}|\,|\mathbf{A} - {}^t\mathbf{A}| = 0,$$

ce qui entraîne la conclusion.

7.2.4 Transformations orthogonales à trois dimensions

Soit E orienté et de dimension 3. Soit en outre φ une transformation orthogonale de E. D'après 7.2.3, φ ou $-\varphi$ admettent un vecteur invariant non nul.

Supposons que φ admette un tel vecteur. La droite vectorielle D engendrée par ce vecteur est alors invariante. Désignons par S le plan vectoriel orthogonal à D. Comme φ conserve l'orthogonalité, S est stable par φ (cf. 6.1.3). En outre, la restriction de φ à S est une transformation orthogonale de S. Prenons un vecteur unitaire $\mathbf{e}_3$ de D et orientons D par ce vecteur. Orientons ensuite S au moyen de D et choisissons une base orthonormale directe $(\mathbf{e}_1, \mathbf{e}_2)$ de S. En raison de 7.2.2, la matrice de φ dans la base $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ s'écrit sous l'une des deux formes

$$\mathbf{A}_1 = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{ou} \quad \mathbf{A}_2 = \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ \sin\theta & -\cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{7.16}$$

où θ est un angle orienté déterminé par φ. Plus précisément,

$$\begin{aligned} \mathbf{A}_\varphi &= \mathbf{A}_1 \ \text{ si } \det\varphi = 1, \\ \mathbf{A}_\varphi &= \mathbf{A}_2 \ \text{ si } \det\varphi = -1. \end{aligned} \tag{7.17}$$

Dans le premier cas, on appelle φ *rotation d'axe D et d'angle orienté θ.* Dans le deuxième cas, φ est une symétrie orthogonale par rapport au plan engendré par $\mathbf{e}_3$ et le vecteur de composantes $\cos\frac{\theta}{2}$, $\sin\frac{\theta}{2}$, 0.

Supposons maintenant que $-\varphi$ admette un vecteur invariant non nul. Dans une base orthonormale directe dont le troisième vecteur est un vecteur invariant par $-\varphi$, la matrice de $-\varphi$ s'écrit sous l'une des deux formes (7.16). Par conséquent, la matrice de φ dans cette base s'exprime sous l'une des deux formes

$$\mathbf{A}_3 = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad \text{ou} \quad \mathbf{A}_4 = \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ \sin\theta & -\cos\theta & 0 \\ 0 & 0 & -1 \end{bmatrix}, \tag{7.18}$$

où θ désigne l'angle orienté déterminé par $-\varphi$ majoré de π. Plus précisément,

$$\begin{aligned} \mathbf{A}_\varphi &= \mathbf{A}_3 \text{ si } \det\varphi = -1 \ (\det(-\varphi) = 1), \\ \mathbf{A}_\varphi &= \mathbf{A}_4 \text{ si } \det\varphi = 1 \ (\det(-\varphi) = -1). \end{aligned} \tag{7.19}$$

Dans le premier cas, φ est la composée d'une rotation d'angle orienté θ et d'une symétrie plane orthogonale, plus exactement la symétrie orthogonale par rapport au plan vectoriel orthogonal à l'axe de rotation. Dans le deuxième cas, φ est une rotation d'angle π dont l'axe est la droite vectorielle engendrée par le vecteur de composantes $\cos\frac{\theta}{2}$, $\sin\frac{\theta}{2}$, 0.

Lorsque E est de dimension 3, nos conclusions se résument donc ainsi: *toute transformation orthogonale de déterminant 1 est une rotation et toute transformation orthogonale de déterminant -1 est la composée d'une rotation et d'une symétrie plane orthogonale, cette composée pouvant se réduire à la seule symétrie (lorsque la rotation se réduit à l'identité).*

7.2.5 Généralisation

Signalons, sans démonstration, le résultat général suivant: E étant de dimension finie non nulle n, pour toute transformation orthogonale φ de E, il existe une base orthonormale dans laquelle la matrice de φ s'écrit sous la forme

$$\begin{bmatrix} \mathbf{I}_k & & & & & \\ & -\mathbf{I}_l & & & 0 & \\ & & \mathbf{R}_1 & & & \\ & & & \cdot & & \\ & 0 & & & \cdot & \\ & & & & & \mathbf{R}_m \end{bmatrix}, \tag{7.20}$$

où k, l, m sont des entiers positifs tels que $k + l + 2m = n$, $\mathbf{I}_k$ et $\mathbf{I}_l$ sont les matrices-unités d'ordre k et l et

$$\mathbf{R}_i = \begin{bmatrix} \cos\theta_i & -\sin\theta_i \\ \sin\theta_i & \cos\theta_i \end{bmatrix}, \quad i = 1, 2, ..., m.$$

Précisons toutefois que les trois types de sous-matrices figurant dans (7.20) ne sont pas nécessairement tous présents en même temps.

Ce résultat affirme, en substance, que l'espace vectoriel E est somme directe de droites et de plans vectoriels orthogonaux deux à deux et stables par φ.

7.2.6 Calcul de l'angle et de l'axe de rotation

Soit $\mathbf{A} = (a_{ij})$ une matrice orthogonale d'ordre 3 et de déterminant 1. L'espace E étant supposé orienté et muni d'une base orthonormale directe $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$, l'application linéaire φ de E dans E associée à $\mathbf{A}$ est une rotation d'angle orienté θ dépendant du choix de l'orientation de l'axe de rotation. Le cosinus de θ est toutefois indépendant de ce choix et peut être calculé ainsi: d'après 7.2.4, la matrice de φ dans une base orthonormale directe dont le troisième vecteur est un des deux vecteurs unitaires engendrant l'axe de rotation est la matrice $\mathbf{A}_1$ de (7.16); d'après (6.36), $\mathbf{A}$ et $\mathbf{A}_1$ sont semblables, donc elles ont la même trace (cf. 6.6.12), autrement dit,

$$\operatorname{tr}\mathbf{A} = \operatorname{tr}\mathbf{A}_1 = 2\cos\theta + 1,$$

ce qui entraîne

$$\cos\theta = \frac{1}{2}(\operatorname{tr}\mathbf{A} - 1) = \frac{1}{2}(a_{11} + a_{22} + a_{33} - 1). \tag{7.21}$$

Supposons maintenant que l'axe de rotation soit orienté par $\mathbf{v}$. L'angle de rotation est alors l'unique nombre θ de l'intervalle $[0, \pi]$ satisfaisant à (7.21), ou l'opposé $-\theta$ de ce nombre, suivant que le déterminant $\det(\mathbf{x}_e, \mathbf{A}\mathbf{x}_e, \mathbf{v}_e)$ est non négatif ou négatif, $\mathbf{x}$ étant ici un vecteur quelconque n'appartenant pas à l'axe de rotation. Dans la pratique, il convient de prendre pour $\mathbf{x}$ un des vecteurs de la base. Par exemple, si le choix de $\mathbf{e}_1$ est possible,

$$\det(\mathbf{x}_e, \mathbf{A}\mathbf{x}_e, \mathbf{v}_e) = \begin{vmatrix} 1 & a_{11} & v_1 \\ 0 & a_{21} & v_2 \\ 0 & a_{31} & v_3 \end{vmatrix} = \begin{vmatrix} a_{21} & v_2 \\ a_{31} & v_3 \end{vmatrix}. \tag{7.22}$$

Lorsque $\mathbf{A}$ n'est pas symétrique, autrement dit $\mathbf{A} \neq {}^t\mathbf{A} = \mathbf{A}^{-1}$, ce qui signifie que l'angle de rotation est différent de 0 et de π, l'axe de rotation est engendré par le vecteur non nul $\mathbf{v}$ défini par

$$\mathbf{v}_e = \begin{pmatrix} a_{32} - a_{23} \\ a_{13} - a_{31} \\ a_{21} - a_{12} \end{pmatrix}. \tag{7.23}$$

En effet, un calcul tout à fait simple montre que

$$(\mathbf{A} - {}^t\mathbf{A})\mathbf{x}_e = (\mathbf{v} \times \mathbf{x})_e$$

pour tout vecteur $\mathbf{x}$ de E; d'autre part, si $\mathbf{x}$ est un vecteur de l'axe de rotation,

$$\mathbf{x}_e = \mathbf{A}\mathbf{x}_e = \mathbf{A}^{-1}\mathbf{x}_e = {}^t\mathbf{A}\mathbf{x}_e,$$

donc

$$(\mathbf{A} - {}^t\mathbf{A})\mathbf{x}_e = \mathbf{0}$$

et, par conséquent,

$$\mathbf{v} \times \mathbf{x} = \mathbf{0},$$

ce qui prouve que $\mathbf{v}$ et $\mathbf{x}$ sont linéairement dépendants et donc que $\mathbf{v}$ appartient à l'axe de rotation.

7.2.7 Calcul de la matrice d'une rotation

Supposons que E soit de dimension 3 et muni d'une base orthonormale ($\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$) définissant son orientation. Soit D la droite vectorielle engendrée et orientée par un vecteur unitaire $\mathbf{v}$ de composantes v_1, v_2, v_3. Nous nous proposons de calculer la matrice $\mathbf{A}$ de la rotation d'axe D et d'angle orienté θ. Le calcul peut être fait au moyen d'un changement de base, comme dans l'exemple 6.6.11. La base auxiliaire ($\mathbf{e}'_1$, $\mathbf{e}'_2$, $\mathbf{e}'_3$) est définie de la manière suivante: $\mathbf{e}'_3$ est le vecteur $\mathbf{v}$; $\mathbf{e}'_1$ est un vecteur unitaire orthogonal à $\mathbf{v}$ que nous choisissons, par exemple, dans le plan vectoriel engendré par $\mathbf{e}_1$ et $\mathbf{e}_2$; $\mathbf{e}'_2$ est le vecteur $\mathbf{e}'_3 \times \mathbf{e}'_1$. Dans la base ($\mathbf{e}'_1$, $\mathbf{e}'_2$, $\mathbf{e}'_3$) la matrice de la rotation est

$$\mathbf{A}' = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

La matrice de passage est la matrice orthogonale

$$\mathbf{P}_{ee'} = \begin{pmatrix} \dfrac{v_2}{r} & \dfrac{v_1 v_3}{r} & v_1 \\ -\dfrac{v_1}{r} & \dfrac{v_2 v_3}{r} & v_2 \\ 0 & -r & v_3 \end{pmatrix},$$

où $r = (v_1^2 + v_2^2)^{1/2}$. D'après (6.36), la matrice cherchée est

$$\mathbf{A} = \mathbf{P}_{ee'}\mathbf{A}'\,{}^t\mathbf{P}_{ee'}.$$

Exemple numérique: si $v_1 = v_2 = v_3 = \dfrac{\sqrt{3}}{3}$ et $\theta = \dfrac{\pi}{6}$,

$$\mathbf{P}_{ee'} = \frac{1}{6}\begin{pmatrix} 3\sqrt{2} & \sqrt{6} & 2\sqrt{3} \\ -3\sqrt{2} & \sqrt{6} & 2\sqrt{3} \\ 0 & -2\sqrt{6} & 2\sqrt{3} \end{pmatrix} \quad \text{et} \quad \mathbf{A} = \frac{1}{3}\begin{pmatrix} 1+\sqrt{3} & 1-\sqrt{3} & 1 \\ 1 & 1+\sqrt{3} & 1-\sqrt{3} \\ 1-\sqrt{3} & 1 & 1+\sqrt{3} \end{pmatrix}.$$

La matrice $\mathbf{A}$ peut également être calculée en passant par l'écriture explicite de l'image $\varphi(\mathbf{x})$ d'un vecteur quelconque $\mathbf{x}$. On constate d'abord que

$$\begin{aligned} \| \mathbf{x} - (\mathbf{x} \mid \mathbf{v})\mathbf{v} \|^2 &= \| \mathbf{x} \|^2 - (\mathbf{x} \mid \mathbf{v})^2 = \| \mathbf{x} \|^2(1 - \cos^2\eta) \\ &= \| \mathbf{x} \|^2 \sin^2\eta = \| \mathbf{v} \times \mathbf{x} \|^2, \end{aligned}$$

où η est l'angle de $\mathbf{x}$ et $\mathbf{v}$. A l'aide de la figure 7.3, on vérifie ensuite que

$$\varphi(\mathbf{x}) = (\mathbf{x} \mid \mathbf{v})\mathbf{v} + \cos\theta(\mathbf{x} - (\mathbf{x} \mid \mathbf{v})\mathbf{v}) + \sin\theta(\mathbf{v} \times \mathbf{x}). \tag{7.24}$$

Traduite en termes matriciels, cette relation s'écrit

$$\mathbf{A}\mathbf{x}_e = ((1 - \cos\theta)\mathbf{v}_e{}^t\mathbf{v}_e + \cos\theta\mathbf{I} + \sin\theta\mathbf{V})\mathbf{x}_e,$$

où

$$\mathbf{V} = \begin{bmatrix} 0 & -v_3 & v_2 \\ v_3 & 0 & -v_1 \\ -v_2 & v_1 & 0 \end{bmatrix}. \tag{7.25}$$

La matrice de la rotation est donc donnée par la formule

$$\mathbf{A} = (1 - \cos\theta)\mathbf{v}_e{}^t\mathbf{v}_e + \cos\theta\mathbf{I} + \sin\theta\mathbf{V}. \tag{7.26}$$

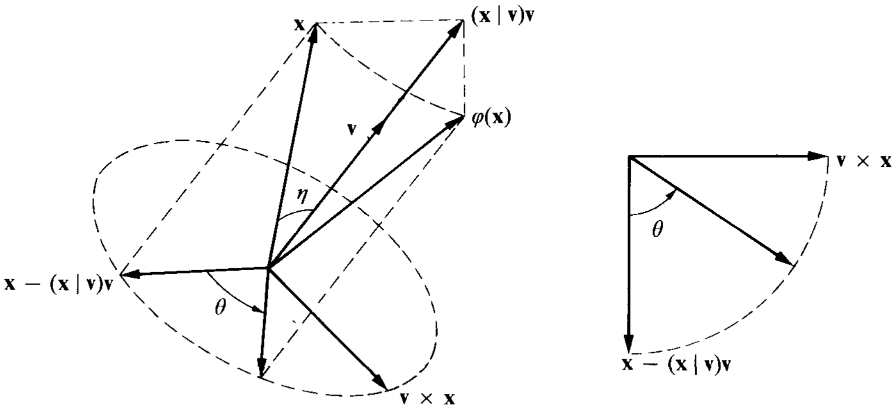


Fig. 7.3

7.3 ISOMÉTRIES, SIMILITUDES

7.3.1 Introduction

Les applications qui conservent la structure métrique d'un espace sont appelées isométries. Il s'avère que toute isométrie d'un espace affine est une application affine.

Dans cette section, $\mathscr{E}$ désignera un espace affine euclidien de direction E.

7.3.2 Isométries

On dit qu'une application Φ de $\mathscr{E}$ dans $\mathscr{E}$ est une *isométrie* si elle *conserve la distance*, c'est-à-dire si

$$\delta(\Phi(P), \Phi(Q)) = \delta(P, Q) \tag{7.27}$$

pour tout couple (P, Q) de points de $\mathscr{E}$.

7.3.3 Proposition. Caractérisation des isométries

Pour qu'une application Φ de $\mathscr{E}$ dans $\mathscr{E}$ soit une isométrie, il faut et il suffit que Φ soit affine et admette comme application linéaire associée une transformation orthogonale.

DÉMONSTRATION

Supposons que Φ conserve la distance. Choisissons un point quelconque O de $\mathscr{E}$ et désignons par φ l'application de E dans E définie par

$$\varphi(\overrightarrow{OP}) = \overrightarrow{\Phi(O)\Phi(P)}.$$

Nous devons montrer que φ est une transformation orthogonale. De toute évidence $\varphi(\mathbf{0}) = \mathbf{0}$. En outre, par (7.27) et (2.53),

$$\begin{aligned} \| \varphi(\overrightarrow{OP}) - \varphi(\overrightarrow{OQ}) \| &= \| \overrightarrow{\Phi(O)\Phi(P)} - \overrightarrow{\Phi(O)\Phi(Q)} \| = \| \overrightarrow{\Phi(Q)\Phi(P)} \| \\ &= \delta(\Phi(Q), \Phi(P)) = \delta(Q, P) = \| \overrightarrow{QP} \| = \| \overrightarrow{OP} - \overrightarrow{OQ} \|, \end{aligned}$$

ce qui montre que φ satisfait à la condition (b) de la proposition 7.1.3 et donc que φ est une transformation orthogonale.

Réciproquement, si Φ est affine et l'application linéaire associée φ est une transformation orthogonale, (2.53), (6.43) et la conservation de la norme entraînent que

$$\delta(\Phi(P), \Phi(Q)) = \| \overrightarrow{\Phi(P)\Phi(Q)} \| = \| \varphi(\overrightarrow{PQ}) \| = \| \overrightarrow{PQ} \| = \delta(P, Q),$$

ce qui montre que Φ conserve la distance.

7.3.4 Remarques

(1) Lorsque $\mathscr{E}$ est de dimension finie non nulle, l'équation matricielle d'une isométrie dans un repère orthonormal quelconque de $\mathscr{E}$ s'écrit

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{b},$$

où $\mathbf{A}$ est une matrice orthogonale.

(2) Toute isométrie est injective (bijective si $\mathscr{E}$ est de dimension finie). C'est une conséquence de (3) de 6.7.13 et 7.1.4.

(3) La composée de deux isométries (non nécessairement distinctes), ainsi que l'inverse de toute isométrie bijective, est encore une isométrie. Cela découle directement de la définition 7.3.2.

(4) Toute isométrie conserve les angles. En effet, $\widehat{\Phi(P)\Phi(Q)\Phi(R)}$ est l'angle des vecteurs $\overrightarrow{\Phi(Q)\Phi(P)} = \varphi(\overrightarrow{QP})$ et $\overrightarrow{\Phi(Q)\Phi(R)} = \varphi(\overrightarrow{QR})$, donc l'angle des vecteurs $\overrightarrow{QP}$ et $\overrightarrow{QR}$, puisque φ conserve les angles, c'est-à-dire $P\hat{Q}R$.

(5) Lorsque $\mathscr{E}$ est de dimension finie non nulle, toute isométrie conserve le volume des parallélépipèdes. Cela résulte de (6.65), vu que $|\det\varphi| = 1$.

7.3.5 Exemples

Les translations et les symétries affines orthogonales sont des isométries.

7.3.6 Isométries à deux dimensions

Supposons que $\mathscr{E}$ soit orienté, de dimension 2 et muni d'un repère orthonormal direct d'origine O. Soit Φ une isométrie de $\mathscr{E}$ et φ la transformation orthogonale associée à Φ.

Si $\varphi = \mathrm{id}_E$, Φ est une translation.

Si $\det\varphi = 1$ et $\varphi \neq \mathrm{id}_E$, Φ admet un point invariant unique P_0, car la matrice $\mathbf{I} - \mathbf{A}$ de l'équation (6.58) est, dans ce cas, inversible, du fait que

$$|\mathbf{I} - \mathbf{A}| = \begin{vmatrix} 1-\cos\theta & \sin\theta \\ -\sin\theta & 1-\cos\theta \end{vmatrix} = (1 - \cos\theta)^2 + \sin^2\theta = 2(1 - \cos\theta) \neq 0.$$

Il s'ensuit que Φ est une rotation du vectorialisé $\mathscr{E}_{P_0}$. On l'appelle *rotation de centre P_0 et d'angle orienté θ*, cet angle orienté n'étant autre que celui de la rotation (vectorielle) φ.

Si $\det\varphi = -1$, grâce aux résultats obtenus dans 7.2.2, nous déduisons de l'équation matricielle (6.56) (ou directement de (6.43)) que Φ est la composée d'une symétrie orthogonale par rapport à une droite $\mathscr{D}$ passant par O et d'une translation de vecteur $\mathbf{v} = \overrightarrow{O\Phi(O)}$.

Supposons d'abord que $\mathbf{v}$ soit orthogonal à la direction de $\mathscr{D}$. Le point $P_0 = O + \frac{1}{2}\mathbf{v}$ est alors un point invariant de Φ, donc Φ est une symétrie orthogonale par rapport à la droite $\mathscr{D}_1$ passant par P_0 et parallèle à $\mathscr{D}$.

Supposons que $\mathbf{v}$ ne soit pas orthogonal à la direction de $\mathscr{D}$. Alors

$$\mathbf{v} = \mathbf{v}_1 + \mathbf{v}_2,$$

où $\mathbf{v}_1$ est un vecteur orthogonal à la direction de $\mathscr{D}$ et $\mathbf{v}_2$ un vecteur de la direction de $\mathscr{D}$. Désignons par $\mathscr{D}_1$ l'image de $\mathscr{D}$ par la translation de vecteur $\frac{1}{2}\mathbf{v}_1$. D'après ce qui vient d'être établi, Φ est la composée de la symétrie orthogonale par rapport à $\mathscr{D}_1$ et de la translation de vecteur $\mathbf{v}_2$ (fig. 7.4).

Une isométrie de $\mathscr{E}$ est donc une translation ou une rotation ou la composée d'une symétrie axiale orthogonale et d'une translation parallèlement à l'axe de symétrie, cette composée pouvant se réduire à la seule symétrie.

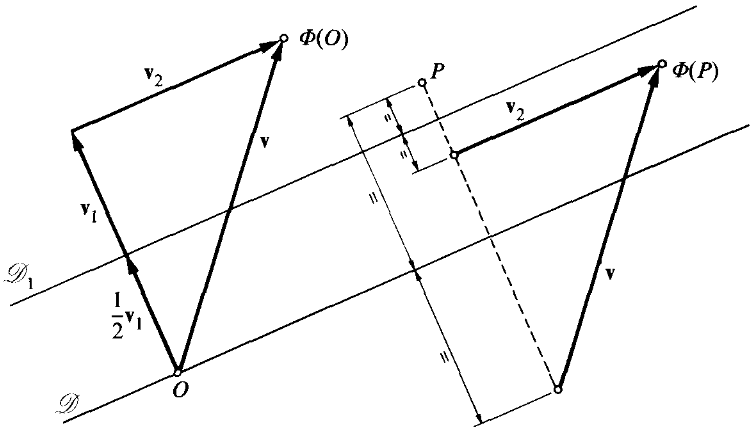


Fig. 7.4

7.3.7 Isométries à trois dimensions

Les isométries à 3 dimensions peuvent être classifiées par des arguments analogues à ceux que nous avons utilisés dans le paragraphe précédent. La discussion étant toutefois plus longue, nous nous limiterons à énoncer les conclusions.

Supposons que $\mathscr{E}$ soit orienté et de dimension 3 et considérons une isométrie Φ de $\mathscr{E}$ de transformation orthogonale associée φ.

Si $\varphi = \mathrm{id}_E$, Φ est une translation.

Si $\det\varphi = 1$ et $\varphi \neq \mathrm{id}_E$, nous distinguons deux cas:

(1) Φ admet un point invariant P_0. Dans ce cas, Φ est une rotation du vectorialisé $\mathscr{E}_{P_0}$. Soit $\mathscr{D}$ et θ respectivement l'axe et l'angle orienté de cette rotation. On appelle Φ *rotation d'axe* $\mathscr{D}$ *et d'angle orienté* θ. Evidemment, la direction de $\mathscr{D}$ n'est autre que l'axe (orienté) de la rotation (vectorielle) φ et θ l'angle orienté de cette rotation.

(2) Φ n'admet aucun point invariant. Dans ce cas, Φ est la composée d'une rotation et d'une translation parallèlement à l'axe de rotation.

Si $\det\varphi = -1$, nous distinguons également deux cas:

(3) Φ admet un point invariant P_0. Dans ce cas, Φ est la composée d'une rotation et d'une symétrie plane orthogonale, plus exactement la symétrie orthogonale par rapport au plan passant par P_0 et orthogonal à l'axe de rotation, cette

composée pouvant se réduire à la seule symétrie (lorsque la rotation se réduit à l'identité).

(4) Φ n'admet aucun point invariant. Dans ce cas, Φ est la composée d'une symétrie plane orthogonale et d'une translation parallèlement au plan de symétrie.

7.3.8 Similitudes

Soit λ un nombre positif. On dit qu'une application Φ de $\mathscr{E}$ dans $\mathscr{E}$ est une *similitude* de rapport λ si elle *multiplie la distance* par λ, c'est-à-dire si

$$\delta(\Phi(P), \Phi(Q)) = \lambda\delta(P, Q) \tag{7.28}$$

pour tout couple (P, Q) de points de $\mathscr{E}$.

Une similitude de rapport 1 est une isométrie. Toute homothétie de rapport non nul μ est une similitude de rapport $|\mu|$.

La composée d'une similitude de rapport λ et d'une homothétie de rapport λ^{-1} est manifestement une isométrie. A l'aide de cette observation, les similitudes peuvent être caractérisées ainsi:

7.3.9 Proposition. Caractérisation des similitudes

Soit Φ une application de $\mathscr{E}$ dans $\mathscr{E}$ et λ un nombre positif. Les conditions suivantes sont équivalentes:

(a) *Φ est une similitude de rapport λ.*

(b) *Φ est la composée d'une isométrie et d'une homothétie de rapport λ.*

(c) *Φ est une application affine dont l'application linéaire associée est λ fois une transformation orthogonale.*

7.3.10 Remarques

(1) Lorsque $\mathscr{E}$ est de dimension finie non nulle, l'équation matricielle d'une similitude de rapport λ dans un repère orthonormal quelconque de $\mathscr{E}$ s'écrit

$$\mathbf{y} = \lambda\mathbf{A}\mathbf{x} + \mathbf{b},$$

où $\mathbf{A}$ est une matrice orthogonale.

(2) Toute similitude est injective (bijective si $\mathscr{E}$ est de dimension finie).

(3) La composée de deux similitudes (non nécessairement distinctes) de rapports respectifs λ et μ est une similitude de rapport $\lambda\mu$. L'inverse d'une similitude bijective de rapport λ est une similitude de rapport λ^{-1}.

(4) Toute similitude conserve les angles.

(5) Lorsque $\mathscr{E}$ est de dimension finie non nulle n, toute similitude de rapport λ multiplie le volume des parallélépipèdes par λ^n. Cela résulte de (6.65), puisque $|\det\lambda\varphi| = \lambda^n$.

7.4 EXERCICES

7.4.1 Soit E l'espace vectoriel des polynômes muni du produit scalaire défini dans l'exercice 2.9.2. On considère l'application $\mathbf{p} \to \varphi(\mathbf{p})$ de E dans E définie par $\varphi(\mathbf{p})(t) \equiv t\mathbf{p}(t)$.

(a) Montrer que φ est une transformation orthogonale.
(b) Montrer que φ n'est pas surjective.

7.4.2 Pour quelles valeurs de a et de b la matrice carrée d'ordre $n \geqslant 2$

$$a\begin{pmatrix} b & 1 & 1 & . & . & . & 1 \\ 1 & b & 1 & . & . & . & . \\ 1 & 1 & b & . & . & . & . \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ . & . & . & . & . & b & 1 \\ 1 & . & . & . & . & 1 & b \end{pmatrix}$$

est-elle orthogonale?

7.4.3 Soit a, b, c et d des nombres non tous nuls. Après avoir constaté que les colonnes de la matrice

$$\mathbf{A} = \begin{pmatrix} a & -b & -c & -d \\ b & a & d & -c \\ c & -d & a & b \\ d & c & -b & a \end{pmatrix}$$

sont orthogonales deux à deux, déterminer les valeurs de α pour lesquelles $\alpha\mathbf{A}$ est une matrice orthogonale. A l'aide des propriétés des matrices orthogonales, déduire ensuite $|\mathbf{A}|$ et $\mathbf{A}^{-1}$.

7.4.4 Soit $\mathbf{A}$ une matrice antisymétrique (cf. exercice 4.7.16).

(a) Montrer que $\mathbf{I} + \mathbf{A}$ est une matrice inversible. (Prouver que l'équation $\mathbf{Ax} = -\mathbf{x}$ n'admet que la solution nulle et appliquer le corollaire 3.5.4 et la proposition 4.2.4.)
(b) Montrer que la matrice $(\mathbf{I} - \mathbf{A})(\mathbf{I} + \mathbf{A})^{-1}$ est orthogonale.

7.4.5 Soit $\mathbf{A}$ une matrice antisymétrique (cf. exercice 4.7.16). Montrer que la matrice $\exp(\mathbf{A})$ est orthogonale.

7.4.6 Soit E un espace vectoriel euclidien, orienté et de dimension 2. Soit $\mathbf{x}$ et $\mathbf{y}$ des vecteurs non nuls de E de même norme. Montrer qu'il existe une unique rotation φ et une unique symétrie axiale ψ telles que $\varphi(\mathbf{x}) = \mathbf{y}$ et $\psi(\mathbf{x}) = \mathbf{y}$.

7.4.7 Soit E un espace vectoriel euclidien, orienté et de dimension 3. Soit φ l'application linéaire de E dans E dont la matrice dans une base orthonormale directe de E est

$$\frac{1}{4}\begin{bmatrix} 1 & \sqrt{3} & 2\sqrt{3} \\ 2\sqrt{3} & -2 & 0 \\ \sqrt{3} & 3 & -2 \end{bmatrix}.$$

(a) Montrer que φ est une rotation.
(b) Trouver l'angle et l'axe de rotation.

7.4.8 Soit E un espace vectoriel euclidien muni d'une base orthonormale définissant son orientation. Trouver la matrice dans cette base de la rotation d'angle orienté $\pi/3$ dont l'axe est la droite vectorielle engendrée et orientée par le vecteur $\mathbf{v}(2, 1, -1)$.

7.4.9 Soit E un espace vectoriel euclidien, orienté, de dimension 3 et muni d'une base orthonormale directe $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$. On considère l'application linéaire $\varphi = \varphi_3 \circ \varphi_2 \circ \varphi_1$ de E dans E, où φ_1, φ_2, φ_3 sont les rotations d'angles orientés respectifs θ_1, θ_2, θ_3 (*angles d'Euler*) et dont les axes respectifs sont les droites vectorielles engendrées et orientées par $\mathbf{e}_3$, $\varphi_1(\mathbf{e}_1)$, $(\varphi_2 \circ \varphi_1)(\mathbf{e}_3)$.
(a) Montrer que φ est une rotation.
(b) Chercher la matrice de φ dans la base $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$. (Poser $\mathbf{e}_i' = \varphi_1(\mathbf{e}_i)$, $\mathbf{e}_i'' = (\varphi_2 \circ \varphi_1)(\mathbf{e}_i)$, $\mathbf{e}_i''' = \varphi(\mathbf{e}_i)$, $i = 1, 2, 3$, utiliser (1) de 6.6.3 et (6.34).)

Exercices sur les isométries et les similitudes

7.4.10 Dans un espace affine euclidien de dimension 3, muni d'un repère orthonormal, on considère l'application affine d'équation matricielle

$$\mathbf{y} = \frac{1}{3}\begin{bmatrix} -1 & 2 & -2 \\ 2 & 2 & 1 \\ -2 & 1 & 2 \end{bmatrix}\mathbf{x} + \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}.$$

Quelle est la nature de cette application?

7.4.11 Dans un espace affine euclidien, orienté, de dimension 3 et muni d'un repère orthonormal direct, on considère l'application affine d'équation matricielle

$$\mathbf{y} = \frac{1}{3}\begin{bmatrix} 2 & 1 & 2 \\ 1 & 2 & -2 \\ -2 & 2 & 1 \end{bmatrix}\mathbf{x} + \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}.$$

Montrer que cette application est une rotation dont on déterminera l'angle et l'axe.

7.4.12 Dans un espace affine euclidien de dimension 4, muni d'un repère orthonormal, on considère l'application affine Φ d'équation matricielle

$$\mathbf{y} = \begin{bmatrix} 3 & 1 & -1 & 1 \\ -1 & 3 & -1 & -1 \\ 1 & 1 & 3 & -1 \\ -1 & 1 & 1 & 3 \end{bmatrix} \mathbf{x} + \begin{bmatrix} 1 \\ 0 \\ 2 \\ 4 \end{bmatrix}.$$

(a) Montrer que Φ est une similitude dont on déterminera le rapport.
(b) Calculer le volume de l'image par Φ du parallélépipède construit sur les points $P_0(1, 0, 0, 0)$, $P_1(0, 1, 1, -1)$, $P_2(2, -1, 0, 0)$, $P_3(0, 1, 2, 0)$ et $P_4(-1, 1, -1, 1)$.

7.4.13 Soit Φ_1 et Φ_2 deux symétries orthogonales par rapport à des plans non parallèles d'un espace affine euclidien de dimension 3. Montrer que $\Phi_1 \circ \Phi_2$ est une rotation dont on déterminera l'angle et l'axe.

7.4.14 Soit $\mathscr{S}_1$ et $\mathscr{S}_2$ deux hyperplans d'un espace affine euclidien de dimension supérieure à 1 dont les vecteurs normaux respectifs sont orthogonaux. Soit Φ_1 et Φ_2 les symétries orthogonales par rapport à $\mathscr{S}_1$ et à $\mathscr{S}_2$. Montrer que Φ_1 et Φ_2 commutent, c'est-à-dire que $\Phi_1 \circ \Phi_2 = \Phi_2 \circ \Phi_1$.

7.4.15 Soit Φ_1 et Φ_2 deux rotations d'un espace affine euclidien, orienté et de dimension 3. A quelles conditions Φ_1 et Φ_2 commutent?

7.4.16 Soit Φ_1 et Φ_2 deux symétries orthogonales par rapport à des hyperplans parallèles d'un espace affine euclidien de dimension finie non nulle. Montrer que $\Phi_1 \circ \Phi_2$ est une translation dont on déterminera le vecteur.

7.4.17 Démontrer le point (4) de 7.3.7.

CHAPITRE 8

Valeurs propres et vecteurs propres

8.1 EXEMPLES PRÉLIMINAIRES

8.1.1 Introduction

Une application linéaire φ d'un espace vectoriel dans lui-même peut être étudiée plus aisément lorsqu'on connaît des vecteurs non nuls $\mathbf{x}$ dont l'image $\varphi(\mathbf{x})$ est un multiple de $\mathbf{x}$. On appelle tout vecteur qui se transforme de cette manière vecteur propre de φ. Similairement, on appelle vecteur propre d'une matrice carrée $\mathbf{A}$ tout vecteur-colonne non nul $\mathbf{x}$ tel que $\mathbf{Ax}$ est un multiple de $\mathbf{x}$.

La notion de vecteur propre sera reprise et approfondie à partir de la section 8.2. Dans cette section, nous nous limiterons à illustrer son utilité par trois exemples.

8.1.2 Exemple. Résolution d'un système différentiel

Considérons deux masses égales m attachées à trois ressorts de même longueur, comme il est indiqué dans la figure 8.1. Nous savons que lorsqu'un

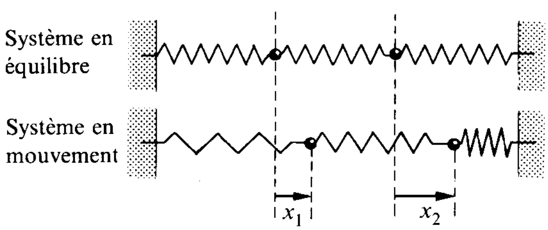


Fig. 8.1

ressort est comprimé ou étiré d'une longueur x, il exerce une force en sens opposé proportionnelle à x. Soit c le facteur de proportionnalité. Les équations qui régissent le mouvement sont alors

$$\begin{aligned} m\ddot{x}_1 &= -cx_1 + c(x_2-x_1) = -2cx_1 + cx_2 \\ m\ddot{x}_2 &= -c(x_2-x_1) - cx_2 = cx_1 - 2cx_2. \end{aligned} \tag{8.1}$$

Bien entendu, x_1 et x_2 dépendent ici du temps t et les deux points qui les surmontent désignent leur deuxième dérivée par rapport à t.

Posons

$$a = \frac{c}{m}, \quad \mathbf{A} = \begin{bmatrix} -2a & a \\ a & -2a \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

Les deux équations (8.1) s'écrivent de manière équivalente sous la forme

$$\ddot{\mathbf{x}} = \mathbf{A}\mathbf{x}, \tag{8.2}$$

où il est convenu que dériver un vecteur-colonne signifie dériver chacun de ses termes.

Nous nous proposons de résoudre cette équation différentielle matricielle à l'aide d'un changement de variables

$$\mathbf{x} = \mathbf{P}\mathbf{y}, \tag{8.3}$$

où

$$\mathbf{P} = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix}$$

est une matrice constante inversible et

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

une matrice-colonne dépendant de t. En substituant $\mathbf{Py}$ à $\mathbf{x}$ dans (8.2), nous obtenons

$$\mathbf{P}\ddot{\mathbf{y}} = \mathbf{APy} \qquad \text{ou} \qquad \ddot{\mathbf{y}} = \mathbf{P}^{-1}\mathbf{APy}. \tag{8.4}$$

Nous pouvons ainsi constater que le changement de variables (8.3) sera avantageux si $\mathbf{P}^{-1}\mathbf{AP}$ est une matrice plus simple que $\mathbf{A}$. Demandons-nous, sans détour, s'il existe une matrice inversible $\mathbf{P}$ telle que

$$\mathbf{P}^{-1}\mathbf{AP} = \operatorname{diag}(\lambda_1, \lambda_2) \qquad \text{ou} \qquad \mathbf{AP} = \mathbf{P}\operatorname{diag}(\lambda_1, \lambda_2). \tag{8.5}$$

Remarquons qu'en vertu de (4.8) cette dernière équation est équivalente aux deux équations

$$\mathbf{A}\begin{bmatrix} p_{11} \\ p_{21} \end{bmatrix} = \lambda_1 \begin{bmatrix} p_{11} \\ p_{21} \end{bmatrix} \qquad \text{et} \qquad \mathbf{A}\begin{bmatrix} p_{12} \\ p_{22} \end{bmatrix} = \lambda_2 \begin{bmatrix} p_{12} \\ p_{22} \end{bmatrix}.$$

Nous voyons ainsi que la réponse à notre question est affirmative si et seulement si $\mathbf{A}$ admet deux vecteurs propres linéairement indépendants. Afin de savoir si tel est le cas, cherchons les valeurs de λ pour lesquelles l'équation matricielle

$$\mathbf{A}\begin{bmatrix} p_1 \\ p_2 \end{bmatrix} = \lambda \begin{bmatrix} p_1 \\ p_2 \end{bmatrix}, \quad \text{c'est-à-dire} \quad \begin{bmatrix} -2a-\lambda & a \\ a & -2a-\lambda \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \tag{8.6}$$

admet au moins une solution non nulle. D'après le corollaire 3.5.4 (vu les propositions 4.2.4 et 5.1.6), ces valeurs sont les solutions de l'équation

$$\begin{vmatrix} -2a-\lambda & a \\ a & -2a-\lambda \end{vmatrix} = (2a + \lambda)^2 - a^2 = 0,$$

soit $\lambda_1 = -a$ et $\lambda_2 = -3a$. Des solutions non nulles correspondantes de l'équation matricielle (8.6) sont, par exemple,

$$\begin{bmatrix} p_{11} \\ p_{21} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \text{et} \quad \begin{bmatrix} p_{12} \\ p_{22} \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

En prenant

$$\mathbf{P} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix},$$

nous voyons donc que l'équation différentielle matricielle (8.4) devient

$$\ddot{\mathbf{y}} = \operatorname{diag}(-a, -3a)\mathbf{y}, \qquad \text{ou encore} \qquad \begin{aligned} \ddot{y}_1 &= -ay_1 \\ \ddot{y}_2 &= -3ay_2. \end{aligned}$$

La solution générale de ces deux équations étant donnée par

$$\begin{aligned} y_1 &= c_1 \sin(\sqrt{a}t + d_1) \\ y_2 &= c_2 \sin(\sqrt{3a}t + d_2), \end{aligned}$$

la solution générale du système différentiel (8.1) s'écrira, compte tenu de (8.3),

$$\begin{aligned} x_1 &= c_1 \sin(\sqrt{a}t + d_1) - c_2 \sin(\sqrt{3a}t + d_2) \\ x_2 &= c_1 \sin(\sqrt{a}t + d_1) + c_2 \sin(\sqrt{3a}t + d_2), \end{aligned} \tag{8.7}$$

où c_1, c_2, d_1 et d_2 sont des constantes déterminées par les conditions initiales, c'est-à-dire par la position initiale $\mathbf{x}(0)$ et la vitesse initiale $\dot{\mathbf{x}}(0)$.

Par exemple, si $a = 1$, dans les conditions initiales

$$\mathbf{x}(0) = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \quad \text{et} \quad \dot{\mathbf{x}}(0) = \begin{bmatrix} -1 \\ 1 \end{bmatrix},$$

la solution du système différentiel (8.1) est

$$\begin{aligned} x_1 &= -\frac{2}{\sqrt{3}} \sin(\sqrt{3}t + \frac{\pi}{3}) \\ x_2 &= \frac{2}{\sqrt{3}} \sin(\sqrt{3}t + \frac{\pi}{3}). \end{aligned}$$

On notera que le procédé de résolution que nous venons de présenter s'applique à toute équation différentielle matricielle

$$\ddot{\mathbf{x}} = \mathbf{A}\mathbf{x} \tag{8.8}$$

dont la matrice carrée $\mathbf{A}$ est d'ordre 2 et admet deux vecteurs linéairement indépendants $\mathbf{p}_1$ et $\mathbf{p}_2$ tels que $\mathbf{A}\mathbf{p}_1 = \lambda_1\mathbf{p}_1$ et $\mathbf{A}\mathbf{p}_2 = \lambda_2\mathbf{p}_2$, où λ_1 et λ_2 sont des nombres négatifs. La solution générale s'écrit, dans ce cas,

$$\mathbf{x} = c_1 \sin(\sqrt{-\lambda_1}t + d_1)\mathbf{p}_1 + c_2 \sin(\sqrt{-\lambda_2}t + d_2)\mathbf{p}_2. \tag{8.9}$$

8.1.3 Exemple. Calcul de la direction des axes d'une ellipse

Dans le plan affine euclidien, muni d'un repère orthonormal $(O; \mathbf{e}_1, \mathbf{e}_2)$, considérons l'ellipse d'équation

$$5x_1^2 - 2\sqrt{3}x_1x_2 + 7x_2^2 = 16. \tag{8.10}$$

Nous nous proposons de déterminer la direction de ses axes de symétrie. Un calcul tout à fait simple montre que l'équation (8.10) peut s'écrire sous la forme matricielle

$$^t\mathbf{x}\mathbf{A}\mathbf{x} = 16,$$

où

$$\mathbf{x} = (\overrightarrow{OP})_e = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad \text{et} \quad \mathbf{A} = \begin{bmatrix} 5 & -\sqrt{3} \\ -\sqrt{3} & 7 \end{bmatrix},$$

P désignant un point générique de l'ellipse. D'après un résultat d'analyse, la direction de la droite perpendiculaire à cette courbe au point P est définie par le gradient du premier membre de l'équation (8.10), à savoir par

$$\begin{bmatrix} \dfrac{\partial(^t\mathbf{x}\mathbf{A}\mathbf{x})}{\partial x_1} \\ \dfrac{\partial(^t\mathbf{x}\mathbf{A}\mathbf{x})}{\partial x_2} \end{bmatrix} = 2 \begin{bmatrix} 5x_1 - \sqrt{3}x_2 \\ \\ -\sqrt{3}x_1 + 7x_2 \end{bmatrix}.$$

Or,

$$\begin{bmatrix} 5x_1 - \sqrt{3}x_2 \\ -\sqrt{3}x_1 + 7x_2 \end{bmatrix} = \mathbf{A}\mathbf{x},$$

donc $\mathbf{Ax}$ définit un vecteur $\mathbf{n}$ normal à l'ellipse au point P (fig. 8.2). Les axes de symétrie étant perpendiculaires à l'ellipse, il s'avère ainsi que leurs directions sont

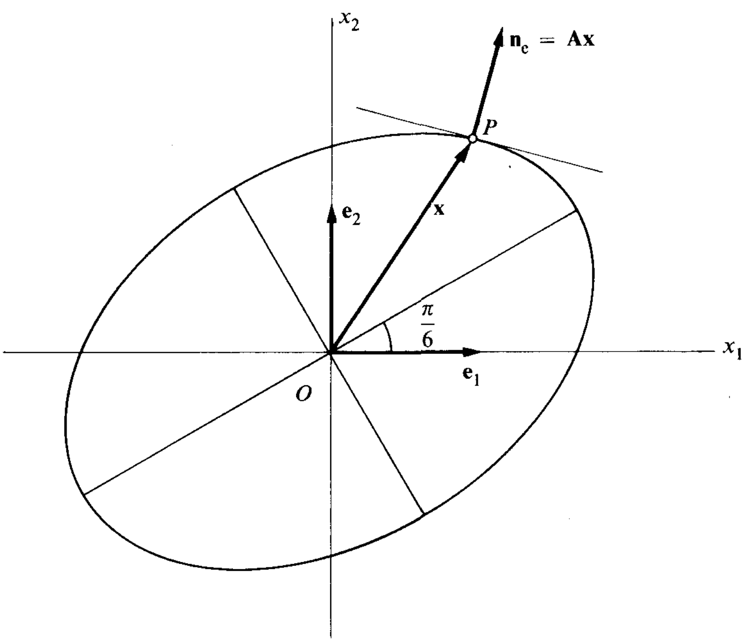


Fig. 8.2

déterminées par les vecteurs non nuls $\mathbf{x}$ tels que $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$, c'est-à-dire par les vecteurs propres de $\mathbf{A}$. Un calcul analogue à celui que nous avons effectué dans l'exemple 8.1.2 nous fournit ces directions, à savoir

$$\begin{pmatrix} \sqrt{3} \\ 1 \end{pmatrix} \quad \text{et} \quad \begin{pmatrix} -1 \\ \sqrt{3} \end{pmatrix}.$$

8.1.4 Exemple. Etude d'une application linéaire

Soit E un espace vectoriel de dimension 2, muni d'une base $(\mathbf{e}_1, \mathbf{e}_2)$. Nous nous proposons d'établir la nature de l'application linéaire φ de E dans E dont la matrice dans la base $(\mathbf{e}_1, \mathbf{e}_2)$ est

$$\mathbf{A} = \begin{pmatrix} 2 & 2 \\ 1 & 3 \end{pmatrix}.$$

A cet effet, vérifions si φ admet des vecteurs propres. L'équation vectorielle $\varphi(\mathbf{x}) = \lambda\mathbf{x}$ étant équivalente à l'équation matricielle $\mathbf{A}\mathbf{x}_e = \lambda\mathbf{x}_e$, il nous faut chercher les valeurs de λ pour lesquelles l'équation

$$\begin{pmatrix} 2-\lambda & 2 \\ 1 & 3-\lambda \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{8.11}$$

admet au moins une solution non nulle. Ces valeurs sont les solutions de l'équation

$$\begin{vmatrix} 2-\lambda & 2 \\ 1 & 3-\lambda \end{vmatrix} = \lambda^2 - 5\lambda + 4 = 0,$$

soit

$$\lambda_1 = 1 \quad \text{et} \quad \lambda_2 = 4.$$

Pour chacune d'entre elles, l'équation matricielle (8.11) admet au moins une solution non nulle, par exemple

$$(\mathbf{e}_1')_e = \begin{pmatrix} 2 \\ -1 \end{pmatrix} \quad \text{et} \quad (\mathbf{e}_2')_e = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Comme $\varphi(\mathbf{e}_1') = \mathbf{e}_1'$ et $\varphi(\mathbf{e}_2') = 4\mathbf{e}_2'$, la matrice de φ dans la base $(\mathbf{e}_1', \mathbf{e}_2')$ est

$$\mathbf{A}' = \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix}.$$

L'application φ est donc une dilatation relativement à D_1, de direction D_2 et de rapport 4, D_1 et D_2 désignant les droites vectorielles engendrées respectivement par $\mathbf{e}_1'$ et $\mathbf{e}_2'$ (fig. 8.3).

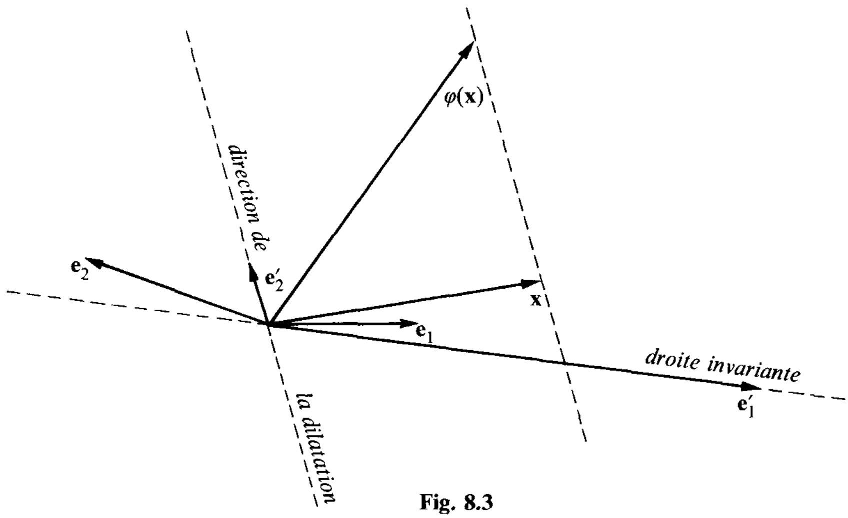


Fig. 8.3

8.2 DÉFINITIONS ET PREMIÈRES CONSÉQUENCES

8.2.1 Introduction

Dans cette section, nous abordons l'étude des différentes notions liées à la notion de vecteur propre.

Tout au long de la section, E désignera un espace vectoriel différent de $\{\mathbf{0}\}$ et φ une application linéaire de E dans E.

8.2.2 Valeurs propres, vecteurs propres, sous-espaces propres

On dit qu'un nombre λ est une *valeur propre* de φ s'il existe un vecteur non nul $\mathbf{x}$ de E tel que $\varphi(\mathbf{x}) = \lambda\mathbf{x}$. Ce vecteur est alors appelé *vecteur propre* associé à la valeur propre λ.

Dire que λ est une valeur propre de φ équivaut donc à dire que le noyau de $\varphi - \lambda \mathrm{id}_E$ ne se réduit pas à $\{\mathbf{0}\}$. Si tel est le cas, on désigne ce noyau par $S(\lambda)$ et on l'appelle *sous-espace propre* associé à la valeur propre λ. Les vecteurs propres associés à la valeur propre λ sont ainsi les vecteurs non nuls de $S(\lambda)$.

On appelle l'ensemble des valeurs propres de φ *spectre* de φ.

8.2.3 Remarques

(1) Si $\mathbf{x}$ est un vecteur propre associé à la valeur propre λ et α un nombre non nul, $\alpha\mathbf{x}$ est également un vecteur propre associé à la valeur propre λ.

(2) Les vecteurs propres associés à la valeur propre 1 sont les vecteurs invariants non nuls.

(3) Les vecteurs propres associés à la valeur propre 0 sont les vecteurs non nuls du noyau de φ.

(4) Si α est un nombre non nul, λ est une valeur propre de φ si et seulement si $\alpha\lambda$ est une valeur propre de $\alpha\varphi$. En outre, le sous-espace propre de φ associé à λ est égal au sous-espace propre de $\alpha\varphi$ associé à $\alpha\lambda$.

(5) Si λ est une valeur propre de φ et k un entier positif, λ^k est une valeur propre de φ^k. En outre, le sous-espace propre de φ associé à λ est inclus dans le sous-espace propre de φ^k associé à λ^k. Si φ est bijective, cette assertion est également vraie pour les entiers k négatifs.

(6) Si λ et μ sont des valeurs propres distinctes de φ, alors $S(\lambda) \cap S(\mu) = \{\mathbf{0}\}$. En d'autres termes, un vecteur propre ne peut être associé à deux valeurs propres.

(7) Supposons que φ désigne une transformation orthogonale d'un espace vectoriel euclidien.

Si λ est une valeur propre de φ, alors $\lambda = 1$ ou $\lambda = -1$. En effet, $\mathbf{x}$ étant un vecteur propre associé à λ, vu que φ conserve la norme,

$$\| \mathbf{x} \| = \| \varphi(\mathbf{x}) \| = \| \lambda\mathbf{x} \| = | \lambda | \| \mathbf{x} \|,$$

d'où l'on tire $| \lambda | = 1$.

Si 1 et -1 sont des valeurs propres de φ, alors $S(1)$ et $S(-1)$ sont orthogonaux. Prenons en effet un vecteur propre $\mathbf{x}$ associé à 1 et un vecteur propre $\mathbf{y}$ associé à -1; comme φ conserve le produit scalaire,

$$(\mathbf{x} \mid \mathbf{y}) = (\varphi(\mathbf{x}) \mid \varphi(\mathbf{y})) = (\mathbf{x} \mid -\mathbf{y}) = -(\mathbf{x} \mid \mathbf{y}),$$

d'où l'on tire $(\mathbf{x} \mid \mathbf{y}) = 0$.

8.2.4 Exemples

(1) L'homothétie de rapport λ admet une seule valeur propre, à savoir λ. Le sous-espace propre associé est l'espace E tout entier. En particulier, l'application identique, l'application nulle et la symétrie centrale admettent comme seule valeur propre respectivement 1, 0 et -1.

(2) Supposons que E soit somme directe de deux sous-espaces vectoriels S et T différents de $\{\mathbf{0}\}$. Les valeurs propres de la dilatation parallèlement à S, de direction T et de rapport λ, sont 1 et λ. Si λ est différent de 1, $S(1) = S$ et $S(\lambda) = T$. Si λ est égal à 1, $S(1) = E$.

(3) Supposons que E soit euclidien et de dimension 3. Une rotation d'angle orienté différent de 0 et π n'admet qu'une seule valeur propre, à savoir 1. Le sous-espace propre associé est l'axe de rotation.

(4) Supposons que E soit euclidien et de dimension 2. Une rotation d'angle orienté différent de 0 et π n'admet aucune valeur propre.

8.2.5 Stabilité des sous-espaces propres

Rappelons qu'un sous-espace vectoriel S de E est dit *stable* si son image $\varphi(S)$ est contenue dans S (cf. 6.1.3).

Soit λ une valeur propre de φ. Pour tout vecteur $\mathbf{x}$ de $S(\lambda)$, $\varphi(\mathbf{x}) = \lambda\mathbf{x}$ est encore un vecteur de $S(\lambda)$. En d'autres termes, $S(\lambda)$ est stable.

On notera que la restriction de φ à $S(\lambda)$ est une homothétie de rapport λ.

8.2.6 Proposition. Sommes directes de sous-espaces propres

Si $\lambda_1, \lambda_2, ..., \lambda_k$ sont des valeurs propres distinctes de φ, la somme des sous-espaces propres $S(\lambda_1)$, $S(\lambda_2)$, ..., $S(\lambda_k)$ est directe.

DÉMONSTRATION

Raisonnons par récurrence sur k. Pour $k = 1$, il n'y a rien à démontrer. Supposons que la proposition soit vraie pour tout ensemble constitué de $k - 1$ valeurs propres ($k > 1$). D'après 1.8.10, il nous faut prouver que si $\mathbf{s}_1, \mathbf{s}_2, ..., \mathbf{s}_k$ sont des vecteurs appartenant respectivement à $S(\lambda_1)$, $S(\lambda_2)$, ..., $S(\lambda_k)$ et tels que

$$\mathbf{s}_1 + \mathbf{s}_2 + ... + \mathbf{s}_k = \mathbf{0}, \tag{8.12}$$

alors $\mathbf{s}_1 = \mathbf{s}_2 = ... = \mathbf{s}_k = \mathbf{0}$. En prenant l'image par φ des deux membres de (8.12), nous voyons d'abord que

$$\lambda_1\mathbf{s}_1 + \lambda_2\mathbf{s}_2 + ... + \lambda_k\mathbf{s}_k = \mathbf{0}. \tag{8.13}$$

En soustrayant (8.13) de (8.12) multipliée par λ_k, nous obtenons ensuite la relation

$$(\lambda_k - \lambda_1)\mathbf{s}_1 + (\lambda_k - \lambda_2)\mathbf{s}_2 + ... + (\lambda_k - \lambda_{k-1})\mathbf{s}_{k-1} = \mathbf{0}. \tag{8.14}$$

Or, par hypothèse de récurrence, la somme de $S(\lambda_1)$, $S(\lambda_2)$, ..., $S(\lambda_{k-1})$ est directe, donc chaque terme de la somme au premier membre de (8.14) est nul. Comme $\lambda_k \neq \lambda_i$ pour $i = 1, 2, ..., k - 1$, il s'ensuit que $\mathbf{s}_1 = \mathbf{s}_2 = ... = \mathbf{s}_{k-1} = \mathbf{0}$. De (8.12) il résulte enfin que $\mathbf{s}_k = \mathbf{0}$, ce qui achève la démonstration.

8.2.7 Corollaire

Si $\lambda_1, \lambda_2, ..., \lambda_k$ sont des valeurs propres distinctes de φ et $\mathbf{x}_1, \mathbf{x}_2, .., \mathbf{x}_k$ des vecteurs propres associés respectivement à ces valeurs propres, alors $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k$ sont linéairement indépendants.

DÉMONSTRATION

Comme la somme de $S(\lambda_1)$, $S(\lambda_2)$, ..., $S(\lambda_k)$ est directe, la relation $\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 + ... + \alpha_k\mathbf{x}_k = \mathbf{0}$ implique $\alpha_1\mathbf{x}_1 = \alpha_2\mathbf{x}_2 = ... = \alpha_k\mathbf{x}_k = \mathbf{0}$ et donc $\alpha_1 = \alpha_2 = ... = \alpha_k = 0$.

8.2.8 Applications linéaires diagonalisables

Supposons que E soit de dimension finie n. Le spectre de φ est alors un ensemble fini. En effet, la proposition 8.2.6 et (1.13) entraînent que ce spectre comprend n éléments au plus. Cela étant établi, on dit que l'application linéaire φ est *diagonalisable* si E est somme directe des sous-espaces propres $S(\lambda_1)$, $S(\lambda_2)$, ..., $S(\lambda_l)$, où $\lambda_1, \lambda_2, ..., \lambda_l$ sont les éléments du spectre de φ, c'est-à-dire les valeurs propres de φ.

Par exemple, les homothéties et les dilatations sont diagonalisables. En particulier, les projections et les symétries sont diagonalisables.

La justification du qualificatif «diagonalisable» apparaîtra dans la proposition 8.4.1.

8.2.9 Décomposition spectrale

Supposons que φ soit diagonalisable. Alors

$$\varphi = \lambda_1\varphi_1 + \lambda_2\varphi_2 + ... + \lambda_l\varphi_l, \tag{8.15}$$

où $\lambda_1, \lambda_2, ..., \lambda_l$ sont les valeurs propres de φ et, pour $i = 1, 2, ..., l$, φ_i est la projection sur $S(\lambda_i)$ parallèlement à $S(\lambda_1) \oplus ... \oplus \widehat{S(\lambda_i)} \oplus ... \oplus S(\lambda_l)$. En effet, si $\mathbf{s}_1 + \mathbf{s}_2 + ... + \mathbf{s}_l$ est la décomposition d'un vecteur quelconque $\mathbf{x}$ de E suivant les sous-espaces propres de φ, alors

$$\begin{aligned}\varphi(\mathbf{x}) &= \varphi(\mathbf{s}_1 + \mathbf{s}_2 + ... + \mathbf{s}_l) = \varphi(\mathbf{s}_1) + \varphi(\mathbf{s}_2) + ... + \varphi(\mathbf{s}_l)\\ &= \lambda_1\mathbf{s}_1 + \lambda_2\mathbf{s}_2 + ... + \lambda_l\mathbf{s}_l = \lambda_1\varphi_1(\mathbf{x}) + \lambda_2\varphi_2(\mathbf{x}) + ... + \lambda_l\varphi_l(\mathbf{x}),\end{aligned}$$

ce qui démontre la relation (8.15). On appelle cette relation *décomposition spectrale* de φ.

8.3 FORMULATION MATRICIELLE, POLYNÔME CARACTÉRISTIQUE

8.3.1 Valeurs propres, vecteurs propres, sous-espaces propres d'une matrice

Soit $\mathbf{A}$ une matrice carrée d'ordre n et φ l'application linéaire associée canoniquement à $\mathbf{A}$. On appelle *valeurs propres, vecteurs propres* et *sous-espaces propres* de $\mathbf{A}$ les valeurs propres, vecteurs propres et sous-espaces propres de φ. Dire qu'un nombre λ est une valeur propre de $\mathbf{A}$ équivaut donc à dire qu'il existe un vecteur-colonne non nul $\mathbf{x}$ de $\mathbb{R}^n$ tel que

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x}. \tag{8.16}$$

Ce vecteur-colonne est un vecteur propre de **A** associé à la valeur propre λ et le sous-espace propre de **A** associé à cette valeur propre est constitué de tous les vecteurs-colonnes $\mathbf{x}$ de $\mathbb{R}^n$ vérifiant (8.16).

L'ensemble des valeurs propres d'une matrice carrée **A** est appelé *spectre* de **A**. Déterminer le spectre de **A** revient à trouver les valeurs réelles du paramètre t telles que l'équation matricielle

$$(\mathbf{A} - t\mathbf{I})\mathbf{x} = \mathbf{0} \tag{8.17}$$

admette au moins une solution non nulle. Cette équation est l'expression matricielle du système homogène de matrice associée $\mathbf{A} - t\mathbf{I}$. D'après le corollaire 3.5.4 (vu les propositions 4.2.4 et 5.1.6), ce système admet au moins une solution non nulle si et seulement si

$$|\mathbf{A} - t\mathbf{I}| = 0. \tag{8.18}$$

La marche à suivre pour calculer les vecteurs propres (sous-espaces propres) d'une matrice carrée **A** est donc celle-ci:

(a) Calculer le déterminant de $\mathbf{A} - t\mathbf{I}$.
(b) Chercher les valeurs réelles de t qui annulent ce déterminant.
(c) Pour chacune des valeurs trouvées, calculer une solution non nulle (une base du sous-espace des solutions) du système homogène de matrice associée $\mathbf{A} - t\mathbf{I}$.

8.3.2 Exemple de calcul des sous-espaces propres

Nous allons calculer les sous-espaces propres de la matrice

$$\mathbf{A} = \begin{pmatrix} -15 & 1 & -9 \\ 0 & 6 & 0 \\ 4 & 1 & -3 \end{pmatrix}.$$

Les valeurs propres sont les solutions de l'équation

$$\begin{vmatrix} -15-t & 1 & -9 \\ 0 & 6-t & 0 \\ 4 & 1 & -3-t \end{vmatrix} = (6 - t)(9 + t)^2 = 0,$$

soit 6 et -9. Le système homogène de matrice associée $\mathbf{A} - 6\mathbf{I}$ est équivalent au système

$$\begin{aligned} -21x_1 + x_2 - 9x_3 &= 0 \\ 4x_1 + x_2 - 9x_3 &= 0. \end{aligned}$$

Le sous-espace propre associé à la valeur propre 6 est la droite vectorielle engendrée par le vecteur-colonne

$$\begin{pmatrix} 0 \\ 9 \\ 1 \end{pmatrix}.$$

Le système homogène de matrice associée $\mathbf{A} + 9\mathbf{I}$ est

$$\begin{aligned} -6x_1 + x_2 - 9x_3 &= 0 \\ 15x_2 &= 0 \\ 4x_1 + x_2 + 6x_3 &= 0. \end{aligned}$$

Le sous-espace propre associé à la valeur propre -9 est la droite vectorielle engendrée par le vecteur-colonne

$$\begin{bmatrix} -3 \\ 0 \\ 2 \end{bmatrix}.$$

8.3.3 Recherche des valeurs propres et des sous-espaces propres d'une application linéaire

Soit E un espace vectoriel de dimension finie non nulle n et φ une application linéaire de E dans E. La recherche des valeurs propres et des sous-espaces propres de φ peut se faire par l'entremise d'une base $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ de E. Désignons par $\mathbf{A}$ la matrice de φ dans cette base et exprimons l'équation $\varphi(\mathbf{x}) = \lambda\mathbf{x}$ sous sa forme matricielle équivalente $\mathbf{A}\mathbf{x}_e = \lambda\mathbf{x}_e$. Nous voyons alors immédiatement que $\mathbf{x}$ est un vecteur propre de φ associé à la valeur propre λ si et seulement si $\mathbf{x}_e$ est un vecteur propre de $\mathbf{A}$ associé à la même valeur propre. Il en résulte que les spectres de φ et de $\mathbf{A}$ sont confondus et que la recherche des sous-espaces propres de φ associés à une valeur propre se réduit à la recherche des sous-espaces propres de $\mathbf{A}$ associés à la même valeur propre.

On notera, au passage, que la j-ième colonne de $\mathbf{A}$ est de la forme

$$\begin{array}{cl} \begin{bmatrix} 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \\ \lambda \\ 0 \\ \cdot \\ \cdot \\ 0 \end{bmatrix} & \begin{array}{l} \\ \\ \\ \\ \\ \leftarrow j\text{-ième} \\ \\ \\ \\ \\ \end{array} \end{array}$$

si et seulement si λ est une valeur propre de φ et $\mathbf{e}_j$ est un vecteur propre associé à cette valeur propre.

8.3.4 Polynôme caractéristique

Soit $\mathbf{A}$ une matrice carrée d'ordre n. Par récurrence sur n, on constate aisément, moyennant (5.16) ou (5.18), que le déterminant de $\mathbf{A} - t\mathbf{I}$ est un polynôme en t de degré n. On l'appelle *polynôme caractéristique* de $\mathbf{A}$. On appelle en outre l'équation (8.18) *équation caractéristique* de $\mathbf{A}$. Ainsi, les valeurs propres de $\mathbf{A}$ sont les racines réelles du polynôme caractéristique de $\mathbf{A}$, c'est-à-dire les solutions réelles de l'équation caractéristique de $\mathbf{A}$.

8.3.5 Remarques

(1) Toute matrice carrée d'ordre impair admet au moins une valeur propre. Cela est dû au fait que les racines complexes d'un polynôme à coefficients réels apparaissent par couples de conjugués.

(2) D'après l'exemple (3) de 5.1.4, le polynôme caractéristique d'une matrice triangulaire d'ordre n $\mathbf{A} = (a_{ij})$ est $(a_{11} - t)(a_{22} - t) \dots (a_{nn} - t)$. Le spectre d'une telle matrice est donc l'ensemble des valeurs prises par les termes diagonaux de $\mathbf{A}$.

(3) Les polynômes caractéristiques d'une matrice carrée $\mathbf{A}$ et de sa transposée ${}^{\tau}\mathbf{A}$ sont égaux (exceptionnellement la transposition est désignée ici par τ). En effet, vu la propriété (b) de la proposition 5.1.8,

$$|{}^{\tau}\mathbf{A} - t\mathbf{I}| = |{}^{\tau}(\mathbf{A} - t\mathbf{I})| = |\mathbf{A} - t\mathbf{I}|,$$

ce qu'il fallait prouver.

8.3.6 Calcul de quelques coefficients du polynôme caractéristique

Soit $\mathbf{A} = (a_{ij})$ une matrice carrée d'ordre n. Pour $n = 2$ et $n = 3$, un calcul facile montre que

$$\begin{vmatrix} a_{11}-t & a_{12} \\ a_{21} & a_{22}-t \end{vmatrix} = t^2 - (a_{11} + a_{22})t + a_{11}a_{22} - a_{12}a_{21} = t^2 - (\operatorname{tr}\mathbf{A})t + |\mathbf{A}|$$

et

$$\begin{vmatrix} a_{11}-t & a_{12} & a_{13} \\ a_{21} & a_{22}-t & a_{23} \\ a_{31} & a_{32} & a_{33}-t \end{vmatrix} = -t^3 + (\operatorname{tr}\mathbf{A})t^2 - (\sum_{i=1}^{3} |\mathbf{A}_{ii}|)t + |\mathbf{A}|,$$

où $\mathbf{A}_{ii}$ est la matrice obtenue en supprimant la i-ième ligne et la i-ième colonne de $\mathbf{A}$. Dans le cas général, la connaissance d'une expression pour chaque coefficient du polynôme caractéristique n'est pas d'un grand intérêt. Relevons néanmoins que, au signe près, le premier et le deuxième coefficient sont encore 1 et $\operatorname{tr}\mathbf{A}$ et que le terme constant est encore $|\mathbf{A}|$. Plus précisément,

$$|\mathbf{A} - t\mathbf{I}| = (-1)^n t^n + (-1)^{n-1}(\operatorname{tr}\mathbf{A})t^{n-1} + \dots + |\mathbf{A}|. \tag{8.19}$$

8.3.7 Polynôme caractéristique d'une application linéaire

Soit E un espace vectoriel de dimension finie non nulle et φ une application linéaire de E dans E. On appelle *polynôme caractéristique* de φ le déterminant de l'application linéaire $\varphi - t\,\mathrm{id}_E$. D'après la définition de déterminant d'une application linéaire (cf. 6.6.13), le polynôme caractéristique de φ est donc égal au polynôme caractéristique de la matrice de φ dans une base quelconque de E.

8.3.8 Polynôme caractéristique de deux matrices semblables

Deux matrices semblables ont le même polynôme caractéristique, à savoir le polynôme caractéristique d'une application linéaire quelconque qui leur est simultanément associée (cf. 6.6.12). En particulier, deux matrices semblables ont le même spectre.

On notera, cependant, que deux matrices peuvent avoir le même polynôme caractéristique sans être semblables. Par exemple, les deux matrices

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{et} \quad \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$$

ont le même polynôme caractéristique, à savoir $(1 - t)^2$, mais elles ne sont pas semblables, car la seule matrice semblable à $\mathbf{I}$ est $\mathbf{I}$.

8.3.9 Multiplicité des valeurs propres

Soit E et φ comme dans 8.3.7 et $\mathbf{A}$ une matrice carrée. Soit en outre λ une valeur propre de φ ou de $\mathbf{A}$. On appelle *multiplicité* de λ, et on note $m(\lambda)$, la multiplicité de λ en tant que racine du polynôme caractéristique (c'est-à-dire l'exposant de $\lambda - t$ dans la décomposition de ce polynôme en facteurs irréductibles). On dit que λ est *simple*, *double* ou *triple*, suivant que $m(\lambda)$ est égal à 1, 2 ou 3.

Par exemple, si $\mathbf{A} = \operatorname{diag}(a_1, a_2, ..., a_n)$, $m(\lambda)$ est le nombre de termes a_i égaux à λ. On notera que le sous-espace propre $S(\lambda)$ est, dans ce cas, engendré par les vecteurs $\mathbf{e}_i$ de la base canonique de même indice que les a_i égaux à λ. La dimension de $S(\lambda)$ est donc égale à $m(\lambda)$.

On peut se demander si à une valeur propre de multiplicité m est toujours associé un sous-espace propre de dimension m. L'exemple 8.3.2 montre que ce n'est pas le cas. En effet, -9 est une valeur propre double, mais le sous-espace propre $S(-9)$ est de dimension 1.

Nous allons démontrer qu'en général

$$\dim S(\lambda) \leqslant m(\lambda). \tag{8.20}$$

A cet effet, désignons par φ soit l'application que nous nous sommes donnée, soit l'application associée canoniquement à $\mathbf{A}$. Choisissons une base quelconque $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_k)$ de $S(\lambda)$ et complétons cette base en une base $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ de E ou de $\mathbb{R}^n$. La matrice de φ dans cette base est de la forme

$$\left[\begin{array}{c:c} \lambda\mathbf{I}_k & \mathbf{C} \\ \hdashline \mathbf{O} & \mathbf{B} \end{array}\right],$$

où $\mathbf{B}$ est une matrice carrée d'ordre $n - k$ et $\mathbf{C}$ une matrice de type $k \times (n - k)$. D'après (5.24), le polynôme caractéristique de φ est égal à $(\lambda - t)^k |\mathbf{B} - t\mathbf{I}_{n-k}|$. Il s'ensuit que la multiplicité de λ est au moins k. Puisque k est la dimension de $S(\lambda)$, l'inégalité (8.20) est démontrée.

8.4 RÉDUCTION À LA FORME DIAGONALE

8.4.1 Théorème. Caractérisation des applications linéaires diagonalisables

Soit E un espace vectoriel de dimension finie non nulle n. Soit en outre φ une application linéaire de E dans E. Les conditions suivantes sont équivalentes:
(a) *φ est diagonalisable.*
(b) *E admet une base de vecteurs propres de φ.*
(c) *E admet une base dans laquelle la matrice de φ est diagonale.*
(d) *Toute racine du polynôme caractéristique de φ est réelle et de multiplicité égale à la dimension du sous-espace propre associé.*

DÉMONSTRATION

(a) entraîne (b). Supposons que E soit somme directe de $S(\lambda_1)$, $S(\lambda_2)$, ..., $S(\lambda_l)$, où $\{\lambda_1, \lambda_2, ..., \lambda_l\}$ est le spectre de φ. Soit $(\mathbf{e}_1^i, \mathbf{e}_2^i, ..., \mathbf{e}_{n_i}^i)$ une base de $S(\lambda_i)$ pour $i = 1, 2, ..., l$. Les termes de la famille

$$(\mathbf{e}_1^1, \mathbf{e}_2^1, ..., \mathbf{e}_{n_1}^1, \mathbf{e}_1^2, \mathbf{e}_2^2, ..., \mathbf{e}_{n_2}^2, ..., \mathbf{e}_1^l, \mathbf{e}_2^l, ..., \mathbf{e}_{n_l}^l) \tag{8.21}$$

sont des vecteurs propres de φ. D'autre part, grâce à (1.13), nous déduisons de l'hypothèse que $n_1 + n_2 + ... + n_l = \dim E$. Pour démontrer que E admet une base de vecteurs propres, il nous reste donc à établir que la famille (8.21) est libre. Or, puisque la somme de $S(\lambda_1)$, $S(\lambda_2)$, ..., $S(\lambda_l)$ est directe, la relation

$$\alpha_1^1\mathbf{e}_1^1 + ... + \alpha_{n_1}^1\mathbf{e}_{n_1}^1 + \alpha_1^2\mathbf{e}_1^2 + ... + \alpha_{n_2}^2\mathbf{e}_{n_2}^2 + ... + \alpha_1^l\mathbf{e}_1^l + ... + \alpha_{n_l}^l\mathbf{e}_{n_l}^l = \mathbf{0}$$

entraîne

$$\alpha_1^i\mathbf{e}_1^i + \alpha_2^i\mathbf{e}_2^i + ... + \alpha_{n_i}^i\mathbf{e}_{n_i}^i = \mathbf{0}$$

pour $i = 1, 2, ..., l$. Mais $(\mathbf{e}_1^i, \mathbf{e}_2^i, ..., \mathbf{e}_{n_i}^i)$ est une famille libre pour chaque i, ce qui nous permet d'en conclure que

$$\alpha_1^i = \alpha_2^i = ... = \alpha_{n_i}^i = 0$$

pour $i = 1, 2, ..., l$ et donc que la famille (8.21) est libre.

(b) entraîne (c). La matrice de φ dans n'importe quelle base de E constituée de vecteurs propres est diagonale. Dans la base (8.21), par exemple, cette matrice est

$$\begin{pmatrix} \lambda_1\mathbf{I}_{n_1} & & & 0 \\ & \lambda_2\mathbf{I}_{n_2} & & \\ & & \ddots & \\ 0 & & & \lambda_l\mathbf{I}_{n_l} \end{pmatrix}. \tag{8.22}$$

(c) entraîne (d). Supposons que E admette une base dans laquelle la matrice de φ est diagonale. D'après la remarque (2) de 8.3.5 et l'exemple présenté dans 8.3.9, la condition (d) est alors remplie.

(d) entraîne (a). Soit $\lambda_1, \lambda_2, ..., \lambda_l$ les racines réelles du polynôme caractéristique de φ. Par hypothèse, $m(\lambda_1) + m(\lambda_2) + ... + m(\lambda_l) = n$. Encore par hypothèse, la dimension de $S(\lambda_i)$ est égale à $m(\lambda_i)$ pour $i = 1, 2, ..., l$. En vertu de (1.13), la dimension de $S(\lambda_1) \oplus S(\lambda_2) \oplus ... \oplus S(\lambda_l)$ est donc n, ce qui entraîne que cette somme directe est égale à E.

8.4.2 Corollaire

Les données étant celles du théorème 8.4.1, *si les racines du polynôme caractéristique de φ sont toutes réelles et simples, alors φ est diagonalisable.*

8.4.3 Matrices diagonalisables

Soit **A** une matrice carrée d'ordre n. On dit que **A** est *diagonalisable* si l'application linéaire associée canoniquement à **A** est diagonalisable. D'après le théorème 8.4.1, les conditions suivantes sont équivalentes:

(a) **A** est diagonalisable.

(b) **A** admet n vecteurs propres linéairement indépendants.

(c) **A** est semblable à une matrice diagonale, autrement dit, il existe une matrice inversible **P** et une matrice diagonale **D** telles que

$$\mathbf{D} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P} \quad \text{ou} \quad \mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^{-1}. \tag{8.23}$$

(d) Toute racine du polynôme caractéristique de **A** est réelle et de multiplicité égale à la dimension du sous-espace propre associé. (Cette condition est remplie si toute racine est réelle et simple.)

La matrice **P** de la relation (8.23) est la matrice de passage de la base canonique à une base de vecteurs propres de **A**. Les colonnes de **P** sont donc des vecteurs propres de **A** et les termes diagonaux de **D** les valeurs propres correspondantes.

L'équivalence des conditions (b) et (c) peut être établie sans l'aide du théorème 8.4.1. Il suffit d'écrire la relation (8.23) sous la forme équivalente

$$\mathbf{A}\mathbf{P} = \mathbf{P}\mathbf{D} \tag{8.24}$$

et d'appliquer aux deux membres de cette égalité la règle (4.8).

8.4.4 Marche à suivre et exemple

La marche à suivre pour diagonaliser une matrice carrée **A** est la suivante:

(a) Former le polynôme caractéristique de **A** et déterminer ses racines (leur numérotation est arbitraire), ainsi que leurs multiplicités. S'il y a des racines complexes, **A** n'est pas diagonalisable.

(b) Chercher une base de chaque sous-espace propre de **A**. Si la dimension d'un sous-espace propre est inférieure à la multiplicité de la valeur propre correspondante (cela ne peut se produire que si la multiplicité est supérieure à 1), **A** n'est pas diagonalisable.
(c) Réunir ces bases en une famille de vecteurs-colonnes constituant la matrice **P**.
(d) Former la matrice diagonale **D**, c'est-à-dire (8.22).
(e) Calculer $\mathbf{P}^{-1}$ et écrire $\mathbf{D} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$.

A titre d'exemple, nous allons diagonaliser la matrice

$$\mathbf{A} = \begin{bmatrix} 1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{bmatrix}.$$

Les valeurs propres sont les racines du polynôme caractéristique

$$|\mathbf{A} - t\mathbf{I}| = -t^3 + 12t + 16 = -(t + 2)^2(t - 4),$$

soit -2 (double) et 4. Le rang de la matrice

$$\mathbf{A} - (-2)\mathbf{I} = \begin{bmatrix} 3 & -3 & 3 \\ 3 & -3 & 3 \\ 6 & -6 & 6 \end{bmatrix}$$

est égal à 1, donc $\dim S(-2) = 3 - 1 = 2$ (par la proposition 3.5.3). Comme $2 = m(-2)$, nous constatons à ce point que **A** est effectivement diagonalisable. Le système homogène associé à la matrice $\mathbf{A} - (-2)\mathbf{I}$ est équivalent à l'équation

$$x_1 - x_2 + x_3 = 0.$$

Deux solutions linéairement indépendantes de cette équation sont

$$\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \quad \text{et} \quad \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}.$$

Le système homogène associé à la matrice

$$\mathbf{A} - 4\mathbf{I} = \begin{bmatrix} -3 & -3 & 3 \\ 3 & -9 & 3 \\ 6 & -6 & 0 \end{bmatrix}$$

est équivalent au système

$$\begin{aligned} x_1 + x_2 - x_3 &= 0 \\ x_1 - x_2 \phantom{{}- x_3} &= 0. \end{aligned}$$

Une solution non nulle de ce système est

$$\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}.$$

Posons

$$\mathbf{P} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & -1 & 2 \end{bmatrix}.$$

Alors

$$\mathbf{P}^{-1} = \frac{1}{2}\begin{bmatrix} -1 & 3 & -1 \\ 2 & -2 & 0 \\ 1 & -1 & 1 \end{bmatrix}$$

et $\mathrm{diag}(-2, -2, 4) = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$.

8.4.5 Calcul des puissances entières d'une matrice diagonalisable

Supposons que $\mathbf{A}$ vérifie la relation (8.23). D'après (6.41), elle vérifie alors également la relation

$$\mathbf{D}^k = \mathbf{P}^{-1}\mathbf{A}^k\mathbf{P} \quad \text{ou} \quad \mathbf{A}^k = \mathbf{P}\mathbf{D}^k\mathbf{P}^{-1}, \tag{8.25}$$

pour tout entier non négatif k (pour tout entier k si $\mathbf{A}$ est inversible). Cela nous fournit un moyen très utile de calcul des puissances entières d'une matrice.

A titre d'exemple, nous allons calculer les puissances entières positives d'une matrice de transition d'ordre 2. D'après 4.6.2, une telle matrice s'exprime sous la forme

$$\mathbf{A} = \begin{bmatrix} 1-a & a \\ b & 1-b \end{bmatrix},$$

où a et b sont des nombres compris entre 0 et 1. Ecartons le cas banal où $a = b = 0$. Comme la somme des termes de chaque ligne de $\mathbf{A}$ est égale à 1, le vecteur

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

est un vecteur propre de $\mathbf{A}$ associé à la valeur propre 1. La deuxième valeur propre est donc $|\mathbf{A}| = 1 - a - b$, puisque le produit des racines du polynôme caractéristique est égal au terme constant de ce polynôme, à savoir $|\mathbf{A}|$. Un vecteur propre associé à la valeur propre $1-a-b$ est

$$\begin{bmatrix} a \\ -b \end{bmatrix}.$$

La relation (8.25) s'écrit donc

$$\begin{aligned}
\mathbf{A}^k &= \begin{bmatrix} 1 & a \\ 1 & -b \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & (1-a-b)^k \end{bmatrix}\begin{bmatrix} 1 & a \\ 1 & -b \end{bmatrix}^{-1} \\
&= \frac{1}{-a-b}\begin{bmatrix} -b-a(1-a-b)^k & -a+a(1-a-b)^k \\ -b+b(1-a-b)^k & -a-b(1-a-b)^k \end{bmatrix} \\
&= \frac{1}{a+b}\begin{bmatrix} b & a \\ b & a \end{bmatrix} + \frac{(1-a-b)^k}{a+b}\begin{bmatrix} a & -a \\ -b & b \end{bmatrix}.
\end{aligned} \tag{8.26}$$

Ce résultat nous permet de répondre affirmativement à la question posée à la fin de l'exemple 4.6.5. En effet, si $a + b < 2$, alors $\lim_{k\to\infty} (1 - a - b)^k = 0$, donc

$$\lim_{k\to\infty} \mathbf{A}^k = \frac{1}{a+b}\begin{bmatrix} b & a \\ b & a \end{bmatrix}$$

et, par conséquent (cf. (4.30)),

$$\lim_{k\to\infty} \mathbf{L}^{(k)} = \mathbf{L}^{(0)} \lim_{k\to\infty} \mathbf{A}^k = \mathbf{L}^{(0)}\frac{1}{a+b}\begin{bmatrix} b & a \\ b & a \end{bmatrix} = \frac{1}{a+b}\begin{bmatrix} b & a \end{bmatrix}, \tag{8.27}$$

ce qui montre que la limite des lois $\mathbf{L}^{(k)}$ est indépendante de la loi initiale $\mathbf{L}^{(0)}$.

Après un grand nombre d'achats, les probabilités d'achat de m_1 et de m_2 sont donc approximativement

$$\frac{b}{a+b} \cong 0{,}6 \qquad \text{et} \qquad \frac{a}{a+b} \cong 0{,}4.$$

On notera que la limite (8.27) est l'unique loi invariante de $\mathbf{A}$, c'est-à-dire l'unique vecteur-ligne $\mathbf{L}$ dont les termes sont non négatifs et de somme égale à 1, tel que

$$\mathbf{L} = \mathbf{LA}.$$

8.4.6 Retour aux fonctions matricielles

Il a été observé dans 6.6.14 que (8.23) entraîne la relation

$$f(\mathbf{D}) = \mathbf{P}^{-1}f(\mathbf{A})\mathbf{P} \qquad \text{ou} \qquad f(\mathbf{A}) = \mathbf{P}f(\mathbf{D})\mathbf{P}^{-1}, \tag{8.28}$$

où f est un polynôme ou, plus généralement, une fonction définie par une série entière appropriée (cf. 4.5.2).

L'utilité de l'expression de $f(\mathbf{A})$ ainsi obtenue est due au fait que

$$f(\mathrm{diag}(a_1, a_2, ..., a_n)) = \mathrm{diag}(f(a_1), f(a_2), ..., f(a_n)), \tag{8.29}$$

ce qui se démontre aisément en partant de (4.17).

Par exemple, si f est la fonction exponentielle (cf. 4.5.4), (8.28) s'écrit sous la forme

$$\exp(\mathbf{A}) = \mathbf{P}\mathrm{diag}(\mathrm{e}^{d_1}, \mathrm{e}^{d_2}, ..., \mathrm{e}^{d_n})\mathbf{P}^{-1}, \tag{8.30}$$

où $d_1, d_2, ..., d_n$ sont les termes diagonaux de $\mathbf{D}$.

Comme application numérique, considérons la matrice

$$\mathbf{A} = \frac{1}{4}\begin{bmatrix} 4 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 1 & 4 \end{bmatrix}.$$

Ses valeurs propres sont 1 (double) et $\frac{1}{2}$. Les sous-espaces propres associés sont engendrés respectivement par

$$\begin{pmatrix}1\\0\\0\end{pmatrix}, \quad \begin{pmatrix}0\\0\\1\end{pmatrix} \quad \text{et} \quad \begin{pmatrix}1\\-2\\1\end{pmatrix}.$$

L'expression (8.30) devient donc

$$\exp(\mathbf{A}) = \frac{1}{2}\begin{pmatrix}1 & 0 & 1\\0 & 0 & -2\\0 & 1 & 1\end{pmatrix}\begin{pmatrix}e & 0 & 0\\0 & e & 0\\0 & 0 & \sqrt{e}\end{pmatrix}\begin{pmatrix}2 & 1 & 0\\0 & 1 & 2\\0 & -1 & 0\end{pmatrix}$$
$$= \frac{1}{2}\begin{pmatrix}2e & e-\sqrt{e} & 0\\0 & 2\sqrt{e} & 0\\0 & e-\sqrt{e} & 2e\end{pmatrix},$$

où $e = \mathrm{e}^1$.

8.5 RÉDUCTION DES APPLICATIONS LINÉAIRES NON DIAGONALISABLES

8.5.1 Introduction

Lorsqu'une application linéaire d'un espace vectoriel de dimension finie non nulle dans lui-même n'est pas diagonalisable, il est utile de savoir s'il existe des bases dans lesquelles sa matrice est plus simple que dans d'autres. Une étude complète de cette question dépasse les objectifs que nous nous sommes fixés. C'est pourquoi nous nous limiterons à établir un résultat intermédiaire, accessible par les moyens dont nous disposons, d'où nous tirerons quelques conséquences d'intérêt théorique et pratique.

Dans cette section, E désignera un espace vectoriel de dimension finie non nulle n.

8.5.2 Théorème. Réduction à une forme simplifiée

Soit φ une application linéaire de E dans E admettant au moins une valeur propre. Soit en outre m la somme des multiplicités des valeurs propres de φ. Il existe une base de E dans laquelle la matrice de φ est de la forme

$$\left(\begin{array}{cccc:c}a_{11} & a_{12} & \cdots & a_{1m} & \\ & a_{22} & \cdots & a_{2m} & \\ & & \ddots & \vdots & \\ & 0 & & \vdots & \mathbf{B}\\ & & & a_{mm} & \\ \hdashline & & \mathbf{O} & & \end{array}\right), \qquad (8.31)$$

où $a_{11}, a_{22}, \ldots, a_{mm}$ sont les valeurs propres de φ répétées un nombre de fois égal à leur multiplicité et rangées dans un ordre arbitraire, mais choisi d'avance, et **B** *est une matrice de type $n \times (n - m)$ (manquante si $m = n$).*

DÉMONSTRATION

Nous raisonnons par récurrence sur la somme des multiplicités des valeurs propres. Lorsque cette somme est égale à 1, il suffit de choisir une base dont le premier terme est un vecteur appartenant à l'unique sous-espace propre de φ. Supposons que le théorème soit vrai pour toute application linéaire d'un espace vectoriel de dimension finie non nulle dans lui-même telle que la somme des multiplicités de ses valeurs propres soit $m - 1$ $(m > 1)$. Considérons l'application linéaire φ de l'énoncé, désignons par $\mathbf{e}_1$ un vecteur propre associé à la valeur propre a_{11} et complétons $\mathbf{e}_1$ en une base $(\mathbf{e}_1, \mathbf{e}'_2, \ldots, \mathbf{e}'_n)$ de E. Désignons la matrice de φ dans cette base par

$$\mathbf{A} = \begin{pmatrix} a_{11} & a'_{12} & a'_{13} & \ldots & a'_{1n} \\ 0 & a'_{22} & a'_{23} & \ldots & a'_{2n} \\ & & \ldots\ldots & & \\ 0 & a'_{n2} & a'_{n3} & \ldots & a'_{nn} \end{pmatrix}.$$

Désignons en outre par $\mathbf{A}'$ la matrice obtenue en supprimant la première ligne et la première colonne de $\mathbf{A}$, par S le sous-espace vectoriel engendré par $(\mathbf{e}'_2, \mathbf{e}'_3, \ldots, \mathbf{e}'_n)$ et par φ' l'application linéaire de S dans S dont la matrice dans la base $(\mathbf{e}'_2, \mathbf{e}'_3, \ldots, \mathbf{e}'_n)$ est $\mathbf{A}'$. Puisque

$$|\mathbf{A} - t\mathbf{I}_n| = (a_{11} - t)|\mathbf{A}' - t\mathbf{I}_{n-1}|,$$

$a_{22}, a_{33}, \ldots, a_{mm}$ sont les valeurs propres de φ' répétées un nombre de fois égal à leur multiplicité. Par hypothèse de récurrence, il existe une base $(\mathbf{e}_2, \mathbf{e}_3, \ldots, \mathbf{e}_n)$ de S dans laquelle la matrice de φ' est de la forme

$$\left(\begin{array}{cccc:c} a_{22} & a_{23} & \ldots & a_{2m} & \\ & a_{33} & \ldots & a_{3m} & \\ & & \ddots & \vdots & \\ 0 & & & \vdots & \mathbf{B}' \\ & & & a_{mm} & \\ \hdashline & & \mathbf{O} & & \end{array}\right), \tag{8.32}$$

où $\mathbf{B}'$ est une matrice de type $(n - 1) \times (n - m)$. Montrons que la matrice de φ dans la base $(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ est de la forme (8.31). Par définition de $\mathbf{A}$ et de φ', pour $j = 2, 3, \ldots, n$,

$$\varphi(\mathbf{e}'_j) = a'_{1j}\mathbf{e}_1 + a'_{2j}\mathbf{e}'_2 + \ldots + a'_{nj}\mathbf{e}'_n = a'_{1j}\mathbf{e}_1 + \varphi'(\mathbf{e}'_j).$$

En passant aux combinaisons linéaires des vecteurs $\mathbf{e}'_2, \mathbf{e}'_3, \ldots, \mathbf{e}'_n$, nous en déduisons que l'image de tout vecteur $\mathbf{x}$ de S s'écrit sous la forme

$$\varphi(\mathbf{x}) = \alpha \mathbf{e}_1 + \varphi'(\mathbf{x}),$$

où α est un nombre dépendant de $\mathbf{x}$. En prenant $\mathbf{x} = \mathbf{e}_j$ et en écrivant a_{1j} au lieu de α pour $j = 2, 3, \ldots, m$, nous en concluons, vu le sens de la matrice (8.32), que

$$\varphi(\mathbf{e}_j) = a_{1j}\mathbf{e}_1 + \varphi'(\mathbf{e}_j) = a_{1j}\mathbf{e}_1 + a_{2j}\mathbf{e}_2 + \ldots + a_{jj}\mathbf{e}_j,$$

ce qui prouve notre assertion.

8.5.3 Corollaire. Existence de sous-espaces stables emboîtés

Soit φ comme dans le théorème et m la somme des multiplicités des valeurs propres de φ. Il existe m sous-espaces vectoriels $S_1, S_2, \ldots, S_m$ de E, stables par φ, de dimensions respectives 1, 2, ..., *m et tels que $S_1 \subset S_2 \subset \ldots \subset S_m$.*

DÉMONSTRATION

Soit $(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ une base dans laquelle la matrice de φ est la matrice (8.31). Pour $j = 1, 2, \ldots, m$, désignons par S_j le sous-espace vectoriel engendré par la famille $(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_j)$. Il est évident que $S_1, S_2, \ldots, S_m$ satisfont aux exigences du corollaire.

8.5.4 Applications linéaires trigonalisables

On dit qu'une application linéaire φ de E dans E est *trigonalisable* s'il existe une base de E dans laquelle la matrice de φ est triangulaire.

On notera que si la matrice de φ dans une base $(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ est triangulaire supérieure, alors dans la base $(\mathbf{e}_n, \mathbf{e}_{n-1}, \ldots, \mathbf{e}_1)$ elle est triangulaire inférieure.

8.5.5 Proposition. Caractérisation des applications linéaires trigonalisables

Pour qu'une application linéaire de E dans E soit trigonalisable il faut et il suffit que les racines de son polynôme caractéristique soient toutes réelles.

DÉMONSTRATION

D'après la remarque (2) de 8.3.5, cette condition est nécessaire. Le théorème 8.5.2, appliqué au cas où $m = n$, nous montre qu'elle est suffisante.

8.5.6 Point de vue matriciel

Soit $\mathbf{A}$ une matrice carrée d'ordre n. Par le biais de l'application linéaire associée canoniquement à $\mathbf{A}$, nous déduisons du théorème 8.5.2 et de la proposition 8.5.5 les énoncés que voici:

(1) Si **A** admet au moins une valeur propre et si la somme des multiplicités des valeurs propres de **A** est m, alors **A** est semblable à une matrice de la forme (8.31).

(2) On dit que **A** est *trigonalisable* si l'application linéaire associée canoniquement à **A** est trigonalisable. Les conditions suivantes sont équivalentes:

(a) **A** est trigonalisable.
(b) **A** est semblable à une matrice triangulaire.
(c) Les racines du polynôme caractéristique de **A** sont toutes réelles.

8.5.7 Exemple d'utilisation de l'énoncé (1).

Soit **A** une matrice carrée d'ordre n. Soit en outre λ une valeur propre de **A** de multiplicité k. Nous nous proposons de démontrer que le rang de la matrice $(\mathbf{A} - \lambda\mathbf{I})^k$ est $n - k$. D'après ce qui précède, **A** est semblable à une matrice de la forme

$$\mathbf{A}' = \left(\begin{array}{c|c} \mathbf{T} & \mathbf{C} \\ \hline \mathbf{O} & \mathbf{B} \end{array}\right),$$

où **T** est une matrice triangulaire supérieure d'ordre k dont les termes diagonaux sont égaux à λ, **B** une matrice carrée d'ordre $n - k$ et **C** une matrice de type $k \times (n - k)$. Par conséquent, $(\mathbf{A} - \lambda\mathbf{I})^k$ est semblable à $(\mathbf{A}' - \lambda\mathbf{I})^k$, qui s'écrit, par la règle de multiplication par blocs,

$$(\mathbf{A}' - \lambda\mathbf{I})^k = \left(\begin{array}{c|c} (\mathbf{T}-\lambda\mathbf{I}_k)^k & \mathbf{C}' \\ \hline \mathbf{O} & (\mathbf{B}-\lambda\mathbf{I}_{n-k})^k \end{array}\right).$$

Or, d'après la remarque (4) de 4.1.7, la matrice $(\mathbf{T} - \lambda\mathbf{I}_k)^k$ est nulle. D'autre part, d'après 8.3.8, **A** et **A**′ ont le même polynôme caractéristique, à savoir $(\lambda - t)^k \,|\, \mathbf{B} - t\mathbf{I}_{n-k} \,|$ (cf. (5.24)). Par la définition de multiplicité d'une valeur propre, $|\, \mathbf{B} - \lambda\mathbf{I}_{n-k} \,|$ est donc non nul, ce qui entraîne que le rang de $(\mathbf{B} - \lambda\mathbf{I}_{n-k})^k$ est $n - k$. Nous en concluons que le rang de $(\mathbf{A}' - \lambda\mathbf{I})^k$, et donc celui de $(\mathbf{A} - \lambda\mathbf{I})^k$, est également $n - k$.

8.5.8 Marche à suivre pour trigonaliser une matrice et exemple

La marche à suivre pour trigonaliser une matrice carrée **A** d'ordre n est la suivante:

(a) Former le polynôme caractéristique de $\mathbf{A}$ et déterminer ses racines, ainsi que leurs multiplicités. S'il y a des racines complexes, $\mathbf{A}$ n'est pas trigonalisable.
(b) Chercher une base de chaque sous-espace propre de $\mathbf{A}$.
(c) Réunir ces bases en une famille de vecteurs-colonnes et prolonger cette famille en une base de $\mathbb{R}^n$ constituant une matrice $\mathbf{P}_1$.
(d) Calculer la matrice $\mathbf{A}_1$ par la formule

$$\mathbf{P}_1^{-1}\mathbf{A}\mathbf{P}_1 = \left(\begin{array}{ccc:c} d_1 & & 0 & \\ & \ddots & & \mathbf{B}_1 \\ 0 & & d_{n_1} & \\ \hdashline & \mathbf{O} & & \mathbf{A}_1 \end{array}\right).$$

(e) Appliquer les règles (b) et (c) à la matrice $\mathbf{A}_1$, tout en observant que les valeurs propres de $\mathbf{A}_1$ sont celles de $\mathbf{A}$ qui figurent un nombre de fois inférieur à leur multiplicité parmi les termes $d_1, d_2, ..., d_{n_1}$. Il en résulte une matrice carrée $\mathbf{P}_2$ d'ordre $n - n_1$ permettant de calculer une matrice carrée $\mathbf{A}_2$ par la formule

$$\mathbf{P}_2^{-1}\mathbf{A}_1\mathbf{P}_2 = \left(\begin{array}{ccc:c} d_{n_1+1} & & 0 & \\ & \ddots & & \mathbf{B}_2 \\ 0 & & d_{n_2} & \\ \hdashline & \mathbf{O} & & \mathbf{A}_2 \end{array}\right).$$

(f) Appliquer de nouveau les règles (b) et (c) aux matrices $\mathbf{A}_2, \mathbf{A}_3, ...$. A noter que les valeurs propres de $\mathbf{A}_i$ sont celles de $\mathbf{A}$ qui figurent un nombre de fois inférieur à leur multiplicité parmi les termes $d_1, d_2, ..., d_{n_i}$.
(g) Lorsque $\mathbf{P}_k^{-1}\mathbf{A}_{k-1}\mathbf{P}_k$ est une matrice triangulaire supérieure (donc au plus tard lorsque $\mathbf{A}_k$ est d'ordre 1), poser

$$\mathbf{P} = \mathbf{P}_1\begin{pmatrix} \mathbf{I}_{n_1} & \mathbf{O} \\ \mathbf{O} & \mathbf{P}_2 \end{pmatrix}\begin{pmatrix} \mathbf{I}_{n_2} & \mathbf{O} \\ \mathbf{O} & \mathbf{P}_3 \end{pmatrix}\cdots\begin{pmatrix} \mathbf{I}_{n_{k-1}} & \mathbf{O} \\ \mathbf{O} & \mathbf{P}_k \end{pmatrix}.$$

(h) $\mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ est alors une matrice triangulaire supérieure.

A titre d'exemple, nous allons trigonaliser la matrice

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

Le polynôme caractéristique de $\mathbf{A}$ décomposé en facteurs irréductibles est $(t - 2)^3(t - 4)$. Les sous-espaces propres associés aux valeurs propres 2 et 4 sont engendrés respectivement par les vecteurs-colonnes

$$\begin{bmatrix} -1 \\ 1 \\ -1 \\ 1 \end{bmatrix} \quad \text{et} \quad \begin{bmatrix} 1 \\ -1 \\ 1 \\ 1 \end{bmatrix}.$$

Posons

$$\mathbf{P}_1 = \begin{bmatrix} -1 & 1 & 1 & 0 \\ 1 & -1 & 0 & 1 \\ -1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix}.$$

Alors

$$\mathbf{P}_1^{-1} = \begin{bmatrix} 0 & 0 & -\frac{1}{2} & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 1 & 0 \end{bmatrix} \quad \text{et} \quad \mathbf{P}_1^{-1}\mathbf{A}\mathbf{P}_1 = \left[\begin{array}{cc|cc} 2 & 0 & \multicolumn{2}{c}{\mathbf{B}_1} \\ 0 & 4 & & \\ \hline \multicolumn{2}{c|}{\mathbf{O}} & 2 & 0 \\ & & -2 & 2 \end{array}\right].$$

Appliquons le procédé de réduction à la matrice

$$\mathbf{A}_1 = \begin{bmatrix} 2 & 0 \\ -2 & 2 \end{bmatrix}.$$

Comme $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ est un vecteur propre de $\mathbf{A}_1$ associé à la valeur propre 2,

$$\mathbf{P}_2^{-1}\mathbf{A}_1\mathbf{P}_2$$

est une matrice triangulaire supérieure si l'on pose

$$\mathbf{P}_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

La matrice de passage réduisant $\mathbf{A}$ à la forme triangulaire supérieure est donc

$$\mathbf{P} = \mathbf{P}_1 \left[\begin{array}{c|c} \mathbf{I}_2 & \mathbf{O} \\ \hline \mathbf{O} & \mathbf{P}_2 \end{array}\right] = \begin{bmatrix} -1 & 1 & 0 & 1 \\ 1 & -1 & 1 & 0 \\ -1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix}.$$

En effet

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \begin{bmatrix} 2 & 0 & -1 & 0 \\ 0 & 4 & -1 & 0 \\ 0 & 0 & 2 & -2 \\ 0 & 0 & 0 & 2 \end{bmatrix}.$$

On peut constater (a posteriori) qu'un seul pas aurait suffi si les deux dernières colonnes de $\mathbf{P}_1$ avaient été choisies dans l'ordre inverse.

8.6 TRANSFORMATIONS ET MATRICES SYMÉTRIQUES

8.6.1 Introduction

Cette section sera consacrée à l'étude des applications linéaires associées aux matrices symétriques par l'intermédiaire d'une base orthonormale. Il sera établi que ces applications admettent une base constituée de vecteurs propres orthogonaux deux à deux.

Tout au long de la section, E désignera un espace vectoriel euclidien de dimension finie non nulle n.

8.6.2 Transformations symétriques

On appelle *transformation symétrique* de E toute application φ de E dans E telle que

$$(\varphi(\mathbf{x}) \mid \mathbf{y}) = (\mathbf{x} \mid \varphi(\mathbf{y})) \tag{8.33}$$

pour tout couple $(\mathbf{x}, \mathbf{y})$ de vecteurs de E.

Relevons aussitôt que toute transformation symétrique φ est linéaire. En effet, pour tout couple $(\mathbf{x}, \mathbf{y})$ de vecteurs de E et tout couple (α, β) de nombres,

$$\begin{aligned}
&(\varphi(\alpha\mathbf{x} + \beta\mathbf{y}) - \alpha\varphi(\mathbf{x}) - \beta\varphi(\mathbf{y}) \mid \mathbf{z}) \\
&= (\varphi(\alpha\mathbf{x} + \beta\mathbf{y}) \mid \mathbf{z}) - \alpha(\varphi(\mathbf{x}) \mid \mathbf{z}) - \beta(\varphi(\mathbf{y}) \mid \mathbf{z}) \\
&= (\alpha\mathbf{x} + \beta\mathbf{y} \mid \varphi(\mathbf{z})) - \alpha(\mathbf{x} \mid \varphi(\mathbf{z})) - \beta(\mathbf{y} \mid \varphi(\mathbf{z})) = 0,
\end{aligned}$$

quel que soit le vecteur $\mathbf{z}$ de E, ce qui nous montre que le vecteur $\varphi(\alpha\mathbf{x} + \beta\mathbf{y}) - \alpha\varphi(\mathbf{x}) - \beta\varphi(\mathbf{y})$ est orthogonal à tout vecteur $\mathbf{z}$ de E, donc qu'il est nul, autrement dit, que $\varphi(\alpha\mathbf{x} + \beta\mathbf{y}) = \alpha\varphi(\mathbf{x}) + \beta\varphi(\mathbf{y})$.

8.6.3 Proposition. Caractérisation des transformations symétriques

La matrice de toute transformation symétrique φ dans une base orthonormale de E arbitrairement choisie est symétrique. Réciproquement, si φ est une application linéaire de E dans E et s'il existe une base orthonormale de E dans laquelle la matrice de φ est symétrique, alors φ est une transformation symétrique.

DÉMONSTRATION

Soit φ une application linéaire de E dans E. Désignons par $\mathbf{A}$ la matrice de φ dans une base orthonormale $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ de E. Vu que

$$(\varphi(\mathbf{x}))_e = \mathbf{A}\mathbf{x}_e \quad \text{et} \quad (\varphi(\mathbf{y}))_e = \mathbf{A}\mathbf{y}_e,$$

la relation (8.33) s'écrit, de manière équivalente, sous l'une des deux formes

$${}^t(\mathbf{A}\mathbf{x}_e)\mathbf{y}_e = {}^t\mathbf{x}_e(\mathbf{A}\mathbf{y}_e)$$

ou

$$^t\mathbf{x}_e(^t\mathbf{A} - \mathbf{A})\mathbf{y}_e = 0.$$

Comme $\mathbf{x}$ et $\mathbf{y}$ sont des vecteurs quelconques, cette relation est vraie si et seulement si $^t\mathbf{A} = \mathbf{A}$, ce qui démontre la proposition.

8.6.4 Exemples

Toute application linéaire φ de E dans E admettant une base constituée de vecteurs propres orthogonaux deux à deux est une transformation symétrique. En effet, la matrice de φ dans une telle base rendue orthonormale (pour pouvoir appliquer la proposition 8.6.3) est diagonale, donc symétrique. Nous verrons ci-après (théorème 8.6.6) qu'il n'y a pas d'autres transformations symétriques.

En particulier, les dilatations orthogonales, donc les projections et les symétries orthogonales, sont des transformations symétriques.

8.6.5 Proposition. Orthogonalité des sous-espaces propres d'une transformation symétrique

Deux vecteurs propres $\mathbf{x}$ *et* $\mathbf{y}$ *d'une transformation symétrique, ou d'une matrice symétrique, associés à des valeurs propres distinctes* λ *et* μ, *sont orthogonaux.*

DÉMONSTRATION

Désignons par φ soit une transformation symétrique, soit l'application linéaire (qui est également une transformation symétrique, d'après la proposition 8.6.3) associée canoniquement à une matrice symétrique. Si $\mathbf{x}$ et $\mathbf{y}$ sont des vecteurs propres de φ associés respectivement aux valeurs propres λ et μ, alors

$$\lambda(\mathbf{x} \mid \mathbf{y}) = (\lambda\mathbf{x} \mid \mathbf{y}) = (\varphi(\mathbf{x}) \mid \mathbf{y}) = (\mathbf{x} \mid \varphi(\mathbf{y})) = (\mathbf{x} \mid \mu\mathbf{y}) = \mu(\mathbf{x} \mid \mathbf{y}),$$

où l'égalité centrale est justifiée par (8.33). Il s'ensuit que

$$(\lambda - \mu)(\mathbf{x} \mid \mathbf{y}) = 0,$$

donc que $(\mathbf{x} \mid \mathbf{y}) = 0$ si $\lambda \neq \mu$.

8.6.6 Théorème. Diagonalisation des transformations symétriques

Si une application linéaire φ *de* E *dans* E *satisfait à l'une des trois conditions suivantes, alors elle satisfait aux deux autres.*

(a) φ *est une transformation symétrique.*

(b) E *admet une base constituée de vecteurs propres de* φ *orthogonaux deux à deux.*

(c) φ *est diagonalisable et ses sous-espaces propres sont orthogonaux deux à deux.*

DÉMONSTRATION

(a) entraîne (b). Nous raisonnons par récurrence sur la dimension de l'espace vectoriel euclidien. Lorsque cette dimension est 1, l'assertion est évidente. Supposons qu'elle soit vraie pour tout espace vectoriel euclidien de dimension $n - 1$ ($n > 1$). Soit f la fonction réelle définie par

$$f(\mathbf{x}) = \frac{(\varphi(\mathbf{x}) \mid \mathbf{x})}{(\mathbf{x} \mid \mathbf{x})},$$

$\mathbf{x}$ parcourant les vecteurs non nuls de E. Il est évident que $f(\alpha\mathbf{x}) = f(\mathbf{x})$ pour tout nombre non nul α. En outre, puisque f est une fonction continue, elle atteint sa borne inférieure sur la sphère unité $\{\mathbf{x} : \| \mathbf{x} \| = 1\}$. Soit $\mathbf{e}$ un vecteur où cette borne est atteinte. Alors, pour tout vecteur non nul $\mathbf{x}$ de E,

$$f(\mathbf{x}) = f\left(\frac{\mathbf{x}}{\| \mathbf{x} \|}\right) \geqslant f(\mathbf{e}). \tag{8.34}$$

Soit $\mathbf{y}$ un vecteur de E arbitrairement choisi et g la fonction de la variable réelle t définie, dans un voisinage de 0, par

$$g(t) = f(\mathbf{e} + t\mathbf{y}).$$

D'après (8.34),

$$g(t) \geqslant g(0)$$

pour tout t de ce voisinage, donc la dérivée de g au point 0 est nulle. Or, un calcul facile que nous omettrons montre que

$$\dot{g}(0) = (\varphi(\mathbf{y}) \mid \mathbf{e}) + (\varphi(\mathbf{e}) \mid \mathbf{y}) - 2(\varphi(\mathbf{e}) \mid \mathbf{e})(\mathbf{e} \mid \mathbf{y}).$$

D'autre part, grâce à (8.33), l'expression au second membre peut s'écrire

$$2(\varphi(\mathbf{e}) \mid \mathbf{y}) - 2(\varphi(\mathbf{e}) \mid \mathbf{e})(\mathbf{e} \mid \mathbf{y}) = 2(\varphi(\mathbf{e}) - (\varphi(\mathbf{e}) \mid \mathbf{e})\mathbf{e} \mid \mathbf{y}).$$

La dérivée $\dot{g}(0)$ étant nulle et le vecteur $\mathbf{y}$ quelconque, nous en déduisons que

$$\varphi(\mathbf{e}) - (\varphi(\mathbf{e}) \mid \mathbf{e})\mathbf{e} = \mathbf{0}.$$

En d'autres termes, $\lambda = (\varphi(\mathbf{e}) \mid \mathbf{e})$ est une valeur propre de φ et $\mathbf{e}$ est un vecteur propre associé à cette valeur propre. Désignons par S le complémentaire orthogonal dans E de la droite vectorielle engendrée par $\mathbf{e}$. Pour tout vecteur $\mathbf{x}$ de S, encore par (8.33),

$$(\varphi(\mathbf{x}) \mid \mathbf{e}) = (\mathbf{x} \mid \varphi(\mathbf{e})) = \lambda(\mathbf{x} \mid \mathbf{e}) = 0,$$

ce qui prouve que $\varphi(\mathbf{x})$ appartient à S, donc que S est un sous-espace vectoriel stable. Il s'ensuit que la restriction φ' de φ à S est une transformation symétrique de S. Comme S est de dimension $n - 1$, il existe, par hypothèse de récurrence, une base de S constituée de vecteurs propres de φ', donc de φ, orthogonaux deux à deux. En adjoignant à cette base le vecteur $\mathbf{e}$, nous obtenons une base de E constituée de vecteurs propres de φ orthogonaux deux à deux.

(b) entraîne (a). Cette implication a été démontrée dans l'exemple 8.6.4.

(b) entraîne (c). D'après le théorème 8.4.1, φ est diagonalisable. En outre, les sous-espaces propres de φ sont orthogonaux deux à deux, puisque chacun d'entre eux est engendré par une famille de vecteurs extraite d'une base orthogonale de E.

(c) entraîne (b). Il suffit de choisir une base orthogonale de chaque sous-espace propre et de réunir ces bases en une base de E.

8.6.7 Décomposition spectrale d'une transformation symétrique

Toute transformation symétrique φ de E admet la décomposition spectrale (cf. 8.2.9)

$$\varphi = \lambda_1\varphi_1 + \lambda_2\varphi_2 + \dots + \lambda_l\varphi_l,$$

où λ_1, λ_2, ..., λ_l sont les valeurs propres de φ et φ_1, φ_2, ..., φ_l les projections orthogonales sur les sous-espaces propres associés respectivement à ces valeurs propres.

8.6.8 Diagonalisation des matrices symétriques

Soit $\mathbf{A}$ une matrice symétrique d'ordre n. En prenant pour φ l'application linéaire associée canoniquement à $\mathbf{A}$, du théorème 8.6.6 nous déduisons les deux conclusions équivalentes suivantes:

(a) $\mathbf{A}$ admet n vecteurs propres orthogonaux deux à deux.
(b) Il existe une matrice orthogonale $\mathbf{P}$ et une matrice diagonale $\mathbf{D}$ telles que

$$\mathbf{D} = {}^t\mathbf{P}\mathbf{A}\mathbf{P} \quad \text{ou} \quad \mathbf{A} = \mathbf{P}\mathbf{D}{}^t\mathbf{P}. \tag{8.35}$$

La marche à suivre pour diagonaliser une matrice symétrique est celle qui a été décrite dans 8.4.4. Il faut seulement noter que la condition (d) de 8.4.3 est automatiquement remplie et qu'il y a avantage à choisir les bases des sous-espaces propres orthonormales, car la matrice de passage $\mathbf{P}$ sera alors orthogonale et, de ce fait, $\mathbf{P}^{-1}$ pourra être remplacée par ${}^t\mathbf{P}$.

8.6.9 Estimation des valeurs propres extrêmes

Soit φ une transformation symétrique de E. D'après la démonstration du théorème 8.6.6, le nombre

$$\lambda_{\min} = \min_{\mathbf{x}\neq\mathbf{0}} \frac{(\varphi(\mathbf{x}) \mid \mathbf{x})}{(\mathbf{x} \mid \mathbf{x})}$$

est une valeur propre de φ et chaque vecteur $\mathbf{x}$ où ce minimum est atteint est un vecteur propre associé à $\lambda_{\min}$. En prenant $-\varphi$ à la place de φ, nous en déduisons, grâce à la remarque (4) de 8.2.3, que le nombre

$$\lambda_{\max} = \max_{\mathbf{x}\neq\mathbf{0}} \frac{(\varphi(\mathbf{x}) \mid \mathbf{x})}{(\mathbf{x} \mid \mathbf{x})}$$

est également une valeur propre de φ et chaque vecteur $\mathbf{x}$ où ce maximum est atteint est un vecteur propre associé à $\lambda_{\max}$.

Nous allons maintenant démontrer que $\lambda_{\min}$ est la plus petite valeur propre de φ et $\lambda_{\max}$ la plus grande. Nous nous limitons à considérer $\lambda_{\min}$, car pour $\lambda_{\max}$ le raisonnement est analogue. Soit λ une valeur propre de φ et $\mathbf{y}$ un vecteur propre associé à λ. Par définition de minimum,

$$\lambda_{\min} = \min_{\mathbf{x}\neq\mathbf{0}} \frac{(\varphi(\mathbf{x}) \mid \mathbf{x})}{(\mathbf{x} \mid \mathbf{x})} \leqslant \frac{(\varphi(\mathbf{y}) \mid \mathbf{y})}{(\mathbf{y} \mid \mathbf{y})} = \frac{(\lambda\mathbf{y} \mid \mathbf{y})}{(\mathbf{y} \mid \mathbf{y})} = \lambda,$$

ce qui démontre que $\lambda_{\min} \leqslant \lambda$.

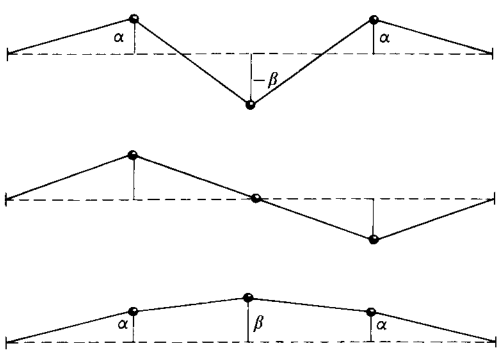
Les résultats auxquels nous sommes parvenus s'avèrent utiles, dans la pratique, pour calculer des estimations de $\lambda_{\min}$ et $\lambda_{\max}$. En effet, on obtient souvent de bonnes évaluations de ces deux valeurs propres, en calculant la valeur du quotient

$$\frac{(\varphi(\mathbf{x}) \mid \mathbf{x})}{(\mathbf{x} \mid \mathbf{x})} \tag{8.36}$$

pour des vecteurs $\mathbf{x}$ particuliers suggérés par le type de problème à résoudre. Par exemple,

$$\ddot{\mathbf{x}} = \mathbf{A}\mathbf{x}, \quad \text{avec} \quad \mathbf{A} = \begin{bmatrix} -2 & 1 & 0 \\ 1 & -2 & 1 \\ 0 & 1 & -2 \end{bmatrix},$$

est l'équation différentielle matricielle régissant le mouvement de trois masses égales attachées aux points de quadrisection d'une corde élastique fixée à ses extrémités. Comme dans l'exemple 8.1.2, les fréquences propres du mouvement sont définies par $\sqrt{-\lambda}$, λ parcourant le spectre de $\mathbf{A}$, et les amplitudes correspondantes par les vecteurs propres associés (cf. (8.9)). Intuitivement, les trois modes d'oscillation harmonique sont représentés dans la figure 8.4. Pour estimer $\lambda_{\min}$ et $\lambda_{\max}$, on calcule la valeur du quotient (8.36) pour des vecteurs $\mathbf{x}$ de la forme indiquée dans cette figure.



fréquence maximale définie par $(-\lambda)_{\max} = -\lambda_{\min}$

vecteur propre associé à $\lambda_{\min}$ $\cong \begin{bmatrix} \alpha \\ -\beta \\ \alpha \end{bmatrix}$ $0 < \alpha < \beta$

fréquence intermédiaire

fréquence minimale définie par $(-\lambda)_{\min} = -\lambda_{\max}$

vecteur propre associé à $\lambda_{\max}$ $\cong \begin{bmatrix} \alpha \\ \beta \\ \alpha \end{bmatrix}$ $0 < \alpha < \beta$

Fig. 8.4

Par exemple,

$$\text{pour } \mathbf{x} = \begin{bmatrix} 2 \\ -3 \\ 2 \end{bmatrix} \left(\begin{bmatrix} 5 \\ -7 \\ 5 \end{bmatrix} \right), \ \lambda_{\min} \cong \frac{(\mathbf{Ax} \mid \mathbf{x})}{(\mathbf{x} \mid \mathbf{x})} = -3{,}411... \ (-3{,}414...);$$

$$\text{pour } \mathbf{x} = \begin{bmatrix} 2 \\ 3 \\ 2 \end{bmatrix} \left(\begin{bmatrix} 5 \\ 7 \\ 5 \end{bmatrix} \right), \ \lambda_{\max} \cong \frac{(\mathbf{Ax} \mid \mathbf{x})}{(\mathbf{x} \mid \mathbf{x})} = -0{,}588... \ (-0{,}585...).$$

Les valeurs exactes sont:

$$\lambda_{\min} = -\sqrt{2} - 2 = -3{,}414..., \quad \lambda_{\max} = \sqrt{2} - 2 = -0{,}585...$$

Les sous-espaces propres associés sont les droites vectorielles engendrées par

$$\begin{bmatrix} 1 \\ -\sqrt{2} \\ 1 \end{bmatrix} \quad \text{et} \quad \begin{bmatrix} 1 \\ \sqrt{2} \\ 1 \end{bmatrix} \quad \text{(cf. exercice 8.8.25).}$$

8.7 APPLICATION AUX SYSTÈMES DIFFÉRENTIELS

8.7.1 But

Nous nous proposons d'appliquer la notion de vecteur propre à la résolution des *systèmes différentiels linéaires du premier ordre à coefficients constants.* Par un tel système, on entend une famille d'équations différentielles de la forme

$$\begin{aligned} \dot{x}_1 &= a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n + b_1 \\ \dot{x}_2 &= a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n + b_2 \\ &\ldots\ldots\ldots\ldots \\ \dot{x}_n &= a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n + b_n, \end{aligned} \tag{8.37}$$

où $x_1, x_2, ..., x_n$ désignent des fonctions inconnues de $C^1_{\mathbb{R}}$, $\dot{x}_1, \dot{x}_2, ..., \dot{x}_n$ leurs dérivées respectives, $a_{11}, a_{12}, ..., a_{nn}$ des constantes (réelles), appelées *coefficients du système*, et $b_1, b_2, ..., b_n$ des fonctions de $C_{\mathbb{R}}$.

Posons

$$\mathbf{A} = (a_{ij}), \ \mathbf{b} = (b_j), \ \mathbf{x} = (x_j) \ \text{et} \ \dot{\mathbf{x}} = (\dot{x}_j).$$

Le système différentiel (8.37) s'écrit alors sous la forme d'une équation différentielle matricielle:

$$\dot{\mathbf{x}} = \mathbf{Ax} + \mathbf{b} \quad (\text{ou } \dot{\mathbf{x}}(t) \equiv \mathbf{Ax}(t) + \mathbf{b}(t)). \tag{8.38}$$

On appelle *solution* de cette équation (ou du système (8.37)) toute fonction à valeurs vectorielles $\mathbf{x} = (x_j)$ vérifiant (8.38).

8.7.2 Conversion d'une équation différentielle linéaire d'ordre n en un système différentiel linéaire du premier ordre

Considérons l'équation différentielle linéaire d'ordre n

$$a_n x^{(n)} + a_{n-1}x^{(n-1)} + \ldots + a_1\dot{x} + a_0 x + b = 0 \quad (a_n \neq 0) \tag{8.39}$$

et posons

$$x_1 = x,\ x_2 = \dot{x},\ \ldots,\ x_n = x^{(n-1)}.$$

Le système différentiel

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= x_3 \\ &\cdots\cdots\cdots\cdots \\ \dot{x}_n &= -\frac{a_0}{a_n}x_1 - \frac{a_1}{a_n}x_2 - \ \ldots\ - \frac{a_{n-1}}{a_n}x_n - \frac{b}{a_n} \end{aligned}$$

est alors équivalent à l'équation (8.39).

8.7.3 Problèmes aux valeurs initiales

Soit $\mathbf{v}$ un vecteur-colonne quelconque de $\mathbb{R}^n$. Pour tout point t_0 de $\mathbb{R}$, il existe une solution $\mathbf{x}$ et une seule de l'équation (8.38) satisfaisant à la condition initiale

$$\mathbf{x}(t_0) = \mathbf{v}.$$

Ce résultat d'analyse est énoncé sans démonstration.

8.7.4 Structure de l'ensemble des solutions

Lorsque la fonction à valeurs vectorielles $\mathbf{b}$ est nulle, on dit que l'équation (8.38) (ou le système (8.37)) est *homogène*. Dans ce cas, toute combinaison linéaire de solutions est encore une solution et il existe au moins une solution, à savoir la solution nulle. En d'autres termes, l'ensemble S des solutions est un sous-espace vectoriel de l'espace vectoriel E constitué des fonctions à valeurs vectorielles continûment dérivables. Montrons que S est de dimension n. A cet effet, désignons par $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ les solutions satisfaisant aux conditions initiales respectives

$$\mathbf{x}_1(0) = \mathbf{e}_1,\ \mathbf{x}_2(0) = \mathbf{e}_2,\ \ldots,\ \mathbf{x}_n(0) = \mathbf{e}_n, \tag{8.40}$$

où $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ sont les vecteurs de la base canonique de $\mathbb{R}^n$. Ces solutions sont linéairement indépendantes, puisque leurs valeurs respectives au point $t = 0$ le sont. En outre, elles engendrent S, car toute solution $\mathbf{x}$ s'écrit sous la forme

$$\mathbf{x} = \alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 + \ldots + \alpha_n\mathbf{x}_n, \tag{8.41}$$

où $\alpha_1, \alpha_2, ..., \alpha_n$ sont les termes du vecteur-colonne $\mathbf{x}(0)$. En effet, la relation (8.41) est vraie pour $t = 0$, donc pour tout t, en raison du résultat d'unicité énoncé dans 8.7.3. Nous en concluons que $(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_n)$ est une base de S et, par suite, que S est de dimension n.

Lorsque la fonction à valeurs vectorielles $\mathbf{b}$ n'est pas nulle, en raisonnant comme dans 3.5.5, nous voyons que l'ensemble des solutions de l'équation (8.38) (ou du système (8.37)) est un sous-espace affine de E de dimension n, plus exactement le sous-espace affine

$$\mathbf{x}_0 + S,$$

où $\mathbf{x}_0$ est une solution particulière de l'équation (8.38) et S le sous-espace vectoriel des solutions de l'*équation homogène associée*, c'est-à-dire de l'équation obtenue en posant $\mathbf{b} = \mathbf{0}$ dans (8.38).

8.7.5 Recherche d'une base de solutions de l'équation homogène associée

Nous n'examinerons que le cas particulier où les racines du polynôme caractéristique de $\mathbf{A}$ sont toutes réelles (le cas général sera traité dans A.4.6).

Supposons d'abord que $\mathbf{A}$ admette n vecteurs propres linéairement indépendants $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n$. Soit $\lambda_1, \lambda_2, ..., \lambda_n$ les valeurs propres correspondantes (non nécessairement différentes entre elles). Les n fonctions à valeurs vectorielles $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_n$ définies par

$$\mathbf{x}_1(t) \equiv e^{\lambda_1 t}\mathbf{v}_1, \quad \mathbf{x}_2(t) \equiv e^{\lambda_2 t}\mathbf{v}_2, \quad ..., \quad \mathbf{x}_n(t) \equiv e^{\lambda_n t}\mathbf{v}_n \tag{8.42}$$

constituent une base de S. En effet, elles sont linéairement indépendantes, puisque leurs valeurs respectives au point $t = 0$ le sont, et elles vérifient l'équation $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$, car, pour $i = 1, 2, ..., n$,

$$\dot{\mathbf{x}}_i(t) \equiv \lambda_i e^{\lambda_i t}\mathbf{v}_i \equiv e^{\lambda_i t}\mathbf{A}\mathbf{v}_i \equiv \mathbf{A}(e^{\lambda_i t}\mathbf{v}_i) \equiv \mathbf{A}\mathbf{x}_i(t).$$

Supposons maintenant que $\mathbf{A}$ n'admette pas n vecteurs propres linéairement indépendants, autrement dit, qu'il y ait des valeurs propres dont la multiplicité est supérieure à la dimension du sous-espace propre associé. Soit λ une telle valeur propre et k sa multiplicité. On peut encore expliciter k solutions linéairement indépendantes concourant à la constitution d'une base de S. Ces solutions sont définies, pour $i = 1, 2, ..., k$, par la formule

$$\begin{aligned}\mathbf{x}_i(t) \equiv e^{\lambda t}\Big(\mathbf{I} + t(\mathbf{A} - \lambda\mathbf{I}) + &\\ \frac{t^2}{2!}(\mathbf{A} - \lambda\mathbf{I})^2 + ... + \frac{t^{k-1}}{(k-1)!}(\mathbf{A} - \lambda\mathbf{I})^{k-1}\Big)\mathbf{v}_i,&\end{aligned} \tag{8.43}$$

où $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k$ sont k solutions linéairement indépendantes de l'équation

$$(\mathbf{A} - \lambda\mathbf{I})^k\mathbf{v} = \mathbf{0}. \tag{8.44}$$

Rappelons que le rang de $(\mathbf{A} - \lambda\mathbf{I})^k$ est $n - k$, d'après 8.5.7, donc que la dimension de l'espace des solutions de (8.44) est k, d'après la proposition 3.5.3.

On peut démontrer par un calcul direct que les fonctions $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k$ définies par (8.43) sont effectivement des solutions de l'équation $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$. Nous omettons toutefois ce calcul, car nous effectuerons la vérification par d'autres moyens dans le paragraphe suivant. Ces solutions sont linéairement indépendantes, puisque leurs valeurs respectives au point $t = 0$ le sont. Nous laissons au lecteur la tâche de démontrer que la réunion des familles de solutions issues des différentes valeurs propres de $\mathbf{A}$ constitue une famille libre, donc une base de S.

Par exemple, lorsque λ est une valeur propre double de $\mathbf{A}$ et le sous-espace propre associé à λ est de dimension 1, deux solutions linéairement indépendantes de l'équation $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$ sont définies par

$$\mathbf{x}_1(t) \equiv e^{\lambda t}\mathbf{v}_1,$$
$$\mathbf{x}_2(t) \equiv e^{\lambda t}(\mathbf{I} + t(\mathbf{A} - \lambda\mathbf{I}))\mathbf{v}_2,$$

où $\mathbf{v}_1$ est un vecteur propre de $\mathbf{A}$, donc une solution de l'équation (8.44) (avec $k = 2$), et $\mathbf{v}_2$ une deuxième solution de cette équation non multiple de $\mathbf{v}_1$. On appelle parfois la solution $\mathbf{x}_2$ *solution additionnelle.*

8.7.6 Quelques compléments

Soit $\mathbf{v}$ un vecteur-colonne quelconque de $\mathbb{R}^n$. La fonction à valeurs vectorielles définie par

$$\mathbf{x}(t) \equiv \exp(t\mathbf{A})\mathbf{v} \tag{8.45}$$

est une solution de l'équation $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$. En effet, d'après (4.29),

$$\dot{\mathbf{x}}(t) \equiv \mathbf{A}\exp(t\mathbf{A})\mathbf{v} \equiv \mathbf{A}\mathbf{x}(t). \tag{8.46}$$

Cette solution satisfait à la condition initiale $\mathbf{x}(0) = \mathbf{v}$. En particulier, lorsque $\mathbf{v}$ parcourt les vecteurs-colonnes $\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n$ de la base canonique de $\mathbb{R}^n$, la formule (8.45) fournit les solutions $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_n$ constituant la base de S exhibée dans 8.7.4. Du fait que $\exp(t\mathbf{A})\mathbf{e}_j$ est la j-ième colonne de $\exp(t\mathbf{A})$, nous constatons que ces solutions sont les colonnes de $\exp(t\mathbf{A})$.

Il en découle que toute autre base $(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_n)$ de S est liée à $\exp(t\mathbf{A})$ par la relation

$$\exp(t\mathbf{A}) \equiv \mathbf{X}(t)\mathbf{X}^{-1}(0), \tag{8.47}$$

où $\mathbf{X}(t)$ est la matrice carrée dont les colonnes sont $\mathbf{x}_1(t), \mathbf{x}_2(t), ..., \mathbf{x}_n(t)$. En effet, les colonnes de $\mathbf{X}(t)\mathbf{X}^{-1}(0)$ sont des combinaisons linéaires des colonnes de $\mathbf{X}(t)$, donc des solutions de l'équation $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$; en outre, pour $t = 0$, ces colonnes se réduisent respectivement à $\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n$, ce qui nous permet de conclure que (8.47) est vraie, en raison du résultat d'unicité énoncé dans 8.7.3.

On notera que lorsque $\mathbf{A}$ admet n vecteurs propres linéairement indépendants et que les colonnes de $\mathbf{X}(t)$ sont les solutions définies dans (8.42), la relation (8.47) n'est autre que la relation (8.30) appliquée à la matrice $t\mathbf{A}$. En effet, dans ce cas,

$$\mathbf{X}(t) \equiv \mathbf{X}(0)\operatorname{diag}(e^{\lambda_1 t}, e^{\lambda_2 t}, ..., e^{\lambda_n t})\mathbf{X}^{-1}(0).$$

Montrons maintenant que les fonctions définies par (8.43) sont des solutions de l'équation $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$. Comme $\mathbf{v}_i$ vérifie l'équation (8.44),

$$(\mathbf{A} - \lambda\mathbf{I})^l\mathbf{v}_i = \mathbf{0}$$

pour tout entier $l \geqslant k$. Grâce à (4.26) et à (4.27), la formule (8.43) peut donc s'écrire sous la forme

$$\mathbf{x}_i(t) \equiv e^{\lambda t}\exp(t(\mathbf{A} - \lambda\mathbf{I}))\mathbf{v}_i \equiv \exp(t\mathbf{A})\mathbf{v}_i.$$

D'après (8.46), $\mathbf{x}_i$ est donc une solution de l'équation $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$.

8.7.7 Recherche d'une solution particulière de l'équation (8.38)

Posons

$$\mathbf{x}(t) \equiv \exp(t\mathbf{A})\mathbf{v}(t),$$

où $\mathbf{v}$ désigne momentanément une fonction inconnue à valeurs dans $\mathbb{R}^n$ (variation de la constante). Nous nous proposons de déterminer $\mathbf{v}$ de telle manière que $\mathbf{x}$ soit une solution de l'équation (8.38) et satisfasse à la condition initiale $\mathbf{x}(0) = \mathbf{c}$, qui équivaut à la condition $\mathbf{v}(0) = \mathbf{c}$. L'insertion de l'expression de $\mathbf{x}(t)$ dans (8.38) conduit à l'équation

$$\mathbf{A}\exp(t\mathbf{A})\mathbf{v}(t) + \exp(t\mathbf{A})\dot{\mathbf{v}}(t) \equiv \mathbf{A}\exp(t\mathbf{A})\mathbf{v}(t) + \mathbf{b}(t).$$

Vu (4.28), cette équation se réduit à

$$\dot{\mathbf{v}}(t) \equiv \exp(-t\mathbf{A})\mathbf{b}(t),$$

et entraîne donc

$$\mathbf{v}(t) - \mathbf{v}(0) \equiv \int_0^t \exp(-s\mathbf{A})\mathbf{b}(s)\mathrm{d}s.$$

Par conséquent, la solution $\mathbf{x}$ de (8.38) qui satisfait à la condition initiale $\mathbf{x}(0) = \mathbf{c}$ est

$$\mathbf{x}(t) \equiv \exp(t\mathbf{A})\Big(\int_0^t \exp(-s\mathbf{A})\mathbf{b}(s)\mathrm{d}s + \mathbf{c}\Big). \tag{8.48}$$

8.7.8 Résolution par trigonalisation

Nous allons décrire une autre méthode de résolution de l'équation (8.38). Vu l'hypothèse faite au début de 8.7.5, il existe, d'après (2) de 8.5.6, une matrice inversible $\mathbf{P}$ telle que

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \tilde{\mathbf{A}} = \begin{pmatrix} \tilde{a}_{11} & \tilde{a}_{12} & \dots & \tilde{a}_{1n} \\ 0 & \tilde{a}_{22} & \dots & \tilde{a}_{2n} \\ & & \dots\dots & \\ 0 & 0 & \dots & \tilde{a}_{nn} \end{pmatrix}.$$

Par le changement de variable

$$\mathbf{x} = \mathbf{P}\mathbf{y},$$

l'équation (8.38) se transforme en

$$\mathbf{P}\dot{\mathbf{y}} = \mathbf{A}\mathbf{P}\mathbf{y} + \mathbf{b},$$

ou encore

$$\dot{\mathbf{y}} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}\mathbf{y} + \mathbf{P}^{-1}\mathbf{b} = \tilde{\mathbf{A}}\mathbf{y} + \tilde{\mathbf{b}}.$$

Cette équation est équivalente au système différentiel

$$\begin{aligned}
\dot{y}_1 &= \tilde{a}_{11}y_1 + \tilde{a}_{12}y_2 + \dots + \tilde{a}_{1n}y_n + \tilde{b}_1\\
\dot{y}_2 &= \qquad\quad\ \tilde{a}_{22}y_2 + \dots + \tilde{a}_{2n}y_n + \tilde{b}_2\\
&\dots\dots\dots\dots\\
\dot{y}_n &= \qquad\qquad\qquad\qquad\ \ \tilde{a}_{nn}y_n + \tilde{b}_n.
\end{aligned}$$

Les fonctions inconnues $y_1, y_2, \dots, y_n$ peuvent être déterminées dans l'ordre inverse, en résolvant chaque équation du système, de la dernière à la première, par substitutions successives. La solution de l'équation (8.38) sera alors $\mathbf{x} = \mathbf{P}\mathbf{y}$.

Si aucune condition initiale n'est posée, cette solution renferme n paramètres réels et doit être considérée comme solution générale de (8.38).

En présence d'une condition initiale $\mathbf{x}(0) = \mathbf{c}$, la manière la plus simple de procéder consiste à déterminer successivement les fonctions $y_n, y_{n-1}, \dots, y_1$ de telle façon que y_i satisfasse à la condition initiale $y_i(0) = \tilde{c}_i$, où $\tilde{c}_i$ est le i-ième terme de $\tilde{\mathbf{c}} = \mathbf{P}^{-1}\mathbf{c}$.

8.7.9 Exemple

Nous nous proposons de résoudre le problème aux valeurs initiales

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}, \quad \mathbf{x}(0) = \mathbf{c},$$

où

$$\mathbf{A} = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 1 & -1 \\ 0 & 2 & 4 \end{bmatrix} \quad \text{et} \quad \mathbf{c} = \begin{bmatrix} 1 \\ 2 \\ -3 \end{bmatrix}.$$

Les valeurs propres de $\mathbf{A}$ sont 3 et 2 (double). Les sous-espaces propres associés à ces deux valeurs propres sont engendrés respectivement par les vecteurs-colonnes

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix} \quad \text{et} \quad \mathbf{v}_2 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}.$$

La multiplicité de la valeur propre 2 étant 2 et la dimension du sous-espace propre qui lui est associé 1, il nous faut calculer une solution additionnelle. A cet effet, cherchons une solution $\mathbf{v}_3$ de l'équation

$$(\mathbf{A} - 2\mathbf{I})^2\mathbf{v} = \begin{bmatrix} 0 & -1 & -1 \\ 0 & -1 & -1 \\ 0 & 2 & 2 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \mathbf{0}$$

qui n'est pas multiple de $\mathbf{v}_2$. Il est évident que le vecteur-colonne

$$\mathbf{v}_3 = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$$

convient, de sorte que la solution générale de l'équation $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$ s'écrit (vu (8.43)) sous la forme

$$\begin{aligned} \mathbf{x}(t) &\equiv \alpha_1 e^{3t}\mathbf{v}_1 + \alpha_2 e^{2t}\mathbf{v}_2 + \alpha_3 e^{2t}(\mathbf{v}_3 + t(\mathbf{A} - 2\mathbf{I})\mathbf{v}_3) \\ &\equiv \alpha_1 e^{3t}\begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix} + \alpha_2 e^{2t}\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \alpha_3 e^{2t}\begin{bmatrix} t \\ 1 \\ -1 \end{bmatrix}, \end{aligned} \tag{8.49}$$

où α_1, α_2, α_3 sont des constantes réelles arbitrairement choisies. Pour compléter la résolution, il nous reste à déterminer les valeurs de α_1, α_2, α_3, de sorte que la condition initiale soit remplie. Cette condition se traduit par l'équation

$$\begin{bmatrix} 1 \\ 2 \\ -3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ -2 & 0 & -1 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix},$$

d'où l'on tire $\alpha_1 = 1$, $\alpha_2 = 0$, $\alpha_3 = 1$. La solution cherchée est donc

$$\mathbf{x}(t) \equiv \begin{bmatrix} e^{3t} + te^{2t} \\ e^{3t} + e^{2t} \\ -2e^{3t} - e^{2t} \end{bmatrix}.$$

8.7.10 Exemple

Considérons le problème aux valeurs initiales

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{b}, \quad \mathbf{x}(0) = \mathbf{c},$$

où

$$\mathbf{A} = \begin{bmatrix} 4 & 5 \\ -2 & -2 \end{bmatrix}, \quad \mathbf{b}(t) \equiv \begin{bmatrix} e^t\cos t \\ 0 \end{bmatrix} \quad \text{et} \quad \mathbf{c} = \frac{1}{4}\begin{bmatrix} 3 \\ -2 \end{bmatrix}.$$

Le lecteur constatera aisément que les deux fonctions à valeurs vectorielles $\mathbf{x}_1$ et $\mathbf{x}_2$ définies par

$$\begin{aligned} \mathbf{x}_1(t) &\equiv e^t(\cos t\begin{bmatrix} -5 \\ 3 \end{bmatrix} - \sin t\begin{bmatrix} 0 \\ -1 \end{bmatrix}), \\ \mathbf{x}_2(t) &\equiv e^t(\sin t\begin{bmatrix} -5 \\ 3 \end{bmatrix} + \cos t\begin{bmatrix} 0 \\ -1 \end{bmatrix}), \end{aligned} \tag{8.50}$$

sont des solutions de l'équation $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$ (nous montrerons dans A.4.6 comment elles ont été calculées). Ces deux solutions sont linéairement indépendantes, puisque leurs valeurs respectives au point $t = 0$

$$\mathbf{x}_1(0) = \begin{bmatrix} -5 \\ 3 \end{bmatrix} \quad \text{et} \quad \mathbf{x}_2(0) = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$$

le sont. En vue d'appliquer la formule (8.48), calculons la matrice $\exp(t\mathbf{A})$ à l'aide de la formule (8.47) et des deux solutions $\mathbf{x}_1$ et $\mathbf{x}_2$:

$$\begin{aligned}\exp(t\mathbf{A}) &\equiv \frac{1}{5}e^t \begin{bmatrix} -5\cos t & -5\sin t \\ 3\cos t + \sin t & 3\sin t - \cos t \end{bmatrix} \begin{bmatrix} -1 & 0 \\ -3 & -5 \end{bmatrix} \\ &\equiv e^t \begin{bmatrix} \cos t + 3\sin t & 5\sin t \\ -2\sin t & -3\sin t + \cos t \end{bmatrix}.\end{aligned}$$

Calculons en outre l'intégrale figurant dans la formule (8.48):

$$\int_0^t \exp(-s\mathbf{A})\mathbf{b}(s)\,\mathrm{d}s \equiv \begin{bmatrix} \int_0^t (\cos^2 s - 3\cos s \sin s)\,\mathrm{d}s \\ \int_0^t 2\cos s \sin s\,\mathrm{d}s \end{bmatrix} \equiv \frac{1}{4}\begin{bmatrix} 3\cos 2t + \sin 2t + 2t - 3 \\ -2\cos 2t + 2 \end{bmatrix}.$$

D'après la formule (8.48) la solution cherchée est

$$\begin{aligned}\mathbf{x}(t) &\equiv e^t \begin{bmatrix} \cos t + 3\sin t & 5\sin t \\ -2\sin t & -3\sin t + \cos t \end{bmatrix} \left(\frac{1}{4}\begin{bmatrix} 3\cos 2t + \sin 2t + 2t - 3 \\ -2\cos 2t + 2 \end{bmatrix} + \frac{1}{4}\begin{bmatrix} 3 \\ -2 \end{bmatrix}\right) \\ &\equiv \frac{1}{4}e^t \begin{bmatrix} \sin t + 3\cos t + 6t\sin t + 2t\cos t \\ -4t\sin t - 2\cos t \end{bmatrix}.\end{aligned}$$

8.7.11 Exemple

L'exemple précédent montre que déjà dans le cas où $n = 2$ l'application de la formule (8.48) exige des calculs assez longs. C'est pourquoi, lorsqu'on sait deviner une solution particulière quelconque de l'équation (8.38), on préfère trouver d'abord la solution générale de cette équation, pour ensuite en déduire la solution particulière qui remplit la condition initiale. A titre d'exemple, considérons le problème aux valeurs initiales

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{b}, \quad \mathbf{x}(0) = \mathbf{c}, \tag{8.51}$$

où $\mathbf{A}$ est la matrice de l'exemple 8.7.9,

$$\mathbf{b}(t) \equiv \begin{bmatrix} 1+t \\ -2t \\ -17-10t \end{bmatrix} \quad \text{et} \quad \mathbf{c} = \begin{bmatrix} -5 \\ 6 \\ 3 \end{bmatrix}.$$

Cherchons une solution particulière de la même forme que $\mathbf{b}(t)$, c'est-à-dire de la forme $\mathbf{x}_0(t) \equiv \mathbf{v}_0 + t\mathbf{v}_1$. La substitution de $\mathbf{x}_0$ à $\mathbf{x}$ dans l'équation (8.51) fournit l'équation

$$\mathbf{v}_1 \equiv \mathbf{A}\mathbf{v}_0 + t\mathbf{A}\mathbf{v}_1 + \mathbf{b} \equiv \mathbf{A}\mathbf{v}_0 + t\mathbf{A}\mathbf{v}_1 + \begin{bmatrix} 1 \\ 0 \\ -17 \end{bmatrix} + t\begin{bmatrix} 1 \\ -2 \\ -10 \end{bmatrix}.$$

En égalant les coefficients des mêmes puissances de t dans les deux membres extrêmes, on obtient les équations

$$\mathbf{A}\mathbf{v}_1 = \begin{bmatrix} -1 \\ 2 \\ 10 \end{bmatrix} \quad \text{et} \quad \mathbf{A}\mathbf{v}_0 = \mathbf{v}_1 + \begin{bmatrix} -1 \\ 0 \\ 17 \end{bmatrix},$$

d'où l'on tire

$$\mathbf{v}_1 = \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix} \quad \text{et} \quad \mathbf{v}_0 = \begin{bmatrix} -4 \\ 5 \\ 2 \end{bmatrix}.$$

D'après 8.7.4, la solution générale de l'équation (8.51) s'obtient en additionnant la solution particulière

$$\mathbf{x}_0(t) \equiv \begin{bmatrix} -4 \\ 5 \\ 2 \end{bmatrix} + t\begin{bmatrix} -2 \\ 3 \\ 1 \end{bmatrix} \equiv \begin{bmatrix} -4-2t \\ 5+3t \\ 2+t \end{bmatrix}$$

à la solution générale de l'équation homogène associée, c'est-à-dire à la solution (8.49). La condition initiale se traduit par l'équation

$$\begin{bmatrix} -5 \\ 6 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ -2 & 0 & -1 \end{bmatrix}\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} + \begin{bmatrix} -4 \\ 5 \\ 2 \end{bmatrix},$$

d'où l'on tire $\alpha_1 = -2$, $\alpha_2 = 1$, $\alpha_3 = 3$. La solution cherchée est donc

$$\mathbf{x}(t) \equiv \begin{bmatrix} -2e^{3t} + (3t+1)e^{2t} - 2t - 4 \\ -2e^{3t} + 3e^{2t} + 3t + 5 \\ 4e^{3t} - 3e^{2t} + t + 2 \end{bmatrix}.$$

8.7.12 Exemple

Considérons le problème aux valeurs initiales

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{b}, \quad \mathbf{x}(0) = \mathbf{c},$$

où

$$\mathbf{A} = \begin{bmatrix} 2 & 0 & 1 & 1 \\ -2 & 2 & 1 & -1 \\ 0 & 0 & 3 & 1 \\ 0 & -2 & -1 & 3 \end{bmatrix}, \quad \mathbf{b}(t) \equiv \begin{bmatrix} 2 \\ 3e^{2t} \\ 0 \\ -2te^{2t} \end{bmatrix} \quad \text{et} \quad \mathbf{c} = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 2 \end{bmatrix}.$$

Nous allons résoudre ce problème par trigonalisation de la matrice **A**. Dans 8.5.8, nous avons calculé une matrice de passage **P** réduisant **A** à la forme triangulaire. D'après 8.7.8, le changement de variable $\mathbf{x} = \mathbf{P}\mathbf{y}$ réduit notre problème au suivant:

$$\begin{array}{llllll} \dot{y}_1 = 2y_1 & & - y_3 & & - te^{2t} & y_1(0) = 1 \\ \dot{y}_2 = & 4y_2 & - y_3 & & - te^{2t} & y_2(0) = 1 \\ \dot{y}_3 = & & 2y_3 & - 2y_4 & + 3e^{2t} & y_3(0) = 1 \\ \dot{y}_4 = & & & 2y_4 & + 2, & y_4(0) = 1. \end{array}$$

Rappelons, au passage, que la matrice de ce système différentiel est le résultat de la multiplication $\mathbf{P}^{-1}\mathbf{A}\mathbf{P}$, que

$$\begin{pmatrix} -te^{2t} \\ -te^{2t} \\ 3e^{2t} \\ 2 \end{pmatrix} \equiv \mathbf{P}^{-1}\mathbf{b}(t) \qquad \text{et} \qquad \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \mathbf{P}^{-1}\mathbf{c}.$$

En résolvant les quatre équations dans l'ordre inverse, nous déterminons successivement y_4, y_3, y_2 et y_1:

$$\begin{aligned} y_4(t) &\equiv 2e^{2t} - 1 \\ y_3(t) &\equiv (2 - t)e^{2t} - 1 \\ y_2(t) &\equiv \frac{1}{4}e^{4t} + e^{2t} - \frac{1}{4} \\ y_1(t) &\equiv (\frac{3}{2} - 2t)e^{2t} - \frac{1}{2}. \end{aligned}$$

Par le calcul de $\mathbf{P}\mathbf{y}$, nous obtenons la solution cherchée:

$$\begin{aligned} x_1(t) &\equiv \frac{1}{4}e^{4t} + (\frac{3}{2} + 2t)e^{2t} - \frac{3}{4} \\ x_2(t) &\equiv -\frac{1}{4}e^{4t} + (\frac{5}{2} - 3t)e^{2t} - \frac{5}{4} \\ x_3(t) &\equiv \frac{1}{4}e^{4t} - (\frac{1}{2} - 2t)e^{2t} + \frac{1}{4} \\ x_4(t) &\equiv \frac{1}{4}e^{4t} + (\frac{5}{2} - 2t)e^{2t} - \frac{3}{4}. \end{aligned}$$

8.8 EXERCICES

8.8.1 Montrer que 3 est une valeur propre de la matrice

$$\begin{pmatrix} 4 & 0 & 2 & 0 & 1 \\ 0 & 4 & 0 & 2 & 1 \\ 1 & 0 & 5 & 0 & 1 \\ 0 & 1 & 0 & 5 & 1 \\ 1 & 0 & 2 & 0 & 4 \end{pmatrix},$$

puis trouver la dimension du sous-espace propre associé.

8.8.2 Chercher la solution de l'équation différentielle $\ddot{\mathbf{x}} = \mathbf{Ax}$, où

$$\mathbf{A} = \begin{bmatrix} -5 & 2 \\ 2 & -2 \end{bmatrix}.$$

(Appliquer la formule (8.9).)

8.8.3 Pour chaque valeur de a, trouver les valeurs propres et les sous-espaces propres associés de la matrice

$$\begin{bmatrix} -a^2 & a^2 & a \\ 0 & a & 0 \\ -4a & -a^2+5a-2 & a^2+2 \end{bmatrix}.$$

8.8.4 Soit $\mathbf{A}$ une matrice carrée d'ordre n. Montrer que si λ est une valeur propre de $\mathbf{A}$, alors $|\lambda| \leqslant \|\mathbf{A}\|$. (Utiliser (d) de l'exercice 4.7.18.)

8.8.5 On dit qu'une application linéaire φ d'un espace vectoriel E dans lui-même est *nilpotente* s'il existe un entier positif k tel que φ^k est l'application nulle. Soit φ une application nilpotente d'un espace vectoriel $E \neq \{\mathbf{0}\}$ dans lui-même.
(a) Montrer que 0 est une valeur propre de φ.
(b) Montrer que 0 est la seule valeur propre de φ.

8.8.6 Soit λ une valeur propre d'une application linéaire φ d'un espace vectoriel E de dimension finie non nulle n dans lui-même. Démontrer les assertions suivantes:
(a) $\dim \operatorname{Im}(\varphi - \lambda \mathrm{id}_E) \leqslant n - 1$. (Appliquer (6.13).)
(b) $\operatorname{Im}(\varphi - \lambda \mathrm{id}_E)$ est un sous-espace stable par φ.
(c) Tout sous-espace vectoriel S de E qui contient $\operatorname{Im}(\varphi - \lambda \mathrm{id}_E)$ est stable par φ.
(d) En déduire qu'il existe un sous-espace vectoriel S de E, de dimension $n - 1$ et qui est stable par φ.

8.8.7 Soit $\mathbf{A} = (a_{ij})$ une matrice de transition.
(a) Montrer que 1 est une valeur propre de $\mathbf{A}$.
(b) Montrer que si λ est une valeur propre de $\mathbf{A}$, alors $|\lambda| \leqslant 1$.

8.8.8 Soit E un espace vectoriel de dimension 2, muni d'une base $(\mathbf{e}_1, \mathbf{e}_2)$. On désigne par D_1, D_2, D_3 et D_4 les droites vectorielles engendrées respectivement par $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_1 + \mathbf{e}_2$ et $\mathbf{e}_1 - \mathbf{e}_2$, par φ la projection sur D_1 parallèlement à D_2 et par ψ la symétrie par rapport à D_3 parallèlement à D_4. Trouver les valeurs propres et les vecteurs propres de $\varphi \circ \psi$ et de $\psi \circ \varphi$
(a) par un raisonnement géométrique,
(b) en introduisant les matrices de φ et de ψ.

8.8.9 Soit φ et ψ des applications linéaires d'un espace vectoriel $E \neq \{\mathbf{0}\}$ dans lui-même. Démontrer les assertions suivantes:
(a) L'ensemble des valeurs propres non nulles de $\varphi \circ \psi$ est égal à l'ensemble des valeurs propres non nulles de $\psi \circ \varphi$.

(b) 0 peut être une valeur propre de $\varphi \circ \psi$ sans en être une de $\psi \circ \varphi$. (Considérer l'espace vectoriel et l'application φ introduits dans l'exercice 7.4.1 et trouver une application ψ appropriée.)
(c) Si E est de dimension finie, les spectres de $\varphi \circ \psi$ et de $\psi \circ \varphi$ sont égaux.

8.8.10 Soit E l'espace vectoriel des polynômes de degré inférieur ou égal à 4. Soit φ l'application linéaire de E dans E définie par

$$\varphi(\mathbf{p})(t) \equiv (t^2 - 1)\dot{\mathbf{p}}(t) - 4t\mathbf{p}(t).$$

(a) Montrer que les polynômes $\mathbf{p}_k(t) \equiv (t-1)^{4-k}(t+1)^k$ $(k = 0, 1, 2, 3, 4)$ sont des vecteurs propres de φ et déterminer les valeurs propres correspondantes.
(b) L'application φ est-elle diagonalisable?

8.8.11 Soit φ une application linéaire d'un espace vectoriel de dimension finie non nulle dans lui-même. Montrer que si φ est diagonalisable, $\operatorname{rg}(\varphi^2) = \operatorname{rg}\varphi$ (cf. exercice 6.8.5).

8.8.12 Soit φ une application linéaire diagonalisable d'un espace vectoriel E de dimension finie non nulle dans lui-même. Montrer que s'il existe un entier $k > 1$ tel que $\varphi^k = \mathrm{id}_E$, alors $\varphi^2 = \mathrm{id}_E$.

8.8.13 Chercher les valeurs propres de la matrice

$$\mathbf{A} = \begin{pmatrix} -14 & 43 & -22 \\ 1 & 20 & -14 \\ 3 & 25 & -21 \end{pmatrix},$$

sachant que l'une d'entre elles est double. (Rappeler que toute racine double d'un polynôme est également une racine de la dérivée de ce polynôme.) $\mathbf{A}$ est-elle diagonalisable?

8.8.14 Chercher les valeurs propres et les sous-espaces propres associés de la matrice $\mathbf{A} = \alpha\mathbf{I}_n + \mathbf{a}{}^t\mathbf{b}$, où α est un nombre et $\mathbf{a}$, $\mathbf{b}$ sont des vecteurs-colonnes de $\mathbb{R}^n$. $\mathbf{A}$ est-elle diagonalisable?

8.8.15 Soit $\mathscr{E}$ un espace affine de dimension 4, muni d'un repère. Indiquer la nature de l'application affine de $\mathscr{E}$ dans $\mathscr{E}$ d'équation matricielle

$$\mathbf{y} = \frac{1}{2}\begin{pmatrix} 3 & -1 & -1 & -1 \\ 1 & 3 & 1 & 1 \\ -1 & 1 & 3 & 1 \\ -1 & 1 & 1 & 3 \end{pmatrix}\mathbf{x} + \begin{pmatrix} -2 \\ 4 \\ 2 \\ 2 \end{pmatrix}.$$

8.8.16 Chercher les valeurs propres des matrices

$$\begin{pmatrix} 0 & 3 & 0 \\ 3 & 0 & -4 \\ 0 & -4 & 0 \end{pmatrix} \quad \text{et} \quad \begin{pmatrix} -5 & 10 & 30 \\ 0 & 0 & -5 \\ 0 & 0 & 5 \end{pmatrix}$$

et montrer que ces matrices sont semblables. (Utiliser (b) et (c) de l'exercice 6.8.19.)

8.8.17 Les matrices

$$\begin{bmatrix} -2 & -3 & 2 \\ 0 & -2 & 0 \\ 0 & 4 & 0 \end{bmatrix} \quad \text{et} \quad \frac{1}{3}\begin{bmatrix} -4 & -4 & 2 \\ -2 & -2 & -2 \\ 0 & 0 & -6 \end{bmatrix}$$

sont-elles semblables? (Raisonner en termes de valeurs propres et de sous-espaces propres.)

8.8.18 Caractériser les matrices carrées d'ordre 4 dont le spectre est $\{1, 2, -1, 3\}$.

8.8.19 Calculer les puissances entières de la matrice

$$\begin{bmatrix} 0 & 1 & -1 \\ 1 & 2 & 1 \\ -1 & 1 & 2 \end{bmatrix}.$$

8.8.20 Calculer exp(**A**) pour

$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 4 & 1 \end{bmatrix}.$$

8.8.21 Soit **A** une matrice diagonalisable. Montrer que le déterminant de exp(**A**) est $e^{\mathrm{tr}\mathbf{A}}$.

8.8.22 Trigonaliser la matrice

$$\begin{bmatrix} -5 & -1 & -3 & -1 \\ -3 & 1 & 3 & 1 \\ -3 & 1 & -1 & 5 \\ -3 & 1 & -1 & -3 \end{bmatrix}.$$

8.8.23 Soit φ une transformation symétrique d'un espace vectoriel euclidien E de dimension finie. Montrer que $\mathrm{Ker}\,\varphi$ et $\mathrm{Im}\,\varphi$ sont orthogonaux et complémentaires dans E.

8.8.24 Soit a un nombre et b un nombre non nul. Soit **A** la matrice d'ordre $n > 1$

$$\begin{bmatrix} a & b & . & . & . & b \\ b & a & . & & & . \\ . & . & . & . & & . \\ . & & . & . & . & . \\ . & & & . & a & b \\ b & . & . & . & b & a \end{bmatrix}.$$

(a) Montrer que $a - b$ est une valeur propre de **A** et trouver le sous-espace propre associé à cette valeur propre.
(b) En s'appuyant sur le fait que **A** est symétrique, trouver un autre sous-espace propre et déterminer la valeur propre qui lui correspond.

8.8.25 Déterminer la plus petite et la plus grande valeur propre de la matrice

$$\mathbf{A} = \begin{bmatrix} -2 & 1 & 0 \\ 1 & -2 & 1 \\ 0 & 1 & -2 \end{bmatrix},$$

ainsi que les vecteurs propres associés à ces valeurs propres, sachant qu'ils sont de la forme

$$\mathbf{v} = \begin{bmatrix} \alpha \\ \beta \\ \alpha \end{bmatrix}.$$

(Trouver le minimum et le maximum du quotient $f(\alpha, \beta) = \dfrac{(\mathbf{Av} \mid \mathbf{v})}{(\mathbf{v} \mid \mathbf{v})}$ en annulant les dérivées partielles de f par rapport à α et à β.)

8.8.26 Soit $\mathbf{A}$ une matrice symétrique d'ordre n, S un sous-espace vectoriel de $\mathbb{R}^n$ de dimension non nulle k, $(\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k)$ une base orthonormale de S et $\mathbf{B}$ la matrice de type $n \times k$ dont les colonnes sont les vecteurs $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k$.
(a) Montrer que

$$\max_{\mathbf{x} \in S - \{\mathbf{0}\}} \frac{(\mathbf{Ax} \mid \mathbf{x})}{(\mathbf{x} \mid \mathbf{x})} = \max_{\mathbf{y} \in \mathbb{R}^k - \{\mathbf{0}\}} \frac{({}^t\mathbf{BABy} \mid \mathbf{y})}{(\mathbf{y} \mid \mathbf{y})}.$$

(Substituer $\mathbf{By}$ à $\mathbf{x}$.)
(b) En déduire que le premier membre est égal à la plus grande valeur propre de la matrice symétrique ${}^t\mathbf{BAB}$.

8.8.27 *Calcul par itération de la valeur propre de valeur absolue maximale.* Soit $\mathbf{A}$ une matrice carrée d'ordre n admettant n vecteurs propres linéairement indépendants $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_n$. On suppose que les valeurs propres correspondantes $\lambda_1, \lambda_2, ..., \lambda_n$ satisfont à la condition

$$|\lambda_1| > |\lambda_i|, \quad i = 2, 3, ..., n.$$

Soit $\mathbf{a}$ un vecteur de $\mathbb{R}^n$ dont la première composante α_1 dans la base $(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_n)$ est non nulle. On pose $\mathbf{a}_k = \mathbf{A}^k\mathbf{a} = (a_i^{(k)})$.
(a) Montrer que $\mathbf{a}_k = \lambda_1^k(\alpha_1\mathbf{x}_1 + \mathbf{r}_k)$, où $\mathbf{r}_k \to \mathbf{0}$ quand $k \to \infty$.
(b) En déduire que pour tout i tel que le i-ième terme de $\mathbf{x}_1$ est non nul,

$$\lim_{k \to \infty} \frac{a_i^{(k+1)}}{a_i^{(k)}} = \lambda_1.$$

Exercices sur les systèmes différentiels

8.8.28 Résoudre les problèmes aux valeurs initiales $\dot{\mathbf{x}} = \mathbf{Ax}$, $\mathbf{x}(0) = \mathbf{c}$, où

(a) $$\mathbf{A} = \begin{bmatrix} 1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{bmatrix} \quad \text{et} \quad \mathbf{c} = \begin{bmatrix} 2 \\ -1 \\ 2 \end{bmatrix}.$$

(b) $$\mathbf{A} = \frac{1}{2}\begin{pmatrix} 4 & -5 & 5 \\ 2 & -7 & 11 \\ 2 & -1 & 5 \end{pmatrix} \quad \text{et} \quad \mathbf{c} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}.$$

8.8.29 Résoudre le problème aux valeurs initiales $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{b}$, $\mathbf{x}(0) = \mathbf{c}$, où $\mathbf{A}$ est la matrice de l'exercice 8.8.22,

$$\mathbf{b}(t) \equiv e^t \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} \quad \text{et} \quad \mathbf{c} = \begin{pmatrix} 1 \\ -1 \\ 1 \\ 1 \end{pmatrix}.$$

(Utiliser la trigonalisation effectuée dans l'exercice cité.)

8.8.30 Trouver l'ensemble des solutions de l'équation différentielle $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{b}$, où

$$\mathbf{A} = \frac{1}{3}\begin{pmatrix} 4 & 0 & 1 \\ 2 & 7 & 0 \\ 0 & -2 & 4 \end{pmatrix} \quad \text{et} \quad \mathbf{b}(t) \equiv e^{-t}\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}.$$

8.8.31 Par le calcul de $\exp(t\mathbf{A})$, trouver une base de solutions de l'équation différentielle $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$, où $\mathbf{A}$ est la matrice carrée d'ordre n

$$\begin{pmatrix} \lambda & 1 & & & & \\ & \lambda & 1 & & 0 & \\ & & \cdot & \cdot & & \\ & & & \cdot & \cdot & \\ & & & & \cdot & \cdot \\ & 0 & & & \lambda & 1 \\ & & & & & \lambda \end{pmatrix}.$$

(Ecrire $\mathbf{A} = \lambda\mathbf{I} + \mathbf{B}$ et en déduire que $\exp(t\mathbf{A}) \equiv e^{\lambda t}\exp(t\mathbf{B})$; chercher ensuite $\exp(t\mathbf{B})$ en calculant directement $\mathbf{B}^2, \mathbf{B}^3, \ldots, \mathbf{B}^n$.)

CHAPITRE 9

Formes bilinéaires symétriques

9.1 RÉDUCTION DES FORMES BILINÉAIRES SYMÉTRIQUES

9.1.1 Introduction

Le présent chapitre est consacré à l'étude des formes bilinéaires symétriques et des formes quadratiques qui leur sont associées. Ces nouveaux objets jouent un rôle important aussi bien en algèbre linéaire qu'en géométrie. On les rencontre également en analyse, en physique et en statistique. Un des problèmes clefs de la théorie des formes bilinéaires symétriques est le problème de la réduction. Nous le résoudrons à l'aide de résultats obtenus dans la section 8.6.

Dans cette section, E désignera un espace vectoriel initialement de dimension quelconque et par la suite de dimension finie.

9.1.2 Formes bilinéaires

On appelle *forme bilinéaire* dans E toute application f de l'ensemble des couples de vecteurs de E dans $\mathbb{R}$ satisfaisant aux deux conditions suivantes:

(a) Pour tout vecteur $\mathbf{y}$ de E, l'application $\mathbf{x} \to f(\mathbf{x}, \mathbf{y})$ est une forme linéaire dans E.

(b) Pour tout vecteur $\mathbf{x}$ de E, l'application $\mathbf{y} \to f(\mathbf{x}, \mathbf{y})$ est une forme linéaire dans E.

On dit qu'une forme bilinéaire f dans E est *symétrique* si $f(\mathbf{x}, \mathbf{y}) = f(\mathbf{y}, \mathbf{x})$ pour tout couple $(\mathbf{x}, \mathbf{y})$ de vecteurs de E.

Voici quelques exemples:

(1) Tout produit scalaire dans E est une forme bilinéaire symétrique.

(2) $(\mathbf{x}, \mathbf{y}) \to 2x_1y_2 + 3x_2y_3 - x_3y_3$ est une forme bilinéaire dans $\mathbb{R}^3$.

(3) $(\mathbf{x}, \mathbf{y}) \to \det(\mathbf{x}, \mathbf{y})$ est une forme bilinéaire dans $\mathbb{R}^2$ (antisymétrique).

9.1.3 Forme quadratique associée à une forme bilinéaire

Soit f une forme bilinéaire dans E. On appelle *forme quadratique* associée à f l'application q de E dans $\mathbb{R}$

$$\mathbf{x} \to q(\mathbf{x}) = f(\mathbf{x}, \mathbf{x}).$$

Par exemple:

(1) La forme quadratique associée au produit scalaire $(\mathbf{x}, \mathbf{y}) \to (\mathbf{x} \mid \mathbf{y})$ est l'application $\mathbf{x} \to (\mathbf{x} \mid \mathbf{x}) = \|\mathbf{x}\|^2$.

(2) La forme quadratique associée à l'exemple (2) ci-dessus est l'application $\mathbf{x} \to 2x_1x_2 + 3x_2x_3 - x_3^2$.

(3) la forme quadratique associée à l'exemple (3) est l'application nulle.

Si f est une forme bilinéaire symétrique dans E et q la forme quadratique qui lui est associée, alors, pour tout couple $(\mathbf{x}, \mathbf{y})$ de vecteurs de E,

$$f(\mathbf{x}, \mathbf{y}) = \frac{1}{2}(q(\mathbf{x} + \mathbf{y}) - q(\mathbf{x}) - q(\mathbf{y})). \tag{9.1}$$

En effet,

$$\begin{aligned} q(\mathbf{x} + \mathbf{y}) &= f(\mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y}) = f(\mathbf{x}, \mathbf{x}) + 2f(\mathbf{x}, \mathbf{y}) + f(\mathbf{y}, \mathbf{y}) \\ &= q(\mathbf{x}) + 2f(\mathbf{x}, \mathbf{y}) + q(\mathbf{y}), \end{aligned}$$

ce qui entraîne la relation (9.1).

La connaissance de q permet ainsi de retrouver f. On notera que cette reconstitution de f à partir de q n'est possible que si f est symétrique (dans l'exemple (3), q est nulle sans que f le soit).

9.1.4 Matrice d'une forme bilinéaire

Jusqu'à la fin de cette section, nous supposerons que E est de dimension finie non nulle n.

Considérons une base quelconque $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ de E et désignons les composantes des vecteurs $\mathbf{x}$ et $\mathbf{y}$ dans cette base respectivement par $x_1, x_2, ..., x_n$ et $y_1, y_2, ..., y_n$. Soit f une forme bilinéaire dans E. D'après (a) et (b) de la définition 9.1.2,

$$\begin{aligned} f(\mathbf{x}, \mathbf{y}) &= f\Big(\sum_{i=1}^{n} x_i\mathbf{e}_i, \sum_{j=1}^{n} y_j\mathbf{e}_j\Big) = \sum_{i=1}^{n} x_i f\Big(\mathbf{e}_i, \sum_{j=1}^{n} y_j\mathbf{e}_j\Big) \\ &= \sum_{i=1}^{n} x_i\Big(\sum_{j=1}^{n} y_j f(\mathbf{e}_i, \mathbf{e}_j)\Big) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_iy_j, \end{aligned} \tag{9.2}$$

où nous avons posé $f(\mathbf{e}_i, \mathbf{e}_j) = a_{ij}$.

On appelle la matrice carrée d'ordre n $\mathbf{A} = (a_{ij})$ *matrice de la forme bilinéaire* f dans la base $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$. Lorsque f est représentée par le dernier membre de (9.2), on appelle les nombres a_{ij} *coefficients* de f.

Il est évident que la matrice d'une forme bilinéaire dépend de la base choisie. En outre, cette matrice est symétrique si et seulement si la forme bilinéaire est symétrique.

A l'aide de la matrice $\mathbf{A}$, l'égalité des deux membres extrêmes de (9.2) s'écrit, plus simplement,

$$f(\mathbf{x}, \mathbf{y}) = {}^t\mathbf{x}_e\mathbf{A}\mathbf{y}_e. \tag{9.3}$$

La forme quadratique q associée à f s'exprime donc sous la forme

$$q(\mathbf{x}) = {}^t\mathbf{x}_e\mathbf{A}\mathbf{x}_e. \tag{9.4}$$

Si f est symétrique, on dit indifféremment que $\mathbf{A}$ est la matrice de la forme bilinéaire f ou de la forme quadratique q associée à f.

Voici l'écriture explicite de $q(\mathbf{x})$ dans le cas où $\mathbf{A}$ est symétrique et $n = 2$ ou 3:

$n = 2$: $q(\mathbf{x}) = a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2$.
$n = 3$: $q(\mathbf{x}) = a_{11}x_1^2 + 2a_{12}x_1x_2 + 2a_{13}x_1x_3 + a_{22}x_2^2 + 2a_{23}x_2x_3 + a_{33}x_3^2$.

Inversement, lorsque $q(\mathbf{x})$ est écrit explicitement, il est facile de former la matrice symétrique $\mathbf{A}$. Par exemple, la matrice de la forme quadratique

$$\mathbf{x} \to q(\mathbf{x}) = 4x_1x_2 - 3x_2^2 - 12x_2x_3$$

est

$$\mathbf{A} = \begin{bmatrix} 0 & 2 & 0 \\ 2 & -3 & -6 \\ 0 & -6 & 0 \end{bmatrix}.$$

9.1.5 Forme bilinéaire associée à une matrice

Soit $\mathbf{A}$ une matrice carrée d'ordre n. On appelle *forme bilinéaire associée canoniquement à* $\mathbf{A}$ la forme bilinéaire f dans $\mathbb{R}^n$ définie par

$$f(\mathbf{x}, \mathbf{y}) = {}^t\mathbf{x}\mathbf{A}\mathbf{y}.$$

On appelle *forme quadratique associée canoniquement à* $\mathbf{A}$ la forme quadratique associée à f.

9.1.6 Transformation de la matrice d'une forme bilinéaire par suite d'un changement de base

Soit f une forme bilinéaire dans E, $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ et $(\mathbf{e}'_1, \mathbf{e}'_2, ..., \mathbf{e}'_n)$ des bases de E, $\mathbf{A}$ la matrice de f dans la base $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ et $\mathbf{A}'$ la matrice de f dans la base $(\mathbf{e}'_1, \mathbf{e}'_2, ..., \mathbf{e}'_n)$. D'après (9.3), d'une part

$$f(\mathbf{x}, \mathbf{y}) = {}^t\mathbf{x}_e\mathbf{A}\mathbf{y}_e = {}^t(\mathbf{P}_{ee'}\mathbf{x}_{e'})\mathbf{A}(\mathbf{P}_{ee'}\mathbf{y}_{e'}) = {}^t\mathbf{x}_{e'}({}^t\mathbf{P}_{ee'}\mathbf{A}\mathbf{P}_{ee'})\mathbf{y}_{e'}$$

et d'autre part

$$f(\mathbf{x}, \mathbf{y}) = {}^t\mathbf{x}_{e'}\mathbf{A}'\mathbf{y}_{e'}.$$

Comme $\mathbf{x}$ et $\mathbf{y}$ sont des vecteurs quelconques, cela entraîne la relation

$$\mathbf{A}' = {}^t\mathbf{P}_{ee'}\mathbf{A}\mathbf{P}_{ee'}. \tag{9.5}$$

9.1.7 Réduction d'une forme bilinéaire symétrique

Réduire une forme bilinéaire symétrique f (ou la forme quadratique q associée à f) signifie trouver une base dans laquelle la matrice de f (ou de q) est diagonale. Dans une telle base, on dit que la forme bilinéaire (ou la forme quadratique) est *réduite*.

Soit f une forme bilinéaire symétrique dans E. Munissons E d'une base $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ arbitrairement choisie et désignons par $\mathbf{A}$ la matrice de f dans cette base. D'après (b) de 8.6.8, il existe une matrice orthogonale $\mathbf{P}$ et une matrice diagonale $\mathbf{D}$ telles que

$$\mathbf{D} = {}^t\mathbf{PAP}. \tag{9.6}$$

Rappelons qu'une telle matrice $\mathbf{P}$ se forme en prenant comme colonnes n vecteurs propres de $\mathbf{A}$ orthogonaux deux à deux et unitaires; les termes diagonaux de $\mathbf{D}$ sont alors les valeurs propres correspondantes. Posons, pour $j = 1, 2, ..., n$,

$$\mathbf{e}'_j = p_{1j}\mathbf{e}_1 + p_{2j}\mathbf{e}_2 + ... + p_{nj}\mathbf{e}_n, \tag{9.7}$$

où $p_{1j}, p_{2j}, ..., p_{nj}$ sont les termes de la j-ième colonne de $\mathbf{P}$. La famille de vecteurs $(\mathbf{e}'_1, \mathbf{e}'_2, ..., \mathbf{e}'_n)$ est une base de E et la matrice de passage $\mathbf{P}_{ee'}$ est égale à $\mathbf{P}$. En outre, d'après (9.5), la matrice de f dans la base $(\mathbf{e}'_1, \mathbf{e}'_2, ..., \mathbf{e}'_n)$ est la matrice $\mathbf{D}$ de (9.6). Nous avons ainsi démontré qu'il existe toujours une base $(\mathbf{e}'_1, \mathbf{e}'_2, ..., \mathbf{e}'_n)$ de E dans laquelle la forme bilinéaire symétrique f est réduite. Dans cette base, $f(\mathbf{x}, \mathbf{y})$ et $q(\mathbf{x})$ s'écrivent sous la forme

$$f(\mathbf{x}, \mathbf{y}) = d_1x'_1y'_1 + d_2x'_2y'_2 + ... + d_nx'_ny'_n,$$
$$q(\mathbf{x}) = d_1x'^2_1 + d_2x'^2_2 + ... + d_nx'^2_n,$$

où $x'_1, x'_2, ..., x'_n$ et $y'_1, y'_2, ..., y'_n$ sont les composantes respectives de $\mathbf{x}$ et de $\mathbf{y}$ et $d_1, d_2, ..., d_n$ les termes diagonaux de la matrice (diagonale) de f.

On notera que si E est euclidien et la base initiale $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ orthonormale, la base $(\mathbf{e}'_1, \mathbf{e}'_2, ..., \mathbf{e}'_n)$ définie ci-dessus est également orthonormale.

9.1.8 Exemples

(1) Nous nous proposons de réduire la forme quadratique q définie, dans une base $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ de E, par

$$q(\mathbf{x}) = 7x_1^2 + 4x_1x_2 + 6x_2^2 + 4x_2x_3 + 5x_3^2.$$

La matrice de q est

$$\mathbf{A} = \begin{bmatrix} 7 & 2 & 0 \\ 2 & 6 & 2 \\ 0 & 2 & 5 \end{bmatrix}.$$

Les racines de son polynôme caractéristique

$$\begin{vmatrix} 7-t & 2 & 0 \\ 2 & 6-t & 2 \\ 0 & 2 & 5-t \end{vmatrix} = -t^3 + 18t^2 - 99t + 162$$

sont 3, 6 et 9. Définissons la base $(\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3)$ par

$$(\mathbf{e}'_1)_e = \frac{1}{3}\begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix} \quad \text{(vecteur propre unitaire de } \mathbf{A} \text{ associé à la valeur propre 3)},$$

$$(\mathbf{e}'_2)_e = \frac{1}{3}\begin{bmatrix} 2 \\ -1 \\ -2 \end{bmatrix} \quad \text{(vecteur propre unitaire de } \mathbf{A} \text{ associé à la valeur propre 6)},$$

$$(\mathbf{e}'_3)_e = \frac{1}{3}\begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix} \quad \text{(vecteur propre unitaire de } \mathbf{A} \text{ associé à la valeur propre 9)}.$$

Dans la base $(\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3)$, la forme quadratique q est réduite:

$$q(\mathbf{x}) = 3x_1'^2 + 6x_2'^2 + 9x_3'^2.$$

(2) La réduction d'une forme quadratique peut également être effectuée par le procédé de *complétion des carrés*. Nous nous limiterons à traiter le cas où $n = 2$ (cf. exercice 9.4.3 pour le cas général). Supposons que a_{11} soit non nul. Alors

$$\begin{aligned} q(\mathbf{x}) &= a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 = a_{11}(x_1^2 + 2\frac{a_{12}}{a_{11}}x_1x_2) + a_{22}x_2^2 \\ &= a_{11}(x_1 + \frac{a_{12}}{a_{11}}x_2)^2 + (a_{22} - \frac{a_{12}^2}{a_{11}})x_2^2. \end{aligned}$$

En posant $x'_1 = x_1 + \dfrac{a_{12}}{a_{11}}x_2$ et $x'_2 = x_2$ (ce qui revient à effectuer un changement de base), nous voyons que

$$q(\mathbf{x}) = d_1x_1'^2 + d_2x_2'^2,$$

où $d_1 = a_{11}$ et $d_2 = a_{22} - \dfrac{a_{12}^2}{a_{11}}$. Si a_{11} est nul et a_{22} n'est pas nul, il suffit d'échanger les rôles de ces deux termes. Si tous les deux sont nuls, en posant $x'_1 = x_1 + x_2$ et $x'_2 = x_1 - x_2$, nous voyons que

$$q(\mathbf{x}) = d_1x_1'^2 + d_2x_2'^2,$$

où $d_1 = -d_2 = \frac{1}{2}a_{12}$.

On remarquera qu'en général la forme réduite obtenue par le procédé de complétion des carrés n'a pas pour coefficients les valeurs propres de la matrice de la forme quadratique.

(3) Nous allons calculer l'intégrale

$$\int_{\mathbb{R}^n} \exp(-\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j)\mathrm{d}x_1\mathrm{d}x_2\ldots\mathrm{d}x_n, \tag{9.8}$$

où $a_{ij} = a_{ji}$ pour tout couple (i, j) d'indices. L'exposant est une forme quadratique que l'on peut réduire par un changement de base dont la matrice est orthogonale. Soit $\mathbf{P}$ la matrice d'un tel changement de base. Alors

$$\sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}x_i x_j = d_1 x_1'^2 + d_2 x_2'^2 + \ldots + d_n x_n'^2,$$

où

$$\begin{pmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{pmatrix} = \mathbf{P} \begin{pmatrix} x_1' \\ x_2' \\ \cdot \\ \cdot \\ \cdot \\ x_n' \end{pmatrix}.$$

Or, la matrice jacobienne de ce changement de base est évidemment

$$\left(\frac{\partial x_i}{\partial x_j'}\right) = \mathbf{P}.$$

L'intégrale (9.8) est donc égale à

$$\int_{\mathbb{R}^n} \exp(-d_1 x_1'^2 - d_2 x_2'^2 - \ldots - d_n x_n'^2) \| \mathbf{P} \| \mathrm{d}x_1' \mathrm{d}x_2' \ldots \mathrm{d}x_n'$$
$$= \int_{\mathbb{R}} \exp(-d_1 x_1'^2) \mathrm{d}x_1' \int_{\mathbb{R}} \exp(-d_2 x_2'^2) \mathrm{d}x_2' \ldots \int_{\mathbb{R}} \exp(-d_n x_n'^2) \mathrm{d}x_n'.$$

Il apparaît ainsi clairement que cette intégrale est finie si et seulement si les coefficients $d_1, d_2, \ldots, d_n$ sont tous positifs. Dans ce cas, sa valeur est

$$\left(\frac{\pi}{d_1}\right)^{\frac{1}{2}} \left(\frac{\pi}{d_2}\right)^{\frac{1}{2}} \ldots \left(\frac{\pi}{d_n}\right)^{\frac{1}{2}} = \left(\frac{\pi^n}{d_1 d_2 \ldots d_n}\right)^{\frac{1}{2}} = \left(\frac{\pi^n}{|\mathbf{A}|}\right)^{\frac{1}{2}},$$

où $\mathbf{A}$ désigne la matrice (a_{ij}). (L'égalité $|\mathbf{A}| = d_1 d_2 \ldots d_n$ est due au fait que $\mathbf{A}$ est semblable à $\operatorname{diag}(d_1, d_2, \ldots, d_n)$.)

9.2 FORMES BILINÉAIRES SYMÉTRIQUES DÉFINIES POSITIVES

9.2.1 Introduction

L'exemple précédent nous apprend que l'intégrale de l'exponentielle négative d'une forme quadratique q est finie si et seulement si $q(\mathbf{x})$ est positif pour tout vecteur non nul $\mathbf{x}$. Nous nous proposons maintenant d'examiner et de caractériser les formes quadratiques qui jouissent de cette propriété.

Tout au long de cette section, E désignera un espace vectoriel de dimension finie non nulle n.

9.2.2 Formes bilinéaires symétriques définies non négatives et définies positives

Soit f une forme bilinéaire symétrique dans E et q la forme quadratique associée à f. On dit que f est *définie non négative* si $f(\mathbf{x}, \mathbf{x}) = q(\mathbf{x}) \geqslant 0$ pour tout vecteur $\mathbf{x}$ de E. On dit que f est *définie positive* si $f(\mathbf{x}, \mathbf{x}) = q(\mathbf{x}) > 0$ pour tout vecteur non nul $\mathbf{x}$ de E.

On notera qu'une forme bilinéaire symétrique définie positive n'est autre qu'un produit scalaire (cf. 2.2.2) et que réduire une telle forme revient à trouver une base orthogonale pour ce produit scalaire. A ce propos, on rappellera que le procédé d'orthogonalisation de Gram-Schmidt décrit précisément les opérations à effectuer pour obtenir une base orthogonale. La méthode générale de réduction exposée dans 9.1.7 fournit maintenant un autre moyen pour parvenir à ce résultat.

9.2.3 Matrices symétriques définies non négatives et définies positives

On dit qu'une matrice symétrique $\mathbf{A}$ d'ordre n est *définie non négative* (*définie positive*) si la forme bilinéaire symétrique associée canoniquement à $\mathbf{A}$ est définie non négative (définie positive). Dire que $\mathbf{A}$ est définie non négative (définie positive) revient donc à dire que ${}^t\mathbf{x}\mathbf{A}\mathbf{x} \geqslant 0$ pour tout vecteur-colonne $\mathbf{x}$ de $\mathbb{R}^n$ (${}^t\mathbf{x}\mathbf{A}\mathbf{x} > 0$ pour tout vecteur-colonne non nul $\mathbf{x}$ de $\mathbb{R}^n$).

En raison de la représentation (9.3), si une forme bilinéaire symétrique f dans E est définie non négative (définie positive), sa matrice dans toute base de E est définie non négative (définie positive). Réciproquement, s'il existe une base de E dans laquelle la matrice de f est définie non négative (définie positive), alors f est définie non négative (définie positive).

On remarquera que les termes diagonaux d'une matrice symétrique définie non négative (définie positive) sont non négatifs (positifs). Ce sont en effet les valeurs que la forme quadratique associée canoniquement à cette matrice attribue aux vecteurs de la base canonique.

La réciproque n'est toutefois vraie que pour les matrices diagonales: si les termes diagonaux d'une telle matrice sont non négatifs (positifs), cette matrice est définie non négative (définie positive).

9.2.4 Proposition. Caractérisation de la positivité par les valeurs propres

Pour qu'une matrice symétrique $\mathbf{A}$ *soit définie non négative (définie positive), il faut et il suffit que ses valeurs propres soient non négatives (positives).*

DÉMONSTRATION

Il existe une base de $\mathbb{R}^n$ dans laquelle la matrice de la forme quadratique associée canoniquement à $\mathbf{A}$ est la matrice $\mathbf{D}$ de (9.6). Les termes diagonaux de $\mathbf{D}$ étant les valeurs propres de $\mathbf{A}$ répétées un nombre de fois égal à leur multiplicité, la proposition découle de la remarque figurant dans 9.2.3.

Voici maintenant un critère pratique permettant de reconnaître les matrices symétriques définies positives sans l'aide de leurs valeurs propres. Nous appellerons *mineurs principaux* d'une matrice carrée $\mathbf{A} = (a_{ij})$ les déterminants

$$a_{11}, \quad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \quad \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}, \ \ldots, \ |\mathbf{A}|.$$

9.2.5 Proposition. Caractérisation de la positivité par les mineurs principaux

Pour qu'une matrice symétrique $\mathbf{A}$ *soit définie positive, il faut et il suffit que ses mineurs principaux soient positifs.*

DÉMONSTRATION

Supposons que $\mathbf{A} = (a_{ij})$ soit définie positive et d'ordre n. Choisissons un entier k de $\{1, 2, ..., n\}$ et posons

$$\tilde{\mathbf{A}} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ & \cdots\cdots & & \\ a_{k1} & a_{k2} & \cdots & a_{kk} \end{pmatrix}. \tag{9.9}$$

Soit $\tilde{\mathbf{x}}$ un vecteur-colonne non nul quelconque de $\mathbb{R}^k$. Désignons par $\mathbf{x}$ le vecteur-colonne de $\mathbb{R}^n$ dont les k premiers termes sont ceux de $\tilde{\mathbf{x}}$ et les restants sont nuls. Par la règle de multiplication par blocs, nous voyons aussitôt que

$${}^t\tilde{\mathbf{x}}\tilde{\mathbf{A}}\tilde{\mathbf{x}} = {}^t\mathbf{x}\mathbf{A}\mathbf{x}.$$

Le second membre étant positif, par hypothèse, il s'ensuit que la matrice $\tilde{\mathbf{A}}$ est définie positive. Son déterminant est donc positif, puisqu'il est le produit des valeurs propres (éventuellement répétées) de $\tilde{\mathbf{A}}$ et que ces valeurs propres sont positives, d'après la proposition 9.2.4.

L'assertion réciproque sera démontrée par récurrence sur l'ordre de $\mathbf{A}$. Lorsque cet ordre est 1, elle est évidente. Supposons qu'elle soit vraie pour toute matrice symétrique d'ordre $n - 1$ ($n > 1$) et considérons une matrice symétrique d'ordre n $\mathbf{A} = (a_{ij})$ dont les mineurs principaux sont positifs. Posons

$$\mathbf{a} = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \cdot \\ \cdot \\ \cdot \\ a_{n-1,n} \end{pmatrix} \quad \text{et} \quad \mathbf{P} = \left(\begin{array}{cccc:c} & & \mathbf{I}_{n-1} & & -\tilde{\mathbf{A}}^{-1}\mathbf{a} \\ \hdashline 0 & 0 & \cdots & 0 & 1 \end{array}\right),$$

où $\tilde{\mathbf{A}}$ est la matrice définie par (9.9) avec $k = n - 1$. Par la règle de multiplication par blocs, nous voyons aisément que

$$
{}^t\mathbf{PAP} = \left[\begin{array}{cccc:c} & & & & 0 \\ & & & & 0 \\ & & \tilde{\mathbf{A}} & & \cdot \\ & & & & \cdot \\ & & & & \cdot \\ & & & & 0 \\ \hdashline 0 & 0 & \ldots & 0 & b \end{array}\right], \tag{9.10}
$$

où b est un nombre qu'il n'est pas nécessaire de préciser. Désignons la matrice du second membre par $\mathbf{A}'$. D'après (9.5), la forme bilinéaire associée canoniquement à $\mathbf{A}$, exprimée dans la base de $\mathbb{R}^n$ constituée des colonnes de $\mathbf{P}$, a pour matrice $\mathbf{A}'$. Il suffit donc de démontrer que $\mathbf{A}'$ est définie positive. Grâce aux propriétés (a) et (b) de la proposition 5.1.8, au fait que $|\mathbf{P}| = 1$ et à (5.16), nous déduisons de (9.10) que

$$|\mathbf{A}| = |{}^t\mathbf{PAP}| = b|\tilde{\mathbf{A}}|.$$

Cela prouve que b est positif, puisque $|\mathbf{A}|$ et $|\tilde{\mathbf{A}}|$ sont positifs, par hypothèse. Considérons maintenant un vecteur-colonne non nul quelconque $\mathbf{x}$ de $\mathbb{R}^n$ et désignons par $\tilde{\mathbf{x}}$ le vecteur-colonne de $\mathbb{R}^{n-1}$ obtenu en supprimant le dernier terme x_n de $\mathbf{x}$. A l'aide de la règle de multiplication par blocs, nous voyons facilement que

$$
{}^t\mathbf{x}\mathbf{A}'\mathbf{x} = {}^t\tilde{\mathbf{x}}\tilde{\mathbf{A}}\tilde{\mathbf{x}} + bx_n^2. \tag{9.11}
$$

Or, si $\tilde{\mathbf{x}}$ n'est pas nul, ${}^t\tilde{\mathbf{x}}\tilde{\mathbf{A}}\tilde{\mathbf{x}}$ est positif, puisque la matrice $\tilde{\mathbf{A}}$ est définie positive, par hypothèse de récurrence. D'autre part, si $\tilde{\mathbf{x}}$ est nul, x_n n'est pas nul et donc bx_n^2 est positif. Nous en concluons que le premier membre de (9.11) est positif pour tout vecteur-colonne non nul $\mathbf{x}$ de $\mathbb{R}^n$, autrement dit, que la matrice $\mathbf{A}'$ est définie positive, ce qu'il fallait démontrer.

9.2.6 Remarque

Les arguments utilisés dans la première partie de la démonstration précédente démontrent que les mineurs principaux d'une matrice symétrique définie non négative sont non négatifs. Toutefois, une matrice symétrique dont les mineurs principaux sont non négatifs n'est pas forcément définie non négative. Par exemple, les mineurs principaux de la matrice

$$
\mathbf{A} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \end{bmatrix}
$$

sont non négatifs, mais

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

9.2.7 Application au calcul des extremums d'une fonction

Soit f une fonction réelle définie dans $\mathbb{R}^n$ admettant, dans un voisinage de $\mathbf{0}$, des dérivées partielles continues d'ordre 2. Posons

$$b_i = \frac{\partial f}{\partial x_i}(\mathbf{0}) \qquad \text{et} \qquad a_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{0}),$$

où i et j parcourent les entiers $1, 2, ..., n$. La matrice symétrique $\mathbf{A} = (a_{ij})$ est appelée *matrice hessienne* de f en $\mathbf{0}$.

Critère. *Supposons que* $\mathbf{0}$ *soit un point stationnaire de* f*, c'est-à-dire que* $b_1 = b_2 = ... = b_n = 0$. *Si* $\mathbf{A}$ *est définie positive,* f *a un minimum relatif strict en* $\mathbf{0}$.

DÉMONSTRATION

Le développement limité de f d'ordre 2 dans une boule B de centre $\mathbf{0}$ et de rayon r suffisamment petit s'écrit

$$f(\mathbf{x}) = f(\mathbf{0}) + \frac{1}{2}{}^t\mathbf{x}\mathbf{A}\mathbf{x} + \| \mathbf{x} \|^2 \varepsilon(\mathbf{x}), \tag{9.12}$$

où $\varepsilon(\mathbf{x}) \to 0$ quand $\| \mathbf{x} \| \to 0$. Compte tenu de 8.6.9, nous en concluons que, pour tout vecteur non nul $\mathbf{x}$ de B,

$$\frac{f(\mathbf{x}) - f(\mathbf{0})}{\| \mathbf{x} \|^2} = \frac{1}{2}\frac{{}^t\mathbf{x}\mathbf{A}\mathbf{x}}{\| \mathbf{x} \|^2} + \varepsilon(\mathbf{x}) \geqslant \frac{\lambda}{2} + \varepsilon(\mathbf{x}),$$

où λ désigne la plus petite valeur propre de $\mathbf{A}$. Comme $\varepsilon(\mathbf{x})$ peut être rendu arbitrairement petit par le choix d'un r suffisamment petit et λ est positif (d'après la proposition 9.2.4), cela montre que $f(\mathbf{x}) - f(\mathbf{0})$ est positif pour tout vecteur non nul $\mathbf{x}$ de B.

A titre d'exemple, cherchons à établir la nature du point stationnaire $\mathbf{0}$ de la fonction f définie par

$$f(x_1, x_2, x_3) = (x_1^2 - x_2x_3 + 1)\exp(-5x_1^2 - 4x_1x_2 + 4x_1x_3 - 5x_2^2 + 4x_2x_3 - 5x_3^2).$$

A l'aide d'un développement limité de la fonction exponentielle ou par le calcul des dérivées partielles du second ordre, on voit aisément que la matrice hessienne de f en $\mathbf{0}$ est

$$\mathbf{A} = \begin{pmatrix} -8 & -4 & 4 \\ -4 & -10 & 3 \\ 4 & 3 & -10 \end{pmatrix}.$$

Les mineurs principaux de $\mathbf{A}$ sont égaux à -8, 64, -504, donc ceux de $-\mathbf{A}$ à 8, 64, 504. Grâce à la proposition 9.2.4 et au critère, il s'ensuit que $-f$ a un minimum relatif strict en $\mathbf{0}$, donc que f a un maximum relatif strict en $\mathbf{0}$.

On remarquera qu'une fonction peut avoir un minimum relatif strict en un point, sans que sa matrice hessienne en ce point soit définie positive. Par exemple, la fonction f définie par

$$f(x_1, x_2) = x_1^4 + x_2^4$$

a un minimum strict en $\mathbf{0}$, mais sa matrice hessienne en $\mathbf{0}$ est nulle.

On remarquera enfin que si f a un minimum relatif en $\mathbf{0}$, le terme ${}^t\mathbf{xAx}$ du développement (9.12) ne peut être négatif, autrement dit, $\mathbf{A}$ est une matrice définie non négative. La réciproque n'est toutefois pas vraie (penser à $-f$, où f est la fonction de l'exemple qui précède).

9.3 RÉDUCTION SIMULTANÉE

9.3.1 Position du problème

Etant donné deux formes bilinéaires symétriques f et g définies dans un espace vectoriel de dimension finie non nulle, existe-t-il une base de cet espace dans laquelle les matrices de f et de g sont diagonales? L'exemple simple suivant montre que la réponse à cette question est généralement négative. Lorsqu'elle est affirmative, on dit que f et g peuvent être *réduites simultanément*.

Considérons les formes bilinéaires symétriques associées canoniquement aux matrices

$$\mathbf{A} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \text{et} \quad \mathbf{B} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

D'après (9.5), réduire simultanément ces deux formes revient à trouver une matrice de passage $\mathbf{P}$ de manière que ${}^t\mathbf{PAP}$ et ${}^t\mathbf{PBP}$ soient des matrices diagonales. Supposons que $\mathbf{P}$ ait été trouvée et désignons ses termes par p_{ij}. Alors

$${}^t\mathbf{PAP} = \begin{pmatrix} p_{11}^2-p_{21}^2 & p_{11}p_{12}-p_{21}p_{22} \\ p_{11}p_{12}-p_{21}p_{22} & p_{12}^2-p_{22}^2 \end{pmatrix},$$

$${}^t\mathbf{PBP} = \begin{pmatrix} 2p_{11}p_{21} & p_{12}p_{21}+p_{11}p_{22} \\ p_{12}p_{21}+p_{11}p_{22} & 2p_{12}p_{22} \end{pmatrix}.$$

Comme les deux seconds membres sont des matrices diagonales,

$$\begin{aligned} p_{11}p_{12} - p_{21}p_{22} &= 0 \\ p_{12}p_{21} + p_{11}p_{22} &= 0, \end{aligned}$$

ou

$$\begin{pmatrix} p_{11} & -p_{21} \\ p_{21} & p_{11} \end{pmatrix} \begin{pmatrix} p_{12} \\ p_{22} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Or, la matrice carrée de cette équation est inversible, car son déterminant est égal au carré de la norme de la première colonne de $\mathbf{P}$. Les termes p_{12} et p_{22} doivent donc être nuls, ce qui est impossible, puisqu'ils constituent la deuxième colonne de **P**. Il est ainsi prouvé que les formes bilinéaires symétriques associées canoniquement aux matrices **A** et **B** ne peuvent pas être réduites simultanément.

9.3.2 Proposition. Réduction simultanée

Soit f et g deux formes bilinéaires symétriques définies dans un espace vectoriel E de dimension finie non nulle. Supposons que g soit définie positive. Il existe alors une base de E dans laquelle f et g sont réduites.

DÉMONSTRATION

Munissons E du produit scalaire g et choisissons une base orthonormale $(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ de E. D'après 9.1.7, il existe une base $(\mathbf{e}'_1, \mathbf{e}'_2, \ldots, \mathbf{e}'_n)$ de E dans laquelle f est réduite. En outre, cette base peut être choisie de telle manière que la matrice de passage $\mathbf{P}_{ee'}$ soit orthogonale. Dans ce cas, la base $(\mathbf{e}'_1, \mathbf{e}'_2, \ldots, \mathbf{e}'_n)$ est encore orthonormale, autrement dit, la matrice de g dans cette base est la matrice-unité.

9.3.3 Calcul de la matrice de passage dans une réduction simultanée

Lorsque les formes bilinéaires f et g de la proposition 9.3.2 sont données par leurs matrices **A** et **B** (dans une base de E), le calcul d'une matrice de passage **P** telle que

$$ {}^t\mathbf{P}\mathbf{A}\mathbf{P} = \mathbf{D} \text{ (diagonale)} \quad \text{et} \quad {}^t\mathbf{P}\mathbf{B}\mathbf{P} = \mathbf{I} \tag{9.13}$$

est assez laborieux. C'est pourquoi nous allons démontrer une proposition fournissant une méthode pour l'effectuer plus aisément.

On appelle *valeur propre de* (**A**, **B**) tout nombre λ tel que

$$|\mathbf{A} - \lambda\mathbf{B}| = 0.$$

Si λ est une valeur propre de (**A**, **B**), on appelle *vecteur propre de* (**A**, **B**) associé à cette valeur propre toute solution non nulle $\mathbf{x}$ de l'équation

$$(\mathbf{A} - \lambda\mathbf{B})\mathbf{x} = \mathbf{0}.$$

On notera que si λ et μ sont des valeurs propres distinctes de (**A**, **B**) et $\mathbf{x}$ et $\mathbf{y}$ des vecteurs propres associés respectivement à ces valeurs propres, alors $\mathbf{x}$ et $\mathbf{y}$ sont orthogonaux pour le produit scalaire défini par **B**, c'est-à-dire la forme bilinéaire associée canoniquement à **B**. En effet,

$$\lambda\,{}^t\mathbf{x}\mathbf{B}\mathbf{y} = {}^t(\lambda\mathbf{B}\mathbf{x})\mathbf{y} = {}^t(\mathbf{A}\mathbf{x})\mathbf{y} = {}^t\mathbf{x}(\mathbf{A}\mathbf{y}) = {}^t\mathbf{x}(\mu\mathbf{B}\mathbf{y}) = \mu\,{}^t\mathbf{x}\mathbf{B}\mathbf{y}, \tag{9.14}$$

ce qui implique ${}^t\mathbf{x}\mathbf{B}\mathbf{y} = 0$.

9.3.4 Proposition

Soit **A** *une matrice symétrique et* **B** *une matrice symétrique définie positive, toutes deux d'ordre n. Une matrice carrée* **P** *d'ordre n vérifie les relations* (9.13) *si* *et seulement si ses colonnes* $\mathbf{p}_1, \mathbf{p}_2, ..., \mathbf{p}_n$ *sont des vecteurs propres de* $(\mathbf{A}, \mathbf{B})$ *et forment une famille orthonormale pour le produit scalaire défini par* $\mathbf{B}$.

DÉMONSTRATION

Commençons par relever qu'une matrice carrée $\mathbf{P}$ vérifie la relation ${}^t\mathbf{PBP} = \mathbf{I}$ si et seulement si ses colonnes forment une famille orthonormale pour le produit scalaire défini par $\mathbf{B}$. Cela dit, considérons une matrice carrée $\mathbf{P}$ vérifiant cette relation. La relation ${}^t\mathbf{PAP} = \mathbf{D}$ s'écrit alors, de manière équivalente, sous la forme

$$ {}^t\mathbf{PAP} = {}^t\mathbf{PBPD}, $$

ou encore

$$ \mathbf{AP} = \mathbf{BPD}. $$

En égalant les colonnes respectives des deux membres de cette égalité, nous concluons que ${}^t\mathbf{PAP} = \mathbf{D}$ si et seulement si

$$ \mathbf{Ap}_j = d_j\mathbf{Bp}_j, \quad \text{ou encore} \quad (\mathbf{A} - d_j\mathbf{B})\mathbf{p}_j = \mathbf{0}, \quad \text{pour } j = 1, 2, ..., n, $$

où $\mathbf{p}_1, \mathbf{p}_2, ..., \mathbf{p}_n$ sont les colonnes de $\mathbf{P}$ et $d_1, d_2, ..., d_n$ les termes diagonaux de $\mathbf{D}$. La proposition est ainsi démontrée.

9.3.5 Exemple

Dans un espace vectoriel de dimension 2, muni d'une base $(\mathbf{e}_1, \mathbf{e}_2)$, deux formes bilinéaires symétriques f et g sont définies par leurs matrices respectives

$$ \mathbf{A} = \begin{pmatrix} 1 & -1 \\ -1 & -2 \end{pmatrix} \quad \text{et} \quad \mathbf{B} = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}. $$

Nous nous proposons de réduire f et g simultanément. Les mineurs principaux de $\mathbf{B}$ sont positifs, donc g est définie positive et la proposition 9.3.4 s'applique. Les valeurs propres de $(\mathbf{A}, \mathbf{B})$ sont les solutions de l'équation

$$ |\mathbf{A} - \lambda\mathbf{B}| = \begin{vmatrix} 1-2\lambda & -1+\lambda \\ -1+\lambda & -2-2\lambda \end{vmatrix} = 3\lambda^2 + 4\lambda - 3 = 0, $$

à savoir

$$ \frac{-2+\sqrt{13}}{3} \quad \text{et} \quad \frac{-2-\sqrt{13}}{3}. $$

Deux vecteurs propres de $(\mathbf{A}, \mathbf{B})$ associés respectivement à ces valeurs propres et unitaires pour le produit scalaire défini par $\mathbf{B}$ sont

$$ \mathbf{p}_1 = \frac{1}{(6(26-7\sqrt{13}))^{1/2}} \begin{pmatrix} \sqrt{13}-5 \\ 2\sqrt{13}-7 \end{pmatrix} \quad \text{et} \quad \mathbf{p}_2 = \frac{1}{(6(26-7\sqrt{13}))^{1/2}} \begin{pmatrix} \sqrt{13}+5 \\ 2\sqrt{13}+7 \end{pmatrix}. $$

Définissons la base $(\mathbf{e}_1', \mathbf{e}_2')$ par

$$(\mathbf{e}_1')_e = \mathbf{p}_1 \quad \text{et} \quad (\mathbf{e}_2')_e = \mathbf{p}_2.$$

La matrice de passage $\mathbf{P}_{ee'}$ vérifie alors les relations (9.13), autrement dit, les matrices respectives de f et de g dans la base $(\mathbf{e}_1', \mathbf{e}_2')$ sont

$$\operatorname{diag}\left(\frac{-2+\sqrt{13}}{3}, \frac{-2-\sqrt{13}}{3}\right) \quad \text{et} \quad \mathbf{I},$$

ce qui équivaut à dire que

$$f(\mathbf{x}, \mathbf{y}) = \frac{-2+\sqrt{13}}{3}x_1'y_1' - \frac{2+\sqrt{13}}{3}x_2'y_2',$$

$$g(\mathbf{x}, \mathbf{y}) = x_1'y_1' + x_2'y_2',$$

où x_1', x_2' et y_1', y_2' sont les composantes respectives de $\mathbf{x}$ et de $\mathbf{y}$ dans la base $(\mathbf{e}_1', \mathbf{e}_2')$.

9.4 EXERCICES

9.4.1 Dans chacun des cas suivants, réduire la forme quadratique q définie, dans une base $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ d'un espace vectoriel E, par
(a) $q(\mathbf{x}) = 2x_1^2 - 2\sqrt{3}x_1x_2 \quad (n = 2)$;
(b) $q(\mathbf{x}) = x_1^2 + 2\sqrt{5}x_1x_2 + 2x_2x_3 + x_3^2 \quad (n = 3)$;
(c) $q(\mathbf{x}) = x_1^2 + 4x_1x_3 - x_2^2 - 2x_3^2 \quad (n = 3)$;
(d) $q(\mathbf{x}) = x_1^2 + 2x_1x_3 + 2x_2^2 + 2x_3x_4 + x_4^2 \quad (n = 4)$.

9.4.2 Soit E un espace vectoriel euclidien, orienté et de dimension 3. Soit en outre $\mathbf{v}$ un vecteur non nul de E. Pour tout couple $(\mathbf{x}, \mathbf{y})$ de vecteurs de E, posons

$$f(\mathbf{x}, \mathbf{y}) = (\mathbf{x} \mid (\mathbf{v} \times \mathbf{y}) \times \mathbf{v}).$$

(a) Montrer que f est une forme bilinéaire symétrique dans E. (Utiliser le produit mixte ou (2.50).)
(b) Montrer que f est définie non négative.
(c) Trouver une base orthonormale de E dans laquelle f est réduite.

9.4.3 *Réduction par complétion des carrés* (cf. exemple (2) de 9.1.8). Soit q la forme quadratique définie, dans une base donnée, par

$$q(\mathbf{x}) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j,$$

où $a_{ij} = a_{ji}$ pour tout couple (i, j) d'indices.

(a) On suppose que $a_{11} \neq 0$ et on pose

$$x_1' = x_1 + \frac{1}{a_{11}}(a_{12}x_2 + ... + a_{1n}x_n), \quad x_2' = x_2, \quad ..., x_n' = x_n.$$

Montrer que

$$q(\mathbf{x}) = a_{11}x_1'^2 + \sum_{i=2}^{n}\sum_{j=2}^{n} a'_{ij}x'_ix'_j,$$

où $a'_{ij} = a'_{ji}$ pour tout couple (i, j) d'indices $\geqslant 2$.

(b) On suppose que $a_{11} = 0$, mais que $a_{1k} \neq 0$, et on pose

$$x'_1 = x_1 + x_k, \quad x'_k = x_1 - x_k, \quad x'_j = x_j \text{ pour } j = 2, ..., \hat{k}, ..., n.$$

Montrer que

$$q(\mathbf{x}) = \sum_{i=1}^{n}\sum_{j=1}^{n} a'_{ij}x'_ix'_j,$$

où $a'_{11} \neq 0$ et $a'_{ij} = a'_{ji}$ pour tout couple (i, j) d'indices.

9.4.4 *Loi d'inertie de Sylvester.* Soit

$$d_1x_1^2 + ... + d_kx_k^2 - d_{k+1}x_{k+1}^2 - ... - d_{k+l}x_{k+l}^2$$

et

$$d'_1x_1'^2 + ... + d'_rx_r'^2 - d'_{r+1}x_{r+1}'^2 - ... - d'_{r+s}x_{r+s}'^2$$

deux formes réduites d'une même forme quadratique, où $d_1, ..., d_k, ..., d_{k+l}$ et $d'_1, ..., d'_r, ..., d'_{r+s}$ sont des nombres positifs. Montrer que $k = r$ et $l = s$ (autrement dit, que le nombre de coefficients positifs et le nombre de coefficients négatifs sont tous deux déterminés par la forme quadratique). (De (9.5) et (a) de l'exercice 6.8.7, déduire que $k + l = r + s$; montrer ensuite que l'hypothèse $k > r$ conduit à une contradiction, en égalant les deux formes réduites et en établissant l'existence d'un vecteur non nul $\mathbf{x}$ dont les composantes $x'_1, ..., x'_r$ et $x_{k+1}, ..., x_n$ sont nulles, où n est la dimension de l'espace vectoriel dans lequel la forme quadratique est définie.)

9.4.5 Montrer que si la matrice d'une forme bilinéaire symétrique dans une base de E est définie non négative (définie positive), alors elle l'est également dans toute autre base de E.

9.4.6 Pour quelles valeurs de a la matrice

$$\begin{pmatrix} 1 & a & a \\ a & 1 & a \\ a & a & 1 \end{pmatrix}$$

est-elle définie positive?

9.4.7 On suppose qu'une fonction réelle f admette, au voisinage de $\mathbf{0} \in \mathbb{R}^3$, le développement limité

$$f(\mathbf{x}) = f(\mathbf{0}) - 3x_1^2 - 2x_1x_2 + 2x_1x_3 - 2x_2^2 + 4x_2x_3 - 4x_3^2 + \|\mathbf{x}\|^2\varepsilon(\mathbf{x}).$$

Montrer que $f(\mathbf{0})$ est un maximum relatif strict.

9.4.8 Montrer qu'une matrice symétrique définie non négative est définie positive si et seulement si son déterminant est non nul.

9.4.9 Soit **A** une matrice de type $m \times n$. Démontrer:
(a) ${}^t\mathbf{A}\mathbf{A}$ est une matrice symétrique définie non négative.
(b) Si **A** est de rang n, ${}^t\mathbf{A}\mathbf{A}$ est une matrice symétrique définie positive.

9.4.10 On appelle *racine carrée* d'une matrice carrée **A** toute matrice carrée **B** telle que $\mathbf{B}^2 = \mathbf{A}$. Montrer que toute matrice symétrique définie non négative (définie positive) admet une unique racine carrée définie non négative (définie positive).

9.4.11 Soit **A** une matrice symétrique définie non négative (définie positive). Montrer que pour tout entier non négatif k (tout entier k), $\mathbf{A}^k$ est définie non négative (définie positive).

9.4.12 Soit **A** une matrice symétrique définie non négative (définie positive). Montrer que tout vecteur propre de $\mathbf{A}^k$, où k est un entier positif (un entier non nul), est un vecteur propre de **A**. Exhiber une matrice symétrique d'ordre 2 dont le carré admet un vecteur propre qui n'est pas un vecteur propre de **A**.

9.4.13 Soit **A** et **B** deux matrices symétriques définies non négatives (définies positives). Montrer que si $\mathbf{A}^k = \mathbf{B}^k$ pour un entier positif k (pour un entier non nul k), alors $\mathbf{A} = \mathbf{B}$.

9.4.14 Soit **A** une matrice symétrique d'ordre n et **B** une matrice symétrique définie positive également d'ordre n. On désigne par $\lambda_{\min}$ et $\lambda_{\max}$ la plus petite et la plus grande valeur propre de (**A**, **B**) (cf. 9.3.3).
(a) Montrer que

$$\min_{\mathbf{x}\in\mathbb{R}^n-\{\mathbf{0}\}} \frac{{}^t\mathbf{x}\mathbf{A}\mathbf{x}}{{}^t\mathbf{x}\mathbf{B}\mathbf{x}} = \lambda_{\min} \quad \text{et} \quad \max_{\mathbf{x}\in\mathbb{R}^n-\{\mathbf{0}\}} \frac{{}^t\mathbf{x}\mathbf{A}\mathbf{x}}{{}^t\mathbf{x}\mathbf{B}\mathbf{x}} = \lambda_{\max}.$$

(Utiliser la proposition 9.3.2.)

(b) A l'aide de (a), trouver le maximum et le minimum de la fonction réelle définie dans $\mathbb{R}^3-\{\mathbf{0}\}$ par

$$\frac{2x_1x_3 + x_2^2 - 3x_3^2}{2x_1^2 + 2x_1x_3 + x_2^2 + 3x_3^2}.$$

CHAPITRE 10

Quadriques

10.1 ÉQUATION GÉNÉRALE D'UNE QUADRIQUE

10.1.1 Introduction

L'ellipse, l'hyperbole et la parabole, en dimension 2, l'ellipsoïde, l'hyperboloïde et le paraboloïde, en dimension 3, sont des exemples de quadriques. On appelle ainsi les lieux géométriques définis par une équation du second degré. Le présent chapitre sera consacré à l'étude de ces lieux.

Tout au long du chapitre, $\mathscr{E}$ désignera un espace affine de dimension finie n supérieure à 1 et de direction E. Nous supposerons que $\mathscr{E}$ est muni d'un repère $(O; \mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n)$.

10.1.2 Quadriques

On dit qu'un sous-ensemble de $\mathscr{E}$ est une *quadrique* s'il est formé par les points P dont les coordonnées $x_1, x_2, \dots, x_n$ satisfont à une équation de la forme

$$\sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j + 2 \sum_{i=1}^{n} b_i x_i = c, \tag{10.1}$$

où les coefficients a_{ij} sont des nombres non tous nuls tels que $a_{ij} = a_{ji}$ pour tout couple (i, j) d'indices et les coefficients b_i, ainsi que c, sont des nombres quelconques.

Une quadrique n'est donc autre qu'un lieu géométrique défini par une équation du second degré (cf. exercice 10.7.1).

Dans le cas des dimensions 2 et 3, l'équation (10.1) s'écrit, sans l'aide du symbole de sommation,

$$a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 + 2b_1x_1 + 2b_2x_2 = c$$

et

$$a_{11}x_1^2 + 2a_{12}x_1x_2 + 2a_{13}x_1x_3 + a_{22}x_2^2 + 2a_{23}x_2x_3 + a_{33}x_3^2 + \\ 2b_1x_1 + 2b_2x_2 + 2b_3x_3 = c.$$

Voici quelques exemples:

(1) $2x_1^2 + x_2^2 = 3$, ellipse si $n = 2$, cylindre elliptique si $n = 3$.
(2) $3x_1^2 - x_2 = 0$, parabole si $n = 2$, cylindre parabolique si $n = 3$.
(3) $4x_1^2 + x_2^2 + 2x_3^2 = 1$, ellipsoïde.
(4) $x_1^2 + 3x_2^2 - 2x_3^2 = 0$, cône du second degré.
(5) $-x_1^2 - 3x_2^2 - x_3^2 = 1$, quadrique vide.

10.1.3 Ecriture matricielle

Désignons par $\mathbf{A}$ la matrice symétrique (a_{ij}), par $\mathbf{b}$ le vecteur-colonne (b_i) et par $\mathbf{x}$ le vecteur-colonne $(\overrightarrow{OP})_e$. L'équation (10.1) s'écrit, de manière équivalente, sous la forme

$$ {}^t\mathbf{x}\mathbf{A}\mathbf{x} + 2\,{}^t\mathbf{b}\mathbf{x} = c. \tag{10.2} $$

Le premier terme du premier membre est appelé *partie quadratique*, le deuxième terme *partie linéaire* et le second membre *terme constant* de l'équation. Nous dirons que $\mathbf{A}$ est la matrice de la partie quadratique et $\mathbf{b}$ le vecteur-colonne de la partie linéaire.

L'objectif principal de ce chapitre est la réduction de l'équation (10.2) à une forme simplifiée dite normale. C'est par cette forme que la configuration du lieu géométrique peut être étudiée le plus aisément.

10.1.4 Hyperplan tangent

Pour former l'équation cartésienne de l'*hyperplan tangent* en un point de la quadrique d'équation (10.2), il nous faut calculer les dérivées partielles du premier membre de (10.2) ou (10.1) par rapport à $x_1, x_2, ..., x_n$. Le lecteur s'assurera facilement que le vecteur-colonne formé de ces dérivées partielles, c'est-à-dire le gradient du premier membre de (10.2) ou (10.1), est (cf. exercice 10.7.17)

$$ \operatorname{grad}({}^t\mathbf{x}\mathbf{A}\mathbf{x} + 2\,{}^t\mathbf{b}\mathbf{x}) = 2(\mathbf{A}\mathbf{x} + \mathbf{b}). \tag{10.3} $$

Fixons un point P_0 de la quadrique et désignons par $\mathbf{x}_0 = (x_i^0)$ le vecteur-colonne $(\overrightarrow{OP_0})_e$. Supposons que le gradient en $\mathbf{x}_0$ soit non nul. L'équation de l'hyperplan tangent en P_0 est alors de la forme

$$ {}^t(\mathbf{A}\mathbf{x}_0 + \mathbf{b})\mathbf{x} = \delta, $$

où $\mathbf{x}$ désigne le vecteur-colonne des coordonnées du point générique de cet hyperplan. Or, P_0 appartient à l'hyperplan tangent et à la quadrique, donc

$$ \delta = {}^t\mathbf{x}_0\mathbf{A}\mathbf{x}_0 + {}^t\mathbf{b}\mathbf{x}_0 = {}^t\mathbf{x}_0\mathbf{A}\mathbf{x}_0 + 2\,{}^t\mathbf{b}\mathbf{x}_0 - {}^t\mathbf{b}\mathbf{x}_0 = c - {}^t\mathbf{b}\mathbf{x}_0. $$

Nous en concluons que l'équation de l'hyperplan tangent à la quadrique en P_0 est

$$ {}^t(\mathbf{A}\mathbf{x}_0 + \mathbf{b})\mathbf{x} = c - {}^t\mathbf{b}\mathbf{x}_0 \quad \text{ou} \quad {}^t\mathbf{x}_0\mathbf{A}\mathbf{x} + {}^t\mathbf{b}(\mathbf{x} + \mathbf{x}_0) = c, \tag{10.4} $$

ou encore,

$$ \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_i^0 x_j + \sum_{i=1}^{n} b_i(x_i + x_i^0) = c. \tag{10.5} $$

Dans le cas des dimensions 2 et 3, l'équation (10.5) s'écrit, sans l'aide du symbole de sommation,

$$ a_{11}x_1^0x_1 + a_{12}(x_1^0x_2 + x_2^0x_1) + a_{22}x_2^0x_2 + b_1(x_1 + x_1^0) + b_2(x_2 + x_2^0) = c, $$

et

$$a_{11}x_1^0x_1 + a_{12}(x_1^0x_2 + x_2^0x_1) + a_{13}(x_1^0x_3 + x_3^0x_1) + a_{22}x_2^0x_2 +$$
$$a_{23}(x_2^0x_3 + x_3^0x_2) + a_{33}x_3^0x_3 +$$
$$b_1(x_1 + x_1^0) + b_2(x_2 + x_2^0) + b_3(x_3 + x_3^0) = c.$$

Par exemple, l'équation du plan tangent à la quadrique d'équation

$$2x_1^2 - x_2^2 + 6x_2x_3 - x_3^2 - 4x_1 + 6x_2 - 2x_3 = -\frac{3}{2} \tag{10.6}$$

au point $P_0(\frac{1}{2}, 1, -1)$ est

$$-x_1 - x_2 + 3x_3 = -\frac{9}{2}.$$

10.1.5 Vecteur normal

Lorsque $\mathscr{E}$ est euclidien et le repère $(O; \mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ est orthonormal, le vecteur défini par

$$\mathbf{Ax}_0 + \mathbf{b}$$

est normal à l'hyperplan tangent à la quadrique au point P_0. On l'appelle *vecteur normal* à la quadrique en P_0.

10.2 CENTRAGE

10.2.1 Centres de symétrie

On dit qu'un point C de $\mathscr{E}$ est un *centre de symétrie* ou simplement un *centre* d'une quadrique si cette quadrique est symétrique par rapport à C, autrement dit, si elle est stable par la symétrie centrale de centre C.

Les exemples de 10.1.2 montrent qu'une quadrique peut avoir une infinité de centres (cylindre elliptique (1)), n'en avoir aucun (parabole (2)) ou en avoir exactement un (ellipsoïde (3) et cône (4)).

Manifestement, une quadrique d'équation

$${}^t\mathbf{x}\mathbf{A}\mathbf{x} = c \tag{10.7}$$

est symétrique par rapport à l'origine du repère. Cette origine est donc un centre de la quadrique.

10.2.2 Effet d'un changement d'origine

Choisissons une nouvelle origine O' et munissons $\mathscr{E}$ du repère $(O'; \mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$. Désignons par $\mathbf{y}$ et $\mathbf{x}'$ les vecteurs-colonnes $(\overrightarrow{OO'})_e$ et $(\overrightarrow{O'P})_e$ (comme auparavant, $\mathbf{x}$ désigne le vecteur-colonne $(\overrightarrow{OP})_e$). Il est évident que

$$\mathbf{x} = \mathbf{x}' + \mathbf{y}.$$

Considérons la quadrique d'équation (10.2) et calculons sa nouvelle équation. Par substitution de $\mathbf{x}' + \mathbf{y}$ à $\mathbf{x}$ dans (10.2), nous obtenons

$$\begin{aligned}&{}^t(\mathbf{x}' + \mathbf{y})\mathbf{A}(\mathbf{x}' + \mathbf{y}) + 2\,{}^t\mathbf{b}(\mathbf{x}' + \mathbf{y})\\ &\qquad = {}^t\mathbf{x}'\mathbf{A}\mathbf{x}' + {}^t\mathbf{x}'\mathbf{A}\mathbf{y} + {}^t\mathbf{y}\mathbf{A}\mathbf{x}' + {}^t\mathbf{y}\mathbf{A}\mathbf{y} + 2\,{}^t\mathbf{b}\mathbf{x}' + 2\,{}^t\mathbf{b}\mathbf{y} = c.\end{aligned}$$

Or, puisqu'une matrice carrée d'ordre 1 est égale à sa transposée,

$${}^t\mathbf{x}'\mathbf{A}\mathbf{y} = {}^t\mathbf{y}\mathbf{A}\mathbf{x}' = {}^t(\mathbf{A}\mathbf{y})\mathbf{x}',$$

ce qui nous permet de conclure que l'équation de la quadrique dans le repère $(O'; \mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ est

$${}^t\mathbf{x}'\mathbf{A}\mathbf{x}' + 2\,{}^t(\mathbf{A}\mathbf{y} + \mathbf{b})\mathbf{x}' = c', \tag{10.8}$$

où

$$c' = c - ({}^t\mathbf{y}\mathbf{A}\mathbf{y} + 2\,{}^t\mathbf{b}\mathbf{y}).$$

10.2.3 Centrage

Supposons que $\mathbf{y}$ vérifie la relation

$$\mathbf{A}\mathbf{y} + \mathbf{b} = \mathbf{0}. \tag{10.9}$$

Alors la partie linéaire de l'équation (10.8) s'annule et la nouvelle origine O' est un centre de la quadrique. En outre, de (10.9) nous déduisons la relation

$${}^t\mathbf{y}\mathbf{A}\mathbf{y} + {}^t\mathbf{y}\mathbf{b} = 0,$$

qui entraîne

$${}^t\mathbf{y}\mathbf{A}\mathbf{y} = -{}^t\mathbf{y}\mathbf{b} = -{}^t\mathbf{b}\mathbf{y}$$

et donc

$$c' = c - {}^t\mathbf{b}\mathbf{y}.$$

Il en résulte que dans un repère $(O'; \mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ tel que $\mathbf{y} = (\overrightarrow{OO'})_{\mathrm{e}}$ satisfait à (10.9), l'équation de la quadrique s'écrit sous la forme

$${}^t\mathbf{x}'\mathbf{A}\mathbf{x}' = c - {}^t\mathbf{b}\mathbf{y}. \tag{10.10}$$

On notera que la matrice de la partie quadratique n'a subi aucune modification.

10.2.4 Quadriques sans centre

Si l'équation (10.9) n'admet aucune solution, la partie linéaire de l'équation (10.2) ne peut être éliminée par un changement de l'origine du repère. Nous verrons dans 10.4.2 que la quadrique n'a, dans ce cas, aucun centre.

10.2.5 Quadriques à centre unique

D'après la proposition 3.5.9, pour que l'équation (10.9) admette une solution unique, il faut et il suffit que le rang de **A** soit n, autrement dit, que **A** soit inversible. Vu ce qui vient d'être dit dans 10.2.4, la quadrique d'équation (10.2) a donc un centre unique si et seulement si la matrice **A** est inversible.

10.2.6 Exemple

La matrice de la partie quadratique de l'équation (10.6) est

$$\mathbf{A} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & -1 & 3 \\ 0 & 3 & -1 \end{bmatrix}.$$

Cette matrice est inversible, donc la quadrique d'équation (10.6) possède un centre unique. Pour le déterminer, nous résolvons le système équivalent à l'équation (10.9)

$$\begin{aligned} 2y_1 \qquad\qquad &= 2 \\ -y_2 + 3y_3 &= -3 \\ 3y_2 - y_3 &= 1. \end{aligned}$$

L'unique solution de ce système est $y_1 = 1$, $y_2 = 0$, $y_3 = -1$. Soit O' le point de coordonnées 1, 0, -1. D'après (10.10), l'équation de la quadrique dans le repère $(O'; \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ est

$$2x_1'^2 - x_2'^2 + 6x_2'x_3' - x_3'^2 = -\frac{3}{2} - \begin{bmatrix} -2 & 3 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} = -\frac{1}{2}.$$

10.3 RÉDUCTION DE L'ÉQUATION D'UNE QUADRIQUE À CENTRE

10.3.1 Avertissement

Dès maintenant, nous admettrons que l'espace affine $\mathscr{E}$ est euclidien (cf. 2.8.2 et 2.5.7). Le repère $(O; \mathbf{e}_1, \mathbf{e}_2 \ldots, \mathbf{e}_n)$ dont il est muni n'est cependant pas supposé orthonormal.

10.3.2 Réduction de la partie quadratique

Considérons une quadrique ayant au moins un centre. Quitte à choisir ce centre comme origine du repère, nous pouvons admettre que son équation dans le repère $(O; \mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ est l'équation (10.7). D'après la proposition 9.3.2, il existe une base orthonormale $(\mathbf{e}'_1, \mathbf{e}'_2, ..., \mathbf{e}'_n)$ dans laquelle la forme quadratique dans E définie par le premier membre de (10.7) est réduite. En d'autres termes, l'équation de la quadrique dans le repère orthonormal $(O; \mathbf{e}'_1, \mathbf{e}'_2, ..., \mathbf{e}'_n)$ s'écrit sous la forme, dite *réduite*,

$$d_1 x_1'^2 + d_2 x_2'^2 + ... + d_n x_n'^2 = c. \tag{10.11}$$

On appelle les droites vectorielles engendrées respectivement par les vecteurs $\mathbf{e}'_1, \mathbf{e}'_2, ..., \mathbf{e}'_n$ *directions principales de la quadrique*.

A noter que si la base initiale $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ est orthonormale, les vecteurs-colonnes $(\mathbf{e}'_1)_e, (\mathbf{e}'_2)_e, ..., (\mathbf{e}'_n)_e$ forment une famille orthonormale de vecteurs propres de la matrice $\mathbf{A}$; en outre, $d_1, d_2, ..., d_n$ sont les valeurs propres correspondantes. En effet, la matrice de passage $\mathbf{P}_{ee'}$ est, dans ce cas, orthogonale, donc la relation (9.5) (avec $\mathbf{A}'$ diagonale) s'identifie aux relations (8.23) ou (8.24) (avec $\mathbf{P} = \mathbf{P}_{ee'}$), d'où les assertions résultent.

10.3.3 Tableau des quadriques non vides à centre dans le cas des dimensions 2 et 3

Quitte à diviser les deux membres de l'équation (10.11) par c, si c n'est pas nul, nous pouvons admettre que son terme constant vaut 1 ou 0. En posant alors $a_i = 1/(|d_i|)^{1/2}$ pour tout i tel que d_i n'est pas nul, nous voyons que cette équation se réduit, après renumérotation éventuelle des vecteurs de la base, à l'une des *formes normales* suivantes:

(1) Cas où $n = 2$ (coniques).

Le(s) centre(s) n'appartien(nen)t pas à la quadrique:

$\dfrac{x_1^2}{a_1^2} + \dfrac{x_2^2}{a_2^2} = 1$ ellipse (cercle si $a_1 = a_2$) (fig. 10.1);

$\dfrac{x_1^2}{a_1^2} - \dfrac{x_2^2}{a_2^2} = 1$ hyperbole (équilatère si $a_1 = a_2$) (fig. 10.2);

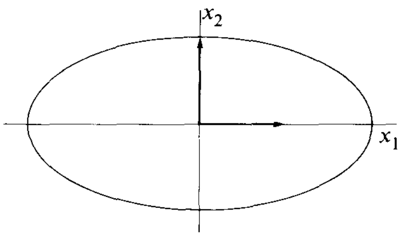
$\dfrac{x_1^2}{a_1^2} = 1$ paire de droites parallèles.

Le(s) centre(s) appartien(nen)t à la quadrique:

$\dfrac{x_1^2}{a_1^2} + \dfrac{x_2^2}{a_2^2} = 0$ point (origine);

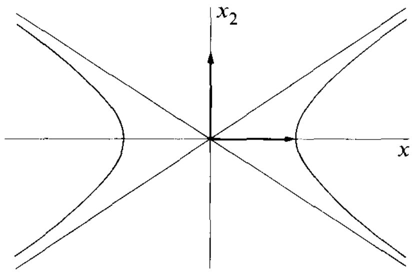
$\dfrac{x_1^2}{a_1^2} - \dfrac{x_2^2}{a_2^2} = 0$ paire de droites concourantes;

$x_1^2 = 0$ droite.



ellipse ($a_1 = 2$, $a_2 = 1$)

Fig. 10.1



hyperbole ($a_1 = 1$, $a_2 = \frac{2}{3}$)

Fig. 10.2

(2) Cas où $n = 3$.

Le(s) centre(s) n'appartien(nen)t pas à la quadrique:

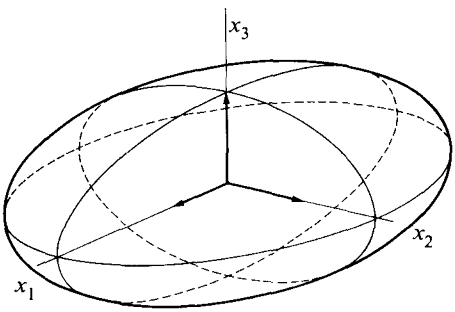
$\frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} + \frac{x_3^2}{a_3^2} = 1$	ellipsoïde (de révolution si $a_i = a_j$ pour un couple (i, j) d'indices; sphère si $a_1 = a_2 = a_3$) (fig. 10.3);
$\frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} - \frac{x_3^2}{a_3^2} = 1$	hyperboloïde à une nappe (de révolution si $a_1 = a_2$) (fig. 10.4);
$-\frac{x_1^2}{a_1^2} - \frac{x_2^2}{a_2^2} + \frac{x_3^2}{a_3^2} = 1$	hyperboloïde à deux nappes (de révolution si $a_1 = a_2$) (fig. 10.5);
$\frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} = 1$	cylindre elliptique (cylindre de révolution si $a_1 = a_2$) (fig. 10.6);
$\frac{x_1^2}{a_1^2} - \frac{x_2^2}{a_2^2} = 1$	cylindre hyperbolique (fig. 10.7);
$\frac{x_1^2}{a_1^2} = 1$	paire de plans parallèles.

Le(s) centre(s) appartien(nen)t à la quadrique:

$\frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} + \frac{x_3^2}{a_3^2} = 0$	point (origine);
$\frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} - \frac{x_3^2}{a_3^2} = 0$	cône du second degré (cône de révolution si $a_1 = a_2$) (fig. 10.8);
$\frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} = 0$	droite;

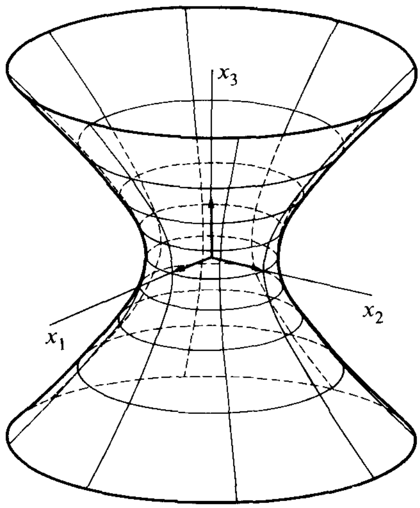
$$\frac{x_1^2}{a_1^2} - \frac{x_2^2}{a_2^2} = 0 \qquad \text{paire de plans;}$$

$$x_1^2 = 0 \qquad \text{plan.}$$



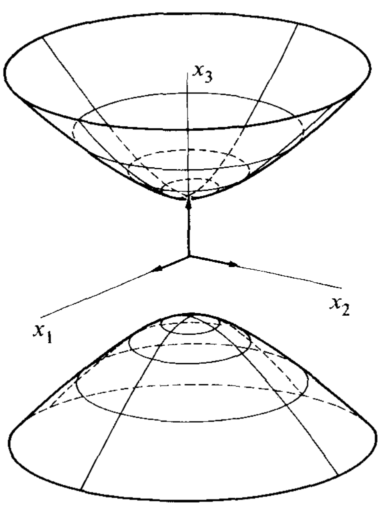
ellipsoïde ($a_1 = 3$, $a_2 = 2$, $a_3 = 1$)

Fig. 10.3



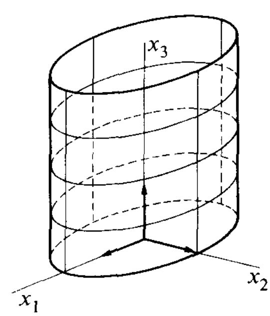
hyperboloïde à une nappe
($a_1 = a_2 = a_3 = 1$)

Fig. 10.4



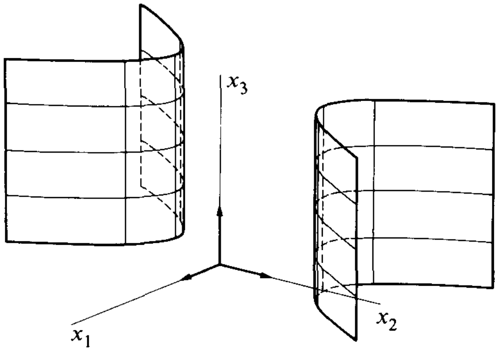
hyperboloïde à deux nappes
($a_1 = a_2 = a_3 = 1$)

Fig. 10.5



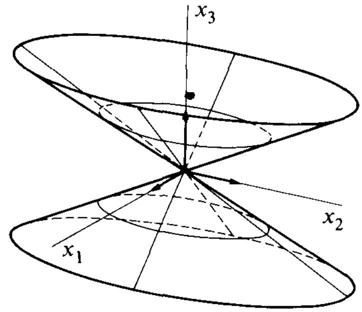
cylindre elliptique ($a_1 = 2$, $a_2 = 1$)

Fig. 10.6



cylindre hyperbolique ($a_1 = 2$, $a_2 = 1$)

Fig. 10.7



cône du second degré
($a_1 = 1$, $a_2 = 2$, $a_3 = 1$)

Fig. 10.8

10.4 RÉDUCTION DE L'ÉQUATION D'UNE QUADRIQUE SANS CENTRE

10.4.1 Réduction de l'équation

Considérons une quadrique n'ayant aucun centre. Par le changement de base effectué dans 10.3.2, nous obtenons un repère orthonormal dans lequel l'équation de la quadrique s'écrit sous la forme (10.2), avec $\mathbf{A}$ diagonale. Sans restreindre la généralité, nous pouvons donc admettre que le repère initial $(O; \mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ est orthonormal et que l'équation de la quadrique dans ce repère est

$$d_1x_1^2 + d_2x_2^2 + ... + d_nx_n^2 + 2b_1x_1 + 2b_2x_2 + ... + 2b_nx_n = c. \qquad (10.12)$$

Posons $\mathbf{D} = \mathrm{diag}(d_1, d_2, ..., d_n)$ et $\mathbf{b} = (b_i)$. D'après 10.2.3, l'équation

$$\mathbf{Dy} + \mathbf{b} = \mathbf{0} \qquad (10.13)$$

n'admet aucune solution, autrement la partie linéaire pourrait être éliminée de l'équation (10.12) et la quadrique aurait au moins un centre, contrairement à l'hypothèse. D'après la proposition 3.2.7, le rang de la matrice augmentée $(\mathbf{D} \mid \mathbf{b})$ est donc supérieur au rang de $\mathbf{D}$. Il en découle deux conséquences. Premièrement, le rang r de $\mathbf{D}$ est inférieur à n, ce qui entraîne que r termes diagonaux de $\mathbf{D}$ sont non nuls, par exemple $d_1, d_2, ..., d_r$, et les restants $d_{r+1}, d_{r+2}, ..., d_n$ sont nuls. Deuxièmement, au moins un des termes $b_{r+1}, b_{r+2}, ..., b_n$ est non nul, par exemple b_{r+1}. Prenons comme nouvelle origine du repère le point O' de coordonnées

$$-\frac{b_1}{d_1}, \; -\frac{b_2}{d_2}, \; ..., \; -\frac{b_r}{d_r}, \; \frac{c'}{2b_{r+1}}, \; 0, ..., \; 0, \qquad (10.14)$$

où

$$c' = c + \frac{b_1^2}{d_1} + \dots + \frac{b_r^2}{d_r}.$$

L'équation de la quadrique dans le repère orthonormal $(O'; \mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n)$ s'écrit alors sous la forme, dite *réduite*,

$$d_1 x_1'^2 + d_2 x_2'^2 + \dots + d_r x_r'^2 + 2b_{r+1}x'_{r+1} + \dots + 2b_n x'_n = 0. \tag{10.15}$$

On appelle les droites vectorielles engendrées respectivement par les vecteurs $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ *directions principales de la quadrique.*

10.4.2 Caractérisation des quadriques sans centre

Nous avons vu que l'élimination de la partie linéaire de l'équation (10.2) par un changement de l'origine du repère n'est possible que dans le cas où l'équation (10.9) admet au moins une solution. Observons maintenant qu'un changement de la base du repère n'entraîne pas la disparition de la partie linéaire, car le vecteur-colonne $\mathbf{b}$ se transforme en ${}^t\mathbf{P}\mathbf{b}$, où $\mathbf{P}$ désigne la matrice de passage. Cela dit, supposons que l'équation (10.9) n'admette aucune solution, mais que la quadrique ait au moins un centre. Choisissons ce centre comme origine du repère et transformons l'équation (10.2) en une équation de la forme (10.12) par un changement de base approprié. La partie linéaire ne pouvant être éliminée, par hypothèse, l'équation (10.13) n'admet aucune solution. Nous en concluons qu'il existe des indices i et j tels que $d_i \neq 0$, $d_j = 0$ et $b_j \neq 0$. Considérons alors la section de la quadrique par le plan passant par l'origine, de vecteurs directeurs $\mathbf{e}_i$ et $\mathbf{e}_j$. Cette section est symétrique par rapport à l'origine, puisque la quadrique l'est. Or, cela n'est pas possible, puisque son équation

$$d_i x_i^2 + 2b_i x_i + 2b_j x_j = c$$

est celle d'une parabole, c'est-à-dire d'une courbe qui n'a aucun centre de symétrie.

Nous avons ainsi démontré qu'en l'absence de solutions de l'équation (10.9), la quadrique d'équation (10.2) n'a aucun centre.

10.4.3 Tableau des quadriques sans centre dans le cas des dimensions 2 et 3

En divisant ses deux membres par $2b_{r+1}$ et en posant $a_i = (2\,|\,b_{r+1}\,|/|\,d_i\,|)^{1/2}$ pour $i = 1, 2, \dots, r$, nous voyons que l'équation (10.15) se réduit, après changement éventuel de $\mathbf{e}_{r+1}$ en $-\mathbf{e}_{r+1}$, à l'une des *formes normales* suivantes:

(1) Cas où $n = 2$ (coniques).

$\dfrac{x_1^2}{a_1^2} - x_2 = 0$ parabole (fig. 10.9).

(2) Cas où $n = 3$.

$\frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} - x_3 = 0$ paraboloïde elliptique (paraboloïde de révolution si $a_1 = a_2$) (fig. 10.10);

$\frac{x_1^2}{a_1^2} - \frac{x_2^2}{a_2^2} - x_3 = 0$ paraboloïde hyperbolique (fig. 10.11);

$\frac{x_1^2}{a_1^2} - x_2 = 0$ cylindre parabolique (fig. 10.12).

Note. Par les opérations indiquées au début du paragraphe, l'équation que l'on obtient en dernier est en fait de la forme

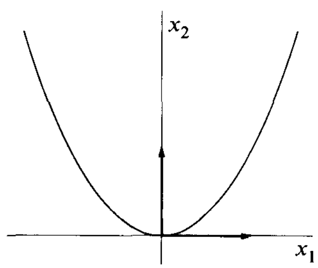
$$\frac{x_1^2}{a_1^2} - x_2 + bx_3 = 0.$$

Pour supprimer le terme bx_3, il faut encore effectuer le changement de base

$$\mathbf{e}_1' = \mathbf{e}_1, \quad \mathbf{e}_2' = \cos\theta\mathbf{e}_2 + \sin\theta\mathbf{e}_3, \quad \mathbf{e}_3' = -\sin\theta\mathbf{e}_2 + \cos\theta\mathbf{e}_3,$$

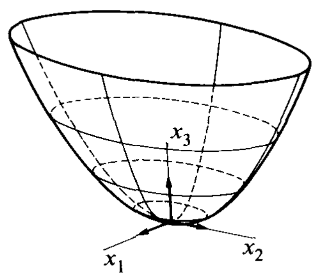
où

$$\cos\theta = \frac{1}{(1+b^2)^{1/2}}, \quad \sin\theta = -\frac{1}{(1+b^2)^{1/2}}.$$



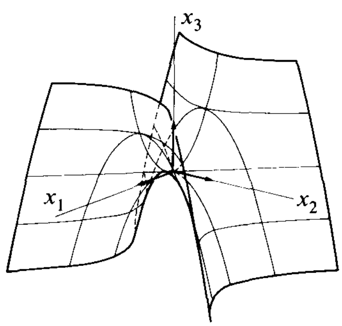
parabole ($a_1 = 1$)

Fig. 10.9



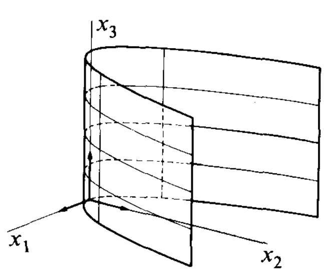
paraboloïde elliptique ($a_1 = 1$, $a_2 = 2$)

Fig. 10.10



paraboloïde hyperbolique ($a_1 = a_2 = 1$)

Fig. 10.11



cylindre parabolique ($a_1 = 1$)

Fig. 10.12

10.5 EXEMPLES DE RÉDUCTION

10.5.1 Introduction

Dans cette section, nous présentons trois exemples de réduction. Dans ces exemples, nous supposerons que l'espace affine euclidien est de dimension 3 et muni d'un repère orthonormal $(O;\, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$.

10.5.2 Exemple

Nous nous proposons d'établir la nature de la quadrique d'équation

$$x_1^2 + 4x_1x_3 + 3x_2^2 + x_3^2 + 2x_1 + 4x_2 + 2x_3 = -1.$$

Centrage. Le système équivalent à l'équation (10.9)

$$\begin{aligned} y_1 \quad\;\; + 2y_3 &= -1 \\ 3y_2 \quad\;\; &= -2 \\ 2y_1 \quad\;\; + \;\; y_3 &= -1 \end{aligned}$$

admet une solution unique, soit $y_1 = -\frac{1}{3}$, $y_2 = -\frac{2}{3}$, $y_3 = -\frac{1}{3}$. Le point $O'(-\frac{1}{3}, -\frac{2}{3}, -\frac{1}{3})$ est donc le seul centre de la quadrique et l'équation de cette quadrique dans le repère $(O';\, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ s'écrit

$$\tilde{x}_1^2 + 4\tilde{x}_1\tilde{x}_3 + 3\tilde{x}_2^2 + \tilde{x}_3^2 = -1 - {}^t\mathbf{by} = 1.$$

Réduction de la partie quadratique. Les racines du polynôme caractéristique

$$\begin{vmatrix} 1-t & 0 & 2 \\ 0 & 3-t & 0 \\ 2 & 0 & 1-t \end{vmatrix} = (3 - t)(t^2 - 2t - 3)$$

sont les entiers 3 (double) et -1. L'équation réduite est donc

$$3x_1'^2 + 3x_2'^2 - x_2'^2 = 1.$$

D'après (2) de 10.3.3, il s'agit d'un hyperboloïde de révolution à une nappe.

10.5.3 Exemple

Nous allons réduire l'équation

$$-2x_1x_2 + 2x_1x_3 - 2x_2^2 - 2x_2x_3 - 2x_3^2 + 2x_1 + 14x_2 + 16x_3 = 38.$$

Centrage. Le système équivalent à l'équation (10.9)

$$\begin{aligned} -y_2 + \;\; y_3 &= -1 \\ -y_1 - 2y_2 - \;\; y_3 &= -7 \\ y_1 - \;\; y_2 - 2y_3 &= -8 \end{aligned}$$

admet une solution unique, soit $y_1 = -1$, $y_2 = 3$, $y_3 = 2$. La quadrique a donc un centre unique $O'(-1, 3, 2)$ et dans le repère $(O'; \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ son équation s'écrit

$$-2\tilde{x}_1\tilde{x}_2 + 2\tilde{x}_1\tilde{x}_3 - 2\tilde{x}_2^2 - 2\tilde{x}_2\tilde{x}_3 - 2\tilde{x}_3^2 = 38 - {}^t\mathbf{b}\mathbf{y} = 2.$$

Réduction de la partie quadratique. Les racines du polynôme caractéristique

$$\begin{vmatrix} -t & -1 & 1 \\ -1 & -2-t & -1 \\ 1 & -1 & -2-t \end{vmatrix} = (t+3)(-t^2 - t + 2)$$

sont les entiers -3, -2 et 1. L'équation réduite est donc

$$-3x_1'^2 - 2x_2'^2 + x_3'^2 = 2.$$

D'après (2) de 10.3.3, il s'agit d'un hyperboloïde à deux nappes.

Repère dans lequel l'équation est réduite. Un repère orthonormal $(O'; \mathbf{e}_1', \mathbf{e}_2', \mathbf{e}_3')$ dans lequel l'équation de la quadrique est réduite est défini par le centre O' et

$$(\mathbf{e}_1')_e = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \quad \text{(vecteur propre unitaire associé à la valeur propre } -3\text{)},$$

$$(\mathbf{e}_2')_e = \frac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} \quad \text{(vecteur propre unitaire associé à la valeur propre } -2\text{)},$$

$$(\mathbf{e}_3')_e = \frac{1}{\sqrt{6}}\begin{pmatrix} -2 \\ 1 \\ -1 \end{pmatrix} \quad \text{(vecteur propre unitaire associé à la valeur propre } 1\text{)}.$$

10.5.4 Exemple

Nous nous proposons de réduire l'équation

$$2x_1^2 + 2x_1x_2 - 2x_1x_3 + 3x_2^2 + 4x_2x_3 + 3x_3^2 + 2x_1 + 4x_2 = 2.$$

Centrage. Le système équivalent à l'équation (10.9)

$$\begin{aligned} 2y_1 + y_2 - y_3 &= -1 \\ y_1 + 3y_2 + 2y_3 &= -2 \\ -y_1 + 2y_2 + 3y_3 &= 0 \end{aligned}$$

n'admet aucune solution, car le rang de la matrice associée est 2 et le rang de la matrice augmentée est 3. La quadrique n'a donc aucun centre.

Réduction de la partie quadratique. Les racines du polynôme caractéristique

$$\begin{vmatrix} 2-t & 1 & -1 \\ 1 & 3-t & 2 \\ -1 & 2 & 3-t \end{vmatrix} = -t(t^2 - 8t + 15)$$

sont les entiers 3, 5 et 0. Une base orthonormale ($\mathbf{e}'_1$, $\mathbf{e}'_2$, $\mathbf{e}'_3$) dans laquelle la partie quadratique est réduite est définie par

$$(\mathbf{e}'_1)_\mathrm{e} = \frac{1}{\sqrt{6}}\begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix} \quad \text{(vecteur propre unitaire associé à la valeur propre 3),}$$

$$(\mathbf{e}'_2)_\mathrm{e} = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \quad \text{(vecteur propre unitaire associé à la valeur propre 5),}$$

$$(\mathbf{e}'_3)_\mathrm{e} = \frac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \quad \text{(vecteur propre unitaire associé à la valeur propre 0).}$$

Le vecteur-colonne de la nouvelle partie linéaire est

$$^t\mathbf{P}_{\mathrm{ee}'}\mathbf{b} = \frac{1}{\sqrt{6}}\begin{pmatrix} 2 & 1 & -1 \\ 0 & \sqrt{3} & \sqrt{3} \\ \sqrt{2} & -\sqrt{2} & \sqrt{2} \end{pmatrix}\begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} = \frac{1}{\sqrt{6}}\begin{pmatrix} 4 \\ 2\sqrt{3} \\ -\sqrt{2} \end{pmatrix}.$$

Dans le repère (O; $\mathbf{e}'_1$, $\mathbf{e}'_2$, $\mathbf{e}'_3$) l'équation de la quadrique est donc

$$3\tilde{x}_1^2 + 5\tilde{x}_2^2 + \frac{8}{\sqrt{6}}\tilde{x}_1 + \frac{4}{\sqrt{2}}\tilde{x}_2 - \frac{2}{\sqrt{3}}\tilde{x}_3 = 2.$$

Equation réduite. Prenons, selon (10.14), le point $O'(-\frac{4}{3\sqrt{6}}, -\frac{\sqrt{2}}{5}, -\frac{74\sqrt{3}}{45})$ (rapporté à (O; $\mathbf{e}'_1$, $\mathbf{e}'_2$, $\mathbf{e}'_3$)) comme nouvelle origine du repère. Dans le repère (O'; $\mathbf{e}'_1$, $\mathbf{e}'_2$, $\mathbf{e}'_3$) l'équation de la quadrique est alors

$$3x_1'^2 + 5x_2'^2 - \frac{2}{\sqrt{3}}x'_3 = 0.$$

D'après (2) de 10.4.3, il s'agit d'un paraboloïde elliptique.

10.6 REPRÉSENTATIONS PARAMÉTRIQUES

10.6.1 Avertissement

Dans cette section, nous supposerons que $\mathscr{E}$ est orienté, de dimension 3 et muni d'un repère orthonormal direct (O; $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$). Nous désignerons par $\mathscr{D}$ la droite passant par O, de direction engendrée et orientée par $\mathbf{e}_3$.

10.6.2 L'hyperboloïde à une nappe comme surface réglée

Considérons la droite d'équations paramétriques

$$\begin{aligned} x_1 &= 1 \\ x_2 &= \alpha \\ x_3 &= \alpha. \end{aligned}$$

L'image de cette droite par la rotation d'axe $\mathscr{D}$ et d'angle orienté β est la droite d'équations paramétriques

$$\begin{aligned} x_1 &= \cos\beta - \alpha\sin\beta \\ x_2 &= \sin\beta + \alpha\cos\beta \\ x_3 &= \alpha. \end{aligned}$$

Lorsque α parcourt $\mathbb{R}$ et β l'intervalle $[0, 2\pi)$, le point de coordonnées x_1, x_2, x_3 décrit une surface engendrée par des droites appelées *génératrices*. Cette surface n'est autre que l'hyperboloïde à une nappe d'équation

$$x_1^2 + x_2^2 - x_3^2 = (\cos\beta - \alpha\sin\beta)^2 + (\sin\beta + \alpha\cos\beta)^2 - \alpha^2 = 1.$$

Plus généralement, l'hyperboloïde à une nappe d'équation

$$\frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} - \frac{x_3^2}{a_3^2} = 1 \tag{10.16}$$

est engendré par les droites de la famille (indexée par β)

$$\begin{aligned} x_1 &= a_1(\cos\beta - \alpha\sin\beta) \\ x_2 &= a_2(\sin\beta + \alpha\cos\beta) \\ x_3 &= a_3\alpha. \end{aligned} \tag{10.17}$$

On appelle une surface engendrée par des droites *surface réglée*. L'hyperboloïde à une nappe est un exemple de surface réglée. Les cônes et les cylindres en sont d'autres.

10.6.3 Equations paramétriques d'une surface de révolution

On appelle *surface de révolution* d'axe $\mathscr{D}$ toute surface stable par les rotations d'axe $\mathscr{D}$.

Soit

$$\begin{aligned} x_1 &= a(\alpha) \\ x_3 &= b(\alpha) \end{aligned}$$

les équations paramétriques de la section d'une telle surface par le plan d'équation $x_2 = 0$. Les équations paramétriques de la surface elle-même sont alors (fig. 10.13)

$$\begin{aligned} x_1 &= a(\alpha)\cos\beta \\ x_2 &= a(\alpha)\sin\beta \\ x_3 &= b(\alpha), \end{aligned} \tag{10.18}$$

où α parcourt un sous-ensemble de $\mathbb{R}$ et β l'intervalle $[0, 2\pi)$.

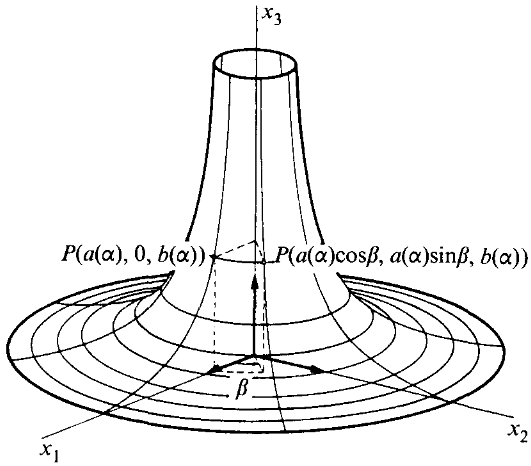


Fig. 10.13

Par exemple, lorsque

$$a(\alpha) = a_1 \mathrm{ch}\alpha$$
$$b(\alpha) = a_3 \mathrm{sh}\alpha,$$

α parcourant $\mathbb{R}$, les équations (10.18) sont celles d'un hyperboloïde de révolution à une nappe:

$$x_1 = a_1 \mathrm{ch}\alpha \cos\beta$$
$$x_2 = a_1 \mathrm{ch}\alpha \sin\beta$$
$$x_3 = a_3 \mathrm{sh}\alpha.$$

On notera qu'en remplaçant a_1 par a_2 dans la deuxième équation, on obtient une paramétrisation de l'hyperboloïde à une nappe d'équation (10.16), différente de celle qui est définie par (10.17).

10.6.4 Equations paramétriques de quelques quadriques

Les équations paramétriques suivantes s'obtiennent, comme dans l'exemple de l'hyperboloïde à une nappe présenté ci-dessus, en supposant d'abord que $a_1 = a_2$. Le paramètre β parcourt l'intervalle $[0, 2\pi)$.

Ellipsoïde:

$$x_1 = a_1 \cos\alpha \cos\beta$$
$$x_2 = a_2 \cos\alpha \sin\beta \quad (\alpha \in [-\frac{\pi}{2}, \frac{\pi}{2}])$$
$$x_3 = a_3 \sin\alpha.$$

Hyperboloïde à deux nappes:

$$x_1 = a_1 \mathrm{sh}\alpha \cos\beta$$
$$x_2 = a_2 \mathrm{sh}\alpha \sin\beta \quad (\alpha \in [0, \infty))$$
$$x_3 = \pm a_3 \mathrm{ch}\alpha.$$

Cône du second degré:

$$\begin{aligned} x_1 &= a_1\alpha\cos\beta \\ x_2 &= a_2\alpha\sin\beta \quad (\alpha \in \mathbb{R}) \\ x_3 &= a_3\alpha. \end{aligned}$$

Paraboloïde elliptique:

$$\begin{aligned} x_1 &= a_1\alpha\cos\beta \\ x_2 &= a_2\alpha\sin\beta \quad (\alpha \in [0, \infty)) \\ x_3 &= \alpha^2. \end{aligned}$$

10.7 EXERCICES

10.7.1 Montrer que toute équation du second degré en $x_1, x_2, ..., x_n$ peut s'écrire sous la forme (10.1) avec $a_{ij} = a_{ji}$ pour tout couple (i, j) d'indices.

10.7.2 Réduire les équations suivantes par complétion des carrés et indiquer la nature des quadriques que ces équations représentent. (Utiliser les exercices 9.4.3 et 9.4.4.)

(a) $x_1^2 + 3x_2^2 - 2x_2x_3 + 3x_3^2 = 1$.
(b) $x_1^2 + 3x_2^2 + 4x_2x_3 = 1$.
(c) $x_1x_2 + x_1x_3 + x_2x_3 = 1$.

10.7.3 Quelle est la nature de la quadrique dont l'équation dans un repère non orthonormal donné est $x_1^2 + x_2^2 + x_3^2 = 1$?

10.7.4 Soit $\mathscr{E}$ un espace affine euclidien de dimension 3, muni d'un repère orthonormal. Pour chacune des quadriques d'équation donnée ci-dessous, trouver un repère orthonormal de $\mathscr{E}$ dans lequel l'équation est réduite, écrire l'équation réduite et indiquer la nature de la quadrique.

(a) $x_1^2 + 6x_1x_3 - 2x_2^2 + x_3^2 + 4x_1 + 8x_2 - 4x_3 = 16$.
(b) $3x_1^2 - 12x_1x_3 - 2x_2^2 + 12x_3^2 + 18x_1 - 36x_3 = -24$.
(c) $x_1^2 + x_2^2 - 4x_2x_3 + x_3^2 - 2x_1 - 6x_3 = 2$.
(d) $4x_1^2 - x_2^2 + 2x_2x_3 - x_3^2 - 8x_1 = -2$.
(e) $x_1^2 + 3x_2^2 - 2x_2x_3 + 3x_3^2 - 2x_1 - 2x_2 + 6x_3 = -3$.
(f) $x_1^2 + 4x_1x_3 + 3x_2^2 + x_3^2 + 2x_1 + 12x_2 - 2x_3 = -9$.
(g) $\frac{1}{5}x_1^2 + \frac{4}{5}x_1x_3 + 3x_2^2 + \frac{4}{5}x_3^2 - 6x_1 - 6x_2 - 12x_3 = -47$.
(h) $2x_1^2 - 2x_1x_2 + 3x_1x_3 - 4x_2^2 + 6x_2x_3 - 2x_3^2 + 5x_1 + 2x_2 = -2$.
(i) $2x_1x_2 - x_1 - x_2 + x_3 = 0$.

10.7.5 Dans un espace affine euclidien de dimension 3, muni d'un repère orthonormal, on considère la quadrique d'équation

$$(1 + \alpha)x_1^2 + 2(1 - \alpha)x_1x_3 + 2x_2^2 + (1 + \alpha)x_3^2 - 2x_1 - 4x_2 - 2x_3$$
$$= -3 + 2(1 - \alpha^2).$$

(a) Trouver un repère orthonormal indépendant de α dans lequel cette équation est réduite.
(b) Indiquer la nature de la quadrique en distinguant les cas correspondant aux différentes valeurs de α.
(c) Lorsque la quadrique est un hyperboloïde, trouver l'équation (dans le repère initial) du cône asymptote. (On appelle *cône asymptote* d'un hyperboloïde d'équation ${}^t\mathbf{x}\mathbf{A}\mathbf{x} = c$ le cône d'équation ${}^t\mathbf{x}\mathbf{A}\mathbf{x} = 0$.)

10.7.6 Dans le cas d'un espace affine euclidien de dimension 2 ou 3, indiquer, pour chaque quadrique, les symétries orthogonales par lesquelles la quadrique est stable.

10.7.7 Soit ${}^t\mathbf{x}\mathbf{A}\mathbf{x} = 1$ l'équation d'une quadrique dans un espace affine $\mathscr{E}$ de dimension $n > 1$, muni d'un repère $(O; \mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$. Soit $\mathscr{S}$ un sous-espace affine de $\mathscr{E}$ passant par O, de vecteurs directeurs $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k$. On désigne par $\mathbf{B}$ la matrice de type $n \times k$ dont les colonnes sont $(\mathbf{v}_1)_e, (\mathbf{v}_2)_e, ..., (\mathbf{v}_k)_e$. Montrer que la section de la quadrique par $\mathscr{S}$ est la quadrique dans $\mathscr{S}$ dont l'équation dans le repère $(O; \mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k)$ est ${}^t\mathbf{x}'\mathbf{A}'\mathbf{x}' = 1$, où $\mathbf{A}' = {}^t\mathbf{B}\mathbf{A}\mathbf{B}$.

10.7.8 Dans un espace affine euclidien de dimension 3, muni d'un repère orthonormal, on considère l'ellipsoïde d'équation $4x_1^2 + 3x_2^2 + 2x_2x_3 + 3x_3^2 = 1$. Trouver l'équation matricielle d'une affinité orthogonale (c'est-à-dire de direction orthogonale au plan d'affinité) transformant l'ellipsoïde en une sphère.

10.7.9 Dans un espace affine euclidien de dimension 3, muni d'un repère orthonormal, on considère l'ellipsoïde d'équation $x_1^2 + 2x_2^2 + 3x_3^2 = 1$.

(a) Trouver un plan passant par l'origine tel que la section de l'ellipsoïde par ce plan soit un cercle. (Utiliser l'exercice 10.7.7.)
(b) Trouver l'équation matricielle d'une affinité (oblique) transformant l'ellipsoïde en une sphère.

10.7.10 Montrer que l'image d'une quadrique par une application affine bijective est encore une quadrique.

10.7.11 Montrer que

$$P_0\left(\frac{\sqrt{2} + 2}{2}, 2, \frac{\sqrt{2} - 2}{2}\right)$$

est un point de la quadrique d'équation (a) de l'exercice 10.7.2 et trouver l'équation du plan tangent à cette quadrique en P_0.

10.7.12 Soit $\mathscr{E}$ un espace affine euclidien de dimension 3, muni d'un repère orthonormal.

(a) Déterminer les valeurs de α telles que la quadrique dans $\mathscr{E}$ d'équation

$$\frac{x_1^2}{2+\alpha} + \frac{x_2^2}{3+\alpha} + \frac{x_3^2}{6+\alpha} = 1$$

passe par le point $P_0(1, 1, 1)$.

(b) Pour chacune des valeurs trouvées, identifier la quadrique.

(c) Montrer que les plans tangents en P_0 aux quadriques trouvées sont orthogonaux deux à deux.

10.7.13 Dans un espace affine euclidien de dimension 3, muni d'un repère orthonormal, on considère le plan $\mathscr{S}$ passant par les points $P_1(2, 4, -6)$, $P_2(1, 3, -4)$, $P_3(-2, 2, 0)$ et la droite $\mathscr{D}$ passant par l'origine du repère, de vecteur directeur $\mathbf{v}(1, 1, 0)$.

(a) Trouver l'équation du lieu géométrique des points équidistants de $\mathscr{S}$ et de $\mathscr{D}$.

(b) Etablir la nature de ce lieu.

10.7.14 En liaison avec la section 10.1, on appelle *hyperplan polaire* d'un point P_0 relativement à la quadrique d'équation (10.2) l'hyperplan d'équation (10.4) ($\mathbf{x}_0$ désigne le vecteur-colonne des coordonnées de P_0). D'après 10.1.4, si P_0 est un point de la quadrique, l'hyperplan polaire de P_0 est l'hyperplan tangent à la quadrique en P_0.

(a) Montrer que l'hyperplan polaire de P_0 n'est défini que si P_0 n'est pas un centre de la quadrique.

(b) La section de la quadrique par l'hyperplan polaire de P_0 peut être vide. Si elle ne l'est pas, montrer qu'en chacun de ses points qui n'est pas un centre de la quadrique, l'hyperplan tangent à la quadrique passe par P_0.

10.7.15 Trouver les équations des plans passant par le point $P_0(1, 0, 0)$ et tangents à la quadrique d'équation

$$5x_1^2 + 4x_1x_2 + 6x_2^2 - 4x_2x_3 + 7x_3^2 = 3$$

aux points de sa section par le plan d'équation $x_2 = 0$.

10.7.16 Montrer que les équations (10.2) et (10.4) s'écrivent, de manière équivalente, sous les formes respectives

$$^t\tilde{\mathbf{x}}\tilde{\mathbf{A}}\tilde{\mathbf{x}} = 0 \quad \text{et} \quad ^t\tilde{\mathbf{x}}_0\tilde{\mathbf{A}}\tilde{\mathbf{x}} = 0,$$

où

$$\tilde{\mathbf{A}} = \left(\begin{array}{c|c} \mathbf{A} & \mathbf{b} \\ \hline ^t\mathbf{b} & -c \end{array}\right), \quad \tilde{\mathbf{x}} = \left(\begin{array}{c} \mathbf{x} \\ \hline 1 \end{array}\right), \quad \tilde{\mathbf{x}}_0 = \left(\begin{array}{c} \mathbf{x}_0 \\ \hline 1 \end{array}\right).$$

(Multiplier par blocs.)

10.7.17 Si f est une fonction réelle de $\mathbf{x} = (x_i)$, dérivable par rapport à chaque x_i, on pose

$$\operatorname{grad} f(\mathbf{x}) = \left(\frac{\partial f}{\partial x_i}(\mathbf{x})\right).$$

(a) Montrer que $\operatorname{grad}({}^t\mathbf{b}\mathbf{x}) = \mathbf{b}$.
(b) Montrer que $\operatorname{grad}({}^t\mathbf{x}\mathbf{A}\mathbf{x}) = (\mathbf{A} + {}^t\mathbf{A})\mathbf{x}$.
(c) En déduire la formule 10.3.

10.7.18 Montrer que le paraboloïde hyperbolique est une surface réglée.

APPENDICE

Extension aux scalaires complexes

A.1 ESPACES VECTORIELS COMPLEXES

A.1.1 Introduction

L'objectif principal de cet appendice est de présenter les nouveautés que comporte l'utilisation des nombres complexes en algèbre linéaire.

Lorsque les scalaires qui apparaissent dans la définition de la notion d'espace vectoriel 1.2.2 sont les nombres complexes, et non plus les nombres réels, on parle d'*espaces vectoriels sur le corps des nombres complexes* ou, plus simplement, d'*espaces vectoriels complexes.* Pour éviter toute ambiguïté, les espaces vectoriels sur le corps des réels (c'est-à-dire ceux dont il a été question tout au long de ce livre) seront appelés dorénavant *espaces vectoriels réels.* Nous aurions pu envisager des espaces vectoriels sur un corps quelconque, mais cette généralisation aurait dépassé le cadre que nous nous sommes imposé.

Voici quelques exemples d'espaces vectoriels complexes (cf. 1.3.3, 1.3.4 et 1.3.5):

(1) L'espace vectoriel $\mathbb{C}^n$ des vecteurs-colonnes à n termes complexes ($\mathbb{C}^1$ sera identifié à $\mathbb{C}$). Plus généralement, l'espace vectoriel des matrices de type $m \times n$ à termes complexes.

(2) L'espace vectoriel des suites de nombres complexes ou des suites convergentes de nombres complexes.

(3) L'espace vectoriel des polynômes à coefficients complexes en une indéterminée t.

(4) L'espace vectoriel des fonctions complexes d'une variable réelle ou complexe.

(5) L'espace vectoriel des fonctions complexes dérivables.

Nous verrons qu'une grande partie du contenu de ce livre peut être transposée au domaine complexe en prenant simplement des espaces vectoriels complexes à la place des espaces vectoriels réels et en attribuant au terme «nombre» le sens de «nombre complexe». Nous verrons également que pour de nombreuses questions le passage aux scalaires complexes nécessite des retouches, voire même des remaniements et des compléments.

Nous allons à présent passer en revue les différents chapitres du livre et expliquer comment le passage au cas complexe doit être effectué. De la discussion qui suivra seront exclues toutes les sections qui touchent aux espaces affines. La raison de cette exclusion est à rechercher dans le fait que les applications pratiques de la théorie des espaces affines se rapportent le plus souvent au cas réel.

A.1.2 Espaces vectoriels complexes

Si l'on exclut les parties concernant les espaces V et $\mathscr{G}$, le premier chapitre s'étend au cas complexe sans modifications. Naturellement, $\mathbb{C}^n$ assume le rôle de $\mathbb{R}^n$.

A.1.3 Espaces vectoriels hermitiens

L'opération définissant le produit scalaire canonique dans $\mathbb{R}^n$ ne peut être appliquée telle quelle à $\mathbb{C}^n$, car le produit scalaire d'un vecteur-colonne par lui-même ne serait en général pas réel et ne pourrait donc pas servir à définir la norme de ce vecteur. Pour parer à cet inconvénient, la formule (2.8) est remplacée par

$$\left(\begin{pmatrix} a_1 \\ a_2 \\ \cdot \\ \cdot \\ \cdot \\ a_n \end{pmatrix} \middle| \begin{pmatrix} b_1 \\ b_2 \\ \cdot \\ \cdot \\ \cdot \\ b_n \end{pmatrix}\right) = a_1\bar{b}_1 + a_2\bar{b}_2 + \ldots + a_n\bar{b}_n, \tag{A.1}$$

où la barre surmontant une lettre désigne, et désignera dans toute la suite, le conjugué complexe du nombre représenté par cette lettre. La norme d'un vecteur-colonne est donc définie par la formule

$$\begin{aligned} \left\| \begin{pmatrix} a_1 \\ a_2 \\ \cdot \\ \cdot \\ \cdot \\ a_n \end{pmatrix} \right\| &= (a_1\bar{a}_1 + a_2\bar{a}_2 + \ldots + a_n\bar{a}_n)^{1/2} \\ &= (|a_1|^2 + |a_2|^2 + \ldots + |a_n|^2)^{1/2}. \end{aligned} \tag{A.2}$$

Par analogie, dans le cadre des espaces vectoriels formés de fonctions complexes continues, la formule (2.10) est remplacée par

$$(\mathbf{f} \mid \mathbf{g}) = \int_a^b \mathbf{f}(t)\bar{\mathbf{g}}(t)\mathrm{d}t, \tag{A.3}$$

où a et b désignent les extrémités d'un intervalle fermé de $\mathbb{R}$. On remarquera que dans le cas où les vecteurs-colonnes ou les fonctions sont réels, (A.1) et (A.3) coïncident respectivement avec (2.8) et (2.10).

Manifestement, les opérations définies par (A.1) et (A.3) satisfont aux conditions (b) et (c) de la définition 2.2.2. Par contre, elles ne satisfont plus à la condition (a), autrement dit, elles ne sont plus symétriques. Cette constatation nous amène à adapter la définition 2.2.2 de la façon suivante. Soit E un espace vectoriel complexe. On appelle *produit scalaire hermitien* dans E toute application $(\mathbf{x}, \mathbf{y}) \to (\mathbf{x} \mid \mathbf{y})$ de l'ensemble des couples de vecteurs de E dans $\mathbb{C}$ satisfaisant aux conditions (b) et (c) de la définition 2.2.2, ainsi qu'à la condition

(a′) $(\mathbf{x} \mid \mathbf{y}) = \overline{(\mathbf{y} \mid \mathbf{x})}$.

Les formules (A.1) et (A.3) définissent clairement des produits scalaires hermitiens. Comme dans le cas réel, on les appelle *produits scalaires canoniques.*

Un espace vectoriel complexe muni d'un produit scalaire hermitien est appelé *espace vectoriel hermitien.*

On relèvera qu'un produit scalaire hermitien n'est pas bilinéaire. L'absence de symétrie empêche la déduction de la linéarité à droite à partir de la condition (b). Néanmoins, la relation suivante, appelée *semi-linéarité* ou *antilinéarité à droite* du produit scalaire hermitien, se réalise:

$$(\mathbf{x} \mid \beta\mathbf{y} + \gamma\mathbf{z}) = \bar{\beta}(\mathbf{x} \mid \mathbf{y}) + \bar{\gamma}(\mathbf{x} \mid \mathbf{z}). \tag{A.4}$$

En effet, grâce à (a′) et à (b), le premier membre est égal à

$$\overline{(\beta\mathbf{y} + \gamma\mathbf{z} \mid \mathbf{x})} = \bar{\beta}\,\overline{(\mathbf{y} \mid \mathbf{x})} + \bar{\gamma}\,\overline{(\mathbf{z} \mid \mathbf{x})} = \bar{\beta}(\mathbf{x} \mid \mathbf{y}) + \bar{\gamma}(\mathbf{x} \mid \mathbf{z}).$$

La condition (a′) assure que le produit scalaire hermitien d'un vecteur $\mathbf{x}$ par lui-même est réel. D'après les conditions (b) et (c), ce produit scalaire est positif ou nul, suivant que $\mathbf{x}$ est non nul ou nul. Cela nous permet, en particulier, de définir la norme de la même manière qu'en 2.2.5, à savoir par

$$\| \mathbf{x} \| = \sqrt{(\mathbf{x} \mid \mathbf{x})}.$$

Comme dans 2.2.5, nous voyons que si α est un nombre complexe,

$$\| \alpha\mathbf{x} \| = | \alpha | \| \mathbf{x} \|.$$

A.1.4 Adaptation du deuxième chapitre

La semi-linéarité à droite entraîne un changement dans l'écriture du produit scalaire hermitien en fonction des composantes. Plus exactement, les formules (2.26), (2.28) et (2.29) deviennent respectivement

$$(\mathbf{x} \mid \mathbf{y}) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i \bar{y}_j,$$

$$(\mathbf{x} \mid \mathbf{y}) = x_1\bar{y}_1 + x_2\bar{y}_2 + \dots + x_n\bar{y}_n, \tag{A.5}$$

$$\| \mathbf{x} \|^2 = | x_1 |^2 + | x_2 |^2 + \dots + | x_n |^2.$$

Grâce à ces modifications et à celles que nous allons indiquer, la validité des sections 2.2–2.6 du chapitre 2 s'étend aux espaces vectoriels hermitiens. Voici les autres modifications: la formule (2.14) devient

$$\| \mathbf{x} + \mathbf{y} \|^2 = \| \mathbf{x} \|^2 + \| \mathbf{y} \|^2 + 2\mathcal{R}(\mathbf{x} \mid \mathbf{y}), \tag{A.6}$$

où $\mathcal{R}(\mathbf{x} \mid \mathbf{y})$ désigne la partie réelle de $(\mathbf{x} \mid \mathbf{y})$; dans (2.21), $|(\mathbf{x} \mid \mathbf{e})|^2$ remplace $(\mathbf{x} \mid \mathbf{e})^2$; la notion d'angle apparaissant dans 2.4.4 ne doit pas être prise en compte; l'isomorphie 2.5.6 de E et $\mathbb{R}^n$ est à remplacer par l'isomorphie de E et $\mathbb{C}^n$.

La section 2.7 ne se généralise pas au cas complexe, car l'orientation d'un espace vectoriel ne peut être définie que dans le cas réel.

A.2 SYSTÈMES LINÉAIRES, MATRICES ET DÉTERMINANTS

A.2.1 Systèmes linéaires complexes

Le chapitre 3 s'étend sans modifications aux scalaires complexes.

A.2.2 Matrices complexes

Le chapitre 4 s'étend également aux scalaires complexes. La seule modification requise concerne la formule (4.22), qui devient

$$\| \mathbf{A} \| = \left(\sum_{i=1}^{n} \sum_{j=1}^{n} | a_{ij} |^2 \right)^{1/2}.$$

Bien entendu, la section 4.6 consacrée aux matrices de transition reste du domaine réel.

Dorénavant, nous appellerons matrice une matrice à termes complexes et matrice réelle une matrice à termes réels.

A.2.3 Conjuguée et adjointe d'une matrice

On appelle *conjuguée* d'une matrice $\mathbf{A} = (a_{ij})$, la matrice $\bar{\mathbf{A}} = (\bar{a}_{ij})$. Si $\mathbf{A}$ est carrée, on appelle *adjointe* de $\mathbf{A}$, et on note $\mathbf{A}^*$, la matrice ${}^t\bar{\mathbf{A}}\ (= \overline{{}^t\mathbf{A}})$. Nous verrons que l'adjointe d'une matrice joue un rôle analogue à celui de la transposée d'une matrice réelle.

Les opérations de passage aux conjuguées et aux adjointes jouissent des propriétés évidentes suivantes:

(a) $\overline{(\bar{\mathbf{A}})} = \mathbf{A}$. (a′) $(\mathbf{A}^*)^* = \mathbf{A}$.

(b) $\overline{(\mathbf{A} + \mathbf{B})} = \bar{\mathbf{A}} + \bar{\mathbf{B}}$. (b′) $(\mathbf{A} + \mathbf{B})^* = \mathbf{A}^* + \mathbf{B}^*$.

(c) $\overline{(\alpha \mathbf{A})} = \bar{\alpha}\bar{\mathbf{A}}$. (c′) $(\alpha \mathbf{A})^* = \bar{\alpha}\mathbf{A}^*$.
(d) $\overline{(\mathbf{AB})} = \bar{\mathbf{A}}\bar{\mathbf{B}}$. (d′) $(\mathbf{AB})^* = \mathbf{B}^*\mathbf{A}^*$.
(e) $\overline{(\mathbf{A}^{-1})} = \bar{\mathbf{A}}^{-1}$ (si $\mathbf{A}$ est inversible). (e′) $(\mathbf{A}^{-1})^* = (\mathbf{A}^*)^{-1}$ (si $\mathbf{A}$ est inversible).
(f) $\mathrm{rg}\bar{\mathbf{A}} = \mathrm{rg}\mathbf{A}$. (f′) $\mathrm{rg}\mathbf{A}^* = \mathrm{rg}\mathbf{A}$.

On notera que le produit scalaire canonique dans $\mathbb{C}^n$ s'écrit sous la forme (cf. 4.3.1)

$$(\mathbf{x} \mid \mathbf{y}) = {}^t\mathbf{x}\bar{\mathbf{y}} = {}^t\bar{\mathbf{y}}\mathbf{x} = \mathbf{y}^*\mathbf{x}. \tag{A.7}$$

Plus généralement, tout produit scalaire hermitien s'exprime, dans une base orthonormale $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$, par

$$(\mathbf{x} \mid \mathbf{y}) = {}^t\mathbf{x}_e\bar{\mathbf{y}}_e = {}^t\bar{\mathbf{y}}_e\mathbf{x}_e = \mathbf{y}_e^*\mathbf{x}_e. \tag{A.8}$$

A.2.4 Matrices hermitiennes et matrices normales

On appelle *matrice hermitienne* (*matrice normale*) toute matrice carrée $\mathbf{A}$ telle que $\mathbf{A}^* = \mathbf{A}$ $(\mathbf{A}^*\mathbf{A} = \mathbf{A}\mathbf{A}^*)$.

Toute matrice hermitienne est normale. Les matrices hermitiennes réelles sont les matrices symétriques. Toute matrice diagonale est normale.

A.2.5 Déterminants complexes

Si l'on exclut les parties en liaison avec les espaces affines, le chapitre 5 s'étend sans modifications aux scalaires complexes.

On notera l'existence de deux propriétés supplémentaires, la première découlant directement de la définition 5.1.3 et la deuxième de la première jointe à (b) de la proposition 5.1.8:

$$\begin{aligned} &\text{(a) } |\bar{\mathbf{A}}| = \overline{|\mathbf{A}|}. \\ &\text{(b) } |\mathbf{A}^*| = \overline{|\mathbf{A}|}. \end{aligned} \tag{A.9}$$

La propriété (b) entraîne que le déterminant de toute matrice hermitienne est réel.

A.3 APPLICATIONS LINÉAIRES

A.3.1 Applications linéaires entre espaces vectoriels complexes

Les sections du chapitre 6 consacrées aux applications linéaires gardent leur validité dans le cas où les espaces vectoriels sont complexes. En raison de (A.8), le vecteur-colonne $\mathbf{n}_e$ de l'exemple 6.5.8 doit être remplacé par $\bar{\mathbf{n}}_e$.

A.3.2 Transformations unitaires

La conservation du produit scalaire et celle de la norme se définissent comme dans 7.1.2. La proposition 7.1.3 vaut pour les espaces vectoriels hermitiens. L'emploi de la semi-linéarité à droite du produit scalaire hermitien à la place de la linéarité à droite et de (A.6) à la place de (2.14) engendre quelques changements dans la démonstration de cette proposition. Nous laissons au lecteur la tâche de les effectuer.

Une application φ d'un espace vectoriel hermitien E dans lui-même qui satisfait à l'une des trois conditions équivalentes (a), (b) et (c) de la proposition 7.1.3 est appelée *transformation unitaire* de E.

A.3.3 Matrices unitaires

Les arguments utilisés dans 7.1.6 montrent que la matrice $\mathbf{A}$ d'une transformation unitaire d'un espace vectoriel hermitien de dimension finie non nulle et muni d'une base orthonormale vérifie la relation

$$^t\mathbf{A}\bar{\mathbf{A}} = \mathbf{I}, \tag{A.10}$$

qui s'écrit également sous la forme

$$\mathbf{A}^*\mathbf{A} = \mathbf{I}. \tag{A.11}$$

On dit qu'une matrice carrée $\mathbf{A}$ est *unitaire* si elle vérifie l'une des relations équivalentes (A.10) ou (A.11).

Il est évident qu'une matrice réelle est unitaire si et seulement si elle est orthogonale.

La proposition 7.1.9 s'applique aux matrices unitaires, à condition de remplacer $^t\mathbf{A}$ par $\mathbf{A}^*$ et «orthogonal» par «unitaire». En particulier, l'équivalence entre (a) et (b) entraîne que toute matrice unitaire est normale.

A l'aide de (A.9), on montre, comme dans 7.1.10, que la valeur absolue du déterminant d'une matrice unitaire vaut 1. Le déterminant lui-même peut être n'importe quel nombre complexe de valeur absolue 1. Par exemple, $|e^{i\theta}\mathbf{I}_n| = e^{in\theta}$.

La matrice de passage entre deux bases orthonormales est unitaire.

La section 7.2 est propre au cas réel.

A.4 VALEURS PROPRES ET VECTEURS PROPRES

A.4.1 Définitions et premières conséquences

Les notions de valeur propre, de vecteur propre et de sous-espace propre d'une application linéaire d'un espace vectoriel complexe dans lui-même se définissent comme dans 8.2.2. Naturellement, les valeurs propres sont, dans ce cas, des nombres complexes.

Les sections 8.2 et 8.3 s'étendent sans modifications aux espaces vectoriels complexes. Le fait important nouveau est que la somme des multiplicités des valeurs propres d'une application linéaire φ d'un espace vectoriel complexe E de dimension finie non nulle n dans lui-même, ou d'une matrice carrée $\mathbf{A}$ d'ordre n, est toujours égale à n (les facteurs irréductibles de la décomposition d'un polynôme à coefficients complexes sont tous de degré 1). En particulier, φ et $\mathbf{A}$ admettent toujours des vecteurs propres. Cette assertion n'est cependant plus vraie lorsque E est de dimension infinie. Par exemple, si E est l'espace vectoriel complexe formé des polynômes à coefficients complexes et φ l'application linéaire de E dans E définie dans l'exercice 7.4.1, φ n'admet clairement aucun vecteur propre.

On remarquera que si λ est une valeur propre d'une matrice $\mathbf{A}$ et $\mathbf{x}$ un vecteur propre associé à λ, alors $\bar{\lambda}$ est une valeur propre de $\bar{\mathbf{A}}$ et $\bar{\mathbf{x}}$ un vecteur propre associé à $\bar{\lambda}$.

On remarquera aussi que $\bar{\mathbf{A}}$ et $\mathbf{A}^*$ ont le même polynôme caractéristique, puisque $\mathbf{A}^*$ est la transposée de $\bar{\mathbf{A}}$, donc les mêmes valeurs propres, avec les mêmes multiplicités.

A.4.2 Réduction à la forme diagonale

La section 8.4 s'étend au cas complexe sans autre modification que celle qui consiste à supprimer, partout où elle apparaît, la condition exigeant que les racines du polynôme caractéristique soient toutes réelles.

A.4.3 Réduction à la forme triangulaire

La section 8.5 s'étend au cas complexe tout en se simplifiant. En effet, la somme des multiplicités des valeurs propres étant automatiquement égale à n, comme nous l'avons déjà remarqué, le théorème 8.5.2 et la proposition 8.5.5 se réduisent à un seul énoncé affirmant que toute application linéaire de E dans E est trigonalisable, E désignant un espace vectoriel complexe de dimension finie non nulle n. En langage matriciel, cette conclusion s'énonce en disant que toute matrice est semblable à une matrice triangulaire.

A.4.4 Transformations hermitiennes

Lorsque l'espace vectoriel E de la section 8.6 est hermitien, les applications φ de E dans E qui vérifient la relation (8.33) pour tout couple de vecteurs de E sont appelées *transformations hermitiennes* de E. Comme dans 8.6.2, on démontre que toute transformation hermitienne est linéaire.

La caractérisation énoncée dans la proposition 8.6.3 reste vraie dans le cas des transformations hermitiennes, à condition de remplacer «symétrique» par «hermitien».

Si φ est une transformation hermitienne de E et $\mathbf{x}$ un vecteur quelconque de E, le produit scalaire hermitien de $\varphi(\mathbf{x})$ et $\mathbf{x}$ est réel. En effet,

$$(\varphi(\mathbf{x}) \mid \mathbf{x}) = (\mathbf{x} \mid \varphi(\mathbf{x})) = \overline{(\varphi(\mathbf{x}) \mid \mathbf{x})},$$

ce qui démontre l'assertion. Il s'ensuit que toute valeur propre λ de φ est réelle. En effet, si $\mathbf{x}$ est un vecteur propre associé à la valeur propre λ, alors

$$\lambda \|\mathbf{x}\|^2 = \lambda(\mathbf{x} \mid \mathbf{x}) = (\lambda\mathbf{x} \mid \mathbf{x}) = (\varphi(\mathbf{x}) \mid \mathbf{x}),$$

ce qui prouve que λ est un nombre réel.

La proposition 8.6.5 s'applique aux transformations et aux matrices hermitiennes. La démonstration est inchangée.

L'énoncé du théorème 8.6.6 doit être légèrement modifié. Soit φ une application linéaire d'un espace vectoriel hermitien de dimension finie non nulle dans lui-même. Les conditions suivantes sont équivalentes:

(a) φ est une transformation hermitienne.
(b) Les valeurs propres de φ sont réelles et il existe une base constituée de vecteurs propres de φ orthogonaux deux à deux.
(c) φ est diagonalisable, ses valeurs propres sont réelles et les sous-espaces propres associés sont orthogonaux deux à deux.

Ces équivalences sont établies par la démonstration du théorème 8.6.6 écourtée de la partie initiale consacrée à la preuve de l'existence d'un vecteur propre. Cette partie garde toutefois toute son importance, car elle prouve la validité de l'estimation des valeurs propres extrêmes $\lambda_{\min}$ et $\lambda_{\max}$ présentée et illustrée dans 8.6.9.

La décomposition spectrale 8.6.7 s'applique aux transformations hermitiennes.

L'interprétation matricielle de l'équivalence des conditions (a), (b) et (c) ci-dessus s'énonce ainsi: une matrice carrée $\mathbf{A}$ est hermitienne si et seulement s'il existe une matrice unitaire $\mathbf{P}$ et une matrice diagonale *réelle* $\mathbf{D}$ telles que

$$\mathbf{D} = \mathbf{P}^*\mathbf{A}\mathbf{P} \quad \text{ou} \quad \mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^*. \tag{A.12}$$

Evidemment, les colonnes de $\mathbf{P}$ sont des vecteurs propres de $\mathbf{A}$ et les termes diagonaux de $\mathbf{D}$ les valeurs propres correspondantes; en outre, ces vecteurs propres forment une famille orthonormale, puisque $\mathbf{P}$ est unitaire. Réciproquement, toute matrice carrée $\mathbf{P}$ dont les colonnes forment une famille orthonormale de vecteurs propres de $\mathbf{A}$ vérifie (A.12).

A.4.5 Systèmes différentiels complexes

Lorsque les coefficients et les fonctions de la variable réelle t sont complexes, les systèmes différentiels étudiés dans la section 8.7 sont appelés *systèmes différentiels complexes* (linéaires, du premier ordre, à coefficients constants).

La section 8.7 s'applique intégralement aux systèmes différentiels complexes. Naturellement, les valeurs et vecteurs propres complexes remplacent les valeurs et vecteurs propres réels.

A.4.6 Retour aux systèmes différentiels réels

Par l'intermédiaire de solutions complexes, nous sommes à présent en mesure de lever la restriction posée au début de 8.7.5. Supposons que les coefficients a_{ij} du système différentiel (8.37) soient réels. Alors le polynôme caractéristique de $\mathbf{A} = (a_{ij})$ est à coefficients réels. Il en découle que si λ est une valeur propre de $\mathbf{A}$ de multiplicité k, $\bar{\lambda}$ est également une valeur propre de $\mathbf{A}$ de multiplicité k. En outre, si $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k$ sont les k solutions complexes linéairement indépendantes de $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$ définies par la formule (8.43), $\bar{\mathbf{x}}_1, \bar{\mathbf{x}}_2, ..., \bar{\mathbf{x}}_k$ sont les k solutions complexes linéairement indépendantes de la même équation définies par la même formule en partant de $\bar{\lambda}$. Or, d'après ce qui est affirmé dans 8.7.5, les solutions $\mathbf{x}_1, ..., \mathbf{x}_k, \bar{\mathbf{x}}_1, ..., \bar{\mathbf{x}}_k$ sont linéairement indépendantes et concourent à la constitution d'une base de solutions. Il en va donc de même pour les solutions $\mathbf{x}_1 + \bar{\mathbf{x}}_1, ...,$ $\mathbf{x}_k + \bar{\mathbf{x}}_k$, $\mathbf{x}_1 - \bar{\mathbf{x}}_1, ..., \mathbf{x}_k - \bar{\mathbf{x}}_k$, qui ne sont rien d'autre que les parties réelles et imaginaires de $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k$ multipliées respectivement par 2 et 2i ($\mathrm{i} = \sqrt{-1}$). Nous avons ainsi montré qu'à toute valeur propre complexe λ de multiplicité k correspondent $2k$ solutions réelles linéairement indépendantes de $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$, à savoir les parties réelles et imaginaires des solutions complexes fournies par (8.43). Ces $2k$ solutions représentent la contribution des valeurs propres λ et $\bar{\lambda}$.

A titre d'exemple, montrons comment les deux solutions (8.50) ont été trouvées. Le polynôme caractéristique de

$$\mathbf{A} = \begin{bmatrix} 4 & 5 \\ -2 & -2 \end{bmatrix}$$

est

$$|\mathbf{A} - t\mathbf{I}| = t^2 - 2t + 2.$$

Les valeurs propres de $\mathbf{A}$ sont donc $1 + \mathrm{i}$ et $1 - \mathrm{i}$. Un vecteur propre associé à la valeur propre $1 + \mathrm{i}$ est

$$\begin{bmatrix} -5 \\ 3-\mathrm{i} \end{bmatrix}.$$

Une solution complexe de $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$ associée à la valeur propre $1 + \mathrm{i}$ est donc

$$\begin{aligned} \mathbf{x}(t) &\equiv e^{(1+\mathrm{i})t}\begin{bmatrix} -5 \\ 3-\mathrm{i} \end{bmatrix} \equiv e^t(\cos t + \mathrm{i}\sin t)\left(\begin{bmatrix} -5 \\ 3 \end{bmatrix} + \mathrm{i}\begin{bmatrix} 0 \\ -1 \end{bmatrix}\right) \\ &\equiv e^t\left(\cos t\begin{bmatrix} -5 \\ 3 \end{bmatrix} - \sin t\begin{bmatrix} 0 \\ -1 \end{bmatrix}\right) + \mathrm{i}e^t\left(\sin t\begin{bmatrix} -5 \\ 3 \end{bmatrix} + \cos t\begin{bmatrix} 0 \\ -1 \end{bmatrix}\right). \end{aligned}$$

La partie réelle et la partie imaginaire de $\mathbf{x}(t)$ sont les solutions (8.50).

A.5 TRANSFORMATIONS NORMALES

A.5.1 Introduction

Cette section sera consacrée à l'étude des applications linéaires d'un espace vectoriel hermitien dans lui-même caractérisées par la présence d'une base orthonormale constituée de vecteurs propres.

Tout au long de la section, E désignera un espace vectoriel hermitien de dimension finie non nulle n.

A.5.2 Adjointe d'une application linéaire

Soit φ une application linéaire de E dans E. Nous allons montrer qu'il existe une et une seule application linéaire φ^* de E dans E telle que

$$(\varphi(\mathbf{x}) \mid \mathbf{y}) = (\mathbf{x} \mid \varphi^*(\mathbf{y})) \tag{A.13}$$

pour tout couple $(\mathbf{x}, \mathbf{y})$ de vecteurs de E. Cette application φ^* est appelée *adjointe* de φ.

Choisissons une base orthonormale $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ de E, désignons par $\mathbf{A}$ la matrice de φ dans cette base et par φ^* l'application linéaire de E dans E associée à la matrice $\mathbf{A}^*$ à partir de cette même base. En vertu de (A.8), pour tout couple $(\mathbf{x}, \mathbf{y})$ de vecteurs de E,

$$(\varphi(\mathbf{x}) \mid \mathbf{y}) = \mathbf{y}_e^* \mathbf{A} \mathbf{x}_e. \tag{A.14}$$

et

$$(\mathbf{x} \mid \varphi^*(\mathbf{y})) = (\mathbf{A}^* \mathbf{y}_e)^* \mathbf{x}_e = \mathbf{y}_e^* \mathbf{A} \mathbf{x}_e. \tag{A.15}$$

Il s'ensuit que φ^* vérifie (A.13). Supposons que $\varphi^{*\prime}$ vérifie également (A.13). Alors

$$(\mathbf{x} \mid \varphi^*(\mathbf{y})) - (\mathbf{x} \mid \varphi^{*\prime}(\mathbf{y})) = (\mathbf{x} \mid \varphi^*(\mathbf{y}) - \varphi^{*\prime}(\mathbf{y})) = 0,$$

ce qui entraîne, $\mathbf{x}$ et $\mathbf{y}$ étant arbitraires, que $\varphi^*(\mathbf{y}) - \varphi^{*\prime}(\mathbf{y}) = \mathbf{0}$ pour tout $\mathbf{y}$, donc que $\varphi^* = \varphi^{*\prime}$.

A.5.3 Remarque

On peut se demander s'il existe des applications non linéaires φ de E dans E admettant une adjointe, c'est-à-dire une application φ^* de E dans E vérifiant (A.13) pour tout couple $(\mathbf{x}, \mathbf{y})$ de vecteurs de E. La réponse à cette question est négative et pour s'en assurer, il suffit d'adapter les arguments utilisés dans 8.6.2.

A.5.4 Caractérisation des transformations hermitiennes

Il est évident que les transformations hermitiennes de E sont les applications linéaires φ de E dans E telles que $\varphi^* = \varphi$. Pour cette raison, on les appelle également *transformations auto-adjointes*.

A.5.5 Matrice de l'adjointe d'une application linéaire

Soit φ une application linéaire de E dans E. La matrice de φ^* dans une base orthonormale quelconque de E est l'adjointe de la matrice de φ dans cette même base; en d'autres termes,

$$\mathbf{A}_{\varphi^*} = \mathbf{A}_{\varphi}^*. \tag{A.16}$$

Cela est vrai pour la base choisie dans A.5.2 en vue de définir φ^*, donc pour toute base orthonormale, puisque φ^* est unique.

A.5.6 Propriétés de l'adjointe

De (A.16) et des propriétés (a')–(f') de A.2.3 (ou directement de (A.13)), nous déduisons les propriétés suivantes:

(a) $(\varphi^*)^* = \varphi$.
(b) $(\varphi + \psi)^* = \varphi^* + \psi^*$.
(c) $(\alpha\varphi)^* = \bar{\alpha}\varphi^*$.
(d) $(\varphi \circ \psi)^* = \psi^* \circ \varphi^*$.
(e) $(\varphi^{-1})^* = (\varphi^*)^{-1}$ (si φ est bijective).
(f) $\operatorname{rg}\varphi^* = \operatorname{rg}\varphi$.

A.5.7 Transformations normales

On dit qu'une application linéaire φ de E dans E est une *transformation normale* de E si

$$\varphi^* \circ \varphi = \varphi \circ \varphi^*. \tag{A.17}$$

Il résulte aussitôt de (A.16) et (A.17) que la matrice d'une transformation normale dans n'importe quelle base orthonormale de E est normale. Réciproquement, s'il existe une base orthonormale de E dans laquelle la matrice de φ est normale, alors φ est une transformation normale.

Puisque les matrices hermitiennes et unitaires sont normales, toute transformation hermitienne ou unitaire est normale. Cette conclusion résulte aussi directement des relations $\varphi^* = \varphi$ (caractérisant les transformations hermitiennes) et $\varphi^* \circ \varphi = \mathrm{id}_E$ (caractérisant les transformations unitaires).

A.5.8 Relation entre les valeurs propres et les sous-espaces propres d'une transformation normale et de son adjointe

Soit φ une transformation normale de E. Pour tout couple $(\mathbf{x}, \mathbf{y})$ de vecteurs de E,

$$(\varphi(\mathbf{x}) \mid \varphi(\mathbf{y})) = (\mathbf{x} \mid (\varphi^* \circ \varphi)(\mathbf{y})) = (\mathbf{x} \mid (\varphi \circ \varphi^*)(\mathbf{y})) = (\varphi^*(\mathbf{x}) \mid \varphi^*(\mathbf{y})),$$

donc, en particulier,

$$\| \varphi(\mathbf{x}) \| = \| \varphi^*(\mathbf{x}) \| . \tag{A.18}$$

D'autre part, pour tout nombre complexe t, $\varphi - t\mathrm{id}_E$ est une transformation normale de E, car $\varphi - t\mathrm{id}_E$ commute avec son adjointe $\varphi^* - \bar{t}\mathrm{id}_E$. Cette remarque, jointe à (A.18), entraîne que

$$\| \varphi(\mathbf{x}) - t\mathbf{x} \| = \| \varphi^*(\mathbf{x}) - \bar{t}\mathbf{x} \|$$

pour tout vecteur $\mathbf{x}$ de E. Nous en concluons que si λ est une valeur propre de φ, $\bar{\lambda}$ est une valeur propre de φ^* (ce qui découle également de (A.16)) et les sous-espaces propres de φ et de φ^* associés respectivement à λ et à $\bar{\lambda}$ sont égaux.

A.5.9 Orthogonalité des sous-espaces propres d'une transformation normale

Deux vecteurs propres $\mathbf{x}$ et $\mathbf{y}$ d'une transformation normale, ou d'une matrice normale, associés à des valeurs propres distinctes λ et μ sont orthogonaux. La démonstration de cette assertion est analogue à celle de la proposition 8.6.5; il suffit d'utiliser (A.13) au lieu de (8.33) et de remarquer que $\varphi^*(\mathbf{y}) = \bar{\mu}\mathbf{y}$ (d'après A.5.8).

A.5.10 Caractérisation des transformations normales

Soit φ une application linéaire de E dans E. Les conditions suivantes sont équivalentes:

(a) φ est une transformation normale.

(b) E admet une base constituée de vecteurs propres de φ orthogonaux deux à deux.

(c) φ est diagonalisable et ses sous-espaces propres sont orthogonaux deux à deux.

Voici la preuve:

(a) entraîne (b). Nous raisonnons par récurrence sur la dimension de l'espace vectoriel hermitien. Lorsque cette dimension est 1, l'assertion est évidente. Supposons qu'elle soit vraie pour tout espace vectoriel hermitien de dimension $n - 1$ ($n > 1$). Considérons un vecteur propre $\mathbf{e}$ de φ associé à une valeur propre quelconque λ. D'après A.5.8, $\mathbf{e}$ est aussi un vecteur propre de φ^* associé à la valeur propre $\bar{\lambda}$. Désignons par S le complémentaire orthogonal dans E de la droite vectorielle engendrée par $\mathbf{e}$. Pour tout vecteur $\mathbf{x}$ de S,

$$(\varphi(\mathbf{x}) \mid \mathbf{e}) = (\mathbf{x} \mid \varphi^*(\mathbf{e})) = (\mathbf{x} \mid \bar{\lambda}\mathbf{e}) = \lambda(\mathbf{x} \mid \mathbf{e}) = 0$$

et

$$(\varphi^*(\mathbf{x}) \mid \mathbf{e}) = (\mathbf{x} \mid \varphi(\mathbf{e})) = (\mathbf{x} \mid \lambda\mathbf{e}) = \bar{\lambda}(\mathbf{x} \mid \mathbf{e}) = 0,$$

ce qui montre que S est stable par φ et par φ^*. Il en résulte aussitôt que la restriction φ' de φ à S est une transformation normale de S. Le reste s'établit comme dans la démonstration du théorème 8.6.6.

(b) entraîne (a). La matrice de φ dans une base constituée de vecteurs propres est diagonale, donc normale. Si cette base est de plus orthonormale, le fait que la matrice soit normale entraîne que φ est une transformation normale.

L'équivalence de (b) et (c) s'établit comme dans la démonstration du théorème 8.6.6.

A.5.11. Décomposition spectrale

La décomposition spectrale 8.6.7 vaut pour les transformations normales.

A.5.12 Interprétation matricielle

Par l'entremise de l'application linéaire associée canoniquement à une matrice, l'équivalence des conditions (a), (b) et (c) de A.5.10 démontre l'assertion suivante: une matrice carrée $\mathbf{A}$ d'ordre n est normale si et seulement si elle admet n vecteurs propres orthogonaux deux à deux, autrement dit, si et seulement s'il existe une matrice unitaire $\mathbf{P}$ et une matrice diagonale $\mathbf{D}$ vérifiant (A.12).

On remarquera que cette assertion n'a pas d'analogue réel, car une matrice normale réelle peut n'admettre aucun vecteur propre réel (il suffit de penser aux matrices orthogonales d'ordre 2 et de déterminant 1).

A.6 FORMES SESQUILINÉAIRES HERMITIENNES

A.6.1 Introduction

La section 9.1, hormis l'exemple (3) de 9.1.8, conserve sa validité dans le cas où les scalaires sont complexes: on parle alors de formes bilinéaires et de formes quadratiques complexes. La réduction ne peut cependant être effectuée par la méthode décrite dans 9.1.7. Elle peut, par exemple, être réalisée par complétion des carrés (cf. exemple (2) de 9.1.8 et exercice 9.4.3). Les sections 9.2 et 9.3 ne s'appliquent pas aux formes bilinéaires complexes, mais elles s'adaptent, tout comme 9.1.7, aux formes sesquilinéaires que nous allons introduire et étudier.

Tout au long de cette section, E désignera un espace vectoriel complexe de dimension finie non nulle n.

A.6.2 Formes semi-linéaires

On appelle *forme semi-linéaire* ou *antilinéaire* dans E toute application f de E dans $\mathbb{C}$ telle que

$$f(\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2) = \bar{\alpha}_1 f(\mathbf{x}_1) + \bar{\alpha}_2 f(\mathbf{x}_2), \tag{A.19}$$

pour tout couple $(\mathbf{x}_1, \mathbf{x}_2)$ de vecteurs de E et tout couple (α_1, α_2) de nombres complexes.

A.6.3 Formes sesquilinéaires

On appelle *forme sesquilinéaire* dans E toute application f de l'ensemble des couples de vecteurs de E dans $\mathbb{C}$ satisfaisant aux deux conditions suivantes:

(a) Pour tout vecteur $\mathbf{y}$ de E, l'application $\mathbf{x} \to f(\mathbf{x}, \mathbf{y})$ est une forme linéaire dans E.

(b) Pour tout vecteur $\mathbf{x}$ de E, l'application $\mathbf{y} \to f(\mathbf{x}, \mathbf{y})$ est une forme semi-linéaire dans E.

A.6.4 Formes sesquilinéaires hermitiennes et formes quadratiques hermitiennes associées

On dit qu'une forme sesquilinéaire f dans E est hermitienne si $f(\mathbf{x}, \mathbf{y}) = \overline{f(\mathbf{y}, \mathbf{x})}$ pour tout couple $(\mathbf{x}, \mathbf{y})$ de vecteurs de E.

Si f est une forme sesquilinéaire hermitienne, $f(\mathbf{x}, \mathbf{x})$ est réel pour tout vecteur $\mathbf{x}$ de E; on appelle *forme quadratique hermitienne* associée à f l'application q de E dans $\mathbb{R}$

$$\mathbf{x} \to q(\mathbf{x}) = f(\mathbf{x}, \mathbf{x}).$$

A.6.5 Matrice d'une forme sesquilinéaire

Soit f une forme sesquilinéaire dans E et $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ une base de E. En raison de la semi-linéarité à droite, l'expression de $f(\mathbf{x}, \mathbf{y})$ en fonction des composantes de $\mathbf{x}$ et de $\mathbf{y}$ n'est plus (9.2), mais l'expression analogue

$$f(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i \bar{y}_j, \tag{A.20}$$

où $a_{ij} = f(\mathbf{e}_i, \mathbf{e}_j)$.

Comme pour les formes bilinéaires, on appelle la matrice $\mathbf{A} = (a_{ij})$ *matrice de la forme sesquilinéaire* f. Il est évident que f est hermitienne si et seulement si $\mathbf{A}$ est hermitienne.

A l'aide de la matrice $\mathbf{A}$, la relation (A.20) s'écrit sous la forme

$$f(\mathbf{x}, \mathbf{y}) = {}^t\mathbf{x}_e \mathbf{A} \bar{\mathbf{y}}_e. \tag{A.21}$$

La notion de *forme sesquilinéaire associée canoniquement à une matrice carrée* $\mathbf{A}$ se définit comme dans 9.1.5, au moyen de (A.21).

A.6.6 Transformation de la matrice d'une forme sesquilinéaire par suite d'un changement de base

Soit f une forme sesquilinéaire dans E, $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ et $(\mathbf{e}'_1, \mathbf{e}'_2, ..., \mathbf{e}'_n)$ des bases de E, $\mathbf{A}$ et $\mathbf{A}'$ les matrices respectives de f dans chacune de ces bases. Les arguments utilisés dans 9.1.6 montrent que $\mathbf{A}$ et $\mathbf{A}'$ sont liées par la relation

$$\mathbf{A}' = {}^t\mathbf{P}_{ee'} \mathbf{A} \bar{\mathbf{P}}_{ee'} = \bar{\mathbf{P}}^*_{ee'} \mathbf{A} \bar{\mathbf{P}}_{ee'}. \tag{A.22}$$

A.6.7 Réduction d'une forme sesquilinéaire hermitienne

Le procédé de réduction décrit dans 9.1.7 s'adapte aux formes sesquilinéaires hermitiennes. Soit en effet f une telle forme et $\mathbf{A}$ la matrice de f dans une base $(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$ de E. D'après A.4.4 et plus particulièrement (A.12), il existe une matrice unitaire $\mathbf{P}$ et une matrice diagonale réelle $\mathbf{D}$ telles que

$$\mathbf{D} = \mathbf{P}^*\mathbf{A}\mathbf{P}. \tag{A.23}$$

Selon ce qui est rappelé dans A.4.4, une telle matrice $\mathbf{P}$ s'obtient en prenant comme colonnes n vecteurs propres de $\mathbf{A}$ orthogonaux deux à deux et unitaires; les termes diagonaux de $\mathbf{D}$ sont alors les valeurs propres correspondantes. Posons, pour $j = 1, 2, ..., n$,

$$\mathbf{e}'_j = \bar{p}_{1j}\mathbf{e}_1 + \bar{p}_{2j}\mathbf{e}_2 + ... + \bar{p}_{nj}\mathbf{e}_n,$$

où $p_{1j}, p_{2j}, ..., p_{nj}$ sont les termes de la j-ième colonne de $\mathbf{P}$. La famille de vecteurs $(\mathbf{e}'_1, \mathbf{e}'_2, ..., \mathbf{e}'_n)$ est une base de E et la matrice de passage $\mathbf{P}_{ee'}$ est la matrice unitaire $\bar{\mathbf{P}}$. En outre, d'après (A.22) et (A.23), la matrice de f dans la base $(\mathbf{e}'_1, \mathbf{e}'_2, ..., \mathbf{e}'_n)$ est $\mathbf{D}$, autrement dit,

$$f(\mathbf{x}, \mathbf{y}) = d_1 x'_1 \bar{y}'_1 + d_2 x'_2 \bar{y}'_2 + ... + d_n x'_n \bar{y}'_n,$$

$$q(\mathbf{x}) = d_1 |x'_1|^2 + d_2 |x'_2|^2 + ... + d_n |x'_n|^2,$$

où $x'_1, x'_2, ..., x'_n$ et $y'_1, y'_2, ..., y'_n$ sont les composantes respectives de $\mathbf{x}$ et de $\mathbf{y}$ dans la base $(\mathbf{e}'_1, \mathbf{e}'_2, ..., \mathbf{e}'_n)$ et $d_1, d_2, ..., d_n$ les termes diagonaux de $\mathbf{D}$.

A.6.8 Formes sesquilinéaires hermitiennes définies non négatives et définies positives

Soit f une forme sesquilinéaire hermitienne dans E et q la forme quadratique hermitienne associée à f. Comme pour les formes bilinéaires symétriques réelles, on dit que f est *définie non négative* (*définie positive*) si $f(\mathbf{x}, \mathbf{x}) = q(\mathbf{x}) \geqslant 0$ (> 0) pour tout vecteur $\mathbf{x}$ (tout vecteur non nul $\mathbf{x}$) de E.

On notera que la notion de forme sesquilinéaire hermitienne définie positive se confond avec celle de produit scalaire hermitien.

A.6.9 Matrices hermitiennes définies non négatives et définies positives

Ces notions se définissent comme dans 9.2.3, tout en notant (cf. A.6.4) que ${}^t\mathbf{x}\mathbf{A}\bar{\mathbf{x}}$ est réel si $\mathbf{A}$ est une matrice hermitienne et $\mathbf{x}$ un vecteur-colonne complexe. Ainsi, une matrice hermitienne $\mathbf{A}$ d'ordre n est *définie non négative* (*définie positive*) si et seulement si ${}^t\mathbf{x}\mathbf{A}\bar{\mathbf{x}} \geqslant 0$ (> 0) pour tout vecteur-colonne $\mathbf{x}$ (tout vecteur-colonne non nul $\mathbf{x}$) de $\mathbb{C}^n$. On remarquera que $\mathbf{x}$ peut être remplacé par $\bar{\mathbf{x}}$, donc que ${}^t\mathbf{x}\mathbf{A}\bar{\mathbf{x}}$ peut être changé en $\mathbf{x}^*\mathbf{A}\mathbf{x}$.

Les propositions 9.2.4 et 9.2.5 gardent leur validité si l'on remplace «symétrique» par «hermitien». On rappellera que les valeurs propres (cf. A.4.4) et les mineurs principaux (cf. A.2.5) d'une matrice hermitienne sont réels.

Dans la démonstration de la proposition 9.2.5, la matrice $\mathbf{P}$ doit être définie autrement, à savoir par

$$\mathbf{P} = \left[\begin{array}{ccc:c} & \mathbf{I}_{n-1} & & -\overline{\tilde{\mathbf{A}}^{-1}\mathbf{a}} \\ \hdashline 0 & 0 \; . \; . \; . & 0 & 1 \end{array}\right].$$

En outre, ${}^t\mathbf{PA\bar{P}}$ remplace ${}^t\mathbf{PAP}$ et la relation (9.11) devient

$${}^t\mathbf{x}\mathbf{A}'\bar{\mathbf{x}} = {}^t\tilde{\mathbf{x}}\tilde{\mathbf{A}}\bar{\tilde{\mathbf{x}}} + b|x_n|^2.$$

Bien entendu, $\mathbb{C}$ prend la place de $\mathbb{R}$.

A.6.10 Réduction simultanée

Deux formes sesquilinéaires hermitiennes dont l'une est définie positive peuvent être réduites simultanément. Cette assertion est établie par la démonstration de la proposition 9.3.2 si l'on change «produit scalaire» en «produit scalaire hermitien» et «orthogonal» en «unitaire».

Soit $\mathbf{A}$ une matrice hermitienne et $\mathbf{B}$ une matrice hermitienne définie positive du même ordre que $\mathbf{A}$. Les notions de valeur propre et de vecteur propre de $(\mathbf{A}, \mathbf{B})$ sont définies comme dans 9.3.3. Si λ est une valeur propre de $(\mathbf{A}, \mathbf{B})$ et $\mathbf{x}$ un vecteur propre associé à λ, alors

$$\mathbf{x}^*\mathbf{A}\mathbf{x} = \lambda\mathbf{x}^*\mathbf{B}\mathbf{x},$$

ce qui montre que λ est réel, puisque $\mathbf{x}^*\mathbf{A}\mathbf{x}$ et $\mathbf{x}^*\mathbf{B}\mathbf{x}$ le sont et que $\mathbf{x}^*\mathbf{B}\mathbf{x}$ est non nul. Compte tenu de ce résultat, on montre, par une suite d'égalités analogues à celles qui figurent sous (9.14), que si $\mathbf{x}$ et $\mathbf{y}$ sont des vecteurs propres de $(\mathbf{A}, \mathbf{B})$ associés à des valeurs propres distinctes, alors $\mathbf{x}$ et $\mathbf{y}$ sont orthogonaux pour le produit scalaire hermitien défini par $\bar{\mathbf{B}}$.

La proposition 9.3.4 doit être modifiée ainsi: $\mathbf{A}$ et $\mathbf{B}$ étant les matrices introduites ci-dessus, une matrice carrée $\mathbf{P}$ vérifie les relations

$${}^t\mathbf{PA\bar{P}} = \mathbf{D} \text{ (diagonale) } \quad \text{et} \quad {}^t\mathbf{PB\bar{P}} = \mathbf{I}$$

si et seulement si les colonnes de $\bar{\mathbf{P}}$ sont des vecteurs propres de $(\mathbf{A}, \mathbf{B})$ et forment une famille orthonormale pour le produit scalaire hermitien défini par $\bar{\mathbf{B}}$. La démonstration est analogue à celle de la proposition 9.3.4.

A.7 EXERCICES

A.7.1 Montrer qu'un système linéaire de m équations à n inconnues, de matrice associée $\mathbf{A}$ et de second membre $\mathbf{b}$, admet au moins une solution si et seulement si $\mathbf{b}$ est orthogonal (pour le produit scalaire canonique dans $\mathbb{C}^m$) à toute solution du système homogène de matrice associée $\mathbf{A}^*$. (Introduire les applications linéaires φ et φ^* associées canoniquement à $\mathbf{A}$ et à $\mathbf{A}^*$ et démontrer successivement: (a) $\mathrm{Ker}\varphi^*$ et $\mathrm{Im}\varphi$ sont des sous-espaces vectoriels orthogonaux de $\mathbb{C}^m$; (b) $\mathrm{Ker}\varphi^* \cap \mathrm{Im}\varphi = \{\mathbf{0}\}$; (c) $\dim\mathrm{Ker}\varphi^* + \mathrm{rg}\varphi = m$; (d) $\mathrm{Ker}\varphi^* \oplus \mathrm{Im}\varphi = \mathbb{C}^m$.)

A.7.2 Soit $\mathbf{A}$ une matrice carrée. Montrer que $(\mathbf{A}^*)^k = (\mathbf{A}^k)^*$ pour tout entier non négatif k (pour tout entier k si $\mathbf{A}$ est inversible).

A.7.3 Soit $\mathbf{A}$ et $\mathbf{B}$ des matrices carrées d'ordre n, $\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_n$ et $\mathbf{b}_1, \mathbf{b}_2, ..., \mathbf{b}_n$ leurs colonnes respectives. Montrer que les conditions suivantes sont équivalentes:

(a) $(\mathbf{a}_i \mid \mathbf{b}_j) = \begin{cases} 1 \text{ si } i = j, \\ 0 \text{ si } i \neq j. \end{cases}$

(b) $\mathbf{A}$ est inversible et $\mathbf{A}^{-1} = \mathbf{B}^*$.

(c) $\mathbf{a}_1\mathbf{b}_1^* + \mathbf{a}_2\mathbf{b}_2^* + ... + \mathbf{a}_n\mathbf{b}_n^* = \mathbf{I}_n$.

(d) $\mathbf{x} = (\mathbf{x} \mid \mathbf{b}_1)\mathbf{a}_1 + (\mathbf{x} \mid \mathbf{b}_2)\mathbf{a}_2 + ... + (\mathbf{x} \mid \mathbf{b}_n)\mathbf{a}_n$, pour tout vecteur-colonne $\mathbf{x}$ de $\mathbb{C}^n$.

Déduire ensuite de (d) que $(\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_n)$ et, par symétrie, $(\mathbf{b}_1, \mathbf{b}_2, ..., \mathbf{b}_n)$ sont des bases de $\mathbb{C}^n$.

A.7.4 Soit $\mathbf{A}$ une matrice carrée n'admettant que des valeurs propres simples, que l'on désigne par $\lambda_1, \lambda_2, ..., \lambda_n$. Soit $\mathbf{p}_1, \mathbf{p}_2, ..., \mathbf{p}_n$ des vecteurs propres de $\mathbf{A}$ associés respectivement à ces valeurs propres et $\mathbf{q}_1, \mathbf{q}_2, ..., \mathbf{q}_n$ des vecteurs propres de $\mathbf{A}^*$ associés respectivement aux valeurs propres $\bar{\lambda}_1, \bar{\lambda}_2, ..., \bar{\lambda}_n$.

(a) Montrer que $(\mathbf{p}_i \mid \mathbf{q}_j)$ est non nul si $i = j$ et nul si $i \neq j$.

(b) Montrer que si $\mathbf{P}$ et $\mathbf{Q}$ sont les matrices ayant pour colonnes respectives $\mathbf{p}_1, \mathbf{p}_2, ..., \mathbf{p}_n$ et $\dfrac{1}{(\mathbf{p}_1 \mid \mathbf{q}_1)}\mathbf{q}_1, \dfrac{1}{(\mathbf{p}_2 \mid \mathbf{q}_2)}\mathbf{q}_2, ..., \dfrac{1}{(\mathbf{p}_n \mid \mathbf{q}_n)}\mathbf{q}_n$, alors $\mathbf{P}^{-1} = \mathbf{Q}^*$ et donc $\mathbf{Q}^*\mathbf{A}\mathbf{P} = \mathrm{diag}(\lambda_1, \lambda_2, ..., \lambda_n)$. (Utiliser l'équivalence de conditions (a) et (b) de l'exercice A.7.3.)

A.7.5 Soit

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

Trouver une matrice unitaire $\mathbf{P}$ et une matrice diagonale $\mathbf{D}$ telles que $\mathbf{P}^*\mathbf{A}\mathbf{P} = \mathbf{D}$.

A.7.6 Trigonaliser la matrice

$$\begin{pmatrix} 3+2i & 2+i & 2-3i & 1-2i \\ i & -5+2i & 1 & 2+5i \\ -2+3i & -1+2i & 3+2i & 2+i \\ -1 & -2-5i & i & -5+2i \end{pmatrix}.$$

A.7.7 Trouver une base de solutions de l'équation différentielle $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$, où

$$\mathbf{A} = \begin{bmatrix} \sqrt{3} & -1 \\ 1 & \sqrt{3} \end{bmatrix}.$$

A.7.8 Dans un espace vectoriel hermitien E de dimension 4, muni d'une base orthonormale $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4)$, on considère la forme sesquilinéaire hermitienne f de matrice

$$\mathbf{A} = \begin{bmatrix} 0 & 0 & 0 & \mathrm{i} \\ 0 & 0 & \mathrm{i} & 0 \\ 0 & -\mathrm{i} & 0 & 0 \\ -\mathrm{i} & 0 & 0 & 0 \end{bmatrix}.$$

Trouver une base orthonormale dans laquelle f est réduite.

A.7.9 S'assurer que la loi d'inertie de Sylvester (exercice 9.4.4) vaut pour les formes quadratiques hermitiennes.

A.7.10 Adapter la méthode de réduction par complétion des carrés (exercice 9.4.3) au cas des formes quadratiques hermitiennes.

A.7.11 Réduire la forme quadratique hermitienne

$$\mathbf{x} \longrightarrow q(\mathbf{x}) = |x_1|^2 + \mathrm{i}(2x_1\bar{x}_2 - 2x_2\bar{x}_1 + x_2\bar{x}_3 - x_3\bar{x}_2)$$

(a) par la méthode trouvée dans l'exercice A.7.10,
(b) par le calcul des valeurs propres de la matrice de q.

A.7.12 Montrer que si $\mathbf{H}$ est une matrice hermitienne, $\mathbf{U} = \exp(\mathrm{i}\mathbf{H})$ est une matrice unitaire. Réciproquement, montrer que si $\mathbf{U}$ est une matrice unitaire, il existe une matrice hermitienne $\mathbf{H}$ telle que $\mathbf{U} = \exp(\mathrm{i}\mathbf{H})$. (S'assurer que $\mathbf{U}$ s'écrit sous la forme $\mathbf{P}\mathrm{diag}(e^{\mathrm{i}\theta_1}, e^{\mathrm{i}\theta_2}, \ldots, e^{\mathrm{i}\theta_n})\mathbf{P}^*$, où $\mathbf{P}$ est unitaire, et poser $\mathbf{H} = \mathbf{P}\mathrm{diag}(\theta_1, \theta_2, \ldots, \theta_n)\mathbf{P}^*$.)

A.7.13 Soit $\mathbf{A}$ une matrice carrée inversible. Montrer qu'il existe une et une seule matrice hermitienne définie positive $\mathbf{R}$ et une et une seule matrice unitaire $\mathbf{U}$ telles que $\mathbf{A} = \mathbf{R}\mathbf{U}$. (Prendre pour $\mathbf{R}$ l'unique racine carrée définie positive de $\mathbf{A}^*\mathbf{A}$ (cf. exercices 9.4.9 et 9.4.10).)

A.7.14 *Représentation polaire.* Combiner les résultats des exercices A.7.12 et A.7.13.

Index

Bibliographie

M. Berger, *Géométrie*, Cedic/Fernand Nathan, Paris, 1977.

A. Calame, *Mathématiques modernes*, Editions du Griffon, Neuchâtel, 1968.

L. Chambadal, J.L. Ovaert, *Algèbre linéaire et algèbre tensorielle*, Dunod, Paris, 1968.

J. Dieudonné, *Algèbre linéaire et géométrie élémentaire*, Hermann, Paris, 1964.

J. Frenkel, *Géométrie pour l'élève-professeur*, Hermann, Paris, 1973.

F.R. Gantmacher, *The theory of matrices*, Chelsea Publishing Company, New York, 1977.

R.J. Goult, *Applied linear algebra*, Ellis Horwood Ltd., Chichester, 1978.

W. Greub, *Linear algebra*, Springer-Verlag, New York, 1981.

P.R. Halmos, *Finite-dimensional vector spaces*, Van Nostrand, New York, 1958.

K. Hoffman, R. Kunze, *Linear algebra*, Prentice-Hall, Englewood Cliffs, New Jersey, 1961.

M. Jeger, B. Eckmann, *Einführung in die vektorielle Geometrie und lineare Algebra*, Birkhäuser Verlag, Basel, 1967.

H.-J. Kowalsky, *Lineare Algebra*, W. de Gruyter, Berlin, 1970.

P. Lancaster, M. Tismenetsky, *The theory of matrices*, Academic Press, Orlando, Florida, 1985.

S. Lang, *Algèbre linéaire*, Inter Editions, Paris, 1976.

H. Liebeck, *Algebra for scientists & engineers*, J. Wiley & Sons, London, New York, 1969.

S. Mac Lane, G. Birkhoff, *Algèbre*, Gauthier-Villars, Paris, 1971.

B. Noble, *Applied linear algebra*, Prentice-Hall, Englewood Cliffs, New Jersey, 1969.

E. Ramis, C. Deschamps, J. Odoux, *Cours de mathématiques spéciales*, Masson, Paris, 1977.

E. Snapper, R.J. Troyer, *Metric affine geometry*, Academic Press, New York, London, 1971.

E. Sperner, *Einführung in die Analytische Geometrie und Algebra*, Vandenhoeck & Ruprecht, Göttingen, 1957.

C. Tisseron, *Géométries affine, projective et euclidienne*, Hermann, Paris, 1983.