Algèbre Linéaire

Cours complet avec exercices

Niveau : Licence 1re–2e année

Destiné aux étudiants en mathématiques et en ingénierie

Notes de cours

Février 2026

Table des matières

Préface vii

Notation ix

1 Préliminaires 1
Introduction . . . 1
1.1 Ensembles . . . 1
1.2 Applications . . . 2
1.3 Relations . . . 3
1.3.1 Relations d'équivalence . . . 3
1.3.2 Relations d'ordre . . . 4
1.4 Structures algébriques . . . 4
1.4.1 Groupes . . . 4
1.4.2 Anneaux . . . 5
1.4.3 Corps . . . 6
1.5 Les axiomes de corps en détail . . . 6
1.5.1 Le corps $\mathbb{R}$ des nombres réels . . . 7
1.5.2 Le corps $\mathbb{C}$ des nombres complexes . . . 7
1.5.3 Les corps finis $\mathbb{F}_p$ . . . 8
1.6 Techniques de démonstration . . . 9
1.6.1 Preuve directe . . . 9
1.6.2 Preuve par contraposée . . . 9
1.6.3 Preuve par l'absurde . . . 10
1.6.4 Raisonnement par récurrence . . . 10
1.7 Exercices . . . 11
Résumé du chapitre . . . 12

2 Espaces vectoriels et sous-espaces 15
Introduction . . . 15
2.1 Espace vectoriel sur un corps . . . 15
2.1.1 Définition axiomatique . . . 15
2.1.2 Exemples fondamentaux . . . 17
2.2 Sous-espaces vectoriels . . . 18
2.3 Intersection et somme de sous-espaces . . . 20
2.3.1 Intersection . . . 20
2.3.2 Somme de sous-espaces . . . 20
2.4 Somme directe et sous-espaces supplémentaires . . . 21
2.5 Combinaisons linéaires et sous-espace engendré . . . 22
2.6 Familles libres . . . 23
2.7 Bases . . . 25
2.8 Dimension . . . 26
2.9 Dimension de la somme — formule de Grassmann . . . 28

2.10 Rang d'une famille de vecteurs . . . 29
2.11 Applications . . . 30
2.11.1 Espaces de solutions . . . 30
2.11.2 Espaces de polynômes . . . 31
2.11.3 Somme directe — illustration géométrique . . . 31
2.12 Exercices . . . 31
Résumé du chapitre . . . 34

3 Applications linéaires et matrices **37**
Introduction . . . 37
3.1 Définition d'une application linéaire . . . 38
3.2 Exemples fondamentaux . . . 39
3.3 Noyau et image . . . 40
3.4 Théorème du rang . . . 41
3.5 Injectivité, surjectivité, isomorphismes . . . 42
3.6 Composition d'applications linéaires . . . 44
3.7 L'espace vectoriel $\mathcal{L}(E, F)$ . . . 45
3.8 Matrice d'une application linéaire . . . 45
3.9 Changement de base . . . 47
3.10 Opérations sur les matrices . . . 48
3.10.1 Addition et multiplication scalaire . . . 48
3.10.2 Produit matriciel . . . 48
3.10.3 Transposition . . . 49
3.11 Matrices inversibles . . . 49
3.12 Espaces duaux (introduction) . . . 50
3.13 Applications . . . 51
3.13.1 Transformations en infographie . . . 51
3.13.2 Aperçu des chaînes de Markov . . . 52
3.14 Exercices . . . 52
Résumé du chapitre . . . 55

4 Systèmes d'équations linéaires et élimination de Gauss **57**
Introduction . . . 57
4.1 Systèmes d'équations linéaires . . . 58
4.1.1 Définition et notation . . . 58
4.1.2 Écriture matricielle . . . 58
4.1.3 Systèmes homogènes et non homogènes . . . 59
4.2 Opérations élémentaires sur les lignes . . . 59
4.3 Forme échelonnée et forme échelonnée réduite . . . 60
4.4 Élimination de Gauss . . . 61
4.4.1 Algorithme . . . 61
4.4.2 Exemple détaillé . . . 61
4.5 Élimination de Gauss–Jordan . . . 63
4.6 Rang d'une matrice . . . 63
4.7 Théorème de Rouché–Capelli . . . 64
4.8 Structure de l'ensemble des solutions . . . 65
4.9 Représentation paramétrique des solutions . . . 65
4.10 Interprétation géométrique . . . 66
4.10.1 Systèmes 2×2 . . . 66
4.10.2 Systèmes 3×3 . . . 66
4.11 Applications . . . 67
4.11.1 Réseaux de flux . . . 67

4.11.2 Circuits électriques . . . 68
4.11.3 Aperçu : moindres carrés . . . 68
Résumé du chapitre . . . 68
Exercices . . . 69

Index **73**

5 Déterminants **75**
Introduction . . . 75
5.1 Permutations et groupe symétrique . . . 76
5.1.1 Définitions fondamentales . . . 76
5.1.2 Transpositions . . . 76
5.1.3 Signature . . . 77
5.2 Définition du déterminant . . . 78
5.2.1 Formule de Leibniz . . . 78
5.2.2 Caractérisation par les formes multilinéaires alternées . . . 78
5.3 Propriétés fondamentales . . . 79
5.3.1 Multilinéarité par rapport aux lignes et colonnes . . . 79
5.3.2 Déterminant d'un produit . . . 80
5.4 Développement selon une ligne ou une colonne . . . 81
5.4.1 Mineur et cofacteur . . . 81
5.4.2 Développement de Laplace . . . 81
5.5 Propriétés complémentaires . . . 82
5.5.1 Déterminant de la transposée . . . 82
5.5.2 Matrices triangulaires . . . 83
5.5.3 Matrices par blocs . . . 83
5.6 Comatrice et règle de Cramer . . . 83
5.6.1 Comatrice (matrice adjointe) . . . 83
5.6.2 Règle de Cramer . . . 84
5.7 Déterminant et inversibilité . . . 85
5.8 Déterminant d'un endomorphisme . . . 86
5.9 Interprétation géométrique . . . 86
5.9.1 Volume signé . . . 86
5.9.2 Cas du plan : aire du parallélogramme . . . 86
5.9.3 Cas de l'espace : volume du parallélépipède . . . 87
5.9.4 Orientation . . . 87
5.10 Applications . . . 88
5.10.1 Produit vectoriel dans $\mathbb{R}^3$ . . . 88
5.10.2 Formule de l'aire d'un triangle . . . 88
5.10.3 Déterminant de Vandermonde . . . 88
5.11 Exercices . . . 90
Résumé du chapitre . . . 93

6 Valeurs propres, vecteurs propres et diagonalisation **95**
Introduction . . . 95
6.1 Valeurs propres et vecteurs propres . . . 95
6.2 Sous-espaces propres . . . 97
6.3 Polynôme caractéristique . . . 98
6.4 Théorème de Cayley-Hamilton . . . 99
6.5 Spectre d'un endomorphisme . . . 100
6.6 Multiplicité algébrique et multiplicité géométrique . . . 101
6.7 Diagonalisabilité . . . 102

6.8 Algorithme de diagonalisation . . . 103
6.9 Trigonalisation . . . 105
6.10 Polynôme minimal . . . 105
6.11 Applications . . . 107
6.11.1 Calcul de A^k . . . 107
6.11.2 Récurrences linéaires : la suite de Fibonacci . . . 107
6.11.3 Chaînes de Markov : état stationnaire . . . 108
6.11.4 Systèmes d'équations différentielles linéaires . . . 108
6.12 Illustrations . . . 108
6.13 Exercices . . . 108
Résumé du chapitre . . . 112

7 Espaces préhilbertiens et orthogonalité 115
Introduction . . . 115
7.1 Formes bilinéaires . . . 116
7.2 Formes quadratiques . . . 117
7.3 Produit scalaire . . . 118
7.3.1 Cas réel . . . 118
7.3.2 Cas complexe . . . 118
7.4 Exemples fondamentaux . . . 119
7.5 Norme, distance et inégalité de Cauchy-Schwarz . . . 120
7.6 Orthogonalité . . . 121
7.7 Projection orthogonale . . . 123
7.8 Procédé de Gram-Schmidt . . . 124
7.9 Bases orthonormées . . . 125
7.10 Matrices orthogonales et unitaires . . . 126
7.11 Opérateur adjoint . . . 128
7.12 Applications . . . 129
7.12.1 Approximation aux moindres carrés . . . 129
7.12.2 Aperçu de la factorisation QR . . . 129
7.12.3 Aperçu des séries de Fourier . . . 129
7.13 Illustrations . . . 130
7.14 Exercices . . . 130
Résumé du chapitre . . . 134

8 Théorème spectral et matrices symétriques 137
Introduction . . . 137
8.1 Opérateurs auto-adjoints . . . 138
8.2 Valeurs propres réelles . . . 139
8.3 Orthogonalité des vecteurs propres . . . 139
8.4 Théorème spectral pour les matrices symétriques réelles . . . 140
8.5 Théorème spectral pour les matrices hermitiennes . . . 141
8.6 Diagonalisation orthogonale : procédure et exemples . . . 142
8.7 Opérateurs normaux . . . 143
8.8 Matrices définies positives et semi-définies positives . . . 144
8.9 Décomposition en valeurs singulières (SVD) . . . 146
8.10 Applications . . . 147
8.10.1 Analyse en composantes principales (ACP) . . . 147
8.10.2 Classification des formes quadratiques . . . 147
8.10.3 Optimisation : min/max des formes quadratiques . . . 147
8.10.4 Vibrations et modes propres . . . 148
8.11 Illustrations . . . 148

8.11.1 Classification des coniques par les valeurs propres . . . 148
8.11.2 Géométrie de la SVD . . . 149
8.12 Exercices . . . 149
Résumé du chapitre . . . 151

9 Forme Normale de Jordan **153**
Introduction . . . 153
9.1 Endomorphismes nilpotents . . . 153
9.2 Sous-espaces invariants . . . 155
9.3 Sous-espaces caractéristiques . . . 156
9.4 Blocs de Jordan et matrices de Jordan . . . 157
9.5 Théorème de la forme normale de Jordan . . . 158
9.6 Calcul de la forme de Jordan . . . 160
9.7 Exemples détaillés . . . 161
9.8 Forme de Jordan et polynôme minimal . . . 163
9.9 Cayley-Hamilton via la forme de Jordan . . . 164
9.10 Exponentielle de matrice . . . 165
9.11 Applications . . . 166
9.11.1 Systèmes d'équations différentielles linéaires . . . 166
9.11.2 Fonctions de matrices . . . 167
9.11.3 Puissances d'une matrice . . . 168
9.12 Exercices . . . 168
Résumé du chapitre . . . 171

Index **173**

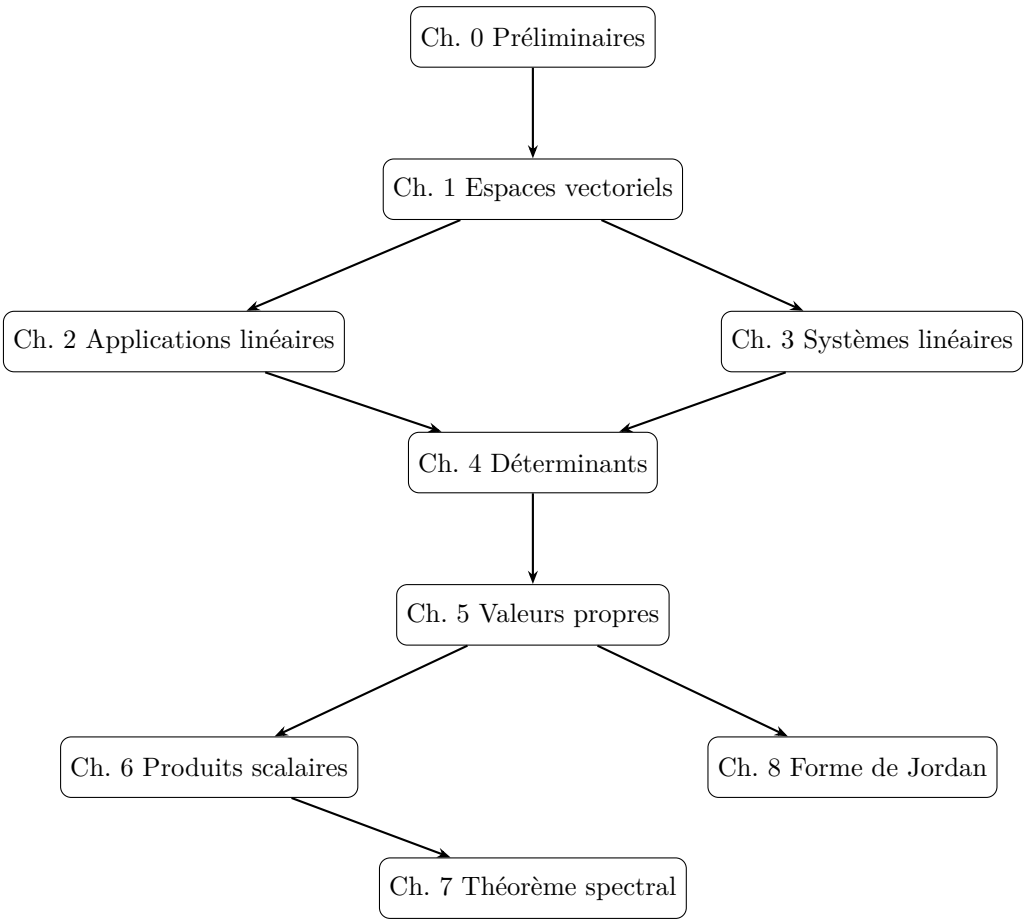
Préface

L'algèbre linéaire est l'une des branches les plus fondamentales et les plus omniprésentes des mathématiques modernes. Ses concepts — espaces vectoriels, applications linéaires, matrices, déterminants, valeurs propres — forment le socle sur lequel reposent de vastes pans de l'analyse, de la géométrie, de la physique, de l'informatique et de l'ingénierie.

Ces notes de cours s'adressent aux étudiants de première et deuxième année de licence en mathématiques ou en école d'ingénieurs. Elles supposent une connaissance solide des mathématiques du lycée (calcul, fonctions, trigonométrie) ainsi que des notions élémentaires de théorie des ensembles et de logique.

Philosophie de ce cours. Nous avons cherché à maintenir un équilibre constant entre *rigueur algébrique* et *intuition géométrique*. Chaque concept abstrait est illustré par des exemples concrets, des figures et des applications. Les démonstrations sont complètes : nous estimons qu'apprendre à lire et à écrire des preuves est un objectif aussi important que la maîtrise des techniques de calcul.

Organisation. Le cours se compose de neuf chapitres, dont voici la structure logique :



Le chapitre 0 (Préliminaires) peut être parcouru rapidement par les étudiants déjà familiers avec les notions de base. Le chapitre 8 (Forme normale de Jordan) est un chapitre avancé, destiné aux étudiants souhaitant approfondir le sujet.

Exercices. Chaque chapitre se termine par une série d'exercices classés par difficulté :

- — exercices de base (vérification directe des définitions),
- — exercices intermédiaires (applications des théorèmes),
- — exercices avancés (problèmes de synthèse, questions ouvertes).

Note historique. L'algèbre linéaire, telle que nous la connaissons aujourd'hui, s'est développée progressivement du XVIII^{e} au XX^{e} siècle. Les contributions majeures incluent : les travaux de Gauss sur l'élimination (1809), la théorie des déterminants de Cauchy (1815), la découverte des valeurs propres par Cauchy et Sylvester, les formes canoniques de Jordan (1870), et la formalisation axiomatique des espaces vectoriels par Peano (1888) et Banach (1920).

Février 2026

Notation

Nous rassemblons ici les notations utilisées tout au long de ce cours. Le lecteur est invité à s'y référer en cas de doute.

Symbole	Signification
Ensembles de nombres	
$\mathbb{N}$	Entiers naturels $\{0, 1, 2, ...\}$
$\mathbb{Z}$	Entiers relatifs
$\mathbb{Q}$	Nombres rationnels
$\mathbb{R}$	Nombres réels
$\mathbb{C}$	Nombres complexes
$\mathbb{K}$	Corps de base (en général $\mathbb{R}$ ou $\mathbb{C}$)
$\mathbb{F}_p$	Corps fini à p éléments
Espaces et structures	
$\mathbb{K}^n$	Espace des n-uplets sur $\mathbb{K}$
$\mathcal{M}_{n,p}(\mathbb{K})$	Espace des matrices $n \times p$ à coefficients dans $\mathbb{K}$
$\mathcal{M}_n(\mathbb{K})$	Matrices carrées $n \times n$
$\mathrm{GL}_n(\mathbb{K})$	Groupe linéaire (matrices inversibles)
$\mathcal{P}_n(\mathbb{K})$	Espace des polynômes de degré $\leq n$
$\mathcal{L}(E, F)$	Espace des applications linéaires de E dans F
$\mathrm{End}(E)$	Endomorphismes de E $(= \mathcal{L}(E, E))$
Applications linéaires et matrices	
$\mathrm{Ker}\, f$	Noyau de f
$\mathrm{Im}\, f$	Image de f
$\mathrm{rg}(f)$, $\mathrm{rg}(A)$	Rang d'une application linéaire ou d'une matrice
Id, I_n	Application identité, matrice identité $n \times n$
A^T	Transposée de A
A^*	Transposée conjuguée (adjointe) de A
A^{-1}	Inverse de A
$\mathrm{tr}(A)$	Trace de A
$\det(A)$	Déterminant de A
Espaces vectoriels	
$\mathrm{Vect}(v_1, \dots, v_k)$	Sous-espace engendré par $v_1, \dots, v_k$
$\dim E$	Dimension de E
$\mathrm{codim}\, F$	Codimension de F dans E
$E \oplus F$	Somme directe
E/F	Espace quotient
E^*	Dual de E

Symbole	Signification
$\mathcal{B} = (e_1, \dots, e_n)$	Base de E
Produit scalaire et normes	
$\langle u, v \rangle$	Produit scalaire de u et v
$\|v\|$	Norme de v
$u \perp v$	u et v sont orthogonaux
$F^\perp$	Orthogonal de F
Valeurs propres	
$\mathrm{Spec}(f)$	Spectre de f (ensemble des valeurs propres)
E_λ	Sous-espace propre associé à λ
$\chi_f(\lambda)$	Polynôme caractéristique de f
$\mu_f(\lambda)$	Polynôme minimal de f
Notations diverses	
$:=$	Égalité par définition
δ_{ij}	Symbole de Kronecker
$\mathrm{sgn}(\sigma)$	Signature d'une permutation σ
$\mathfrak{S}_n$	Groupe symétrique sur $\{1, \dots, n\}$

Chapitre 1

Préliminaires

Introduction

L'algèbre linéaire repose sur un socle de notions que l'on suppose généralement acquises : ensembles, applications, relations, et surtout les structures algébriques élémentaires que sont les groupes, les anneaux et les corps. Ce chapitre zéro a pour vocation de *rappeler* et de *préciser* ces outils, afin que le lecteur dispose d'un langage commun et rigoureux pour la suite du cours.

Nous y revisiterons également les principales techniques de démonstration — preuve directe, contraposée, raisonnement par l'absurde, récurrence — qui seront utilisées de manière intensive dans les chapitres suivants.

Le lecteur déjà à l'aise avec ces prérequis pourra parcourir ce chapitre rapidement, en s'attardant sur les exercices pour vérifier sa maîtrise.

1.1 Ensembles

Définition 1 (Ensemble)

Un *ensemble* est une collection d'objets, appelés *éléments*, réunis selon un critère bien défini. On note $x \in E$ le fait que x est un élément de E, et $x \notin E$ le contraire.

On rappelle les ensembles de nombres classiques :

$$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}.$$

Définition 2 (Opérations ensemblistes)

Soient A et B deux sous-ensembles d'un ensemble E.

(i) **Réunion :** $A \cup B := \{\, x \in E \mid x \in A \text{ ou } x \in B \,\}$.

(ii) **Intersection :** $A \cap B := \{\, x \in E \mid x \in A \text{ et } x \in B \,\}$.

(iii) **Différence :** $A \setminus B := \{\, x \in E \mid x \in A \text{ et } x \notin B \,\}$.

(iv) **Complémentaire :** $A^c := E \setminus A$.

(v) **Produit cartésien :** $A \times B := \{\, (a, b) \mid a \in A,\ b \in B \,\}$.

(vi) **Ensemble des parties :** $\mathcal{P}(E) := \{\, A \mid A \subseteq E \,\}$.

Exemple 3 (Parties de $\{1,2,3\}$)

Soit $E = \{1,2,3\}$. Alors

$$\mathcal{P}(E) = \{\emptyset,\ \{1\},\ \{2\},\ \{3\},\ \{1,2\},\ \{1,3\},\ \{2,3\},\ \{1,2,3\}\}.$$

On a $|\mathcal{P}(E)| = 2^3 = 8$. Plus généralement, si E est fini de cardinal n, alors $|\mathcal{P}(E)| = 2^n$.

Proposition 4 (Lois de De Morgan)

Soient $A, B \subseteq E$. Alors :

$$(A \cup B)^c = A^c \cap B^c \qquad \text{et} \qquad (A \cap B)^c = A^c \cup B^c.$$

Démonstration. Montrons la première égalité. Soit $x \in E$.

$$\begin{aligned} x \in (A \cup B)^c &\iff x \notin A \cup B \\ &\iff x \notin A \text{ et } x \notin B \\ &\iff x \in A^c \text{ et } x \in B^c \\ &\iff x \in A^c \cap B^c. \end{aligned}$$

La seconde se démontre de manière analogue (ou par application de la première au complémentaire). □

1.2 Applications

Définition 5 (Application)

Soient E et F deux ensembles. Une *application* (ou *fonction*) de E dans F est une règle f qui, à chaque élément $x \in E$, associe un unique élément $f(x) \in F$. On note

$$f : E \to F, \qquad x \mapsto f(x).$$

L'ensemble E s'appelle l'*ensemble de départ* (ou *domaine*), et F l'*ensemble d'arrivée* (ou *codomaine*).

Définition 6 (Image et image réciproque)

Soit $f : E \to F$ une application.

(i) L'*image directe* d'une partie $A \subseteq E$ est $f(A) := \{\, f(x) \mid x \in A \,\}$.

(ii) L'*image réciproque* d'une partie $B \subseteq F$ est $f^{-1}(B) := \{\, x \in E \mid f(x) \in B \,\}$.

(iii) L'*image* de f est $\operatorname{Im} f := f(E)$.

Définition 7 (Injectivité, surjectivité, bijectivité)

Soit $f : E \to F$ une application.

(i) f est *injective* si $\forall\, x_1, x_2 \in E,\ f(x_1) = f(x_2) \implies x_1 = x_2$.

(ii) f est *surjective* si $\forall\, y \in F,\ \exists\, x \in E,\ f(x) = y$ (autrement dit, $\operatorname{Im} f = F$).

(iii) f est *bijective* si elle est à la fois injective et surjective.

Exemple 8 (Exemples fondamentaux)

(a) L'application $f\colon \mathbb{R} \to \mathbb{R}$, $x \mapsto x^2$ n'est ni injective (car $f(-1) = f(1) = 1$) ni surjective (car $-1 \notin \operatorname{Im} f$).

(b) L'application $g\colon \mathbb{R} \to \mathbb{R}_+$, $x \mapsto x^2$ est surjective mais pas injective.

(c) L'application $h\colon \mathbb{R}_+ \to \mathbb{R}_+$, $x \mapsto x^2$ est bijective, de réciproque $h^{-1}\colon y \mapsto \sqrt{y}$.

Définition 9 (Composition)

Soient $f\colon E \to F$ et $g\colon F \to G$ deux applications. La *composée* de f et g est l'application

$$g \circ f\colon E \to G, \qquad x \mapsto g(f(x)).$$

Proposition 10 (Composition et injectivité/surjectivité)

Soient $f\colon E \to F$ et $g\colon F \to G$.

(i) Si $g \circ f$ est injective, alors f est injective.

(ii) Si $g \circ f$ est surjective, alors g est surjective.

(iii) $g \circ f$ est bijective si et seulement si f est injective et g est surjective *et* $\operatorname{Im} f = F$ (c'est-à-dire f bijective et g bijective).

Démonstration. *(i)* Supposons $g \circ f$ injective. Soient $x_1, x_2 \in E$ tels que $f(x_1) = f(x_2)$. Alors $g(f(x_1)) = g(f(x_2))$, donc $(g \circ f)(x_1) = (g \circ f)(x_2)$, d'où $x_1 = x_2$ par injectivité de $g \circ f$.

(ii) Supposons $g \circ f$ surjective. Soit $z \in G$. Il existe $x \in E$ tel que $g(f(x)) = z$. En posant $y = f(x) \in F$, on obtient $g(y) = z$, donc g est surjective.

Le point *(iii)* se déduit des deux premiers (exercice). □

1.3 Relations

Définition 11 (Relation binaire)

Une *relation binaire* sur un ensemble E est un sous-ensemble $\mathcal{R} \subseteq E \times E$. On note $x \,\mathcal{R}\, y$ pour $(x, y) \in \mathcal{R}$.

1.3.1 Relations d'équivalence

Définition 12 (Relation d'équivalence)

Une relation $\sim$ sur E est une *relation d'équivalence* si elle vérifie :

(i) **Réflexivité :** $\forall\, x \in E,\ x \sim x$.

(ii) **Symétrie :** $\forall\, x, y \in E,\ x \sim y \implies y \sim x$.

(iii) **Transitivité :** $\forall\, x, y, z \in E,\ (x \sim y \text{ et } y \sim z) \implies x \sim z$.

La *classe d'équivalence* de x est $\bar{x} := \{\, y \in E \mid y \sim x \,\}$. L'*ensemble quotient* est $E/\sim \; := \{\, \bar{x} \mid x \in E \,\}$.

Exemple 13 (Congruence modulo n)

Sur $\mathbb{Z}$, la relation définie par

$$a \equiv b \pmod{n} \iff n \mid (a-b)$$

est une relation d'équivalence. L'ensemble quotient est $\mathbb{Z}/n\mathbb{Z} = \{\bar{0}, \bar{1}, \dots, \overline{n-1}\}$, l'ensemble des classes de congruence modulo n. Par exemple, pour $n = 3$:

$$\bar{0} = \{\dots, -6, -3, 0, 3, 6, \dots\}, \quad \bar{1} = \{\dots, -5, -2, 1, 4, 7, \dots\}, \quad \bar{2} = \{\dots, -4, -1, 2, 5, 8, \dots\}.$$

1.3.2 Relations d'ordre

Définition 14 (Relation d'ordre)

Une relation $\leq$ sur E est une *relation d'ordre* si elle vérifie :

(i) **Réflexivité :** $\forall\, x \in E,\ x \leq x$.

(ii) **Antisymétrie :** $\forall\, x, y \in E,\ (x \leq y \text{ et } y \leq x) \implies x = y$.

(iii) **Transitivité :** $\forall\, x, y, z \in E,\ (x \leq y \text{ et } y \leq z) \implies x \leq z$.

L'ordre est *total* si $\forall\, x, y \in E$, on a $x \leq y$ ou $y \leq x$; sinon il est *partiel*.

Exemple 15 (Ordres classiques)

(a) L'ordre usuel $\leq$ sur $\mathbb{R}$ est un ordre total.

(b) L'inclusion $\subseteq$ sur $\mathcal{P}(E)$ est un ordre partiel (dès que $|E| \geq 2$, il existe des parties incomparables).

(c) La divisibilité $a \mid b$ sur $\mathbb{N}^*$ est un ordre partiel.

1.4 Structures algébriques

Cette section introduit les trois structures fondamentales — groupe, anneau, corps — en insistant sur les axiomes et les exemples concrets. En algèbre linéaire, le corps de base (souvent $\mathbb{R}$ ou $\mathbb{C}$) joue un rôle central ; il est donc essentiel d'en comprendre les propriétés.

1.4.1 Groupes

Définition 16 (Loi de composition interne)

Une *loi de composition interne* sur un ensemble E est une application $\star : E \times E \to E$. On note $x \star y$ l'image du couple (x, y).

Définition 17 (Groupe)

Un *groupe* est un couple $(G, \star)$ où G est un ensemble non vide et $\star$ est une loi de composition interne sur G vérifiant :

(G1) Associativité : $\forall\, a, b, c \in G,\ (a \star b) \star c = a \star (b \star c)$.

(G2) Élément neutre : $\exists\, e \in G,\ \forall\, a \in G,\ a \star e = e \star a = a$.

(G3) Symétrique : $\forall\, a \in G,\ \exists\, a' \in G,\ a \star a' = a' \star a = e$.

Le groupe est dit *abélien* (ou *commutatif*) si de plus :

(G4) Commutativité : $\forall\, a, b \in G,\ a \star b = b \star a$.

Exemple 18 (Groupes classiques)

(a) $(\mathbb{Z}, +)$ est un groupe abélien, d'élément neutre 0. Le symétrique de a est $-a$.

(b) $(\mathbb{Q}^*, \times)$ est un groupe abélien, d'élément neutre 1. Le symétrique de $a \neq 0$ est $1/a$.

(c) $(\mathfrak{S}_n, \circ)$, le groupe des permutations de $\{1, \dots, n\}$, est un groupe non abélien dès que $n \geq 3$, d'ordre $n!$.

(d) $(\mathrm{GL}_n(\mathbb{R}), \times)$, le groupe des matrices inversibles $n \times n$ à coefficients réels, est un groupe non abélien dès que $n \geq 2$.

Proposition 19 (Unicité du neutre et des symétriques)

Dans un groupe $(G, \star)$:

(i) l'élément neutre est unique ;

(ii) pour tout $a \in G$, le symétrique de a est unique.

Démonstration. *(i)* Soient e et e' deux éléments neutres. Alors $e = e \star e' = e'$, puisque e' est neutre à droite et e est neutre à gauche.

(ii) Soient a' et a'' deux symétriques de a. Alors $a' = a' \star e = a' \star (a \star a'') = (a' \star a) \star a'' = e \star a'' = a''$. □

1.4.2 Anneaux

Définition 20 (Anneau)

Un *anneau* est un triplet $(A, +, \times)$ où A est un ensemble muni de deux lois de composition internes vérifiant :

(A1) $(A, +)$ est un groupe abélien (de neutre 0_A).

(A2) Associativité de $\times$: $\forall\, a, b, c \in A,\ (a \times b) \times c = a \times (b \times c)$.

(A3) Élément unité : $\exists\, 1_A \in A,\ \forall\, a \in A,\ 1_A \times a = a \times 1_A = a$.

(A4) Distributivité : $\forall\, a, b, c \in A,$

$$a \times (b + c) = a \times b + a \times c \qquad \text{et} \qquad (a + b) \times c = a \times c + b \times c.$$

L'anneau est *commutatif* si de plus $a \times b = b \times a$ pour tous $a, b \in A$.

Remarque 21 (Conventions)

Certains auteurs n'exigent pas l'existence d'un élément unité ; on parle alors d'*anneau unitaire* lorsque 1_A existe. Dans ce cours, tous nos anneaux seront unitaires.

Exemple 22 (Anneaux classiques)

(a) $(\mathbb{Z}, +, \times)$ est un anneau commutatif.

(b) $(\mathcal{M}_n(\mathbb{R}), +, \times)$ est un anneau non commutatif pour $n \geq 2$.

(c) $(\mathbb{Z}/n\mathbb{Z}, +, \times)$ est un anneau commutatif.

(d) L'anneau des polynômes $\mathbb{R}[X]$ est un anneau commutatif.

1.4.3 Corps

Définition 23 (Corps)

Un *corps* est un triplet $(\mathbb{K}, +, \times)$ tel que :

(C1) $(\mathbb{K}, +, \times)$ est un anneau commutatif.

(C2) $\mathbb{K}$ possède au moins deux éléments : $0_{\mathbb{K}} \neq 1_{\mathbb{K}}$.

(C3) Inverse multiplicatif : $\forall\, a \in \mathbb{K} \setminus \{0_{\mathbb{K}}\},\ \exists\, a^{-1} \in \mathbb{K},\ a \times a^{-1} = 1_{\mathbb{K}}$.

Autrement dit, $(\mathbb{K} \setminus \{0_{\mathbb{K}}\}, \times)$ est un groupe abélien.

Remarque 24 (Corps et algèbre linéaire)

Dans ce cours, la lettre $\mathbb{K}$ désignera toujours un corps, le plus souvent $\mathbb{R}$ ou $\mathbb{C}$. Tous les espaces vectoriels, toutes les matrices, toutes les applications linéaires seront définis *sur un corps* $\mathbb{K}$. La compréhension des axiomes de corps est donc indispensable.

1.5 Les axiomes de corps en détail

Récapitulons de manière systématique. Un corps $(\mathbb{K}, +, \times)$ satisfait les axiomes suivants, pour tous $a, b, c \in \mathbb{K}$:

Axiome	Addition	Multiplication
Associativité	$(a+b)+c = a+(b+c)$	$(a \times b) \times c = a \times (b \times c)$
Commutativité	$a+b = b+a$	$a \times b = b \times a$
Élément neutre	$a+0 = a$	$a \times 1 = a$
Inverse	$a+(-a) = 0$	$a \times a^{-1} = 1 \quad (a \neq 0)$
Distributivité	$a \times (b+c) = a \times b + a \times c$	
Non-trivialité	$0 \neq 1$	

Proposition 25 (Propriétés élémentaires d'un corps)

Dans un corps $(\mathbb{K}, +, \times)$:

(i) $\forall\, a \in \mathbb{K},\ 0 \times a = 0$.

(ii) $\forall\, a, b \in \mathbb{K},\ (-a) \times b = -(a \times b) = a \times (-b)$.

(iii) $\forall\, a, b \in \mathbb{K},\ a \times b = 0 \implies a = 0$ ou $b = 0$ *(intégrité)*.

Démonstration. *(i)* On a $0 \times a = (0+0) \times a = 0 \times a + 0 \times a$. En ajoutant $-(0 \times a)$ des deux côtés, on obtient $0 = 0 \times a$.

(ii) On a $a \times b + (-a) \times b = (a + (-a)) \times b = 0 \times b = 0$, donc $(-a) \times b = -(a \times b)$.

(iii) Supposons $a \neq 0$. Alors a^{-1} existe, et $b = 1 \times b = (a^{-1} \times a) \times b = a^{-1} \times (a \times b) = a^{-1} \times 0 = 0$. □

1.5.1 Le corps $\mathbb{R}$ des nombres réels

Exemple 26 (Le corps $\mathbb{R}$)

L'ensemble $\mathbb{R}$ des nombres réels, muni de l'addition et de la multiplication usuelles, forme un corps. C'est le corps le plus familier ; il possède en outre une structure d'ordre compatible : pour tous $a, b, c \in \mathbb{R}$,

$$a \leq b \implies a + c \leq b + c, \qquad (a \leq b \text{ et } c \geq 0) \implies ac \leq bc.$$

On dit que $\mathbb{R}$ est un *corps ordonné*. De plus, $\mathbb{R}$ vérifie l'*axiome de la borne supérieure* : toute partie non vide et majorée de $\mathbb{R}$ admet une borne supérieure. Cette propriété le distingue de $\mathbb{Q}$ et sera cruciale en analyse.

1.5.2 Le corps $\mathbb{C}$ des nombres complexes

Définition 27 (Nombres complexes)

L'ensemble des *nombres complexes* est

$$\mathbb{C} := \{\, a + bi \mid a, b \in \mathbb{R} \,\},$$

où i est un symbole vérifiant $i^2 = -1$. Les opérations sont :

$$(a+bi) + (c+di) = (a+c) + (b+d)i,$$
$$(a+bi) \times (c+di) = (ac - bd) + (ad + bc)i.$$

Exemple 28 (Inverse dans $\mathbb{C}$)

Soit $z = 2 + 3i \in \mathbb{C}$, $z \neq 0$. Son inverse est

$$z^{-1} = \frac{1}{2+3i} = \frac{2-3i}{(2+3i)(2-3i)} = \frac{2-3i}{4+9} = \frac{2}{13} - \frac{3}{13}i.$$

On utilise la formule générale : si $z = a + bi \neq 0$, alors $z^{-1} = \frac{a-bi}{a^2+b^2}$.

Remarque 29 ($\mathbb{C}$ n'est pas ordonné)

Contrairement à $\mathbb{R}$, le corps $\mathbb{C}$ ne peut pas être muni d'un ordre total compatible avec les opérations. En revanche, $\mathbb{C}$ est *algébriquement clos* : tout polynôme non constant à coefficients complexes admet au moins une racine dans $\mathbb{C}$. Ce résultat fondamental, le *théorème de d'Alembert–Gauss*, sera au cœur de la théorie spectrale (chapitre 1 et chapitres ultérieurs).

1.5.3 Les corps finis $\mathbb{F}_p$

Définition 30 (Corps fini $\mathbb{F}_p$)

Soit p un nombre premier. Le *corps fini* $\mathbb{F}_p$ est l'ensemble $\mathbb{Z}/p\mathbb{Z} = \{\bar{0}, \bar{1}, \ldots, \overline{p-1}\}$ muni de l'addition et de la multiplication modulo p.

Proposition 31 ($\mathbb{Z}/p\mathbb{Z}$ est un corps si et seulement si p est premier)

L'anneau $(\mathbb{Z}/n\mathbb{Z}, +, \times)$ est un corps si et seulement si n est un nombre premier.

Démonstration. ($\Rightarrow$) Si n n'est pas premier, écrivons $n = ab$ avec $1 < a, b < n$. Alors $\bar{a} \neq \bar{0}$ et $\bar{b} \neq \bar{0}$, mais $\bar{a} \times \bar{b} = \overline{ab} = \bar{0}$. L'anneau possède des diviseurs de zéro, donc n'est pas un corps.

($\Leftarrow$) Si p est premier, soit $\bar{a} \in \mathbb{Z}/p\mathbb{Z}$ avec $\bar{a} \neq \bar{0}$, c'est-à-dire $p \nmid a$. Par le théorème de Bézout, il existe $u, v \in \mathbb{Z}$ tels que $au + pv = 1$, d'où $\bar{a} \times \bar{u} = \bar{1}$. Ainsi $\bar{a}$ est inversible. □

Exemple 32 (Tables d'opérations de $\mathbb{F}_3$)

Dans $\mathbb{F}_3 = \{0, 1, 2\}$:

$+$	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

$\times$	0	1	2
0	0	0	0
1	0	1	2
2	0	2	1

On vérifie que $1^{-1} = 1$ et $2^{-1} = 2$ (car $2 \times 2 = 4 \equiv 1 \pmod 3$).

Exemple 33 (Le corps $\mathbb{F}_5$)

Dans $\mathbb{F}_5 = \{0, 1, 2, 3, 4\}$, calculons les inverses multiplicatifs :

$$1^{-1} = 1, \quad 2^{-1} = 3 \ (2 \times 3 = 6 \equiv 1), \quad 3^{-1} = 2, \quad 4^{-1} = 4 \ (4 \times 4 = 16 \equiv 1).$$

On peut faire de l'algèbre linéaire sur $\mathbb{F}_5$: par exemple, résoudre un système d'équations linéaires à coefficients dans $\mathbb{F}_5$.

Remarque 34 (Caractéristique d'un corps)

La *caractéristique* d'un corps $\mathbb{K}$, notée $\mathrm{car}(\mathbb{K})$, est le plus petit entier $p > 0$ tel que $\underbrace{1_{\mathbb{K}} + \cdots + 1_{\mathbb{K}}}_{p} = 0_{\mathbb{K}}$, ou 0 si un tel entier n'existe pas. On a $\mathrm{car}(\mathbb{R}) = \mathrm{car}(\mathbb{C}) = \mathrm{car}(\mathbb{Q}) = 0$ et $\mathrm{car}(\mathbb{F}_p) = p$. Lorsque $\mathrm{car}(\mathbb{K}) > 0$, elle est nécessairement un nombre premier.

1.6 Techniques de démonstration

Les démonstrations sont le cœur des mathématiques. Dans ce cours, nous rédigerons des preuves complètes et nous attendons du lecteur qu'il en fasse autant dans les exercices. Rappelons les techniques les plus courantes.

1.6.1 Preuve directe

Remarque 35 (Méthode)

Pour montrer $P \implies Q$, on suppose P vraie et on en déduit Q par une chaîne d'implications logiques.

Exemple 36 (Preuve directe)

Énoncé. Si n est pair, alors n^2 est pair.
Preuve. Supposons n pair : il existe $k \in \mathbb{Z}$ tel que $n = 2k$. Alors $n^2 = (2k)^2 = 4k^2 = 2(2k^2)$. Comme $2k^2 \in \mathbb{Z}$, on a que n^2 est pair.

1.6.2 Preuve par contraposée

Remarque 37 (Méthode)

L'implication $P \implies Q$ est logiquement équivalente à sa *contraposée* $\neg Q \implies \neg P$. On peut donc montrer $P \implies Q$ en supposant $\neg Q$ et en déduisant $\neg P$.

Exemple 38 (Preuve par contraposée)

Énoncé. Si n^2 est impair, alors n est impair.
Preuve (par contraposée). On montre : si n est pair, alors n^2 est pair. Or c'est exactement le résultat de l'exemple 36. □

1.6.3 Preuve par l'absurde

Remarque 39 (Méthode)

Pour montrer une proposition P, on suppose $\neg P$ et on en déduit une contradiction. On conclut alors que P est vraie.

Exemple 40 ($\sqrt{2}$ est irrationnel)

Énoncé. $\sqrt{2} \notin \mathbb{Q}$.
Preuve. Supposons par l'absurde que $\sqrt{2} \in \mathbb{Q}$. Il existe alors $p, q \in \mathbb{Z}$, $q \neq 0$, avec $\text{pgcd}(p,q) = 1$, tels que $\sqrt{2} = p/q$. En élevant au carré : $2q^2 = p^2$. Donc p^2 est pair, ce qui implique que p est pair (par contraposée du fait que le carré d'un impair est impair). Écrivons $p = 2k$. Alors $2q^2 = 4k^2$, soit $q^2 = 2k^2$. Ainsi q est pair. Mais alors p et q sont tous deux pairs, ce qui contredit $\text{pgcd}(p,q) = 1$.

1.6.4 Raisonnement par récurrence

Remarque 41 (Méthode)

Pour montrer qu'une propriété $P(n)$ est vraie pour tout $n \geq n_0$:

(1) Initialisation : Vérifier que $P(n_0)$ est vraie.

(2) Hérédité : Montrer que pour tout $n \geq n_0$, $P(n) \implies P(n+1)$.

On conclut par le *principe de récurrence* que $P(n)$ est vraie pour tout $n \geq n_0$.

Exemple 42 (Somme des entiers)

Énoncé. Pour tout $n \geq 1$, $\displaystyle\sum_{k=1}^{n} k = \frac{n(n+1)}{2}$.
Preuve par récurrence.
Initialisation ($n = 1$) : $\sum_{k=1}^{1} k = 1 = \frac{1 \cdot 2}{2}$.
Hérédité. Supposons la formule vraie au rang $n \geq 1$. Alors

$$\sum_{k=1}^{n+1} k = \underbrace{\sum_{k=1}^{n} k}_{= n(n+1)/2} + (n+1) = \frac{n(n+1)}{2} + \frac{2(n+1)}{2} = \frac{(n+1)(n+2)}{2},$$

ce qui est la formule au rang $n+1$.

Remarque 43 (Récurrence forte)

Dans la *récurrence forte* (ou *complète*), l'hypothèse de récurrence porte sur *tous* les rangs précédents : on suppose $P(n_0), P(n_0+1), \ldots, P(n)$ pour en déduire $P(n+1)$. Cette variante est parfois indispensable (par exemple pour la factorisation en nombres premiers).

1.7 Exercices

Exercice 44 (Opérations ensemblistes)

Soient $A = \{1, 2, 3, 4, 5\}$, $B = \{3, 4, 5, 6, 7\}$ et $E = \{1, 2, \dots, 10\}$. Déterminer :

(a) $A \cup B$, $A \cap B$, $A \setminus B$, $B \setminus A$;

(b) A^c, B^c, $(A \cup B)^c$, $(A \cap B)^c$;

(c) Vérifier les lois de De Morgan (proposition 4) sur cet exemple.

Exercice 45 (Injectivité et surjectivité)

Pour chacune des applications suivantes, dire si elle est injective, surjective, bijective. Justifier.

(a) $f : \mathbb{Z} \to \mathbb{Z}$, $n \mapsto 2n$.

(b) $g : \mathbb{Z} \to \mathbb{Z}$, $n \mapsto n + 3$.

(c) $h : \mathbb{R} \to \mathbb{R}$, $x \mapsto x^3 - x$.

(d) $\varphi : \mathbb{R} \to \mathbb{R}_+$, $x \mapsto e^x$.

Exercice 46 (Tables de $\mathbb{F}_7$)

(a) Construire la table de multiplication de $\mathbb{F}_7^* = \{1, 2, 3, 4, 5, 6\}$.

(b) En déduire l'inverse de chaque élément non nul.

(c) Résoudre l'équation $3x + 5 = 2$ dans $\mathbb{F}_7$.

Exercice 47 (Composition de bijections)

Soient $f : E \to F$ et $g : F \to G$ deux bijections. Montrer que $g \circ f$ est une bijection et que $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.

Exercice 48 (Sous-groupes de $(\mathbb{Z}, +)$)

Montrer que les sous-groupes de $(\mathbb{Z}, +)$ sont exactement les ensembles de la forme $n\mathbb{Z} := \{\, nk \mid k \in \mathbb{Z} \,\}$, pour $n \in \mathbb{N}$.
Indication : utiliser la division euclidienne.

Exercice 49 (Récurrence — inégalité de Bernoulli)

Montrer par récurrence que pour tout $x \geq -1$ et tout $n \in \mathbb{N}$:

$$(1 + x)^n \geq 1 + nx.$$

Exercice 50 (Corps fini $\mathbb{F}_2$ et parité)

Dans le corps $\mathbb{F}_2 = \{0, 1\}$, on a $1 + 1 = 0$.

(a) Écrire les tables d'addition et de multiplication de $\mathbb{F}_2$.

(b) Résoudre le système linéaire suivant sur $\mathbb{F}_2$:

$$\begin{cases} x + y + z = 1, \\ x + z = 0, \\ y + z = 1. \end{cases}$$

(c) Interpréter la résolution en termes de parité.

Exercice 51 ($\mathbb{Q}(\sqrt{2})$ est un corps)

Soit $\mathbb{Q}(\sqrt{2}) := \left\{ a + b\sqrt{2} \mid a, b \in \mathbb{Q} \right\}$.

(a) Montrer que $\mathbb{Q}(\sqrt{2})$ est stable par addition et multiplication.

(b) Montrer que tout élément non nul de $\mathbb{Q}(\sqrt{2})$ admet un inverse dans $\mathbb{Q}(\sqrt{2})$.

Indication : montrer que si $a + b\sqrt{2} \neq 0$ avec $a, b \in \mathbb{Q}$, alors $a^2 - 2b^2 \neq 0$, puis conjuguer.

(c) En déduire que $(\mathbb{Q}(\sqrt{2}), +, \times)$ est un corps. Quelle est sa caractéristique ?

Exercice 52 (Automorphismes de groupes)

Soit $(G, \star)$ un groupe. Un *automorphisme* de G est une bijection $\varphi \colon G \to G$ telle que $\varphi(a \star b) = \varphi(a) \star \varphi(b)$ pour tous $a, b \in G$.

(a) Montrer que l'ensemble $\mathrm{Aut}(G)$ des automorphismes de G, muni de la composition, est un groupe.

(b) Déterminer $\mathrm{Aut}(\mathbb{Z}/n\mathbb{Z}, +)$ pour $n = 2, 3, 4$.

Résumé du chapitre

Ce qu'il faut retenir.

1. **Ensembles.** Maîtriser les opérations $\cup$, $\cap$, $\setminus$, $\times$, $\mathcal{P}(E)$, ainsi que les lois de De Morgan.
2. **Applications.** Savoir caractériser injectivité, surjectivité et bijectivité. Savoir composer des applications et utiliser les propriétés de la composition.
3. **Relations.** Distinguer relation d'équivalence (classes, quotient) et relation d'ordre (total ou partiel).
4. **Structures algébriques.**
   - *Groupe* : un ensemble avec une loi associative, un neutre et des symétriques.

- *Anneau* : deux lois (addition formant un groupe abélien, multiplication associative avec unité, distributivité).
- *Corps* : un anneau commutatif où tout élément non nul est inversible.

5. Exemples de corps. $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ (caractéristique 0), $\mathbb{F}_p$ pour p premier (caractéristique p). En algèbre linéaire, $\mathbb{K}$ désigne le corps de base.

6. Techniques de démonstration.

- *Preuve directe* : supposer P, déduire Q.
- *Contraposée* : montrer $\neg Q \implies \neg P$.
- *Absurde* : supposer $\neg P$, obtenir une contradiction.
- *Récurrence* : initialisation + hérédité.

Chapitre 2

Espaces vectoriels et sous-espaces

Introduction

L'algèbre linéaire est née de l'étude des vecteurs du plan et de l'espace. Dans $\mathbb{R}^2$, un vecteur $\mathbf{v} = (x, y)$ peut être *additionné* à un autre vecteur et *multiplié* par un scalaire réel, et ces deux opérations satisfont des règles naturelles : l'addition est commutative et associative, la multiplication par un scalaire se distribue sur les sommes, etc. Les mêmes règles s'appliquent dans $\mathbb{R}^3$.

Mais ces opérations ne sont pas l'apanage de la géométrie. Les matrices, les polynômes, les fonctions continues, les suites réelles — tous ces objets peuvent être additionnés entre eux et multipliés par un scalaire, et les mêmes règles algébriques sont satisfaites. L'idée fondamentale de l'algèbre linéaire est de *dégager un cadre abstrait* — celui d'*espace vectoriel* — qui unifie tous ces exemples et permet de les étudier avec les mêmes outils.

Ce chapitre constitue le fondement sur lequel repose l'ensemble du cours. Nous y définissons les espaces vectoriels et leurs sous-espaces, puis nous développons les notions de famille libre, famille génératrice, base et dimension. Ces concepts sont indispensables pour tout ce qui suit : applications linéaires, matrices, déterminants, valeurs propres.

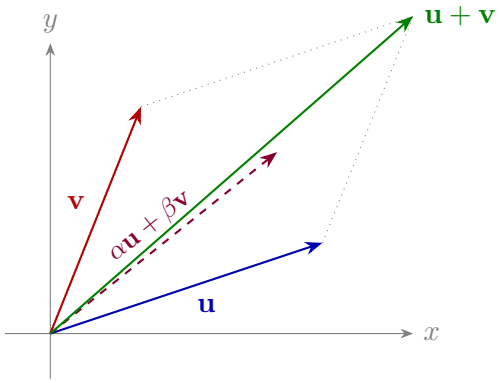


Fig. 2.1 : Combinaison linéaire de deux vecteurs dans $\mathbb{R}^2$.

2.1 Espace vectoriel sur un corps

2.1.1 Définition axiomatique

Dans tout ce chapitre, $\mathbb{K}$ désigne un corps (voir section 1.4.3). Les éléments de $\mathbb{K}$ sont appelés *scalaires*. En pratique, le lecteur pourra penser à $\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$.

Définition 1 (Espace vectoriel)

Un *espace vectoriel* sur $\mathbb{K}$ (ou $\mathbb{K}$-espace vectoriel) est un triplet $(E, +, \cdot)$ où E est un ensemble non vide, $+ \colon E \times E \to E$ est une loi de composition interne (l'*addition vectorielle*) et $\cdot \colon \mathbb{K} \times E \to E$ est une loi de composition externe (la *multiplication par un scalaire*), satisfaisant les axiomes suivants :

Axiomes d'addition.

(EV1) Associativité : $\forall\, u, v, w \in E,\ (u + v) + w = u + (v + w)$.

(EV2) Commutativité : $\forall\, u, v \in E,\ u + v = v + u$.

(EV3) Élément neutre : $\exists\, \mathbf{0} \in E,\ \forall\, u \in E,\ u + \mathbf{0} = u$.

(EV4) Opposé : $\forall\, u \in E,\ \exists\, (-u) \in E,\ u + (-u) = \mathbf{0}$.

Autrement dit, $(E, +)$ est un groupe abélien.

Axiomes de multiplication scalaire.

(EV5) Compatibilité : $\forall\, \alpha, \beta \in \mathbb{K},\ \forall\, u \in E,\ \alpha \cdot (\beta \cdot u) = (\alpha\beta) \cdot u$.

(EV6) Élément unité : $\forall\, u \in E,\ 1_{\mathbb{K}} \cdot u = u$.

(EV7) Distributivité (scalaire) : $\forall\, \alpha \in \mathbb{K},\ \forall\, u, v \in E,\ \alpha \cdot (u + v) = \alpha \cdot u + \alpha \cdot v$.

(EV8) Distributivité (vecteur) : $\forall\, \alpha, \beta \in \mathbb{K},\ \forall\, u \in E,\ (\alpha + \beta) \cdot u = \alpha \cdot u + \beta \cdot u$.

Les éléments de E sont appelés *vecteurs*.

Remarque 2 (Notation)

On omet généralement le point $\cdot$ et on écrit αu au lieu de $\alpha \cdot u$. La multiplication scalaire n'est *pas* une loi de composition interne : elle prend un scalaire dans $\mathbb{K}$ et un vecteur dans E, et renvoie un vecteur dans E.

Proposition 3 (Propriétés immédiates)

Soit E un $\mathbb{K}$-espace vectoriel. Pour tous $\alpha \in \mathbb{K}$ et $u \in E$:

(i) $0_{\mathbb{K}} \cdot u = \mathbf{0}$.

(ii) $\alpha \cdot \mathbf{0} = \mathbf{0}$.

(iii) $(-1_{\mathbb{K}}) \cdot u = -u$.

(iv) $\alpha \cdot u = \mathbf{0} \implies \alpha = 0_{\mathbb{K}}$ ou $u = \mathbf{0}$.

Démonstration. *(i)* On a $0_{\mathbb{K}} \cdot u = (0_{\mathbb{K}} + 0_{\mathbb{K}}) \cdot u = 0_{\mathbb{K}} \cdot u + 0_{\mathbb{K}} \cdot u$ par **(EV8)**. En ajoutant l'opposé de $0_{\mathbb{K}} \cdot u$ des deux côtés, on obtient $\mathbf{0} = 0_{\mathbb{K}} \cdot u$.

(ii) On a $\alpha \cdot \mathbf{0} = \alpha \cdot (\mathbf{0} + \mathbf{0}) = \alpha \cdot \mathbf{0} + \alpha \cdot \mathbf{0}$ par **(EV7)**, d'où $\mathbf{0} = \alpha \cdot \mathbf{0}$ en simplifiant.

(iii) On calcule $u + (-1_{\mathbb{K}}) \cdot u = 1_{\mathbb{K}} \cdot u + (-1_{\mathbb{K}}) \cdot u = (1_{\mathbb{K}} + (-1_{\mathbb{K}})) \cdot u = 0_{\mathbb{K}} \cdot u = \mathbf{0}$ par (i). Donc $(-1_{\mathbb{K}}) \cdot u$ est l'opposé de u.

(iv) Supposons $\alpha \neq 0_{\mathbb{K}}$. Alors α^{-1} existe dans $\mathbb{K}$, et $u = 1_{\mathbb{K}} \cdot u = (\alpha^{-1}\alpha) \cdot u = \alpha^{-1} \cdot (\alpha \cdot u) = \alpha^{-1} \cdot \mathbf{0} = \mathbf{0}$ par (ii). □

2.1.2 Exemples fondamentaux

Exemple 4 (L'espace $\mathbb{K}^n$)

L'ensemble des n-uplets

$$\mathbb{K}^n = \{\, (x_1, \dots, x_n) \mid x_i \in \mathbb{K} \,\}$$

muni de l'addition composante par composante

$$(x_1, \dots, x_n) + (y_1, \dots, y_n) = (x_1 + y_1, \dots, x_n + y_n)$$

et de la multiplication scalaire

$$\alpha(x_1, \dots, x_n) = (\alpha x_1, \dots, \alpha x_n)$$

est un $\mathbb{K}$-espace vectoriel. Le vecteur nul est $\mathbf{0} = (0, \dots, 0)$. C'est l'exemple le plus fondamental ; pour $n = 2$ et $n = 3$, il correspond aux vecteurs du plan et de l'espace.

Exemple 5 (L'espace des matrices $\mathcal{M}_{n,p}(\mathbb{K})$)

L'ensemble $\mathcal{M}_{n,p}(\mathbb{K})$ des matrices à n lignes et p colonnes à coefficients dans $\mathbb{K}$, muni de l'addition matricielle et de la multiplication par un scalaire, est un $\mathbb{K}$-espace vectoriel. Le vecteur nul est la matrice nulle $0_{n\times p}$.

Exemple 6 (L'espace des polynômes)

(a) L'ensemble $\mathbb{K}[X]$ de tous les polynômes à coefficients dans $\mathbb{K}$, muni de l'addition des polynômes et de la multiplication par un scalaire, est un $\mathbb{K}$-espace vectoriel.

(b) Pour tout $n \in \mathbb{N}$, le sous-ensemble

$$\mathbb{K}_n[X] := \{\, P \in \mathbb{K}[X] \mid \deg P \leq n \,\} \cup \{0\}$$

des polynômes de degré au plus n (incluant le polynôme nul) est un $\mathbb{K}$-espace vectoriel.

Exemple 7 (Espaces de fonctions)

Soit X un ensemble non vide. L'ensemble $\mathcal{F}(X, \mathbb{K}) := \{\, f \mid f \colon X \to \mathbb{K} \,\}$ de toutes les fonctions de X dans $\mathbb{K}$, muni des opérations

$$(f + g)(x) = f(x) + g(x), \qquad (\alpha f)(x) = \alpha f(x),$$

est un $\mathbb{K}$-espace vectoriel. De même, si $\mathbb{K} = \mathbb{R}$, les ensembles suivants sont des $\mathbb{R}$-espaces vectoriels :

- $\mathcal{C}^0(\mathbb{R}, \mathbb{R})$, l'espace des fonctions continues ;
- $\mathcal{C}^k(\mathbb{R}, \mathbb{R})$, l'espace des fonctions k fois dérivables à dérivée k-ième continue ;
- $\mathcal{C}^\infty(\mathbb{R}, \mathbb{R})$, l'espace des fonctions indéfiniment dérivables.

Exemple 8 (Espace des solutions d'un système homogène)

Soit $A \in \mathcal{M}_{n,p}(\mathbb{K})$. L'ensemble des solutions du système linéaire homogène $AX = 0$:

$$\mathcal{S} = \{ X \in \mathbb{K}^p \mid AX = 0 \}$$

est un $\mathbb{K}$-espace vectoriel (c'est un sous-espace de $\mathbb{K}^p$, comme nous le verrons dans la section 2.2).

Exemple 9 (L'espace des suites)

L'ensemble $\mathbb{K}^{\mathbb{N}}$ de toutes les suites $(u_n)_{n\geq 0}$ à valeurs dans $\mathbb{K}$, muni de l'addition terme à terme et de la multiplication scalaire, est un $\mathbb{K}$-espace vectoriel.

Remarque 10 (Contre-exemples)

L'ensemble $\mathbb{N}$ muni de l'addition usuelle n'est *pas* un espace vectoriel sur $\mathbb{R}$: d'une part, il n'y a pas d'opposé ($-3 \notin \mathbb{N}$), et d'autre part, la multiplication scalaire par un réel ne reste pas dans $\mathbb{N}$ ($\frac{1}{2} \cdot 3 = \frac{3}{2} \notin \mathbb{N}$).

2.2 Sous-espaces vectoriels

Définition 11 (Sous-espace vectoriel)

Soit E un $\mathbb{K}$-espace vectoriel. Une partie $F \subseteq E$ est un *sous-espace vectoriel* de E (en abrégé : *s.e.v.*) si F, muni des opérations induites par E, est lui-même un $\mathbb{K}$-espace vectoriel.

La vérification directe de tous les axiomes est lourde. Le théorème suivant fournit un critère simple et efficace.

Théorème 12 (Caractérisation des sous-espaces vectoriels)

Soit E un $\mathbb{K}$-espace vectoriel et $F \subseteq E$. Alors F est un sous-espace vectoriel de E si et seulement si :

(i) $F \neq \emptyset$ (ou, de manière équivalente, $\mathbf{0} \in F$) ;

(ii) $\forall\, u, v \in F,\ u + v \in F$ *(stabilité par addition)* ;

(iii) $\forall\, \alpha \in \mathbb{K},\ \forall\, u \in F,\ \alpha u \in F$ *(stabilité par multiplication scalaire).*

De manière équivalente, les conditions (ii) et (iii) peuvent être remplacées par la condition unique :

(ii') $\forall\, \alpha, \beta \in \mathbb{K},\ \forall\, u, v \in F,\ \alpha u + \beta v \in F$ *(stabilité par combinaison linéaire).*

Démonstration. ($\Rightarrow$) Si F est un s.e.v. de E, alors F contient son vecteur nul, qui doit être $\mathbf{0}$ (car $\mathbf{0}_F = 0_{\mathbb{K}} \cdot u = \mathbf{0}_E$ pour tout $u \in F$). Les propriétés (ii) et (iii) découlent du fait que $+$ et $\cdot$ sont des lois sur F.

($\Leftarrow$) Supposons (i), (ii), (iii). Puisque $F \neq \emptyset$, il existe $u \in F$. Par (iii), $0_{\mathbb{K}} \cdot u = \mathbf{0} \in F$. Par (iii), $(-1_{\mathbb{K}}) \cdot u = -u \in F$ pour tout $u \in F$, donc tout vecteur de F admet un opposé dans F.

Par (ii), F est stable par addition. Les axiomes **(EV1)**–**(EV8)** sont satisfaits dans F car ils le sont dans E et $F \subseteq E$. Ainsi $(F, +, \cdot)$ est un $\mathbb{K}$-espace vectoriel.

L'équivalence entre (ii)+(iii) et (ii') est immédiate : (ii') avec $\alpha = \beta = 1$ donne (ii), et (ii') avec $\beta = 0$ donne (iii) ; réciproquement, si (ii) et (iii) sont vérifiées, alors $\alpha u \in F$ et $\beta v \in F$ par (iii), puis $\alpha u + \beta v \in F$ par (ii). □

Exemple 13 (Sous-espaces de $\mathbb{R}^3$)

Les sous-espaces vectoriels de $\mathbb{R}^3$ sont :

- $\{\mathbf{0}\}$, le sous-espace nul (un point) ;
- les droites vectorielles passant par l'origine ;
- les plans vectoriels passant par l'origine ;
- $\mathbb{R}^3$ lui-même.

Par exemple, $D = \{\, t(1,2,3) \mid t \in \mathbb{R} \,\}$ est une droite vectorielle, et $P = \{\, (x,y,z) \in \mathbb{R}^3 \mid x + y + z = 0 \,\}$ est un plan vectoriel.

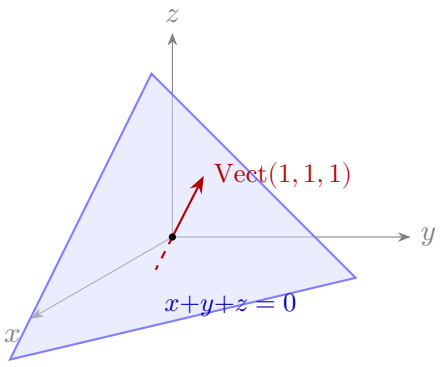


Fig. 2.2 : Un plan vectoriel et une droite vectorielle dans $\mathbb{R}^3$.

Exemple 14 (Sous-espaces de $\mathbb{K}[X]$)

Soit $n \in \mathbb{N}$. L'ensemble $\mathbb{K}_n[X]$ des polynômes de degré $\leq n$ est un sous-espace vectoriel de $\mathbb{K}[X]$:

- $\mathbf{0} = 0 \in \mathbb{K}_n[X]$;
- si $\deg P \leq n$ et $\deg Q \leq n$, alors $\deg(P+Q) \leq n$;
- si $\deg P \leq n$, alors $\deg(\alpha P) \leq n$ pour tout $\alpha \in \mathbb{K}$.

En revanche, $\{\, P \in \mathbb{K}[X] \mid \deg P = n \,\}$ n'est *pas* un sous-espace : il ne contient pas le polynôme nul, et n'est pas stable par addition (la somme de deux polynômes de degré n peut avoir un degré inférieur).

Exemple 15 (Noyau d'une application linéaire)

Nous verrons au chapitre 2 que le noyau et l'image d'une application linéaire sont des sous-espaces vectoriels. Pour l'instant, vérifions-le dans un cas particulier : soit $\varphi \colon \mathbb{R}^3 \to \mathbb{R}$

définie par $\varphi(x, y, z) = 2x - y + 3z$. Alors

$$\operatorname{Ker}\varphi = \{\,(x, y, z) \in \mathbb{R}^3 \mid 2x - y + 3z = 0\,\}$$

est un sous-espace vectoriel de $\mathbb{R}^3$ (c'est le plan d'équation $2x - y + 3z = 0$). En effet, $\mathbf{0} \in \operatorname{Ker}\varphi$, et si $u, v \in \operatorname{Ker}\varphi$ et $\alpha \in \mathbb{R}$, alors $\varphi(\alpha u + v) = \alpha\varphi(u) + \varphi(v) = 0$, donc $\alpha u + v \in \operatorname{Ker}\varphi$.

Proposition 16 (Propriétés élémentaires des sous-espaces)

Soit E un $\mathbb{K}$-espace vectoriel.

(i) $\{\mathbf{0}\}$ et E sont des sous-espaces vectoriels de E (appelés *sous-espaces triviaux*).

(ii) Tout sous-espace vectoriel F de E contient $\mathbf{0}$.

(iii) Si F est un s.e.v. de E et G est un s.e.v. de F, alors G est un s.e.v. de E.

Démonstration. Les points (i) et (ii) sont immédiats d'après le théorème 12. Pour (iii), G est non vide et stable par combinaison linéaire dans F, donc aussi dans E puisque les opérations sont les mêmes. □

2.3 Intersection et somme de sous-espaces

2.3.1 Intersection

Proposition 17 (Intersection de sous-espaces)

Soit E un $\mathbb{K}$-espace vectoriel et $(F_i)_{i\in I}$ une famille de sous-espaces vectoriels de E. Alors l'intersection $\bigcap_{i\in I} F_i$ est un sous-espace vectoriel de E.

Démonstration. Posons $F = \bigcap_{i\in I} F_i$. Puisque $\mathbf{0} \in F_i$ pour tout $i \in I$, on a $\mathbf{0} \in F$, donc $F \neq \emptyset$. Soient $u, v \in F$ et $\alpha, \beta \in \mathbb{K}$. Pour tout $i \in I$, $u, v \in F_i$, donc $\alpha u + \beta v \in F_i$ (car F_i est un s.e.v.). Ainsi $\alpha u + \beta v \in F$. Par le théorème 12, F est un s.e.v. de E. □

Remarque 18 (La réunion n'est pas un sous-espace)

En général, la réunion de deux sous-espaces vectoriels n'est *pas* un sous-espace vectoriel. Par exemple, dans $\mathbb{R}^2$, soient $F = \{\,(t, 0) \mid t \in \mathbb{R}\,\}$ (l'axe des abscisses) et $G = \{\,(0, t) \mid t \in \mathbb{R}\,\}$ (l'axe des ordonnées). Alors $(1, 0) \in F \cup G$ et $(0, 1) \in F \cup G$, mais $(1, 0) + (0, 1) = (1, 1) \notin F \cup G$.

2.3.2 Somme de sous-espaces

Définition 19 (Somme de sous-espaces)

Soient F et G deux sous-espaces vectoriels d'un $\mathbb{K}$-espace vectoriel E. La *somme* de F et G est l'ensemble

$$F + G := \{\,u + v \mid u \in F,\ v \in G\,\}.$$

Plus généralement, si $F_1, \dots, F_r$ sont des sous-espaces de E,

$$F_1 + \cdots + F_r = \{ u_1 + \cdots + u_r \mid u_i \in F_i \}.$$

Proposition 20 (La somme est un sous-espace)

$F + G$ est un sous-espace vectoriel de E. De plus, c'est le plus petit sous-espace de E contenant à la fois F et G.

Démonstration. *$F + G$ est un s.e.v.* Le vecteur $\mathbf{0} = \mathbf{0} + \mathbf{0} \in F + G$, donc $F + G \neq \emptyset$. Soient $w_1 = u_1 + v_1$ et $w_2 = u_2 + v_2$ dans $F + G$ (avec $u_i \in F$, $v_i \in G$) et $\alpha, \beta \in \mathbb{K}$. Alors

$$\alpha w_1 + \beta w_2 = \underbrace{(\alpha u_1 + \beta u_2)}_{\in F} + \underbrace{(\alpha v_1 + \beta v_2)}_{\in G} \in F + G.$$

Minimalité. On a $F \subseteq F + G$ (car tout $u \in F$ s'écrit $u = u + \mathbf{0}$) et $G \subseteq F + G$. Si H est un s.e.v. contenant F et G, alors pour tout $u \in F$ et $v \in G$, on a $u, v \in H$, donc $u + v \in H$, d'où $F + G \subseteq H$. □

2.4 Somme directe et sous-espaces supplémentaires

Définition 21 (Somme directe (interne))

Soient F et G deux sous-espaces vectoriels de E. On dit que la somme $F + G$ est *directe* si tout vecteur $w \in F + G$ s'écrit de manière *unique* sous la forme $w = u + v$ avec $u \in F$ et $v \in G$. On note alors $F \oplus G$.

Théorème 22 (Caractérisation de la somme directe)

Soient F et G deux sous-espaces vectoriels de E. Les assertions suivantes sont équivalentes :

(i) La somme $F + G$ est directe.

(ii) $F \cap G = \{\mathbf{0}\}$.

(iii) Pour tous $u \in F$ et $v \in G$, si $u + v = \mathbf{0}$, alors $u = \mathbf{0}$ et $v = \mathbf{0}$.

Démonstration. *(i)⇒(ii).* Soit $w \in F \cap G$. Alors w s'écrit $w = w + \mathbf{0}$ (avec $w \in F$, $\mathbf{0} \in G$) et $w = \mathbf{0} + w$ (avec $\mathbf{0} \in F$, $w \in G$). Par unicité de la décomposition, $w = \mathbf{0}$.

(ii)⇒(iii). Si $u + v = \mathbf{0}$ avec $u \in F$, $v \in G$, alors $u = -v \in G$. Donc $u \in F \cap G = \{\mathbf{0}\}$, d'où $u = \mathbf{0}$ et $v = \mathbf{0}$.

(iii)⇒(i). Supposons $w = u_1 + v_1 = u_2 + v_2$ avec $u_i \in F$, $v_i \in G$. Alors $(u_1 - u_2) + (v_1 - v_2) = \mathbf{0}$ avec $u_1 - u_2 \in F$ et $v_1 - v_2 \in G$. Par hypothèse, $u_1 - u_2 = \mathbf{0}$ et $v_1 - v_2 = \mathbf{0}$, donc $u_1 = u_2$ et $v_1 = v_2$. La décomposition est unique. □

Définition 23 (Sous-espaces supplémentaires)

Deux sous-espaces F et G de E sont dits *supplémentaires* (dans E) si $E = F \oplus G$, c'est-à-dire si :

$$E = F + G \qquad \text{et} \qquad F \cap G = \{\mathbf{0}\}.$$

On dit aussi que G est un *supplémentaire* de F dans E.

Exemple 24 (Supplémentaires dans $\mathbb{R}^3$)

Soient $F = \{(x, y, 0) \mid x, y \in \mathbb{R}\}$ (le plan $z = 0$) et $G = \{(0, 0, z) \mid z \in \mathbb{R}\}$ (l'axe Oz). Alors $\mathbb{R}^3 = F \oplus G$: tout $(x, y, z) \in \mathbb{R}^3$ s'écrit de manière unique comme $(x, y, 0) + (0, 0, z)$, et $F \cap G = \{(0, 0, 0)\}$.

Remarque 25 (Non-unicité du supplémentaire)

Un sous-espace admet en général *plusieurs* supplémentaires. Par exemple, dans $\mathbb{R}^2$, la droite $F = \text{Vect}((1, 0))$ admet pour supplémentaire toute droite $G = \text{Vect}((a, b))$ avec $b \neq 0$.

La notion de somme directe se généralise à un nombre fini quelconque de sous-espaces.

Définition 26 (Somme directe de plusieurs sous-espaces)

Soient $F_1, \dots, F_r$ des sous-espaces de E. On dit que la somme $F_1 + \cdots + F_r$ est *directe*, et on note $F_1 \oplus \cdots \oplus F_r$, si tout vecteur $w \in F_1 + \cdots + F_r$ s'écrit de manière unique $w = u_1 + \cdots + u_r$ avec $u_i \in F_i$.

Proposition 27 (Caractérisation de la somme directe multiple)

La somme $F_1 + \cdots + F_r$ est directe si et seulement si pour tout $j \in \{1, \dots, r\}$:

$$F_j \cap (F_1 + \cdots + F_{j-1} + F_{j+1} + \cdots + F_r) = \{\mathbf{0}\}.$$

De manière équivalente, si $u_1 + \cdots + u_r = \mathbf{0}$ avec $u_i \in F_i$, alors $u_1 = \cdots = u_r = \mathbf{0}$.

Démonstration. La preuve suit le même schéma que celle du théorème 22. Montrons l'équivalence avec la condition sur l'unicité.

($\Rightarrow$) Si $u_1 + \cdots + u_r = \mathbf{0}$ avec $u_i \in F_i$, on a aussi $\mathbf{0} + \cdots + \mathbf{0} = \mathbf{0}$. Par unicité de la décomposition, $u_i = \mathbf{0}$ pour tout i.

($\Leftarrow$) Si $u_1 + \cdots + u_r = u'_1 + \cdots + u'_r$ avec $u_i, u'_i \in F_i$, alors $(u_1 - u'_1) + \cdots + (u_r - u'_r) = \mathbf{0}$ avec $u_i - u'_i \in F_i$. Par hypothèse, $u_i - u'_i = \mathbf{0}$ pour tout i, soit $u_i = u'_i$. □

2.5 Combinaisons linéaires et sous-espace engendré

Définition 28 (Combinaison linéaire)

Soit E un $\mathbb{K}$-espace vectoriel. Soient $v_1, \dots, v_n \in E$ et $\alpha_1, \dots, \alpha_n \in \mathbb{K}$. Le vecteur

$$w = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n = \sum_{i=1}^{n} \alpha_i v_i$$

est appelé *combinaison linéaire* des vecteurs $v_1, \dots, v_n$ à coefficients $\alpha_1, \dots, \alpha_n$.

Définition 29 (Sous-espace engendré)

Soit $S \subseteq E$ une partie (éventuellement infinie). Le *sous-espace engendré par* S, noté

$\mathrm{Vect}(S)$, est le plus petit sous-espace vectoriel de E contenant S. De manière équivalente :

$$\mathrm{Vect}(S) = \{\, \alpha_1 v_1 + \cdots + \alpha_n v_n \mid n \in \mathbb{N},\ v_i \in S,\ \alpha_i \in \mathbb{K} \,\}.$$

Par convention, $\mathrm{Vect}(\emptyset) = \{\mathbf{0}\}$.
Si $S = \{v_1, \dots, v_n\}$ est finie, on note $\mathrm{Vect}(v_1, \dots, v_n)$.

Proposition 30 ($\mathrm{Vect}(S)$ est l'intersection de tous les sous-espaces contenant S)

Pour toute partie $S \subseteq E$:

$$\mathrm{Vect}(S) = \bigcap_{\substack{F \text{ s.e.v. de } E \\ S \subseteq F}} F.$$

Démonstration. Notons V l'intersection du membre droit. Par la proposition 17, V est un s.e.v. de E, et $S \subseteq V$ (car $S \subseteq F$ pour chaque F dans l'intersection). Donc V est un s.e.v. contenant S. Par définition, $\mathrm{Vect}(S) \subseteq V$ (car $\mathrm{Vect}(S)$ est le plus petit tel s.e.v.).

Inversement, $\mathrm{Vect}(S)$ est un s.e.v. contenant S, donc il apparaît parmi les F de l'intersection, d'où $V \subseteq \mathrm{Vect}(S)$. □

Exemple 31 (Sous-espace engendré dans $\mathbb{R}^3$)

Soient $v_1 = (1, 0, 1)$ et $v_2 = (0, 1, -1)$. Alors

$$\mathrm{Vect}(v_1, v_2) = \{\, \alpha(1,0,1) + \beta(0,1,-1) \mid \alpha, \beta \in \mathbb{R} \,\} = \{\, (\alpha, \beta, \alpha - \beta) \mid \alpha, \beta \in \mathbb{R} \,\}.$$

C'est un plan passant par l'origine dans $\mathbb{R}^3$, d'équation $x - y - z = 0$ (on vérifie que $\alpha - \beta - (\alpha - \beta) = 0$).

Définition 32 (Famille génératrice)

Soit E un $\mathbb{K}$-espace vectoriel. Une famille $(v_i)_{i \in I}$ de vecteurs de E est dite *génératrice* de E si $\mathrm{Vect}((v_i)_{i \in I}) = E$, c'est-à-dire si tout vecteur de E est combinaison linéaire de la famille.

Exemple 33 (Famille génératrice de $\mathbb{R}^3$)

La famille $((1,0,0), (0,1,0), (0,0,1))$ est génératrice de $\mathbb{R}^3$: tout $(x, y, z) \in \mathbb{R}^3$ s'écrit $x(1,0,0) + y(0,1,0) + z(0,0,1)$. La famille $((1,1,0), (1,0,1), (0,1,1))$ est également génératrice (le vérifier en exercice).

2.6 Familles libres

Définition 34 (Famille libre (linéairement indépendante))

Une famille $(v_1, \dots, v_n)$ de vecteurs de E est dite *libre* (ou *linéairement indépendante*) si la seule combinaison linéaire nulle est la combinaison triviale :

$$\alpha_1 v_1 + \cdots + \alpha_n v_n = \mathbf{0} \implies \alpha_1 = \cdots = \alpha_n = 0_{\mathbb{K}}.$$

Une famille qui n'est pas libre est dite *liée* (ou *linéairement dépendante*).

Remarque 35 (Famille liée)

La famille $(v_1, \dots, v_n)$ est liée si et seulement s'il existe des scalaires $\alpha_1, \dots, \alpha_n \in \mathbb{K}$, *non tous nuls*, tels que $\alpha_1 v_1 + \cdots + \alpha_n v_n = \mathbf{0}$. De manière équivalente, l'un des vecteurs est combinaison linéaire des autres.

Proposition 36 (Propriétés des familles libres)

(i) La famille vide est libre (par convention).

(ii) Toute sous-famille d'une famille libre est libre.

(iii) Toute famille contenant le vecteur nul $\mathbf{0}$ est liée.

(iv) Toute famille contenant un vecteur en double est liée.

(v) Une famille (v) réduite à un seul vecteur est libre si et seulement si $v \neq \mathbf{0}$.

(vi) Une famille (v_1, v_2) est liée si et seulement si v_1 et v_2 sont colinéaires (l'un est multiple scalaire de l'autre).

Démonstration. *(i)* Par convention, il n'existe pas de combinaison linéaire non triviale de la famille vide.

(ii) Si une sous-famille admet une relation de dépendance linéaire non triviale, on peut la compléter en ajoutant des coefficients 0 pour les vecteurs absents, ce qui donne une relation de dépendance non triviale de la famille entière, contredisant sa liberté.

(iii) Si $v_j = \mathbf{0}$, la relation $0 \cdot v_1 + \cdots + 1 \cdot v_j + \cdots + 0 \cdot v_n = \mathbf{0}$ est non triviale.

(iv) Si $v_i = v_j$ avec $i \neq j$, la relation avec $\alpha_i = 1$, $\alpha_j = -1$ et les autres coefficients nuls est non triviale.

(v) $\alpha v = \mathbf{0}$ implique $\alpha = 0$ ou $v = \mathbf{0}$. Donc (v) est libre si et seulement si $v \neq \mathbf{0}$.

(vi) (v_1, v_2) est liée ssi il existe $(\alpha_1, \alpha_2) \neq (0, 0)$ avec $\alpha_1 v_1 + \alpha_2 v_2 = \mathbf{0}$. Si $\alpha_1 \neq 0$, alors $v_1 = -(\alpha_2/\alpha_1) v_2$. Si $\alpha_2 \neq 0$, alors $v_2 = -(\alpha_1/\alpha_2) v_1$. □

Exemple 37 (Famille libre dans $\mathbb{R}^3$)

Montrons que $\mathcal{F} = (e_1, e_2, e_3)$ avec $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$, $e_3 = (0, 0, 1)$ est libre. Supposons

$$\alpha_1(1, 0, 0) + \alpha_2(0, 1, 0) + \alpha_3(0, 0, 1) = (0, 0, 0).$$

Alors $(\alpha_1, \alpha_2, \alpha_3) = (0, 0, 0)$, d'où la liberté.

Exemple 38 (Famille libre de polynômes)

Dans $\mathbb{R}[X]$, la famille $(1, X, X^2, \dots, X^n)$ est libre : si $\alpha_0 + \alpha_1 X + \cdots + \alpha_n X^n = 0$ (le polynôme nul), alors $\alpha_0 = \alpha_1 = \cdots = \alpha_n = 0$ par identification des coefficients.

Théorème 39 (Caractérisation des familles liées)

La famille $(v_1, \dots, v_n)$ $(n \geq 2)$ est liée si et seulement si l'un des vecteurs est combinaison linéaire des *précédents* : il existe $j \in \{2, \dots, n\}$ tel que $v_j \in \mathrm{Vect}(v_1, \dots, v_{j-1})$.

Démonstration. $(\Leftarrow)$ Immédiat : si $v_j = \sum_{i=1}^{j-1} \alpha_i v_i$, alors $\sum_{i=1}^{j-1} \alpha_i v_i - v_j = \mathbf{0}$ est une relation non triviale.

$(\Rightarrow)$ Soit $\alpha_1 v_1 + \cdots + \alpha_n v_n = \mathbf{0}$ une relation non triviale. Soit j le plus grand indice tel que $\alpha_j \neq 0$. Si $j = 1$, alors $\alpha_1 v_1 = \mathbf{0}$ avec $\alpha_1 \neq 0$, donc $v_1 = \mathbf{0} \in \mathrm{Vect}(\emptyset)$, ce qui n'est pas tout à fait la forme souhaitée. Cependant, si $v_1 = \mathbf{0}$, la famille est liée de manière triviale. Pour $j \geq 2$, on a $\alpha_j v_j = -\sum_{i=1}^{j-1} \alpha_i v_i$, d'où $v_j = -\sum_{i=1}^{j-1} (\alpha_i/\alpha_j) v_i \in \mathrm{Vect}(v_1, \dots, v_{j-1})$. □

2.7 Bases

Définition 40 (Base)

Une famille $(v_1, \dots, v_n)$ de vecteurs d'un $\mathbb{K}$-espace vectoriel E est une *base* de E si elle est à la fois *libre* et *génératrice*.

Théorème 41 (Caractérisation d'une base)

Soit $\mathcal{B} = (v_1, \dots, v_n)$ une famille de vecteurs de E. Les assertions suivantes sont équivalentes :

(i) $\mathcal{B}$ est une base de E.

(ii) Tout vecteur $v \in E$ s'écrit de manière *unique* comme combinaison linéaire de $v_1, \dots, v_n$:

$$v = \alpha_1 v_1 + \cdots + \alpha_n v_n, \qquad \alpha_1, \dots, \alpha_n \in \mathbb{K} \text{ uniques.}$$

Démonstration. *(i)⇒(ii).* La famille est génératrice : tout $v \in E$ s'écrit $v = \sum_i \alpha_i v_i$. Supposons deux décompositions : $\sum_i \alpha_i v_i = \sum_i \beta_i v_i$. Alors $\sum_i (\alpha_i - \beta_i) v_i = \mathbf{0}$, et la liberté de $\mathcal{B}$ impose $\alpha_i - \beta_i = 0$ pour tout i, d'où l'unicité.

(ii)⇒(i). L'existence de la décomposition implique que $\mathcal{B}$ est génératrice. L'unicité implique la liberté : si $\sum_i \alpha_i v_i = \mathbf{0}$, c'est la même décomposition que $\sum_i 0 \cdot v_i = \mathbf{0}$, donc $\alpha_i = 0$ pour tout i. □

Définition 42 (Coordonnées)

Si $\mathcal{B} = (v_1, \dots, v_n)$ est une base de E et si $v = \alpha_1 v_1 + \cdots + \alpha_n v_n$, les scalaires $(\alpha_1, \dots, \alpha_n) \in \mathbb{K}^n$ sont les *coordonnées* (ou *composantes*) de v dans la base $\mathcal{B}$.

Exemple 43 (Base canonique de $\mathbb{K}^n$)

La *base canonique* de $\mathbb{K}^n$ est $\mathcal{E} = (\mathbf{e}_1, \dots, \mathbf{e}_n)$ où $\mathbf{e}_i = (0, \dots, 0, 1, 0, \dots, 0)$ (le 1 est en position i). Les coordonnées de $v = (x_1, \dots, x_n)$ dans $\mathcal{E}$ sont précisément $(x_1, \dots, x_n)$.

Exemple 44 (Base canonique de $\mathbb{K}_n[X]$)

La base canonique de $\mathbb{K}_n[X]$ est $(1, X, X^2, \dots, X^n)$. Les coordonnées du polynôme $P = a_0 + a_1X + \cdots + a_nX^n$ sont $(a_0, a_1, \dots, a_n)$.

Exemple 45 (Base de $\mathcal{M}_2(\mathbb{R})$)

L'espace $\mathcal{M}_2(\mathbb{R})$ des matrices 2×2 admet pour base canonique :

$$E_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \; E_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \; E_{21} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \; E_{22} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Les coordonnées de $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ sont (a, b, c, d).

2.8 Dimension

Le résultat fondamental suivant garantit que toutes les bases d'un espace vectoriel ont le même nombre d'éléments.

Lemme 46 (Lemme de Steinitz (lemme d'échange))

Soit E un $\mathbb{K}$-espace vectoriel engendré par une famille $(g_1, \dots, g_n)$ de n vecteurs. Soit $(v_1, \dots, v_p)$ une famille libre de E. Alors :

(i) $p \leq n$.

(ii) On peut remplacer p des g_j par les v_i pour obtenir une famille génératrice de E.

Démonstration. On procède par récurrence sur p.

Initialisation ($p = 0$) : il n'y a rien à montrer.

Hérédité. Supposons le résultat vrai pour $p-1$ vecteurs libres, et soit $(v_1, \dots, v_p)$ une famille libre de E. Par hypothèse de récurrence appliquée à $(v_1, \dots, v_{p-1})$, on peut supposer (quitte à renuméroter) que $(v_1, \dots, v_{p-1}, g_p, \dots, g_n)$ est génératrice de E, et $p - 1 \leq n$.

Puisque cette famille engendre E, il existe $\alpha_1, \dots, \alpha_{p-1}, \beta_p, \dots, \beta_n \in \mathbb{K}$ tels que

$$v_p = \alpha_1 v_1 + \cdots + \alpha_{p-1} v_{p-1} + \beta_p g_p + \cdots + \beta_n g_n.$$

Affirmons qu'au moins un β_j est non nul. En effet, si tous les β_j étaient nuls, on aurait $v_p = \alpha_1 v_1 + \cdots + \alpha_{p-1} v_{p-1}$, ce qui contredirait la liberté de $(v_1, \dots, v_p)$.

Soit $j_0 \in \{p, \dots, n\}$ tel que $\beta_{j_0} \neq 0$. On peut exprimer g_{j_0} en fonction de $v_1, \dots, v_p$ et des autres g_j ($j \neq j_0$). En renumérotant, on suppose $j_0 = p$, et on obtient que $(v_1, \dots, v_p, g_{p+1}, \dots, g_n)$ engendre E. De plus, $p \leq n$ (car on a besoin d'au moins un β_j non nul, ce qui requiert $n \geq p$). □

Théorème 47 (Invariance de la dimension)

Si un $\mathbb{K}$-espace vectoriel E admet une base finie, alors toutes ses bases ont le même nombre d'éléments.

Démonstration. Soient $\mathcal{B} = (v_1, \dots, v_n)$ et $\mathcal{B}' = (w_1, \dots, w_m)$ deux bases de E. La famille $\mathcal{B}'$ est libre et $\mathcal{B}$ est génératrice, donc $m \leq n$ par le lemme 46. Inversement, $\mathcal{B}$ est libre et $\mathcal{B}'$ est génératrice, donc $n \leq m$. D'où $m = n$. □

Définition 48 (Dimension)

Un $\mathbb{K}$-espace vectoriel E est dit *de dimension finie* s'il admet une famille génératrice finie. Le nombre d'éléments de toute base de E s'appelle la *dimension* de E et se note $\dim_{\mathbb{K}} E$ ou simplement $\dim E$.
Par convention, $\dim\{\mathbf{0}\} = 0$.

Théorème 49 (Existence de bases)

Tout espace vectoriel de dimension finie admet une base. Plus précisément :

(i) Toute famille génératrice finie contient une base (on peut en extraire une sous-famille qui est une base).

(ii) Toute famille libre peut être complétée en une base.

Démonstration. *(i)* Soit $(g_1, \dots, g_n)$ une famille génératrice de E. Si elle est libre, c'est une base. Sinon, l'un des g_j est combinaison linéaire des autres (par le théorème 39) ; on peut le retirer et la famille restante engendre encore E. On itère ce procédé (qui termine car la famille diminue strictement à chaque étape) jusqu'à obtenir une famille libre, qui est alors une base.

(ii) Soit $(v_1, \dots, v_p)$ une famille libre de E, et $(g_1, \dots, g_n)$ une famille génératrice de E (qui existe car E est de dimension finie). On peut, par le procédé d'échange du lemme 46, compléter $(v_1, \dots, v_p)$ en une famille génératrice en lui adjoignant certains des g_j. La famille obtenue est génératrice ; on en extrait une base par (i) en ne retirant aucun des v_i (car ils forment une famille libre). □

Exemple 50 (Dimensions classiques)

(a) $\dim \mathbb{K}^n = n$ (base canonique $(\mathbf{e}_1, \dots, \mathbf{e}_n)$).

(b) $\dim \mathbb{K}_n[X] = n+1$ (base $(1, X, \dots, X^n)$).

(c) $\dim \mathcal{M}_{n,p}(\mathbb{K}) = np$ (base $(E_{ij})_{1\le i\le n,\, 1\le j\le p}$).

(d) $\mathbb{K}[X]$ est de dimension *infinie* (la famille $(1, X, X^2, \dots)$ est libre, donc $\dim \mathbb{K}[X] \ge n$ pour tout n).

Théorème 51 (Théorème de la dimension pour les sous-espaces)

Soit E un $\mathbb{K}$-espace vectoriel de dimension finie n, et soit F un sous-espace vectoriel de E. Alors :

(i) F est de dimension finie et $\dim F \le \dim E$.

(ii) $\dim F = \dim E$ si et seulement si $F = E$.

Démonstration. *(i)* Toute famille libre de F est aussi une famille libre de E, donc a au plus $n = \dim E$ éléments (par le lemme 46). Soit $(v_1, \dots, v_p)$ une famille libre de F de cardinal maximal (un tel maximum existe car $p \le n$). Montrons qu'elle engendre F : si un vecteur $w \in F$ n'était pas dans $\mathrm{Vect}(v_1, \dots, v_p)$, alors $(v_1, \dots, v_p, w)$ serait libre dans F (car $w \notin \mathrm{Vect}(v_1, \dots, v_p)$), ce qui contredirait la maximalité de p. Donc $\mathrm{Vect}(v_1, \dots, v_p) = F$, et $\mathcal{B}$ est une base de F avec $\dim F = p \le n$.

(ii) Si $\dim F = \dim E = n$, soit $(v_1, \dots, v_n)$ une base de F. C'est une famille libre de n vecteurs dans E de dimension n, donc c'est une base de E (par la proposition 52 ci-dessous). D'où $F = \text{Vect}(v_1, \dots, v_n) = E$. □

Proposition 52 (Critères de base en dimension n)

Soit E un $\mathbb{K}$-espace vectoriel de dimension n. Soit $(v_1, \dots, v_n)$ une famille de n vecteurs de E. Alors les assertions suivantes sont équivalentes :

(i) $(v_1, \dots, v_n)$ est une base de E.

(ii) $(v_1, \dots, v_n)$ est libre.

(iii) $(v_1, \dots, v_n)$ est génératrice.

Démonstration. *(i)⇒(ii)* et *(i)⇒(iii)* sont évidents.

(ii)⇒(i). Soit $\mathcal{B}'$ une base de E (de cardinal n). La famille $(v_1, \dots, v_n)$ est libre et $\mathcal{B}'$ est génératrice de cardinal n. Par le lemme 46, on peut remplacer les n vecteurs de $\mathcal{B}'$ par les v_i (puisque $p = n$) pour obtenir une famille génératrice. Mais cette famille est précisément $(v_1, \dots, v_n)$, qui est donc à la fois libre et génératrice : c'est une base.

(iii)⇒(i). Si $(v_1, \dots, v_n)$ est génératrice, on peut en extraire une base par le théorème 49 (i). Cette base a $\dim E = n$ éléments. Puisqu'on part de n vecteurs et qu'on ne peut en retirer aucun (la base extraite a aussi n éléments), la famille de départ est elle-même une base. □

2.9 Dimension de la somme — formule de Grassmann

Théorème 53 (Formule de Grassmann)

Soient F et G deux sous-espaces vectoriels de dimension finie d'un $\mathbb{K}$-espace vectoriel E. Alors $F + G$ est de dimension finie et

$$\dim(F + G) = \dim F + \dim G - \dim(F \cap G).$$

Démonstration. Posons $r = \dim(F \cap G)$, $p = \dim F$ et $q = \dim G$. Soit $(w_1, \dots, w_r)$ une base de $F \cap G$.

Puisque $F \cap G$ est un s.e.v. de F, on peut compléter cette famille en une base de F : soit $(w_1, \dots, w_r, u_1, \dots, u_s)$ une base de F, avec $s = p - r$.

De même, soit $(w_1, \dots, w_r, v_1, \dots, v_t)$ une base de G, avec $t = q - r$.

Affirmons que $\mathcal{B} = (w_1, \dots, w_r, u_1, \dots, u_s, v_1, \dots, v_t)$ est une base de $F + G$.

$\mathcal{B}$ est génératrice de $F + G$. Tout $x \in F + G$ s'écrit $x = f + g$ avec $f \in F$ et $g \in G$. Or f est combinaison linéaire des w_i et des u_j, et g est combinaison linéaire des w_i et des v_k. Donc x est combinaison linéaire des w_i, u_j, v_k.

$\mathcal{B}$ est libre. Supposons

$$\sum_{i=1}^{r} \gamma_i w_i + \sum_{j=1}^{s} \alpha_j u_j + \sum_{k=1}^{t} \beta_k v_k = \mathbf{0}.$$

Posons $h = \sum_{k=1}^{t} \beta_k v_k = -\sum_{i=1}^{r} \gamma_i w_i - \sum_{j=1}^{s} \alpha_j u_j$. Le membre droit appartient à F (car les w_i et u_j sont dans F). Le membre gauche appartient à G (car les v_k sont dans G). Donc $h \in F \cap G$.

Puisque $(w_1, \dots, w_r)$ est une base de $F \cap G$, il existe $\delta_1, \dots, \delta_r \in \mathbb{K}$ tels que $h = \sum_{i=1}^{r} \delta_i w_i$. D'où

$$\sum_{k=1}^{t} \beta_k v_k - \sum_{i=1}^{r} \delta_i w_i = \mathbf{0}.$$

Puisque $(w_1, \dots, w_r, v_1, \dots, v_t)$ est une base de G (donc libre), on en déduit $\beta_1 = \cdots = \beta_t = 0$ et $\delta_1 = \cdots = \delta_r = 0$.

Il reste $\sum_{i=1}^{r} \gamma_i w_i + \sum_{j=1}^{s} \alpha_j u_j = \mathbf{0}$. Puisque $(w_1, \dots, w_r, u_1, \dots, u_s)$ est une base de F (donc libre), on a $\gamma_1 = \cdots = \gamma_r = 0$ et $\alpha_1 = \cdots = \alpha_s = 0$.

Tous les coefficients sont nuls : $\mathcal{B}$ est libre.

Ainsi $\mathcal{B}$ est une base de $F + G$, et

$$\dim(F+G) = r + s + t = r + (p-r) + (q-r) = p + q - r = \dim F + \dim G - \dim(F \cap G). \quad \square$$

Corollaire 54 (Dimension de la somme directe)

Si $F \cap G = \{\mathbf{0}\}$ (somme directe), alors $\dim(F \oplus G) = \dim F + \dim G$.

Démonstration. C'est le cas $\dim(F \cap G) = 0$ de la formule de Grassmann. $\square$

Corollaire 55 (Supplémentaire et dimension)

Soient E un espace de dimension finie n et F un sous-espace de E de dimension p. Tout supplémentaire G de F a pour dimension $n - p$.

Démonstration. Si $E = F \oplus G$, alors $n = \dim E = \dim F + \dim G = p + \dim G$, d'où $\dim G = n - p$. $\square$

Corollaire 56 (Condition de somme directe en dimension finie)

Soient F et G deux sous-espaces d'un espace E de dimension finie. Alors $E = F \oplus G$ si et seulement si l'une des conditions suivantes est satisfaite :

(i) $E = F + G$ et $\dim E = \dim F + \dim G$;

(ii) $F \cap G = \{\mathbf{0}\}$ et $\dim E = \dim F + \dim G$.

Démonstration. Dans les deux cas, la formule de Grassmann donne la condition manquante.

(i) Si $E = F + G$ et $\dim E = \dim F + \dim G$, la formule de Grassmann donne $\dim(F \cap G) = \dim F + \dim G - \dim(F+G) = \dim E - \dim E = 0$, donc $F \cap G = \{\mathbf{0}\}$.

(ii) Si $F \cap G = \{\mathbf{0}\}$ et $\dim F + \dim G = \dim E$, alors $\dim(F+G) = \dim F + \dim G = \dim E$ par Grassmann, donc $F + G = E$ par le théorème 51 (ii). $\square$

2.10 Rang d'une famille de vecteurs

Définition 57 (Rang d'une famille de vecteurs)

Soit $(v_1, \dots, v_p)$ une famille de vecteurs d'un $\mathbb{K}$-espace vectoriel E de dimension finie. Le *rang* de cette famille est la dimension du sous-espace engendré :

$$\operatorname{rg}(v_1, \dots, v_p) := \dim \operatorname{Vect}(v_1, \dots, v_p).$$

Proposition 58 (Propriétés du rang)

Soit $(v_1, \dots, v_p)$ une famille de vecteurs de E.

(i) $0 \leq \operatorname{rg}(v_1, \dots, v_p) \leq \min(p, \dim E)$.

(ii) $\operatorname{rg}(v_1, \dots, v_p) = p$ si et seulement si la famille est libre.

(iii) $\operatorname{rg}(v_1, \dots, v_p) = \dim E$ si et seulement si la famille est génératrice.

(iv) $\operatorname{rg}(v_1, \dots, v_p) = p = \dim E$ si et seulement si la famille est une base.

Démonstration. *(i)* Le rang est la dimension de $\operatorname{Vect}(v_1, \dots, v_p)$, qui est un s.e.v. de E. Donc $\operatorname{rg} \leq \dim E$. De plus, $(v_1, \dots, v_p)$ est une famille génératrice de $\operatorname{Vect}(v_1, \dots, v_p)$, donc toute base extraite a au plus p éléments.

(ii) La famille est libre ssi elle est une base de $\operatorname{Vect}(v_1, \dots, v_p)$, ce qui arrive ssi $\dim \operatorname{Vect}(v_1, \dots, v_p) = p$.

(iii) La famille est génératrice ssi $\operatorname{Vect}(v_1, \dots, v_p) = E$, soit $\operatorname{rg} = \dim E$.

(iv) Combine (ii) et (iii). □

2.11 Applications

2.11.1 Espaces de solutions

Proposition 59 (Dimension de l'espace des solutions)

Soit $A \in \mathcal{M}_{n,p}(\mathbb{K})$ une matrice de rang r. L'ensemble des solutions du système homogène $AX = 0$ est un sous-espace vectoriel de $\mathbb{K}^p$ de dimension $p - r$.

Ce résultat fondamental, appelé *théorème du rang*, sera démontré au chapitre 2 dans le cadre général des applications linéaires. Pour l'instant, admettons-le et voyons-en une application.

Exemple 60 (Espace des solutions d'un système)

Considérons le système homogène sur $\mathbb{R}$:

$$\begin{cases} x + 2y - z = 0, \\ 2x + 4y - 2z = 0. \end{cases}$$

La matrice du système est $A = \left(\begin{smallmatrix} 1 & 2 & -1 \\ 2 & 4 & -2 \end{smallmatrix}\right)$, de rang 1 (la seconde ligne est le double de la première). L'espace des solutions est donc de dimension $3-1 = 2$. En résolvant, on trouve $x = -2y + z$, d'où :

$$\mathcal{S} = \{ (-2y + z,\ y,\ z) \mid y, z \in \mathbb{R} \} = \operatorname{Vect}((-2, 1, 0),\ (1, 0, 1)).$$

La famille $((-2, 1, 0), (1, 0, 1))$ est une base de $\mathcal{S}$.

2.11.2 Espaces de polynômes

Exemple 61 (Polynômes pairs et impairs)

Soit $E = \mathbb{R}_n[X]$ ($n \in \mathbb{N}$). Posons

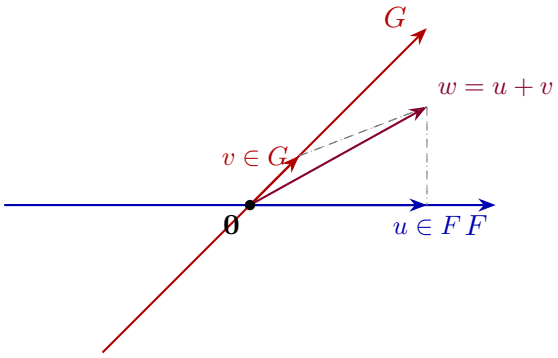
$$F = \{\, P \in E \mid P(-X) = P(X) \,\} \qquad \text{(polynômes pairs)},$$
$$G = \{\, P \in E \mid P(-X) = -P(X) \,\} \qquad \text{(polynômes impairs)}.$$

Alors F et G sont des sous-espaces de E, et $E = F \oplus G$.
En effet, $F \cap G = \{0\}$ (si P est à la fois pair et impair, $P(X) = -P(X)$ pour tout X, donc $2P = 0$, d'où $P = 0$ car $\mathrm{car}(\mathbb{R}) = 0$). De plus, tout $P \in E$ s'écrit

$$P(X) = \underbrace{\frac{P(X) + P(-X)}{2}}_{\in F} + \underbrace{\frac{P(X) - P(-X)}{2}}_{\in G}.$$

Si $n = 2m$, on a $\dim F = m + 1$ et $\dim G = m$, d'où $\dim F + \dim G = (m + 1) + m = 2m + 1 = n + 1 = \dim E$. Si $n = 2m + 1$, on a $\dim F = m + 1$ et $\dim G = m + 1$, d'où $\dim F + \dim G = 2m + 2 = n + 1 = \dim E$.

2.11.3 Somme directe — illustration géométrique



$\mathbb{R}^2 = F \oplus G$

Fig. 2.3 : Décomposition d'un vecteur selon deux sous-espaces supplémentaires dans $\mathbb{R}^2$.

2.12 Exercices

Exercice 62 (Vérification des axiomes)

Montrer que l'ensemble $E = \mathbb{R}^2$ muni des opérations

$$(x_1, y_1) \oplus (x_2, y_2) = (x_1 + x_2,\ y_1 + y_2), \qquad \alpha \odot (x, y) = (\alpha x,\ \alpha y)$$

est un $\mathbb{R}$-espace vectoriel. Identifier le vecteur nul et l'opposé de (x, y).

Exercice 63 (Espace vectoriel ou non ?)

Pour chacun des ensembles suivants, munis des opérations usuelles d'addition et de multiplication scalaire, dire s'il s'agit d'un $\mathbb{R}$-espace vectoriel. Justifier.

(a) $E_1 = \{ (x,y) \in \mathbb{R}^2 \mid x+y=0 \}$.

(b) $E_2 = \{ (x,y) \in \mathbb{R}^2 \mid x+y=1 \}$.

(c) $E_3 = \{ (x,y) \in \mathbb{R}^2 \mid xy=0 \}$.

(d) $E_4 = \{ P \in \mathbb{R}[X] \mid P(0)=0 \}$.

(e) $E_5 = \{ f \in \mathcal{C}^0(\mathbb{R},\mathbb{R}) \mid f(0)=0 \text{ et } f(1)=0 \}$.

Exercice 64 (Sous-espace vectoriel)

Montrer que les ensembles suivants sont des sous-espaces vectoriels de l'espace ambiant indiqué :

(a) $F = \{ (x,y,z) \in \mathbb{R}^3 \mid x-2y+z=0 \}$ dans $\mathbb{R}^3$.

(b) $G = \{ A \in \mathcal{M}_2(\mathbb{R}) \mid \mathrm{tr}(A)=0 \}$ dans $\mathcal{M}_2(\mathbb{R})$.

(c) $H = \{ P \in \mathbb{R}_3[X] \mid P(1)=0 \}$ dans $\mathbb{R}_3[X]$.

Exercice 65 (Famille libre ou liée)

Déterminer si les familles suivantes sont libres ou liées. Si la famille est liée, exhiber une relation de dépendance linéaire.

(a) $((1,2),(3,6))$ dans $\mathbb{R}^2$.

(b) $((1,1,0),(0,1,1),(1,0,-1))$ dans $\mathbb{R}^3$.

(c) $(1, 1+X, 1+X+X^2)$ dans $\mathbb{R}_2[X]$.

(d) (e^x, e^{2x}, e^{3x}) dans $\mathcal{C}^0(\mathbb{R},\mathbb{R})$.

Exercice 66 (Base et coordonnées)

(a) Vérifier que $\mathcal{B} = ((1,1),(1,-1))$ est une base de $\mathbb{R}^2$.

(b) Exprimer le vecteur $v=(3,5)$ dans la base $\mathcal{B}$ (trouver ses coordonnées).

(c) Montrer que $\mathcal{B}' = ((1,1,1),(1,1,0),(1,0,0))$ est une base de $\mathbb{R}^3$ et trouver les coordonnées de $(2,3,5)$ dans $\mathcal{B}'$.

Exercice 67 (Intersection et somme)

Dans $\mathbb{R}^4$, soient

$$F = \mathrm{Vect}((1,0,1,0),\ (0,1,0,1)),$$
$$G = \mathrm{Vect}((1,1,0,0),\ (0,0,1,1)).$$

(a) Déterminer $\dim F$ et $\dim G$.

(b) Déterminer une base de $F \cap G$.

(c) En déduire $\dim(F+G)$ par la formule de Grassmann.

(d) La somme $F + G$ est-elle directe ?

Exercice 68 (Somme directe de sous-espaces)

Dans $\mathbb{R}^3$, soient $F = \{ (x, y, z) \mid x + y = 0 \}$ et $G = \text{Vect}((1, 1, 1))$.

(a) Montrer que F est un sous-espace de dimension 2 et déterminer une base de F.

(b) Montrer que $F \cap G = \{\mathbf{0}\}$.

(c) En déduire que $\mathbb{R}^3 = F \oplus G$.

(d) Décomposer le vecteur $(2, 3, 5) \in \mathbb{R}^3$ en $u + v$ avec $u \in F$ et $v \in G$.

Exercice 69 (Dimension d'un espace de matrices)

Soit $E = \mathcal{M}_n(\mathbb{R})$, et soient

$$\mathcal{S}_n(\mathbb{R}) = \{ A \in E \mid A^\mathsf{T} = A \} \qquad \text{(matrices symétriques)},$$
$$\mathcal{A}_n(\mathbb{R}) = \{ A \in E \mid A^\mathsf{T} = -A \} \qquad \text{(matrices antisymétriques)}.$$

(a) Montrer que $\mathcal{S}_n(\mathbb{R})$ et $\mathcal{A}_n(\mathbb{R})$ sont des sous-espaces de E.

(b) Calculer $\dim \mathcal{S}_n(\mathbb{R})$ et $\dim \mathcal{A}_n(\mathbb{R})$.

(c) Montrer que $E = \mathcal{S}_n(\mathbb{R}) \oplus \mathcal{A}_n(\mathbb{R})$.

Exercice 70 (Sous-espace engendré et rang)

On considère les vecteurs de $\mathbb{R}^4$:

$$v_1 = (1, 2, 1, 0), \quad v_2 = (1, 1, 0, 1), \quad v_3 = (3, 5, 2, 1), \quad v_4 = (0, 1, 1, -1).$$

(a) Montrer que (v_1, v_2, v_3, v_4) est liée en exhibant une relation de dépendance.

(b) Déterminer le rang de cette famille.

(c) En extraire une base de $\text{Vect}(v_1, v_2, v_3, v_4)$.

Exercice 71 (Formule de Grassmann)

Dans $\mathbb{R}^4$, soient

$$F = \{ (x, y, z, t) \in \mathbb{R}^4 \mid x + y = 0,\ z - t = 0 \},$$
$$G = \{ (x, y, z, t) \in \mathbb{R}^4 \mid x - z = 0,\ y + t = 0 \}.$$

(a) Déterminer des bases de F et G, et calculer leurs dimensions.

(b) Déterminer une base de $F \cap G$ et calculer sa dimension.

(c) Vérifier la formule de Grassmann.

(d) La somme $F + G$ est-elle directe ? Est-elle égale à $\mathbb{R}^4$?

Exercice 72 (Supplémentaire dans $\mathbb{R}_n[X]$)

Soient $n \geq 1$ et $F = \{ P \in \mathbb{R}_n[X] \mid P(0) = P(1) = 0 \}$.

(a) Montrer que F est un sous-espace de $\mathbb{R}_n[X]$ et calculer sa dimension.

(b) Trouver un supplémentaire G de F dans $\mathbb{R}_n[X]$ et en donner une base.

(c) Pour $n = 3$, décomposer le polynôme $P = 1 + X + X^2 + X^3$ en $f + g$ avec $f \in F$ et $g \in G$.

Exercice 73 (Dimension infinie)

(a) Montrer que $\mathbb{K}[X]$ est de dimension infinie.

(b) Montrer que l'espace $\mathcal{C}^0([0,1], \mathbb{R})$ des fonctions continues sur $[0,1]$ est de dimension infinie.

Indication : montrer que les fonctions $1, x, x^2, \ldots$ forment une famille libre.

(c) Montrer que l'espace $\mathbb{R}^{\mathbb{N}}$ des suites réelles est de dimension infinie.

Exercice 74 (Une somme qui n'est pas directe)

Dans $\mathbb{R}^3$, soient $F_1 = \text{Vect}((1,0,0),(0,1,0))$, $F_2 = \text{Vect}((0,1,0),(0,0,1))$ et $F_3 = \text{Vect}((1,0,0),(0,0,1))$.

(a) Montrer que $F_1 + F_2 + F_3 = \mathbb{R}^3$.

(b) Montrer que $F_i \cap F_j = \{\mathbf{0}\}$ n'est *pas* vérifié pour tous $i \neq j$.

(c) La somme $F_1 + F_2 + F_3$ est-elle directe ? Justifier.

(d) Comparer avec la condition de la proposition 27.

Exercice 75 (Lemme de Steinitz appliqué)

Soit $E = \mathbb{R}^3$ et $\mathcal{B} = (e_1, e_2, e_3)$ la base canonique. Soit $v_1 = (1,1,0)$, $v_2 = (0,1,1)$.

(a) Vérifier que (v_1, v_2) est libre.

(b) Appliquer le procédé du lemme 46 pour compléter (v_1, v_2) en une base de $\mathbb{R}^3$ en utilisant les vecteurs de $\mathcal{B}$.

(c) Combien de façons y a-t-il de choisir le troisième vecteur parmi $\{e_1, e_2, e_3\}$ pour obtenir une base ?

Résumé du chapitre

Ce qu'il faut retenir.

1. **Espace vectoriel.** Un $\mathbb{K}$-espace vectoriel est un ensemble E muni d'une addition (qui en fait un groupe abélien) et d'une multiplication par un scalaire de $\mathbb{K}$ satisfaisant les axiomes de compatibilité, d'élément unité et de distributivité.

2. **Sous-espace vectoriel.** $F \subseteq E$ est un s.e.v. ssi $F \neq \emptyset$ et F est stable par combinaison linéaire.

3. **Intersection et somme.** L'intersection de sous-espaces est un s.e.v. La somme $F+G$ est le plus petit s.e.v. contenant F et G. La réunion n'est en général *pas* un s.e.v.

4. **Somme directe.** $F+G$ est directe ssi $F \cap G = \{\mathbf{0}\}$. On note $F \oplus G$. Si $E = F \oplus G$, on dit que F et G sont supplémentaires.

5. **Famille libre.** $(\alpha_1 v_1 + \cdots + \alpha_n v_n = \mathbf{0}) \implies (\alpha_i = 0 \ \forall i)$.

6. **Famille génératrice.** Tout vecteur de E est combinaison linéaire de la famille.

7. **Base.** Famille à la fois libre et génératrice. Tout vecteur s'y décompose de manière unique (coordonnées). En dimension finie n : famille libre de n vecteurs $\iff$ famille génératrice de n vecteurs $\iff$ base.

8. **Dimension.** Nombre d'éléments de toute base (invariant). $\dim \mathbb{K}^n = n$, $\dim \mathbb{K}_n[X] = n+1$, $\dim \mathcal{M}_{n,p}(\mathbb{K}) = np$.

9. **Formule de Grassmann.** $\dim(F+G) = \dim F + \dim G - \dim(F \cap G)$.

10. **Rang.** Le rang d'une famille $(v_1, \dots, v_p)$ est $\dim \mathrm{Vect}(v_1, \dots, v_p)$.

11. **Théorèmes-clés.**

    - Lemme de Steinitz (échange) — une famille libre a au plus autant d'éléments qu'une famille génératrice.
    - Toute famille libre peut être complétée en une base.
    - $\dim F \leq \dim E$, avec égalité ssi $F = E$.

Chapitre 3

Applications linéaires et matrices

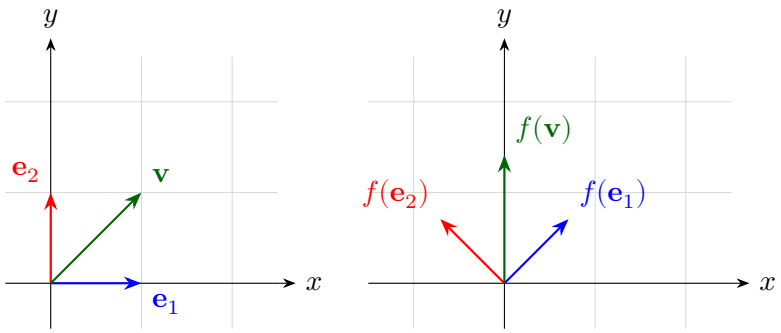
Introduction

L'algèbre linéaire est, en un sens, l'étude des *applications linéaires* : ces transformations qui respectent la structure d'espace vectoriel. Avant d'en donner la définition formelle, observons quelques exemples géométriques dans le plan $\mathbb{R}^2$.

Considérons les transformations suivantes du plan :

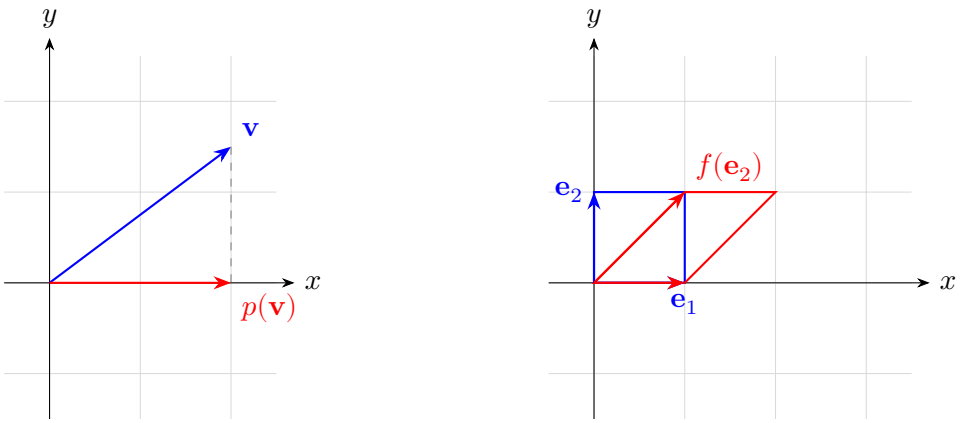
- la **rotation** d'angle θ autour de l'origine, qui envoie le vecteur $\begin{pmatrix} x \\ y \end{pmatrix}$ sur $\begin{pmatrix} x\cos\theta - y\sin\theta \\ x\sin\theta + y\cos\theta \end{pmatrix}$;
- la **réflexion** par rapport à l'axe des abscisses, qui envoie $\begin{pmatrix} x \\ y \end{pmatrix}$ sur $\begin{pmatrix} x \\ -y \end{pmatrix}$;
- la **projection** orthogonale sur l'axe des abscisses, qui envoie $\begin{pmatrix} x \\ y \end{pmatrix}$ sur $\begin{pmatrix} x \\ 0 \end{pmatrix}$;
- l'**homothétie** de rapport λ, qui envoie $\begin{pmatrix} x \\ y \end{pmatrix}$ sur $\begin{pmatrix} \lambda x \\ \lambda y \end{pmatrix}$.

Ces quatre transformations partagent deux propriétés remarquables : elles envoient la somme de deux vecteurs sur la somme des images, et le produit d'un vecteur par un scalaire sur le produit du scalaire par l'image. C'est précisément cette double compatibilité qui caractérise les applications linéaires.



Vecteurs originaux Rotation de $\pi/4$

Fig. 3.1 : Effet d'une rotation de $\pi/4$ sur les vecteurs du plan.



Projection sur Ox Cisaillement $(k = 1)$

Fig. 3.2 : Projection orthogonale et cisaillement : deux transformations linéaires.

3.1 Définition d'une application linéaire

Définition 1 (Application linéaire)

Soient E et F deux $\mathbb{K}$-espaces vectoriels. Une application $f : E \to F$ est dite *linéaire* si elle vérifie les deux conditions suivantes :

(i) **Additivité :** $\forall\, \mathbf{u}, \mathbf{v} \in E, \quad f(\mathbf{u}+\mathbf{v}) = f(\mathbf{u}) + f(\mathbf{v})$.

(ii) **Homogénéité :** $\forall\, \lambda \in \mathbb{K},\ \forall\, \mathbf{u} \in E, \quad f(\lambda \mathbf{u}) = \lambda\, f(\mathbf{u})$.

De manière équivalente, f est linéaire si et seulement si :

$$\forall\, \lambda, \mu \in \mathbb{K},\ \forall\, \mathbf{u}, \mathbf{v} \in E, \quad f(\lambda \mathbf{u} + \mu \mathbf{v}) = \lambda\, f(\mathbf{u}) + \mu\, f(\mathbf{v}).$$

On note $\mathcal{L}(E, F)$ l'ensemble des applications linéaires de E dans F. Lorsque $E = F$, on parle d'*endomorphisme* et l'on note $\mathrm{End}(E) := \mathcal{L}(E, E)$.

Remarque 2 (Terminologie)

On emploie parfois les termes *morphisme d'espaces vectoriels*, *opérateur linéaire* (surtout lorsque $E = F$) ou *homomorphisme*. En anglais, on dit *linear map* ou *linear transformation*.

Proposition 3 (Propriétés élémentaires)

Soit $f : E \to F$ une application linéaire. Alors :

(i) $f(\mathbf{0}_E) = \mathbf{0}_F$.

(ii) $\forall\, \mathbf{u} \in E, \quad f(-\mathbf{u}) = -f(\mathbf{u})$.

(iii) Pour toute combinaison linéaire finie : $f\left(\sum_{i=1}^{n} \lambda_i \mathbf{u}_i\right) = \sum_{i=1}^{n} \lambda_i\, f(\mathbf{u}_i)$.

Démonstration. *(i)* On a $f(\mathbf{0}_E) = f(0 \cdot \mathbf{0}_E) = 0 \cdot f(\mathbf{0}_E) = \mathbf{0}_F$.

(ii) $f(-\mathbf{u}) = f((-1)\cdot\mathbf{u}) = (-1)\cdot f(\mathbf{u}) = -f(\mathbf{u})$.
(iii) Par récurrence immédiate à partir de l'additivité et de l'homogénéité. □

Remarque 4 (Détermination par les images d'une base)

Si E est de dimension finie et si $\mathcal{B} = (\mathbf{e}_1, \dots, \mathbf{e}_n)$ est une base de E, alors une application linéaire $f\colon E \to F$ est entièrement déterminée par les images $f(\mathbf{e}_1), \dots, f(\mathbf{e}_n)$. En effet, pour $\mathbf{u} = \sum_{i=1}^{n} x_i\,\mathbf{e}_i$, on a $f(\mathbf{u}) = \sum_{i=1}^{n} x_i\, f(\mathbf{e}_i)$. Réciproquement, la donnée de n vecteurs quelconques $\mathbf{f}_1, \dots, \mathbf{f}_n \in F$ définit une unique application linéaire $f \in \mathcal{L}(E,F)$ telle que $f(\mathbf{e}_i) = \mathbf{f}_i$ pour tout i.

3.2 Exemples fondamentaux

Exemple 5 (Rotation dans $\mathbb{R}^2$)

Soit $R_\theta\colon \mathbb{R}^2 \to \mathbb{R}^2$ la rotation d'angle θ autour de l'origine. Pour $\mathbf{u} = \begin{pmatrix} x \\ y \end{pmatrix}$, on a

$$R_\theta(\mathbf{u}) = \begin{pmatrix} x\cos\theta - y\sin\theta \\ x\sin\theta + y\cos\theta \end{pmatrix}.$$

Vérifions la linéarité. Soient $\mathbf{u} = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$, $\mathbf{v} = \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}$ et $\lambda \in \mathbb{R}$. Alors :

$$R_\theta(\mathbf{u}+\mathbf{v}) = \begin{pmatrix} (x_1+x_2)\cos\theta - (y_1+y_2)\sin\theta \\ (x_1+x_2)\sin\theta + (y_1+y_2)\cos\theta \end{pmatrix} = R_\theta(\mathbf{u}) + R_\theta(\mathbf{v}),$$
$$R_\theta(\lambda\mathbf{u}) = \begin{pmatrix} \lambda x_1\cos\theta - \lambda y_1\sin\theta \\ \lambda x_1\sin\theta + \lambda y_1\cos\theta \end{pmatrix} = \lambda\, R_\theta(\mathbf{u}).$$

Donc R_θ est un endomorphisme de $\mathbb{R}^2$.

Exemple 6 (Projection orthogonale)

La projection orthogonale $p\colon \mathbb{R}^2 \to \mathbb{R}^2$ sur l'axe des abscisses est définie par $p\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ 0 \end{pmatrix}$. C'est une application linéaire : on vérifie aisément que $p(\mathbf{u}+\mathbf{v}) = p(\mathbf{u}) + p(\mathbf{v})$ et $p(\lambda\mathbf{u}) = \lambda\, p(\mathbf{u})$. De plus, $p \circ p = p$: c'est un *projecteur* (ou *idempotent*).

Exemple 7 (Dérivation)

Soit $E = \mathbb{R}_n[X]$ l'espace des polynômes de degré au plus n à coefficients réels. L'application *dérivation*

$$D\colon \mathbb{R}_n[X] \to \mathbb{R}_{n-1}[X], \qquad P \mapsto P'$$

est linéaire : $(P+Q)' = P' + Q'$ et $(\lambda P)' = \lambda P'$. Son noyau est l'ensemble des polynômes constants et son image est $\mathbb{R}_{n-1}[X]$ tout entier.

Exemple 8 (Intégration)

L'application

$$I : \mathcal{C}([0,1],\mathbb{R}) \to \mathbb{R}, \qquad f \mapsto \int_0^1 f(t)\,dt$$

est linéaire (par linéarité de l'intégrale de Riemann). Ici, l'espace de départ est de dimension infinie et l'espace d'arrivée est de dimension 1.

Exemple 9 (Multiplication par une matrice)

Soit $A \in \mathcal{M}_{m,n}(\mathbb{K})$ une matrice à m lignes et n colonnes. L'application

$$f_A : \mathbb{K}^n \to \mathbb{K}^m, \qquad \mathbf{x} \mapsto A\mathbf{x}$$

est linéaire : $A(\mathbf{x}+\mathbf{y}) = A\mathbf{x} + A\mathbf{y}$ et $A(\lambda\mathbf{x}) = \lambda\, A\mathbf{x}$. Nous verrons que, réciproquement, toute application linéaire entre espaces de dimension finie peut se représenter par une telle multiplication matricielle (section 3.8).

Exemple 10 (Trace d'une matrice)

L'application

$$\mathrm{tr} : \mathcal{M}_n(\mathbb{K}) \to \mathbb{K}, \qquad A = (a_{ij}) \mapsto \sum_{i=1}^{n} a_{ii}$$

est linéaire : $\mathrm{tr}(A+B) = \mathrm{tr}(A) + \mathrm{tr}(B)$ et $\mathrm{tr}(\lambda A) = \lambda\,\mathrm{tr}(A)$.

3.3 Noyau et image

Définition 11 (Noyau)

Soit $f : E \to F$ une application linéaire. Le *noyau* de f est

$$\mathrm{Ker}\, f := \{\, \mathbf{u} \in E \mid f(\mathbf{u}) = \mathbf{0}_F \,\}.$$

Définition 12 (Image)

Soit $f : E \to F$ une application linéaire. L'*image* de f est

$$\mathrm{Im}\, f := \{\, f(\mathbf{u}) \mid \mathbf{u} \in E \,\} = \{\, \mathbf{w} \in F \mid \exists\, \mathbf{u} \in E,\ f(\mathbf{u}) = \mathbf{w} \,\}.$$

Théorème 13 (Le noyau et l'image sont des sous-espaces vectoriels)

Soit $f : E \to F$ une application linéaire.

(i) $\mathrm{Ker}\, f$ est un sous-espace vectoriel de E.

(ii) $\mathrm{Im}\, f$ est un sous-espace vectoriel de F.

Démonstration. *(i)* Le noyau est non vide car $f(\mathbf{0}_E) = \mathbf{0}_F$, donc $\mathbf{0}_E \in \mathrm{Ker}\, f$. Soient $\mathbf{u}, \mathbf{v} \in \mathrm{Ker}\, f$ et $\lambda \in \mathbb{K}$. Alors

$$f(\mathbf{u} + \lambda\mathbf{v}) = f(\mathbf{u}) + \lambda\, f(\mathbf{v}) = \mathbf{0}_F + \lambda\, \mathbf{0}_F = \mathbf{0}_F,$$

donc $\mathbf{u} + \lambda\mathbf{v} \in \operatorname{Ker} f$. Ainsi $\operatorname{Ker} f$ est un sous-espace vectoriel de E.

(ii) L'image est non vide car $\mathbf{0}_F = f(\mathbf{0}_E) \in \operatorname{Im} f$. Soient $\mathbf{w}_1, \mathbf{w}_2 \in \operatorname{Im} f$ et $\lambda \in \mathbb{K}$. Il existe $\mathbf{u}_1, \mathbf{u}_2 \in E$ tels que $f(\mathbf{u}_1) = \mathbf{w}_1$ et $f(\mathbf{u}_2) = \mathbf{w}_2$. Alors

$$\mathbf{w}_1 + \lambda\mathbf{w}_2 = f(\mathbf{u}_1) + \lambda\, f(\mathbf{u}_2) = f(\mathbf{u}_1 + \lambda\mathbf{u}_2) \in \operatorname{Im} f.$$

Donc $\operatorname{Im} f$ est un sous-espace vectoriel de F. □

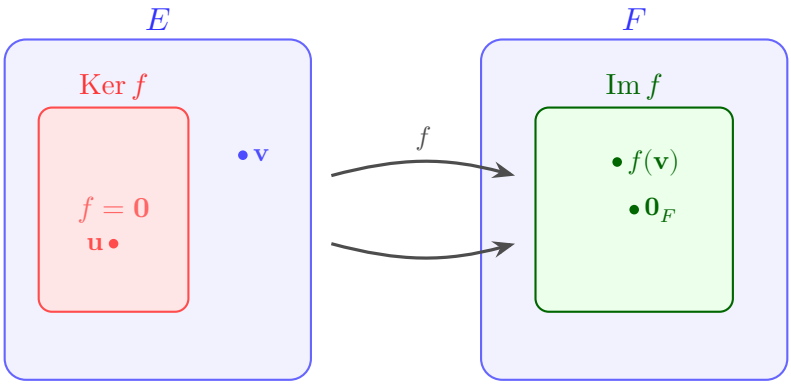


Fig. 3.3 : Noyau et image d'une application linéaire $f : E \to F$.

Définition 14 (Rang)

Soit $f : E \to F$ une application linéaire. Le *rang* de f est la dimension de son image :

$$\operatorname{rg}(f) := \dim(\operatorname{Im} f).$$

Proposition 15 (Image d'une famille génératrice)

Soit $f : E \to F$ linéaire. Si $(u_1, \dots, u_p)$ engendre E, alors $(f(u_1), \dots, f(u_p))$ engendre $\operatorname{Im} f$. En particulier, si $\mathcal{B} = (e_1, \dots, e_n)$ est une base de E, alors

$$\operatorname{Im} f = \operatorname{Vect}(f(e_1), \dots, f(e_n)).$$

Démonstration. Soit $\mathbf{w} \in \operatorname{Im} f$. Il existe $\mathbf{u} \in E$ tel que $f(\mathbf{u}) = \mathbf{w}$. Comme $(u_1, \dots, u_p)$ engendre E, on peut écrire $\mathbf{u} = \sum_{i=1}^{p} \lambda_i u_i$, d'où $\mathbf{w} = f(\mathbf{u}) = \sum_{i=1}^{p} \lambda_i f(u_i)$. □

3.4 Théorème du rang

Le résultat suivant est l'un des théorèmes les plus importants de l'algèbre linéaire. Il relie la dimension de l'espace de départ, celle du noyau et celle de l'image.

Théorème 16 (Théorème du rang)

Soit $f : E \to F$ une application linéaire avec E de dimension finie. Alors $\operatorname{Ker} f$ et $\operatorname{Im} f$ sont de dimension finie et

$$\boxed{\dim E = \dim \operatorname{Ker} f + \dim \operatorname{Im} f = \dim \operatorname{Ker} f + \operatorname{rg}(f).}$$

Démonstration. Posons $n = \dim E$ et $r = \dim \operatorname{Ker} f$.

Étape 1. Le noyau $\operatorname{Ker} f$ est un sous-espace vectoriel de E (théorème 13), donc il est de dimension finie avec $r \leq n$. Choisissons une base $(\mathbf{u}_1, \dots, \mathbf{u}_r)$ de $\operatorname{Ker} f$.

Étape 2. D'après le théorème de la base incomplète, on peut compléter cette base en une base de E :

$$\mathcal{B} = (\mathbf{u}_1, \dots, \mathbf{u}_r, \mathbf{v}_1, \dots, \mathbf{v}_s)$$

où $r + s = n$, c'est-à-dire $s = n - r$.

Étape 3. Montrons que $(f(\mathbf{v}_1), \dots, f(\mathbf{v}_s))$ est une base de $\operatorname{Im} f$. Cela établira $\dim \operatorname{Im} f = s = n - r$, d'où le résultat.

Étape 3a — Famille génératrice. Soit $\mathbf{w} \in \operatorname{Im} f$. Il existe $\mathbf{x} \in E$ tel que $f(\mathbf{x}) = \mathbf{w}$. Décomposons $\mathbf{x}$ dans la base $\mathcal{B}$:

$$\mathbf{x} = \sum_{i=1}^{r} \alpha_i \, \mathbf{u}_i + \sum_{j=1}^{s} \beta_j \, \mathbf{v}_j.$$

Par linéarité de f :

$$\mathbf{w} = f(\mathbf{x}) = \sum_{i=1}^{r} \alpha_i \underbrace{f(\mathbf{u}_i)}_{=\mathbf{0}_F} + \sum_{j=1}^{s} \beta_j \, f(\mathbf{v}_j) = \sum_{j=1}^{s} \beta_j \, f(\mathbf{v}_j).$$

Donc $(f(\mathbf{v}_1), \dots, f(\mathbf{v}_s))$ engendre $\operatorname{Im} f$.

Étape 3b — Liberté. Supposons que $\sum_{j=1}^{s} \beta_j \, f(\mathbf{v}_j) = \mathbf{0}_F$. Par linéarité, cela signifie $f\left(\sum_{j=1}^{s} \beta_j \, \mathbf{v}_j\right) = \mathbf{0}_F$, c'est-à-dire $\sum_{j=1}^{s} \beta_j \, \mathbf{v}_j \in \operatorname{Ker} f$. Il existe donc des scalaires $\alpha_1, \dots, \alpha_r \in \mathbb{K}$ tels que

$$\sum_{j=1}^{s} \beta_j \, \mathbf{v}_j = \sum_{i=1}^{r} \alpha_i \, \mathbf{u}_i,$$

soit

$$\sum_{i=1}^{r} (-\alpha_i) \, \mathbf{u}_i + \sum_{j=1}^{s} \beta_j \, \mathbf{v}_j = \mathbf{0}_E.$$

Comme $(\mathbf{u}_1, \dots, \mathbf{u}_r, \mathbf{v}_1, \dots, \mathbf{v}_s)$ est une base de E, tous les coefficients sont nuls : en particulier, $\beta_1 = \cdots = \beta_s = 0$. La famille $(f(\mathbf{v}_1), \dots, f(\mathbf{v}_s))$ est donc libre.

Conclusion. $(f(\mathbf{v}_1), \dots, f(\mathbf{v}_s))$ est une base de $\operatorname{Im} f$, d'où $\dim \operatorname{Im} f = s = n - r = \dim E - \dim \operatorname{Ker} f$. □

Corollaire 17 (Formule du rang)

Si $\dim E = n$ et $\dim F = m$, alors pour toute application linéaire $f : E \to F$:

$$\operatorname{rg}(f) \leq \min(n, m).$$

Démonstration. On a $\operatorname{rg}(f) = \dim \operatorname{Im} f \leq \dim F = m$ car $\operatorname{Im} f \subseteq F$, et $\operatorname{rg}(f) = n - \dim \operatorname{Ker} f \leq n$ car $\dim \operatorname{Ker} f \geq 0$. □

3.5 Injectivité, surjectivité, isomorphismes

Proposition 18 (Caractérisation de l'injectivité)

Soit $f : E \to F$ une application linéaire. Alors

$$f \text{ est injective} \iff \operatorname{Ker} f = \{\mathbf{0}_E\}.$$

Démonstration. ($\Rightarrow$) Supposons f injective. Soit $\mathbf{u} \in \operatorname{Ker} f$, alors $f(\mathbf{u}) = \mathbf{0}_F = f(\mathbf{0}_E)$. Par injectivité, $\mathbf{u} = \mathbf{0}_E$.

($\Leftarrow$) Supposons $\operatorname{Ker} f = \{\mathbf{0}_E\}$. Soient $\mathbf{u}, \mathbf{v} \in E$ tels que $f(\mathbf{u}) = f(\mathbf{v})$. Par linéarité, $f(\mathbf{u}-\mathbf{v}) = \mathbf{0}_F$, donc $\mathbf{u} - \mathbf{v} \in \operatorname{Ker} f = \{\mathbf{0}_E\}$, d'où $\mathbf{u} = \mathbf{v}$. □

Proposition 19 (Caractérisation de la surjectivité)

Soit $f : E \to F$ une application linéaire. Alors

$$f \text{ est surjective} \iff \operatorname{Im} f = F.$$

En dimension finie, cela équivaut à $\operatorname{rg}(f) = \dim F$.

Définition 20 (Isomorphisme)

Un *isomorphisme* d'espaces vectoriels est une application linéaire bijective. Si un isomorphisme $f : E \to F$ existe, on dit que E et F sont *isomorphes* et l'on note $E \simeq F$ (ou $E \cong F$).

Théorème 21 (Bijectivité en dimension finie)

Soient E et F deux $\mathbb{K}$-espaces vectoriels de **même** dimension finie n, et soit $f : E \to F$ linéaire. Les conditions suivantes sont équivalentes :

(i) f est injective.

(ii) f est surjective.

(iii) f est bijective (isomorphisme).

Démonstration. Posons $n = \dim E = \dim F$. Par le théorème du rang (théorème 16), $n = \dim \operatorname{Ker} f + \operatorname{rg}(f)$.

(i)⇒(iii) : Si f est injective, $\dim \operatorname{Ker} f = 0$, donc $\operatorname{rg}(f) = n = \dim F$, et f est surjective, donc bijective.

(ii)⇒(iii) : Si f est surjective, $\operatorname{rg}(f) = \dim F = n$, donc $\dim \operatorname{Ker} f = 0$, et f est injective, donc bijective.

(iii)⇒(i) et (iii)⇒(ii) sont triviales. □

Corollaire 22 (Isomorphisme et dimension)

Deux $\mathbb{K}$-espaces vectoriels de dimension finie sont isomorphes si et seulement s'ils ont la même dimension. En particulier, tout $\mathbb{K}$-espace vectoriel de dimension n est isomorphe à $\mathbb{K}^n$.

Démonstration. ($\Rightarrow$) Si $f : E \xrightarrow{\sim} F$ est un isomorphisme, alors $\operatorname{rg}(f) = \dim F$ et $\dim \operatorname{Ker} f = 0$, d'où $\dim E = \dim F$.

($\Leftarrow$) Si $\dim E = \dim F = n$, choisissons une base $(\mathbf{e}_1, \dots, \mathbf{e}_n)$ de E et une base $(\mathbf{f}_1, \dots, \mathbf{f}_n)$ de F. L'application f définie par $f(\mathbf{e}_i) = \mathbf{f}_i$ pour tout i est un isomorphisme (exercice : le vérifier). □

Proposition 23 (L'inverse d'un isomorphisme est linéaire)

Si $f : E \to F$ est un isomorphisme, alors $f^{-1} : F \to E$ est aussi une application linéaire (et donc un isomorphisme).

Démonstration. Soient $\mathbf{w}_1, \mathbf{w}_2 \in F$ et $\lambda \in \mathbb{K}$. Posons $\mathbf{u}_i = f^{-1}(\mathbf{w}_i)$, de sorte que $f(\mathbf{u}_i) = \mathbf{w}_i$. Alors

$$f(\mathbf{u}_1 + \lambda \mathbf{u}_2) = f(\mathbf{u}_1) + \lambda\, f(\mathbf{u}_2) = \mathbf{w}_1 + \lambda \mathbf{w}_2.$$

Donc $f^{-1}(\mathbf{w}_1 + \lambda \mathbf{w}_2) = \mathbf{u}_1 + \lambda \mathbf{u}_2 = f^{-1}(\mathbf{w}_1) + \lambda\, f^{-1}(\mathbf{w}_2)$. □

Définition 24 (Automorphisme)

Un *automorphisme* de E est un isomorphisme de E dans lui-même. L'ensemble des automorphismes de E est noté $\mathrm{Aut}(E)$ ou $\mathrm{GL}(E)$.

3.6 Composition d'applications linéaires

Proposition 25 (La composée d'applications linéaires est linéaire)

Soient $f : E \to F$ et $g : F \to G$ deux applications linéaires. Alors $g \circ f : E \to G$ est linéaire.

Démonstration. Soient $\mathbf{u}, \mathbf{v} \in E$ et $\lambda \in \mathbb{K}$. Alors

$$\begin{aligned}(g \circ f)(\mathbf{u} + \lambda \mathbf{v}) &= g(f(\mathbf{u} + \lambda \mathbf{v})) = g(f(\mathbf{u}) + \lambda\, f(\mathbf{v})) \\ &= g(f(\mathbf{u})) + \lambda\, g(f(\mathbf{v})) = (g \circ f)(\mathbf{u}) + \lambda\, (g \circ f)(\mathbf{v}).\end{aligned}$$

□

Proposition 26 (Propriétés de la composition)

Soient $f : E \to F$, $g : F \to G$ et $h : G \to H$ des applications linéaires.

(i) **Associativité :** $h \circ (g \circ f) = (h \circ g) \circ f$.

(ii) **Élément neutre :** $f \circ \mathrm{Id}_E = f$ et $\mathrm{Id}_F \circ f = f$.

(iii) **Inégalité de Sylvester pour le rang :** $\mathrm{rg}(g \circ f) \leq \min(\mathrm{rg}(f), \mathrm{rg}(g))$.

Démonstration. Les points (i) et (ii) sont immédiats. Pour (iii), on a $\mathrm{Im}(g \circ f) = g(f(E)) \subseteq g(F) = \mathrm{Im}\, g$, donc $\mathrm{rg}(g \circ f) \leq \mathrm{rg}(g)$. Par ailleurs, $\mathrm{Im}(g \circ f) = g(\mathrm{Im}\, f)$ et la restriction de g à $\mathrm{Im}\, f$ a un rang au plus $\dim(\mathrm{Im}\, f) = \mathrm{rg}(f)$, donc $\mathrm{rg}(g \circ f) \leq \mathrm{rg}(f)$. □

3.7 L'espace vectoriel $\mathcal{L}(E,F)$

Théorème 27 ($\mathcal{L}(E,F)$ est un espace vectoriel)

Soient E et F deux $\mathbb{K}$-espaces vectoriels. L'ensemble $\mathcal{L}(E,F)$ des applications linéaires de E dans F, muni des opérations :

$$(f+g)(\mathbf{u}) := f(\mathbf{u}) + g(\mathbf{u}),$$
$$(\lambda f)(\mathbf{u}) := \lambda\, f(\mathbf{u}),$$

est un $\mathbb{K}$-espace vectoriel. L'élément nul est l'application nulle $\mathbf{u} \mapsto \mathbf{0}_F$.

Démonstration. Il suffit de vérifier que $f+g$ et λf sont linéaires lorsque $f, g \in \mathcal{L}(E,F)$ et $\lambda \in \mathbb{K}$, puis que les axiomes d'espace vectoriel sont satisfaits. La vérification est directe à partir de la structure d'espace vectoriel de F. □

Proposition 28 (Dimension de $\mathcal{L}(E,F)$)

Si $\dim E = n$ et $\dim F = m$, alors

$$\dim \mathcal{L}(E,F) = nm.$$

Démonstration. Choisissons des bases $(\mathbf{e}_1, \dots, \mathbf{e}_n)$ de E et $(\mathbf{f}_1, \dots, \mathbf{f}_m)$ de F. Pour $1 \le i \le n$ et $1 \le j \le m$, définissons $f_{ij} \in \mathcal{L}(E,F)$ par $f_{ij}(\mathbf{e}_k) = \delta_{ik}\mathbf{f}_j$ (où δ_{ik} est le symbole de Kronecker). La famille $(f_{ij})_{1\le i\le n,\, 1\le j\le m}$ forme une base de $\mathcal{L}(E,F)$, donc $\dim \mathcal{L}(E,F) = nm$. □

Remarque 29 (Structure de $\mathrm{End}(E)$)

Lorsque $E = F$, l'ensemble $\mathrm{End}(E) = \mathcal{L}(E,E)$ est non seulement un espace vectoriel, mais aussi un *anneau* pour la composition : la composition joue le rôle de la multiplication, Id_E est l'élément unité, et la distributivité vient de la linéarité. C'est un anneau non commutatif en général (dès que $\dim E \ge 2$).

3.8 Matrice d'une application linéaire

Définition 30 (Matrice associée à une application linéaire)

Soient E et F deux $\mathbb{K}$-espaces vectoriels de dimensions finies n et m respectivement, munis de bases $\mathcal{B} = (\mathbf{e}_1, \dots, \mathbf{e}_n)$ et $\mathcal{C} = (\mathbf{f}_1, \dots, \mathbf{f}_m)$. Soit $f \in \mathcal{L}(E,F)$.
Pour chaque $j \in \{1, \dots, n\}$, on décompose $f(\mathbf{e}_j)$ dans la base $\mathcal{C}$:

$$f(\mathbf{e}_j) = \sum_{i=1}^{m} a_{ij}\, \mathbf{f}_i.$$

La *matrice de f relativement aux bases $\mathcal{B}$ et $\mathcal{C}$* est la matrice $A = (a_{ij}) \in \mathcal{M}_{m,n}(\mathbb{K})$ dont la j-ème colonne contient les coordonnées de $f(\mathbf{e}_j)$ dans la base $\mathcal{C}$. On la note

$$\mathcal{M}_{\mathcal{B},\mathcal{C}}(f) := A.$$

Lorsque $E = F$ et $\mathcal{B} = \mathcal{C}$, on écrit simplement $\mathcal{M}_{\mathcal{B}}(f)$.

Proposition 31 (Lien entre application linéaire et produit matriciel)

Avec les notations précédentes, si $\mathbf{u} = \sum_{j=1}^{n} x_j \, \mathbf{e}_j$ a pour vecteur de coordonnées $X = (x_1, \dots, x_n)^{\mathsf{T}} \in \mathbb{K}^n$ dans la base $\mathcal{B}$, alors $f(\mathbf{u})$ a pour vecteur de coordonnées $Y = AX \in \mathbb{K}^m$ dans la base $\mathcal{C}$, où $A = \mathcal{M}_{\mathcal{B},\mathcal{C}}(f)$.

Démonstration. On a $f(\mathbf{u}) = f\left(\sum_{j=1}^{n} x_j \, \mathbf{e}_j\right) = \sum_{j=1}^{n} x_j \, f(\mathbf{e}_j) = \sum_{j=1}^{n} x_j \sum_{i=1}^{m} a_{ij} \, \mathbf{f}_i = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij} x_j\right) \mathbf{f}_i$. La i-ème coordonnée de $f(\mathbf{u})$ dans $\mathcal{C}$ est donc $y_i = \sum_{j=1}^{n} a_{ij} x_j$, ce qui est exactement la i-ème composante du produit AX. □

Exemple 32 (Matrice de la rotation)

La rotation R_θ de $\mathbb{R}^2$ dans la base canonique $\mathcal{E} = (\mathbf{e}_1, \mathbf{e}_2)$ a pour matrice :

$$\mathcal{M}_{\mathcal{E}}(R_\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

En effet, $R_\theta(\mathbf{e}_1) = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}$ et $R_\theta(\mathbf{e}_2) = \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix}$.

Exemple 33 (Matrice de la projection)

La projection orthogonale p sur l'axe des abscisses dans $\mathbb{R}^2$ a pour matrice dans la base canonique :

$$\mathcal{M}_{\mathcal{E}}(p) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

Exemple 34 (Matrice de la dérivation)

Soit $D \colon \mathbb{R}_3[X] \to \mathbb{R}_3[X]$ la dérivation, dans la base $\mathcal{B} = (1, X, X^2, X^3)$. On a $D(1) = 0$, $D(X) = 1$, $D(X^2) = 2X$, $D(X^3) = 3X^2$. Donc :

$$\mathcal{M}_{\mathcal{B}}(D) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Théorème 35 (Isomorphisme $\mathcal{L}(E,F) \simeq \mathcal{M}_{m,n}(\mathbb{K})$)

L'application

$$\Phi \colon \mathcal{L}(E,F) \to \mathcal{M}_{m,n}(\mathbb{K}), \qquad f \mapsto \mathcal{M}_{\mathcal{B},\mathcal{C}}(f)$$

est un isomorphisme d'espaces vectoriels. De plus :

(i) $\mathcal{M}_{\mathcal{B},\mathcal{C}}(f+g) = \mathcal{M}_{\mathcal{B},\mathcal{C}}(f) + \mathcal{M}_{\mathcal{B},\mathcal{C}}(g)$.

(ii) $\mathcal{M}_{\mathcal{B},\mathcal{C}}(\lambda f) = \lambda \, \mathcal{M}_{\mathcal{B},\mathcal{C}}(f)$.

(iii) Si $g \colon F \to G$ et $\mathcal{D}$ est une base de G, alors $\mathcal{M}_{\mathcal{B},\mathcal{D}}(g \circ f) = \mathcal{M}_{\mathcal{C},\mathcal{D}}(g) \cdot \mathcal{M}_{\mathcal{B},\mathcal{C}}(f)$.

Démonstration. La linéarité de Φ découle des points (i) et (ii), qui se vérifient par identification des coefficients. L'injectivité vient du fait que si $\mathcal{M}_{\mathcal{B},\mathcal{C}}(f) = 0$, alors $f(\mathbf{e}_j) = \mathbf{0}$ pour tout j, donc

$f = 0$. La surjectivité vient du fait que toute matrice A définit une application linéaire f par $f(\mathbf{e}_j) = \sum_i a_{ij}\mathbf{f}_i$.

Pour (iii), les colonnes de $\mathcal{M}_{\mathcal{B},\mathcal{D}}(g\circ f)$ sont les coordonnées de $(g\circ f)(\mathbf{e}_j) = g(f(\mathbf{e}_j))$ dans $\mathcal{D}$. Si $\mathcal{M}_{\mathcal{B},\mathcal{C}}(f) = A$ et $\mathcal{M}_{\mathcal{C},\mathcal{D}}(g) = B$, alors $f(\mathbf{e}_j) = \sum_k a_{kj}\mathbf{f}_k$ et $g(f(\mathbf{e}_j)) = \sum_k a_{kj} g(\mathbf{f}_k) = \sum_k a_{kj}\sum_i b_{ik}\mathbf{d}_i = \sum_i \left(\sum_k b_{ik}a_{kj}\right)\mathbf{d}_i$, ce qui donne la matrice BA. □

3.9 Changement de base

Définition 36 (Matrice de passage)

Soient $\mathcal{B} = (\mathbf{e}_1, \dots, \mathbf{e}_n)$ et $\mathcal{B}' = (\mathbf{e}'_1, \dots, \mathbf{e}'_n)$ deux bases de E. La *matrice de passage* de $\mathcal{B}$ à $\mathcal{B}'$ est la matrice

$$P_{\mathcal{B}\to\mathcal{B}'} := \mathcal{M}_{\mathcal{B}}(\mathrm{Id}_E)_{\mathcal{B},\mathcal{B}'}$$

dont la j-ème colonne contient les coordonnées de $\mathbf{e}'_j$ dans la base $\mathcal{B}$. Autrement dit, si $\mathbf{e}'_j = \sum_{i=1}^n p_{ij}\,\mathbf{e}_i$, alors $P = (p_{ij})$.

Proposition 37 (Propriétés des matrices de passage)

(i) $P_{\mathcal{B}\to\mathcal{B}'}$ est inversible et $(P_{\mathcal{B}\to\mathcal{B}'})^{-1} = P_{\mathcal{B}'\to\mathcal{B}}$.

(ii) Si X est le vecteur de coordonnées de $\mathbf{u}$ dans $\mathcal{B}$ et X' est celui dans $\mathcal{B}'$, alors $X = P_{\mathcal{B}\to\mathcal{B}'}\,X'$, soit $X' = (P_{\mathcal{B}\to\mathcal{B}'})^{-1}\,X$.

(iii) Pour trois bases $\mathcal{B}, \mathcal{B}', \mathcal{B}''$: $P_{\mathcal{B}\to\mathcal{B}''} = P_{\mathcal{B}\to\mathcal{B}'}\cdot P_{\mathcal{B}'\to\mathcal{B}''}$.

Démonstration. *(i)* La matrice $P_{\mathcal{B}\to\mathcal{B}'}$ est la matrice de l'identité Id_E de la base $\mathcal{B}'$ vers la base $\mathcal{B}$. Comme Id_E est un isomorphisme, sa matrice est inversible. De plus, $P_{\mathcal{B}\to\mathcal{B}'}\cdot P_{\mathcal{B}'\to\mathcal{B}}$ est la matrice de $\mathrm{Id}_E\circ\mathrm{Id}_E = \mathrm{Id}_E$ dans la base $\mathcal{B}$, qui est la matrice identité I_n.

(ii) Soit $\mathbf{u} = \sum_i x_i\mathbf{e}_i = \sum_j x'_j\mathbf{e}'_j$. En remplaçant $\mathbf{e}'_j = \sum_i p_{ij}\mathbf{e}_i$, on obtient $\mathbf{u} = \sum_i \left(\sum_j p_{ij}x'_j\right)\mathbf{e}_i$, d'où $x_i = \sum_j p_{ij}x'_j$, c'est-à-dire $X = PX'$.

(iii) Résulte de la multiplicativité du passage : $X = P_{\mathcal{B}\to\mathcal{B}'}\,X'$ et $X' = P_{\mathcal{B}'\to\mathcal{B}''}\,X''$, d'où $X = P_{\mathcal{B}\to\mathcal{B}'}\cdot P_{\mathcal{B}'\to\mathcal{B}''}\,X''$. □

Théorème 38 (Formule de changement de base)

Soit $f \in \mathrm{End}(E)$. Soient $\mathcal{B}$ et $\mathcal{B}'$ deux bases de E, et soit $P = P_{\mathcal{B}\to\mathcal{B}'}$ la matrice de passage. Si $A = \mathcal{M}_{\mathcal{B}}(f)$ est la matrice de f dans $\mathcal{B}$, alors la matrice de f dans $\mathcal{B}'$ est

$$\boxed{A' = \mathcal{M}_{\mathcal{B}'}(f) = P^{-1}\,A\,P.}$$

Démonstration. La matrice de f dans $\mathcal{B}'$ est $\mathcal{M}_{\mathcal{B}'}(f) = \mathcal{M}_{\mathcal{B}',\mathcal{B}'}(\mathrm{Id}_E\circ f\circ\mathrm{Id}_E)$. Or Id_E de $\mathcal{B}'$ vers $\mathcal{B}$ a pour matrice P et de $\mathcal{B}$ vers $\mathcal{B}'$ a pour matrice P^{-1}. Par la formule de composition des matrices (théorème 35, (iii)) :

$$\mathcal{M}_{\mathcal{B}'}(f) = P^{-1}\cdot\mathcal{M}_{\mathcal{B}}(f)\cdot P = P^{-1}AP.$$

□

Définition 39 (Matrices semblables)

Deux matrices $A, B \in \mathcal{M}_n(\mathbb{K})$ sont dites *semblables* s'il existe une matrice inversible $P \in \mathrm{GL}_n(\mathbb{K})$ telle que $B = P^{-1}AP$. Les matrices semblables représentent le même endo-

morphisme dans des bases différentes.

Remarque 40 (Cas général : deux espaces différents)

Si $f : E \to F$ avec $\mathcal{B}, \mathcal{B}'$ bases de E et $\mathcal{C}, \mathcal{C}'$ bases de F, et si $P = P_{\mathcal{B}\to\mathcal{B}'}$, $Q = P_{\mathcal{C}\to\mathcal{C}'}$, alors :

$$\mathcal{M}_{\mathcal{B}',\mathcal{C}'}(f) = Q^{-1} \cdot \mathcal{M}_{\mathcal{B},\mathcal{C}}(f) \cdot P.$$

3.10 Opérations sur les matrices

Dans cette section, on rappelle les opérations algébriques fondamentales sur les matrices et leurs propriétés.

3.10.1 Addition et multiplication scalaire

Définition 41 (Addition et multiplication scalaire)

Soient $A = (a_{ij}), B = (b_{ij}) \in \mathcal{M}_{m,n}(\mathbb{K})$ et $\lambda \in \mathbb{K}$.

(i) **Addition :** $A + B := (a_{ij} + b_{ij})$.

(ii) **Multiplication scalaire :** $\lambda A := (\lambda\, a_{ij})$.

L'ensemble $\mathcal{M}_{m,n}(\mathbb{K})$ muni de ces opérations est un $\mathbb{K}$-espace vectoriel de dimension mn.

3.10.2 Produit matriciel

Définition 42 (Produit de deux matrices)

Soient $A = (a_{ij}) \in \mathcal{M}_{m,p}(\mathbb{K})$ et $B = (b_{jk}) \in \mathcal{M}_{p,n}(\mathbb{K})$. Le *produit* AB est la matrice $C = (c_{ik}) \in \mathcal{M}_{m,n}(\mathbb{K})$ définie par

$$c_{ik} = \sum_{j=1}^{p} a_{ij} b_{jk}, \qquad 1 \le i \le m,\ 1 \le k \le n.$$

Proposition 43 (Propriétés du produit matriciel)

Pour des matrices de tailles compatibles :

(i) **Associativité :** $(AB)C = A(BC)$.

(ii) **Distributivité :** $A(B + C) = AB + AC$ et $(A + B)C = AC + BC$.

(iii) **Élément neutre :** $I_n A = A$ et $AI_m = A$ (pour $A \in \mathcal{M}_{n,m}(\mathbb{K})$), où I_k est la matrice identité.

(iv) **Compatibilité scalaire :** $\lambda(AB) = (\lambda A)B = A(\lambda B)$.

(v) **Non-commutativité en général :** $AB \neq BA$ en général.

Exemple 44 (Non-commutativité)

Soient $A = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$ et $B = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}$. Alors

$$AB = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \qquad BA = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}.$$

On a $AB \neq BA$.

3.10.3 Transposition

Définition 45 (Transposée)

Soit $A = (a_{ij}) \in \mathcal{M}_{m,n}(\mathbb{K})$. La *transposée* de A est la matrice $A^\mathsf{T} = (a_{ji}) \in \mathcal{M}_{n,m}(\mathbb{K})$ obtenue en échangeant lignes et colonnes.

Proposition 46 (Propriétés de la transposée)

Pour A, B de tailles compatibles et $\lambda \in \mathbb{K}$:

(i) $(A+B)^\mathsf{T} = A^\mathsf{T} + B^\mathsf{T}$.

(ii) $(\lambda A)^\mathsf{T} = \lambda\, A^\mathsf{T}$.

(iii) $(A^\mathsf{T})^\mathsf{T} = A$.

(iv) $(AB)^\mathsf{T} = B^\mathsf{T} \cdot A^\mathsf{T}$ (inversion de l'ordre).

Démonstration. Les points (i)–(iii) sont immédiats par définition. Pour (iv), le coefficient (i,j) de $(AB)^\mathsf{T}$ est $(AB)_{ji} = \sum_k a_{jk} b_{ki} = \sum_k (B^\mathsf{T})_{ik} (A^\mathsf{T})_{kj} = (B^\mathsf{T} \cdot A^\mathsf{T})_{ij}$. □

3.11 Matrices inversibles

Définition 47 (Matrice inversible)

Une matrice $A \in \mathcal{M}_n(\mathbb{K})$ est dite *inversible* s'il existe une matrice $B \in \mathcal{M}_n(\mathbb{K})$ telle que $AB = BA = I_n$. On note alors $B = A^{-1}$. L'ensemble des matrices inversibles d'ordre n est noté $\mathrm{GL}_n(\mathbb{K})$.

Proposition 48 (Caractérisations de l'inversibilité)

Pour $A \in \mathcal{M}_n(\mathbb{K})$, les conditions suivantes sont équivalentes :

(i) A est inversible.

(ii) L'application $\mathbf{x} \mapsto A\mathbf{x}$ de $\mathbb{K}^n$ dans $\mathbb{K}^n$ est bijective.

(iii) Les colonnes de A forment une base de $\mathbb{K}^n$.

(iv) $\mathrm{rg}(A) = n$.

(v) $\mathrm{Ker}(A) = \{\mathbf{0}\}$ (où l'on identifie A à l'application linéaire correspondante).

Proposition 49 (Propriétés de l'inverse)

Soient $A, B \in \mathrm{GL}_n(\mathbb{K})$ et $\lambda \in \mathbb{K}^*$.

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

(ii) $(A^{-1})^{-1} = A$.

(iii) $(\lambda A)^{-1} = \lambda^{-1} A^{-1}$.

(iv) $(A^{\mathsf{T}})^{-1} = (A^{-1})^{\mathsf{T}}$.

Démonstration. *(i)* $(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = A\,I_n\,A^{-1} = AA^{-1} = I_n$. De même pour le produit dans l'autre sens.
(ii) Résulte de $A^{-1} \cdot A = I_n$ et $A \cdot A^{-1} = I_n$.
(iii) $(\lambda A)(\lambda^{-1}A^{-1}) = (\lambda \cdot \lambda^{-1})(AA^{-1}) = I_n$.
(iv) $A^{\mathsf{T}} \cdot (A^{-1})^{\mathsf{T}} = (A^{-1} \cdot A)^{\mathsf{T}} = {I_n}^{\mathsf{T}} = I_n$. □

Exemple 50 (Inverse d'une matrice 2×2)

Soit $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ avec $ad - bc \neq 0$. Alors

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

La quantité $\det A = ad - bc$ est le *déterminant* de A (nous l'étudierons en détail au chapitre suivant).

Remarque 51 (Le groupe $\mathrm{GL}_n(\mathbb{K})$)

L'ensemble $\mathrm{GL}_n(\mathbb{K})$, muni de la multiplication matricielle, est un groupe (non abélien pour $n \geq 2$). C'est le *groupe linéaire général*.

3.12 Espaces duaux (introduction)

Définition 52 (Espace dual)

Soit E un $\mathbb{K}$-espace vectoriel. L'*espace dual* de E est l'espace

$$E^* := \mathcal{L}(E, \mathbb{K}).$$

Les éléments de E^* sont appelés *formes linéaires* sur E.

Exemple 53 (Formes linéaires sur $\mathbb{R}^n$)

Toute forme linéaire sur $\mathbb{R}^n$ est de la forme $\varphi(\mathbf{x}) = a_1x_1 + a_2x_2 + \cdots + a_nx_n$ pour des scalaires $a_1, \ldots, a_n \in \mathbb{R}$ fixés. On peut l'écrire $\varphi(\mathbf{x}) = \mathbf{a}^{\mathsf{T}}\,\mathbf{x}$ où $\mathbf{a} = (a_1, \ldots, a_n)^{\mathsf{T}}$.

Définition 54 (Base duale)

Soit $\mathcal{B} = (\mathbf{e}_1, \dots, \mathbf{e}_n)$ une base de E (de dimension finie). La *base duale* de $\mathcal{B}$ est la famille $\mathcal{B}^* = (\mathbf{e}_1^*, \dots, \mathbf{e}_n^*)$ de formes linéaires définies par

$$\mathbf{e}_i^*(\mathbf{e}_j) = \delta_{ij} = \begin{cases} 1 & \text{si } i = j, \\ 0 & \text{si } i \neq j. \end{cases}$$

Proposition 55 (La base duale est une base de E^*)

La famille $\mathcal{B}^* = (\mathbf{e}_1^*, \dots, \mathbf{e}_n^*)$ est une base de E^*. En particulier, $\dim E^* = \dim E$, et donc $E^* \simeq E$.
Toute forme linéaire $\varphi \in E^*$ se décompose de manière unique :

$$\varphi = \sum_{i=1}^{n} \varphi(\mathbf{e}_i)\, \mathbf{e}_i^*.$$

Démonstration. Liberté. Supposons $\sum_{i=1}^n \lambda_i\, \mathbf{e}_i^* = 0$ (la forme nulle). En évaluant en $\mathbf{e}_j$: $0 = \sum_{i=1}^n \lambda_i\, \mathbf{e}_i^*(\mathbf{e}_j) = \sum_{i=1}^n \lambda_i \delta_{ij} = \lambda_j$. Donc $\lambda_j = 0$ pour tout j.

Famille génératrice. Soit $\varphi \in E^*$ et posons $\psi = \sum_{i=1}^n \varphi(\mathbf{e}_i)\, \mathbf{e}_i^*$. Pour tout j : $\psi(\mathbf{e}_j) = \sum_{i=1}^n \varphi(\mathbf{e}_i)\, \delta_{ij} = \varphi(\mathbf{e}_j)$. Comme φ et ψ coïncident sur une base de E, elles sont égales. □

Remarque 56 (Bidual)

Le *bidual* $E^{**} = (E^*)^*$ contient les formes linéaires sur les formes linéaires. En dimension finie, l'application canonique $\iota : E \to E^{**}$, définie par $\iota(\mathbf{u})(\varphi) = \varphi(\mathbf{u})$, est un isomorphisme naturel (qui ne dépend pas du choix d'une base). On identifie souvent E à E^{**} par cet isomorphisme.

3.13 Applications

3.13.1 Transformations en infographie

En infographie 2D et 3D, les transformations géométriques (translations, rotations, homothéties, projections) sont réalisées par multiplication matricielle. Pour pouvoir exprimer les translations (qui ne sont pas linéaires) sous forme matricielle, on utilise les *coordonnées homogènes* : un point (x, y) du plan est représenté par le vecteur $(x, y, 1)^\mathsf{T} \in \mathbb{R}^3$.

Exemple 57 (Transformations 2D en coordonnées homogènes)

(a) **Translation** de vecteur (t_x, t_y) : $T = \begin{pmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{pmatrix}$.

(b) **Rotation** d'angle θ : $R = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

(c) **Homothétie** de facteurs (s_x, s_y) : $S = \begin{pmatrix} s_x & 0 & 0 \\ 0 & s_y & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

La composition de ces transformations revient à multiplier les matrices. Dans un moteur graphique, l'ensemble des transformations appliquées à un objet est stocké dans une seule matrice.

3.13.2 Aperçu des chaînes de Markov

Exemple 58 (Chaîne de Markov simple)

Un système peut se trouver dans deux états S_1 ou S_2. À chaque étape, la probabilité de transition de S_i à S_j est donnée par la *matrice de transition* :

$$P = \begin{pmatrix} 0{,}7 & 0{,}4 \\ 0{,}3 & 0{,}6 \end{pmatrix},$$

où p_{ij} est la probabilité de passer de l'état j à l'état i. Chaque colonne somme à 1 (matrice *stochastique*). Si $\mathbf{x}_k$ est le vecteur de probabilités d'état à l'étape k, alors $\mathbf{x}_{k+1} = P\mathbf{x}_k$. Les puissances P^k décrivent l'évolution à long terme. La recherche d'un *état stationnaire* $\mathbf{x}$ tel que $P\mathbf{x} = \mathbf{x}$ revient à chercher un vecteur propre associé à la valeur propre 1 — sujet que nous approfondirons dans un chapitre ultérieur.

3.14 Exercices

Exercice 59 (Vérification de linéarité)

Pour chacune des applications suivantes, dire si elle est linéaire. Justifier.

(a) $f : \mathbb{R}^2 \to \mathbb{R}$, $(x, y) \mapsto 2x - 3y$.

(b) $g : \mathbb{R}^2 \to \mathbb{R}$, $(x, y) \mapsto x^2 + y$.

(c) $h : \mathbb{R}^2 \to \mathbb{R}^2$, $(x, y) \mapsto (x + 1,\ y)$.

(d) $\varphi : \mathbb{R}^3 \to \mathbb{R}^2$, $(x, y, z) \mapsto (x - z,\ 2y + z)$.

(e) $\psi : \mathbb{R}_2[X] \to \mathbb{R}_2[X]$, $P \mapsto P - P(0)$.

Exercice 60 (Noyau et image)

Soit $f : \mathbb{R}^3 \to \mathbb{R}^2$ définie par $f(x, y, z) = (x + y - z,\ 2x + y + z)$.

(a) Montrer que f est linéaire.

(b) Déterminer $\operatorname{Ker} f$ (en donner une base).

(c) Déterminer $\operatorname{Im} f$.

(d) Vérifier le théorème du rang.

Exercice 61 (Matrice d'une application linéaire)

Soit $f\colon \mathbb{R}^3 \to \mathbb{R}^2$ définie par $f(x,y,z) = (x+2y,\ 3y-z)$.

(a) Écrire la matrice de f dans les bases canoniques.

(b) Calculer $f(1,-1,2)$ de deux manières : directement et par produit matriciel.

(c) Déterminer le rang de f.

Exercice 62 (Projections)

Soit $p\colon \mathbb{R}^3 \to \mathbb{R}^3$ définie par $p(x,y,z) = (x,y,0)$.

(a) Montrer que p est linéaire.

(b) Vérifier que $p \circ p = p$.

(c) Déterminer $\operatorname{Ker} p$ et $\operatorname{Im} p$.

(d) Montrer que $\mathbb{R}^3 = \operatorname{Ker} p \oplus \operatorname{Im} p$.

Exercice 63 (Injectivité et surjectivité)

Soit $f\colon \mathbb{R}^3 \to \mathbb{R}^3$ l'endomorphisme défini par $f(x,y,z) = (x+y,\ y+z,\ x+z)$.

(a) Écrire la matrice de f dans la base canonique.

(b) f est-elle injective ? Surjective ? Bijective ?

(c) Si f est bijective, trouver f^{-1}.

Exercice 64 (Dérivation et intégration)

Soit $E = \mathbb{R}_3[X]$ muni de la base $\mathcal{B} = (1, X, X^2, X^3)$.

(a) Écrire la matrice de $D\colon P \mapsto P'$ dans $\mathcal{B}$.

(b) Écrire la matrice de $I\colon P \mapsto \int_0^X P(t)\,dt$ dans $\mathcal{B}$ (l'image est dans $\mathbb{R}_4[X]$; prendre la base $\mathcal{C} = (1, X, X^2, X^3, X^4)$).

(c) Calculer le produit DI (comme endomorphisme de E en restreignant au degré ≤ 3). Que constate-t-on ?

Exercice 65 (Changement de base)

Soit $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ l'endomorphisme de matrice $A = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$ dans la base canonique $\mathcal{E}$. Soit $\mathcal{B}' = \left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right)$.

(a) Écrire la matrice de passage P de $\mathcal{E}$ à $\mathcal{B}'$.

(b) Calculer P^{-1}.

(c) Calculer $A' = P^{-1}AP$. Que remarque-t-on ?

(d) Interpréter géométriquement l'action de f.

Exercice 66 (Espace dual)

Soit $E = \mathbb{R}^3$ muni de la base canonique $\mathcal{E} = (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$.

(a) Écrire la base duale $\mathcal{E}^* = (\mathbf{e}_1^*, \mathbf{e}_2^*, \mathbf{e}_3^*)$ explicitement (comme formes linéaires).

(b) Soit $\varphi : \mathbb{R}^3 \to \mathbb{R}$, $(x, y, z) \mapsto 2x - y + 3z$. Exprimer φ dans la base duale.

(c) Soit $\mathcal{B} = ((1,1,0), (1,0,1), (0,1,1))$. Calculer la base duale $\mathcal{B}^*$.

Exercice 67 (Théorème du rang — applications)

(a) Soit $f \in \mathrm{End}(E)$ avec $\dim E = n$. Montrer que $\mathrm{Ker}\, f \cap \mathrm{Im}\, f = \{\mathbf{0}\}$ si et seulement si $\mathrm{Ker}\, f \oplus \mathrm{Im}\, f = E$.

(b) Soit $f \in \mathrm{End}(E)$ tel que $f \circ f = f$ (projecteur). Montrer que $E = \mathrm{Ker}\, f \oplus \mathrm{Im}\, f$.

(c) Soit $f \in \mathrm{End}(E)$ tel que $\mathrm{rg}(f) = 1$. Montrer qu'il existe $\lambda \in \mathbb{K}$ tel que $f^2 = \lambda f$.
Indication : choisir une base adaptée au noyau et à l'image.

Exercice 68 (Matrices et applications linéaires)

Soient $A \in \mathcal{M}_{m,n}(\mathbb{K})$ et $B \in \mathcal{M}_{n,p}(\mathbb{K})$.

(a) Montrer que $\mathrm{rg}(AB) \leq \min(\mathrm{rg}(A), \mathrm{rg}(B))$.

(b) Montrer que $\mathrm{rg}(A) + \mathrm{rg}(B) - n \leq \mathrm{rg}(AB)$ (*inégalité de Sylvester*).
Indication : appliquer le théorème du rang à la restriction de f_A à $\mathrm{Im}(f_B)$.

(c) En déduire que si $A \in \mathcal{M}_n(\mathbb{K})$ avec $\mathrm{rg}(A) = n - 1$, alors $\mathrm{rg}(A^2) \geq n - 2$.

Exercice 69 (Endomorphismes nilpotents)

Un endomorphisme $f \in \mathrm{End}(E)$ est dit *nilpotent* s'il existe $k \in \mathbb{N}^*$ tel que $f^k = 0$ (l'application nulle). Le plus petit tel k est l'*indice de nilpotence*.

(a) Montrer que si f est nilpotent d'indice k, alors $k \leq \dim E$.
Indication : considérer la suite $\{\mathbf{0}\} \subseteq \mathrm{Ker}\, f \subseteq \mathrm{Ker}\, f^2 \subseteq \cdots$

(b) Montrer que si f est nilpotent, alors $\mathrm{Id} + f$ est un automorphisme.
Indication : écrire l'inverse comme une somme finie $\sum_{i=0}^{k-1} (-1)^i f^i$.

(c) Soit $E = \mathbb{R}^3$ et f défini par $f(x, y, z) = (y, z, 0)$. Vérifier que f est nilpotent, déterminer son indice, et calculer $(\mathrm{Id} + f)^{-1}$.

Résumé du chapitre

Ce qu'il faut retenir.

1. **Application linéaire.** $f : E \to F$ est linéaire si $f(\lambda \mathbf{u} + \mu \mathbf{v}) = \lambda f(\mathbf{u}) + \mu f(\mathbf{v})$. Elle est entièrement déterminée par les images d'une base.
2. **Noyau et image.** $\operatorname{Ker} f$ et $\operatorname{Im} f$ sont des sous-espaces vectoriels. f est injective ssi $\operatorname{Ker} f = \{\mathbf{0}\}$.
3. **Théorème du rang.** $\dim E = \dim \operatorname{Ker} f + \operatorname{rg}(f)$. C'est l'outil principal pour déterminer injectivité et surjectivité en dimension finie.
4. **Isomorphismes.** En dimension finie égale, injectivité $\Leftrightarrow$ surjectivité $\Leftrightarrow$ bijectivité. Deux espaces de même dimension sont isomorphes.
5. **Matrice d'une application linéaire.** La j-ème colonne de $\mathcal{M}_{\mathcal{B},\mathcal{C}}(f)$ contient les coordonnées de $f(\mathbf{e}_j)$ dans $\mathcal{C}$. La composition correspond au produit matriciel.
6. **Changement de base.** Si P est la matrice de passage, $A' = P^{-1}AP$. Deux matrices sont semblables ssi elles représentent le même endomorphisme.
7. **Matrices inversibles.** $A \in \mathrm{GL}_n(\mathbb{K})$ ssi $\operatorname{rg}(A) = n$ ssi $\operatorname{Ker} A = \{\mathbf{0}\}$ ssi les colonnes de A forment une base.
8. **Espace dual.** $E^* = \mathcal{L}(E, \mathbb{K})$ est l'espace des formes linéaires. En dimension finie, $\dim E^* = \dim E$.
9. **Formules clés.**
   - $\dim \mathcal{L}(E, F) = \dim E \cdot \dim F$.
   - $\operatorname{rg}(g \circ f) \leq \min(\operatorname{rg} f, \operatorname{rg} g)$.
   - $(AB)^{-1} = B^{-1} \cdot A^{-1}$, $\quad (AB)^{\mathsf{T}} = B^{\mathsf{T}} \cdot A^{\mathsf{T}}$.

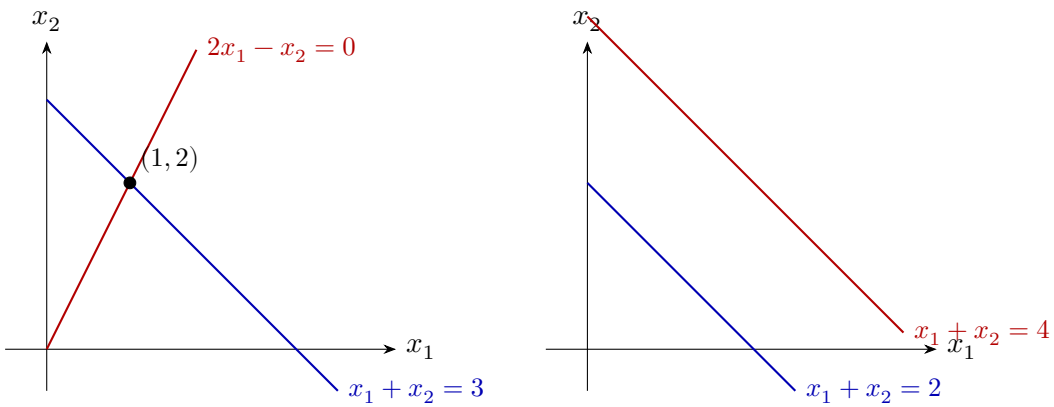
Chapitre 4

Systèmes d'équations linéaires et élimination de Gauss

Introduction

Les systèmes d'équations linéaires apparaissent dans pratiquement tous les domaines des mathématiques appliquées et des sciences. En physique, l'analyse des circuits électriques conduit à des systèmes linéaires via les lois de Kirchhoff ; en économie, les modèles d'équilibre de Leontief se ramènent à la résolution d'un système $A\mathbf{x} = \mathbf{b}$; en informatique, les moteurs de recherche, la compression d'images et l'apprentissage automatique reposent tous, in fine, sur la résolution efficace de systèmes linéaires de grande taille.

Ce chapitre introduit un outil fondamental pour l'étude de ces systèmes : l'*élimination de Gauss*, un algorithme systématique qui ramène tout système linéaire à une forme échelonnée facile à résoudre. Nous établirons le lien entre le rang d'une matrice et l'existence de solutions (théorème de Rouché–Capelli), et nous montrerons que l'ensemble des solutions d'un système linéaire possède une structure géométrique précise : c'est un sous-espace affine.



Système compatible déterminé — Système incompatible

Fig. 4.1 : Interprétation géométrique d'un système de deux équations à deux inconnues : solution unique (gauche) ou absence de solution (droite).

4.1 Systèmes d'équations linéaires

4.1.1 Définition et notation

Définition 1 (Système d'équations linéaires)

Un *système de m équations linéaires à n inconnues* sur un corps $\mathbb{K}$ est un ensemble d'équations de la forme :

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2, \\ \qquad\qquad\qquad\qquad\quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m, \end{cases} \tag{4.1}$$

où les *coefficients* $a_{ij} \in \mathbb{K}$ et les *seconds membres* $b_i \in \mathbb{K}$ sont donnés, et les x_j sont les *inconnues*.

4.1.2 Écriture matricielle

Le système (4.1) se réécrit de manière compacte sous la forme

$$A\mathbf{x} = \mathbf{b}, \tag{4.2}$$

où

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \in \mathcal{M}_{m,n}(\mathbb{K}), \qquad \mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{K}^n, \qquad \mathbf{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \in \mathbb{K}^m.$$

La matrice A est la *matrice des coefficients* du système. La *matrice augmentée* est la matrice $(A \mid \mathbf{b}) \in \mathcal{M}_{m,n+1}(\mathbb{K})$ obtenue en juxtaposant $\mathbf{b}$ comme colonne supplémentaire :

$$(A \mid \mathbf{b}) = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right). \tag{4.3}$$

Définition 2 (Solution d'un système)

Un vecteur $\mathbf{s} = (s_1, \ldots, s_n) \in \mathbb{K}^n$ est une *solution* du système (4.2) si $A\mathbf{s} = \mathbf{b}$. L'*ensemble des solutions* est

$$\mathcal{S} := \{\, \mathbf{x} \in \mathbb{K}^n \mid A\mathbf{x} = \mathbf{b} \,\}.$$

Le système est dit *compatible* (ou *consistant*) si $\mathcal{S} \neq \emptyset$, et *incompatible* sinon.

4.1.3 Systèmes homogènes et non homogènes

Définition 3 (Système homogène)

Le système (4.2) est dit *homogène* si $\mathbf{b} = \mathbf{0}$, c'est-à-dire si toutes les constantes du second membre sont nulles. Le système homogène associé à (4.2) est :

$$A\mathbf{x} = \mathbf{0}. \tag{4.4}$$

Remarque 4 (Compatibilité d'un système homogène)

Un système homogène est toujours compatible car $\mathbf{x} = \mathbf{0}$ est solution. Cette solution est dite *triviale*. La question est donc de savoir s'il existe des solutions *non triviales*.

Proposition 5 (Solutions d'un système homogène)

L'ensemble des solutions du système homogène $A\mathbf{x} = \mathbf{0}$ est un sous-espace vectoriel de $\mathbb{K}^n$. C'est exactement le noyau $\mathrm{Ker}(A)$ de l'application linéaire $\mathbf{x} \mapsto A\mathbf{x}$.

Démonstration. Posons $f : \mathbb{K}^n \to \mathbb{K}^m$, $\mathbf{x} \mapsto A\mathbf{x}$. L'application f est linéaire (cf. chapitre 3). L'ensemble des solutions de $A\mathbf{x} = \mathbf{0}$ est $\mathrm{Ker}(f)$, qui est un sous-espace vectoriel de $\mathbb{K}^n$ d'après la chapitre 3. □

4.2 Opérations élémentaires sur les lignes

L'idée clé de l'élimination de Gauss est de transformer un système en un système *équivalent* (c'est-à-dire ayant le même ensemble de solutions) mais de forme plus simple. Les transformations autorisées sont les *opérations élémentaires sur les lignes*.

Définition 6 (Opérations élémentaires sur les lignes)

Soit $M \in \mathcal{M}_{m,n}(\mathbb{K})$. On appelle *opérations élémentaires sur les lignes* les trois types de transformations suivants :

(i) **Échange de deux lignes :** $L_i \leftrightarrow L_j \quad (i \neq j)$.

(ii) **Multiplication d'une ligne par un scalaire non nul :** $L_i \leftarrow \lambda L_i \quad (\lambda \in \mathbb{K}^*)$.

(iii) **Ajout d'un multiple d'une ligne à une autre :** $L_i \leftarrow L_i + \lambda L_j \quad (\lambda \in \mathbb{K}, i \neq j)$.

Proposition 7 (Conservation de l'ensemble des solutions)

Les opérations élémentaires sur les lignes ne modifient pas l'ensemble des solutions du système linéaire. Autrement dit, si la matrice augmentée $(A' \mid \mathbf{b}')$ est obtenue à partir de $(A \mid \mathbf{b})$ par une suite d'opérations élémentaires, alors les systèmes $A\mathbf{x} = \mathbf{b}$ et $A'\mathbf{x} = \mathbf{b}'$ ont le même ensemble de solutions.

Démonstration. Chaque opération élémentaire correspond à la multiplication à gauche par une matrice inversible $E \in \mathrm{GL}_m(\mathbb{K})$ (matrice élémentaire). Ainsi $A'\mathbf{x} = \mathbf{b}'$ s'écrit $EA\mathbf{x} = E\mathbf{b}$. Comme E est inversible, $EA\mathbf{x} = E\mathbf{b}$ si et seulement si $A\mathbf{x} = \mathbf{b}$ (il suffit de multiplier à gauche par E^{-1}). Par composition, toute suite finie d'opérations élémentaires préserve l'ensemble des solutions. □

Exemple 8 (Matrices élémentaires)

Dans $\mathcal{M}_{3,3}(\mathbb{R})$, les trois types de matrices élémentaires sont :

$$E_{L_1 \leftrightarrow L_2} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad E_{L_2 \leftarrow 3L_2} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad E_{L_3 \leftarrow L_3 + 2L_1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix}.$$

Chacune est inversible ; par exemple, $E^{-1}_{L_3 \leftarrow L_3 + 2L_1} = E_{L_3 \leftarrow L_3 - 2L_1}$.

4.3 Forme échelonnée et forme échelonnée réduite

Définition 9 (Pivot)

Soit $M = (m_{ij})$ une matrice. Pour une ligne L_i non nulle, le *pivot* de L_i est le premier coefficient non nul de cette ligne, c'est-à-dire m_{ij} où $j = \min\{\, k \mid m_{ik} \neq 0 \,\}$.

Définition 10 (Forme échelonnée (en lignes))

Une matrice est sous *forme échelonnée* (ou *forme échelonnée en lignes*, *row echelon form*) si :

(i) toutes les lignes nulles sont en bas de la matrice ;

(ii) le pivot de chaque ligne non nulle est strictement à droite du pivot de la ligne précédente.

Si, de plus :

(iii) chaque pivot vaut 1 ;

(iv) chaque pivot est le seul élément non nul de sa colonne ;

la matrice est sous *forme échelonnée réduite* (*reduced row echelon form*, abrégée RREF).

Exemple 11 (Formes échelonnées)

La matrice suivante est sous forme échelonnée (les pivots sont encadrés) :

$$\begin{pmatrix} \boxed{2} & 3 & -1 & 7 \\ 0 & \boxed{1} & 4 & 2 \\ 0 & 0 & 0 & \boxed{5} \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

La forme échelonnée réduite correspondante est :

$$\begin{pmatrix} \boxed{1} & 0 & -13 & 0 \\ 0 & \boxed{1} & 4 & 0 \\ 0 & 0 & 0 & \boxed{1} \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Théorème 12 (Unicité de la forme échelonnée réduite)

Toute matrice $A \in \mathcal{M}_{m,n}(\mathbb{K})$ est équivalente par lignes à une unique matrice sous forme échelonnée réduite.

Démonstration. Existence. L'algorithme de Gauss–Jordan (section 4.5) produit une forme échelonnée réduite en un nombre fini d'opérations élémentaires.

Unicité. Soient R et R' deux formes échelonnées réduites de A, de sorte que $R = PA$ et $R' = QA$ pour des matrices inversibles P et Q. Comme R et R' sont équivalentes par lignes, il existe S inversible telle que $R' = SR$. Montrons que $R = R'$.

Les colonnes pivots de R et R' sont identiques : en effet, la colonne j de R (resp. R') est une colonne pivot si et seulement si elle n'est pas combinaison linéaire des colonnes précédentes de R (resp. R'), propriété qui ne dépend que de l'espace des colonnes, lequel est le même pour R et R' (car elles sont équivalentes par lignes et l'espace des lignes est préservé, ce qui entraîne que les relations de dépendance entre colonnes sont identiques).

Notons $j_1 < j_2 < \cdots < j_r$ les indices des colonnes pivots communs. Pour $k = 1, \dots, r$, la colonne j_k de R est le vecteur $\mathbf{e}_k$ de $\mathbb{R}^m$ (vecteur avec 1 en position k et 0 ailleurs), et de même pour R'. Pour une colonne non pivot j, les coefficients de la colonne j dans R sont déterminés de façon unique par les relations de dépendance linéaire entre les colonnes de A, qui sont les mêmes pour R et R'. Donc $R = R'$. □

4.4 Élimination de Gauss

L'*élimination de Gauss* (ou *méthode du pivot de Gauss*) est un algorithme qui transforme la matrice augmentée $(A \mid \mathbf{b})$ en une forme échelonnée par des opérations élémentaires sur les lignes.

4.4.1 Algorithme

Étape 1. Choix du pivot. On considère la colonne la plus à gauche qui contient un élément non nul sous la ligne courante (ou sur cette ligne). Si un tel élément se trouve sur une ligne L_k avec $k > i$ (où i est la ligne courante), on effectue l'échange $L_i \leftrightarrow L_k$.

Étape 2. Élimination. Pour chaque ligne L_j avec $j > i$, on effectue $L_j \leftarrow L_j - \frac{a_{ji}}{a_{ii}} L_i$, ce qui annule le coefficient en position (j, i).

Étape 3. Itération. On passe à la ligne suivante $i \leftarrow i + 1$ et on recommence à l'étape **Étape 1.** tant qu'il reste des lignes et des colonnes à traiter.

Le résultat est une matrice sous forme échelonnée. On résout alors le système par *substitution arrière* (*back substitution*).

4.4.2 Exemple détaillé

Exemple 13 (Élimination de Gauss sur un système 3×3)

Résolvons le système :

$$\begin{cases} 2x_1 + x_2 - x_3 = 8, \\ -3x_1 - x_2 + 2x_3 = -11, \\ -2x_1 + x_2 + 2x_3 = -3. \end{cases}$$

Matrice augmentée :

$$\left(\begin{array}{ccc|c} 2 & 1 & -1 & 8 \\ -3 & -1 & 2 & -11 \\ -2 & 1 & 2 & -3 \end{array}\right).$$

Étape 1 : $L_2 \leftarrow L_2 + \frac{3}{2}L_1, \quad L_3 \leftarrow L_3 + L_1$:

$$\left(\begin{array}{ccc|c} 2 & 1 & -1 & 8 \\ 0 & \frac{1}{2} & \frac{1}{2} & 1 \\ 0 & 2 & 1 & 5 \end{array}\right).$$

Étape 2 : $L_3 \leftarrow L_3 - 4L_2$:

$$\left(\begin{array}{ccc|c} 2 & 1 & -1 & 8 \\ 0 & \frac{1}{2} & \frac{1}{2} & 1 \\ 0 & 0 & -1 & 1 \end{array}\right).$$

Substitution arrière :

- L_3 : $-x_3 = 1$, d'où $x_3 = -1$.
- L_2 : $\frac{1}{2}x_2 + \frac{1}{2}(-1) = 1$, d'où $x_2 = 3$.
- L_1 : $2x_1 + 3 - (-1) = 8$, d'où $x_1 = 2$.

L'unique solution est $(x_1, x_2, x_3) = (2, 3, -1)$.

Exemple 14 (Système avec infinité de solutions)

Résolvons :

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

Matrice augmentée :

$$\left(\begin{array}{ccc|c} 1 & -2 & 1 & 0 \\ 2 & 1 & -1 & 3 \\ 4 & -3 & 1 & 3 \end{array}\right) \xrightarrow{L_2 \leftarrow L_2 - 2L_1} \left(\begin{array}{ccc|c} 1 & -2 & 1 & 0 \\ 0 & 5 & -3 & 3 \\ 4 & -3 & 1 & 3 \end{array}\right) \xrightarrow{L_3 \leftarrow L_3 - 4L_1} \left(\begin{array}{ccc|c} 1 & -2 & 1 & 0 \\ 0 & 5 & -3 & 3 \\ 0 & 5 & -3 & 3 \end{array}\right).$$

Puis $L_3 \leftarrow L_3 - L_2$:

$$\left(\begin{array}{ccc|c} 1 & -2 & 1 & 0 \\ 0 & 5 & -3 & 3 \\ 0 & 0 & 0 & 0 \end{array}\right).$$

On a deux pivots et trois inconnues : x_3 est une variable libre. En posant $x_3 = t \in \mathbb{R}$:

$$x_2 = \frac{3 + 3t}{5}, \qquad x_1 = 2 \cdot \frac{3 + 3t}{5} - t = \frac{6 + 6t - 5t}{5} = \frac{6 + t}{5}.$$

L'ensemble des solutions est :

$$\mathcal{S} = \left\{ \frac{1}{5}\begin{pmatrix} 6 \\ 3 \\ 0 \end{pmatrix} + t\,\frac{1}{5}\begin{pmatrix} 1 \\ 3 \\ 5 \end{pmatrix} \,\middle|\, t \in \mathbb{R} \right\} = \left\{ \mathbf{x}_p + t\,\mathbf{v} \mid t \in \mathbb{R} \right\},$$

qui est une droite affine de $\mathbb{R}^3$.

4.5 Élimination de Gauss–Jordan

L'*élimination de Gauss–Jordan* prolonge l'élimination de Gauss pour obtenir la forme échelonnée réduite (RREF). Après avoir obtenu une forme échelonnée par l'algorithme de Gauss, on effectue les étapes supplémentaires suivantes :

Étape 1. Normalisation des pivots. Pour chaque ligne non nulle L_i de pivot a_{ij}, on effectue $L_i \leftarrow \frac{1}{a_{ij}} L_i$ pour rendre le pivot égal à 1.

Étape 2. Élimination vers le haut. En partant du dernier pivot et en remontant, pour chaque pivot en position (i, j), on annule tous les coefficients au-dessus du pivot : $L_k \leftarrow L_k - a_{kj} L_i$ pour $k < i$.

Exemple 15 (Gauss–Jordan complet)

Reprenons la matrice échelonnée de l'exemple 13 :

$$\left(\begin{array}{ccc|c} 2 & 1 & -1 & 8 \\ 0 & \frac{1}{2} & \frac{1}{2} & 1 \\ 0 & 0 & -1 & 1 \end{array}\right).$$

Normalisation : $L_1 \leftarrow \frac{1}{2} L_1$, $L_2 \leftarrow 2L_2$, $L_3 \leftarrow -L_3$:

$$\left(\begin{array}{ccc|c} 1 & \frac{1}{2} & -\frac{1}{2} & 4 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 1 & -1 \end{array}\right).$$

Élimination vers le haut : $L_2 \leftarrow L_2 - L_3$, puis $L_1 \leftarrow L_1 + \frac{1}{2} L_3$, puis $L_1 \leftarrow L_1 - \frac{1}{2} L_2$:

$$\left(\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & -1 \end{array}\right).$$

On lit directement la solution : $(x_1, x_2, x_3) = (2, 3, -1)$.

4.6 Rang d'une matrice

Définition 16 (Rang d'une matrice)

Le *rang* d'une matrice $A \in \mathcal{M}_{m,n}(\mathbb{K})$, noté $\operatorname{rg}(A)$, est le nombre de pivots dans toute forme échelonnée de A. De manière équivalente :

(i) $\operatorname{rg}(A)$ est la dimension de l'espace vectoriel engendré par les lignes de A (rang ligne) ;

(ii) $\operatorname{rg}(A)$ est la dimension de l'espace vectoriel engendré par les colonnes de A (rang colonne) ;

(iii) $\operatorname{rg}(A) = \dim(\operatorname{Im}(f))$, où $f : \mathbb{K}^n \to \mathbb{K}^m$, $\mathbf{x} \mapsto A\mathbf{x}$.

Remarque 17 (Rang ligne = rang colonne)

L'égalité entre le rang ligne et le rang colonne est un résultat fondamental. On l'établit en remarquant que les opérations élémentaires sur les lignes ne changent pas les relations de dépendance linéaire entre les colonnes, et que dans une forme échelonnée, les colonnes pivots forment une famille libre de cardinal égal au nombre de lignes non nulles.

Proposition 18 (Propriétés du rang)

Soit $A \in \mathcal{M}_{m,n}(\mathbb{K})$. Alors :

(i) $0 \leqslant \operatorname{rg}(A) \leqslant \min(m, n)$.

(ii) $\operatorname{rg}(A) = \operatorname{rg}(A^{\mathsf{T}})$.

(iii) Si $P \in \mathrm{GL}_m(\mathbb{K})$ et $Q \in \mathrm{GL}_n(\mathbb{K})$, alors $\operatorname{rg}(PAQ) = \operatorname{rg}(A)$.

(iv) Pour A carrée d'ordre n : A est inversible si et seulement si $\operatorname{rg}(A) = n$.

Démonstration. (i) est immédiat. (ii) découle de l'égalité rang ligne = rang colonne. (iii) : la multiplication à gauche par P inversible correspond à des opérations élémentaires sur les lignes (qui ne changent pas le rang), et la multiplication à droite par Q inversible correspond à des opérations élémentaires sur les colonnes (qui ne changent pas le rang non plus). (iv) : A est inversible si et seulement si sa forme échelonnée réduite est I_n, c'est-à-dire si et seulement si elle a n pivots. □

4.7 Théorème de Rouché–Capelli

Le théorème suivant, connu sous le nom de *théorème de Rouché–Capelli* (ou *Kronecker–Capelli* dans la tradition allemande), fournit le critère fondamental d'existence de solutions.

Théorème 19 (Théorème de Rouché–Capelli)

Soit $A\mathbf{x} = \mathbf{b}$ un système linéaire avec $A \in \mathcal{M}_{m,n}(\mathbb{K})$ et $\mathbf{b} \in \mathbb{K}^m$. Alors :

(i) Le système est compatible si et seulement si $\operatorname{rg}(A) = \operatorname{rg}(A \mid \mathbf{b})$.

(ii) Lorsque le système est compatible, la dimension de l'ensemble des solutions du système homogène associé $A\mathbf{x} = \mathbf{0}$ est $n - \operatorname{rg}(A)$.

(iii) En particulier, la solution est unique si et seulement si $\operatorname{rg}(A) = n$.

Démonstration. *(i)* Le système $A\mathbf{x} = \mathbf{b}$ est compatible si et seulement si $\mathbf{b}$ appartient à l'espace des colonnes de A, c'est-à-dire si $\mathbf{b} \in \operatorname{Im}(A)$. Or $\mathbf{b} \in \operatorname{Im}(A)$ si et seulement si l'ajout de $\mathbf{b}$ comme colonne supplémentaire ne change pas la dimension de l'espace des colonnes, c'est-à-dire $\operatorname{rg}(A \mid \mathbf{b}) = \operatorname{rg}(A)$.

Inversement, si $\mathbf{b} \notin \operatorname{Im}(A)$, alors $\mathbf{b}$ est linéairement indépendant de l'espace des colonnes de A, et $\operatorname{rg}(A \mid \mathbf{b}) = \operatorname{rg}(A) + 1 > \operatorname{rg}(A)$.

(ii) Posons $r = \operatorname{rg}(A)$ et considérons l'application linéaire $f : \mathbb{K}^n \to \mathbb{K}^m$, $\mathbf{x} \mapsto A\mathbf{x}$. On a $\operatorname{Im}(f)$ de dimension r, et le théorème du rang (cf. chapitre 3) donne :

$$\dim(\operatorname{Ker}(f)) = n - \dim(\operatorname{Im}(f)) = n - r.$$

Or $\operatorname{Ker}(f)$ est exactement l'ensemble des solutions du système homogène.

(iii) La solution est unique si et seulement si l'ensemble des solutions du système homogène est réduit à $\{\mathbf{0}\}$, c'est-à-dire $\dim(\mathrm{Ker}(f)) = 0$, ce qui équivaut à $r = n$. □

Corollaire 20 (Nombre de paramètres libres)

Si le système $A\mathbf{x} = \mathbf{b}$ ($A \in \mathcal{M}_{m,n}(\mathbb{K})$) est compatible, l'ensemble des solutions est paramétré par $n - \mathrm{rg}(A)$ paramètres libres.

4.8 Structure de l'ensemble des solutions

Théorème 21 (Structure affine des solutions)

Soit $A\mathbf{x} = \mathbf{b}$ un système linéaire compatible, et soit $\mathbf{x}_p$ une solution particulière. Alors l'ensemble des solutions est le sous-espace affine :

$$\mathcal{S} = \mathbf{x}_p + \mathrm{Ker}(A) = \left\{ \mathbf{x}_p + \mathbf{h} \;\middle|\; \mathbf{h} \in \mathrm{Ker}(A) \right\}. \tag{4.5}$$

Autrement dit, toute solution $\mathbf{x}$ du système s'écrit de manière unique :

$$\mathbf{x} = \mathbf{x}_p + \mathbf{x}_h,$$

où $\mathbf{x}_p$ est une solution particulière de $A\mathbf{x} = \mathbf{b}$ et $\mathbf{x}_h$ est une solution du système homogène $A\mathbf{x} = \mathbf{0}$.

Démonstration. Montrons la double inclusion.

(⊇) Soit $\mathbf{h} \in \mathrm{Ker}(A)$. Alors $A(\mathbf{x}_p + \mathbf{h}) = A\mathbf{x}_p + A\mathbf{h} = \mathbf{b} + \mathbf{0} = \mathbf{b}$, donc $\mathbf{x}_p + \mathbf{h} \in \mathcal{S}$.

(⊆) Soit $\mathbf{x} \in \mathcal{S}$. Posons $\mathbf{h} = \mathbf{x} - \mathbf{x}_p$. Alors $A\mathbf{h} = A\mathbf{x} - A\mathbf{x}_p = \mathbf{b} - \mathbf{b} = \mathbf{0}$, donc $\mathbf{h} \in \mathrm{Ker}(A)$, et $\mathbf{x} = \mathbf{x}_p + \mathbf{h} \in \mathbf{x}_p + \mathrm{Ker}(A)$.

Unicité de la décomposition. Si $\mathbf{x} = \mathbf{x}_p + \mathbf{h}_1 = \mathbf{x}_p + \mathbf{h}_2$ avec $\mathbf{h}_1, \mathbf{h}_2 \in \mathrm{Ker}(A)$, alors $\mathbf{h}_1 = \mathbf{h}_2$. □

Remarque 22 (Interprétation géométrique)

Le théorème 21 dit que l'ensemble des solutions est une « copie translatée » du noyau de A. En dimension finie :

- Si $\mathrm{rg}(A) = n$: $\mathcal{S} = \{\mathbf{x}_p\}$ est un point.
- Si $\mathrm{rg}(A) = n - 1$: $\mathcal{S}$ est une droite affine.
- Si $\mathrm{rg}(A) = n - 2$: $\mathcal{S}$ est un plan affine.
- En général : $\mathcal{S}$ est un sous-espace affine de dimension $n - \mathrm{rg}(A)$.

4.9 Représentation paramétrique des solutions

Lorsqu'un système compatible admet des variables libres, on exprime les *variables liées* (correspondant aux colonnes pivots) en fonction des *variables libres* (correspondant aux colonnes non pivots).

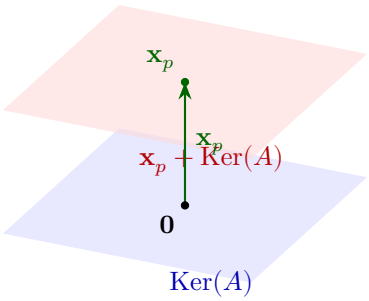


Fig. 4.2 : L'ensemble des solutions $\mathcal{S} = \mathbf{x}_p + \mathrm{Ker}(A)$ est un translaté du noyau.

Exemple 23 (Représentation paramétrique)

Considérons le système de rang 2 à 4 inconnues :

$$\left(\begin{array}{cccc|c} 1 & 2 & 0 & -1 & 3 \\ 0 & 0 & 1 & 3 & -2 \end{array}\right).$$

Les colonnes pivots sont les colonnes 1 et 3 ; les variables libres sont x_2 et x_4. En posant $x_2 = s$, $x_4 = t$:

$$\mathbf{x} = \begin{pmatrix} 3-2s+t \\ s \\ -2-3t \\ t \end{pmatrix} = \underbrace{\begin{pmatrix} 3 \\ 0 \\ -2 \\ 0 \end{pmatrix}}_{\mathbf{x}_p} + s \underbrace{\begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \end{pmatrix}}_{\mathbf{v}_1} + t \underbrace{\begin{pmatrix} 1 \\ 0 \\ -3 \\ 1 \end{pmatrix}}_{\mathbf{v}_2}, \qquad s,t \in \mathbb{K}.$$

Les vecteurs $\mathbf{v}_1$ et $\mathbf{v}_2$ forment une base de $\mathrm{Ker}(A)$, et $\mathcal{S}$ est un plan affine de $\mathbb{K}^4$.

4.10 Interprétation géométrique

4.10.1 Systèmes 2×2

Chaque équation linéaire en deux inconnues $a_1x_1 + a_2x_2 = b$ définit une droite dans $\mathbb{R}^2$. Un système de deux équations correspond à l'intersection de deux droites :

- **Un point** si les droites sont sécantes ($\mathrm{rg}(A) = 2$).
- **Une droite** si les droites sont confondues ($\mathrm{rg}(A) = 1$ et $\mathrm{rg}(A \mid \mathbf{b}) = 1$).
- **L'ensemble vide** si les droites sont parallèles distinctes ($\mathrm{rg}(A) = 1$ et $\mathrm{rg}(A \mid \mathbf{b}) = 2$).

4.10.2 Systèmes 3×3

Chaque équation linéaire en trois inconnues définit un plan dans $\mathbb{R}^3$. Un système de trois équations correspond à l'intersection de trois plans.

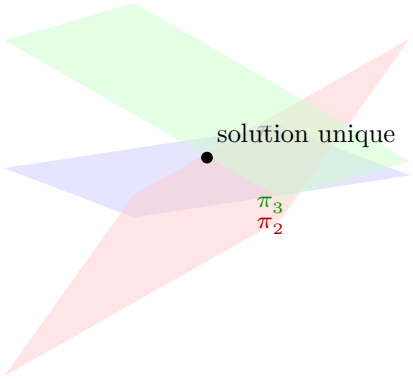


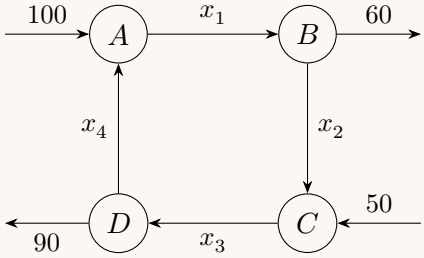
Fig. 4.3 : Trois plans en position générale dans $\mathbb{R}^3$ se coupent en un unique point.

4.11 Applications

4.11.1 Réseaux de flux

Exemple 24 (Réseau routier)

Considérons un réseau de quatre carrefours A, B, C, D reliés par des routes à sens unique. À chaque carrefour, le flux entrant doit égaler le flux sortant (conservation du flux).



Les équations de conservation sont :

$$\begin{cases} A : & 100 + x_4 = x_1, \\ B : & x_1 = x_2 + 60, \\ C : & x_2 + 50 = x_3, \\ D : & x_3 = x_4 + 90. \end{cases} \iff \begin{cases} x_1 - x_4 = 100, \\ x_1 - x_2 = 60, \\ x_2 - x_3 = -50, \\ x_3 - x_4 = 90. \end{cases}$$

La matrice augmentée se réduit en :

$$\left(\begin{array}{cccc|c} 1 & 0 & 0 & -1 & 100 \\ 0 & 1 & 0 & -1 & 40 \\ 0 & 0 & 1 & -1 & 90 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

Le rang est 3 avec 4 inconnues : il y a un paramètre libre $x_4 = t$. L'ensemble des solutions est :

$$(x_1, x_2, x_3, x_4) = (100 + t,\ 40 + t,\ 90 + t,\ t), \qquad t \geqslant 0.$$

La contrainte $t \geqslant 0$ (flux positifs) n'est pas une condition linéaire, mais elle limite le paramètre.

4.11.2 Circuits électriques

Exemple 25 (Lois de Kirchhoff)

Dans un circuit à deux mailles partageant une branche commune, les lois de Kirchhoff (loi des nœuds et loi des mailles) donnent un système linéaire dont les inconnues sont les courants I_1, I_2, I_3. Par exemple, avec des résistances $R_1 = 2\,\Omega$, $R_2 = 4\,\Omega$, $R_3 = 6\,\Omega$ et des tensions $V_1 = 12\,\mathrm{V}$, $V_2 = 8\,\mathrm{V}$:

$$\begin{cases} I_1 - I_2 - I_3 = 0 & \text{(loi des nœuds)},\\ 2I_1 + 6I_3 = 12 & \text{(maille 1)},\\ 4I_2 - 6I_3 = 8 & \text{(maille 2)}. \end{cases}$$

La résolution par Gauss donne $(I_1, I_2, I_3) = (\frac{18}{5}, \frac{13}{5}, \frac{1}{1}) = (3{,}6\ ;\ 2{,}6\ ;\ 1{,}0)$ A.

4.11.3 Aperçu : moindres carrés

Lorsqu'un système surdéterminé $A\mathbf{x} = \mathbf{b}$ (avec $m > n$) n'a pas de solution exacte, on cherche le vecteur $\mathbf{x}^*$ qui minimise $\|A\mathbf{x} - \mathbf{b}\|^2$. Ce problème de *moindres carrés* conduit aux *équations normales* :

$$A^{\mathsf{T}} A\,\mathbf{x}^* = A^{\mathsf{T}}\mathbf{b}, \tag{4.6}$$

qui forment un système linéaire carré de taille $n \times n$. Ce sujet sera développé dans le chapitre sur les espaces euclidiens.

Résumé du chapitre

1. Un **système linéaire** $A\mathbf{x} = \mathbf{b}$ se représente par sa matrice augmentée $(A \mid \mathbf{b})$.

2. Les **opérations élémentaires** sur les lignes (échange, multiplication par un scalaire non nul, ajout d'un multiple) ne changent pas l'ensemble des solutions.

3. L'**élimination de Gauss** transforme la matrice augmentée en **forme échelonnée** ; le prolongement de Gauss–Jordan produit la **forme échelonnée réduite** (unique).

4. Le **rang** $\mathrm{rg}(A)$ est le nombre de pivots dans une forme échelonnée.

5. **Théorème de Rouché–Capelli :** le système est compatible si et seulement si $\mathrm{rg}(A) = \mathrm{rg}(A \mid \mathbf{b})$; dans ce cas, l'ensemble des solutions est paramétré par $n - \mathrm{rg}(A)$ variables libres.

6. **Structure des solutions :** si $\mathbf{x}_p$ est une solution particulière, alors $\mathcal{S} = \mathbf{x}_p + \mathrm{Ker}(A)$ (sous-espace affine).

7. Les systèmes linéaires modélisent des problèmes concrets : réseaux de flux, circuits électriques, moindres carrés, etc.

Exercices

Exercice 26 (Résolution par Gauss)

Résoudre le système suivant par élimination de Gauss :

$$\begin{cases} x_1 + 2x_2 + 3x_3 = 9, \\ 2x_1 - x_2 + x_3 = 8, \\ 3x_1 + x_2 - 2x_3 = 3. \end{cases}$$

Exercice 27 (Forme échelonnée réduite)

Mettre la matrice suivante sous forme échelonnée réduite :

$$A = \begin{pmatrix} 1 & 3 & 5 & 7 \\ 2 & 4 & 6 & 8 \\ 3 & 5 & 7 & 9 \end{pmatrix}.$$

En déduire le rang de A.

Exercice 28 (Système homogène)

Déterminer l'ensemble des solutions du système homogène :

$$\begin{cases} x_1 - x_2 + 2x_3 - x_4 = 0, \\ 2x_1 + x_2 - x_3 + x_4 = 0, \\ x_1 + 2x_2 - 3x_3 + 2x_4 = 0. \end{cases}$$

Donner une base du noyau de la matrice des coefficients.

Exercice 29 (Compatibilité d'un système)

Déterminer pour quelles valeurs du paramètre $a \in \mathbb{R}$ le système suivant est compatible :

$$\begin{cases} x_1 + x_2 + x_3 = 1, \\ x_1 + 2x_2 + 3x_3 = 3, \\ x_1 + 3x_2 + ax_3 = 5. \end{cases}$$

Exercice 30 (Système paramétrique)

Résoudre en fonction du paramètre $\lambda \in \mathbb{R}$:

$$\begin{cases} \lambda x_1 + x_2 + x_3 = 1, \\ x_1 + \lambda x_2 + x_3 = 1, \\ x_1 + x_2 + \lambda x_3 = 1. \end{cases}$$

Distinguer les cas selon les valeurs de λ.

Exercice 31 (Calcul de rang)

Calculer le rang des matrices suivantes en fonction de $\alpha \in \mathbb{R}$:

$$A_1 = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}, \qquad A_2 = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 1 & 3 & \alpha & 10 \end{pmatrix}.$$

Exercice 32 (Structure affine)

On considère le système $A\mathbf{x} = \mathbf{b}$ avec :

$$A = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 2 & 1 & 2 \end{pmatrix}, \qquad \mathbf{b} = \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix}.$$

(a) Vérifier que le système est compatible.

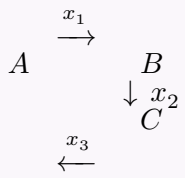
(b) Trouver une solution particulière.

(c) Déterminer une base de $\mathrm{Ker}(A)$.

(d) Écrire l'ensemble des solutions sous la forme $\mathcal{S} = \mathbf{x}_p + \mathrm{Vect}(...)$.

Exercice 33 (Réseau de flux)

Dans le réseau ci-dessous, les flux entrants et sortants aux nœuds sont indiqués :



Flux externe entrant en A : 200 ; flux externe sortant en B : 80 ; flux externe sortant en C : 120.

(a) Écrire le système de conservation du flux.

(b) Résoudre le système et interpréter.

Exercice 34 (Inversibilité et rang)

Soit $A \in \mathcal{M}_n(\mathbb{K})$ une matrice carrée. Montrer que les assertions suivantes sont équivalentes :

(i) A est inversible.

(ii) $\mathrm{rg}(A) = n$.

(iii) Le système $A\mathbf{x} = \mathbf{b}$ a une unique solution pour tout $\mathbf{b} \in \mathbb{K}^n$.

(iv) Le système homogène $A\mathbf{x} = \mathbf{0}$ n'a que la solution triviale.

Exercice 35 (Gauss–Jordan et inverse)

Utiliser l'élimination de Gauss–Jordan pour calculer l'inverse de la matrice :

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

Indication : appliquer Gauss–Jordan à la matrice augmentée $(A \mid I_3)$.

Exercice 36 (Systèmes et applications linéaires)

Soit $f\colon \mathbb{R}^4 \to \mathbb{R}^3$ l'application linéaire définie par :

$$f(x_1, x_2, x_3, x_4) = (x_1 + x_2 + x_3 + x_4,\ 2x_1 + x_2 - x_3 + 2x_4,\ 3x_1 + 2x_2 + x_4).$$

(a) Écrire la matrice de f dans les bases canoniques.

(b) Calculer $\operatorname{rg}(f)$ et $\dim(\operatorname{Ker}(f))$.

(c) Déterminer une base de $\operatorname{Ker}(f)$ et une base de $\operatorname{Im}(f)$.

(d) L'équation $f(\mathbf{x}) = (1, 2, 3)$ a-t-elle des solutions ? Si oui, les déterminer toutes.

Exercice 37 (Théorème de Rouché–Capelli)

Soit $A \in \mathcal{M}_{m,n}(\mathbb{K})$ et $\mathbf{b} \in \mathbb{K}^m$. On suppose que $\operatorname{rg}(A) = \operatorname{rg}(A \mid \mathbf{b}) = r$.

(a) Montrer que si $r = n$, la solution est unique.

(b) Montrer que si $r < n$ et si $\mathbf{v}_1, \dots, \mathbf{v}_{n-r}$ est une base de $\operatorname{Ker}(A)$, alors toute solution s'écrit :

$$\mathbf{x} = \mathbf{x}_p + t_1\mathbf{v}_1 + \cdots + t_{n-r}\mathbf{v}_{n-r}, \qquad t_1, \dots, t_{n-r} \in \mathbb{K}.$$

(c) En déduire que le nombre de solutions est soit 0, soit 1, soit infini (lorsque $\mathbb{K}$ est infini).

Exercice 38 (Système surdéterminé et moindres carrés)

On mesure les points $(0, 1)$, $(1, 2)$, $(2, 4)$, $(3, 5)$ et on cherche la droite $y = a + bx$ au sens des moindres carrés.

(a) Écrire le système surdéterminé $A\mathbf{x} = \mathbf{b}$ correspondant.

(b) Former les équations normales $A^{\mathsf{T}}A\,\mathbf{x}^* = A^{\mathsf{T}}\mathbf{b}$.

(c) Résoudre et déterminer la droite de régression.

Index

$\mathbb{K}^n$, 17
$\oplus$, 21

anneau, 5
    commutatif, 6
application, 2
application linéaire, 37, 38
automorphisme, 44

base, 25
base canonique, 25
base duale, 51
bidual, 51
bijection, 3

caractéristique, 9
changement de base, 47
combinaison de lignes, 59
combinaison linéaire, 22
composition, 3
conservation du flux, 67
coordonnées, 25
coordonnées homogènes, 51
corps, 6
corps fini, 8
critère de base, 28

dimension, 27
    invariance, 26
dimension finie, 27
droite, 66
dual, 50
dérivation, 39
déterminant, 50

endomorphisme, 38
ensemble, 1
ensemble des solutions, 58
espace de fonctions, 17
espace des matrices, 17
espace des polynômes, 17
espace des solutions, 30
espace des suites, 18
espace dual, 50
espace vectoriel, 16

famille génératrice, 23
famille libre, 24
famille liée, 24
forme linéaire, 50
forme échelonnée, 60
forme échelonnée réduite, 60
formule de Grassmann, 28

$\mathrm{GL}_n(\mathbb{K})$, 49
$\mathrm{GL}(E)$, 44
Grassmann, formule de, 28
groupe, 5
    abélien, 5
groupe linéaire, 50

homomorphisme, 38
homothétie, 37

image, 40
indépendance linéaire, 24
infographie, 51
injection, 2
injectivité
    application linéaire, 43
intersection de sous-espaces, 20
inégalité de Sylvester, 54
isomorphe, 43
isomorphisme, 43

Kirchhoff (lois de), 57
Kronecker–Capelli (théorème de), 64

$\mathcal{L}(E, F)$, 38
lemme d'échange, 26
lemme de Steinitz, 26
Leontief (modèle de), 57
loi de composition interne, 4

matrice

augmentée, 58
d'une application linéaire, 45
des coefficients, 58
inversible, 59
matrice stochastique, 52
matrice de passage, 47
matrice de transition, 52
matrice inversible, 49
matrices semblables, 47
moindres carrés, 68
multiplication d'une ligne, 59
multiplication scalaire, 16

nilpotent, 54
nombres complexes, 8
noyau, 40

opérations élémentaires, 59
sur les lignes, 59

pivot, 60
pivot de Gauss, 61
plan, 66
polynôme impair, 31
polynôme pair, 31
produit matriciel, 48
projecteur, 39
projection, 37

rang, 41, 63
colonne, 63
d'une famille de vecteurs, 29
ligne, 63
relation binaire, 3
relation d'ordre, 4
relation d'équivalence, 3
rotation, 37
Rouché–Capelli (théorème de), 64
RREF, 60

réflexion, 37

scalaire, 15
solution
triviale, 59
solution d'un système, 58
somme de sous-espaces, 20
somme directe, 21
sous-espace
trivial, 20
sous-espace affine, 65
sous-espace engendré, 22
sous-espace vectoriel, 18, 59
sous-espaces supplémentaires, 21
substitution arrière, 61
supplémentaire, 21
surjection, 2
surjectivité
application linéaire, 43
système homogène, 18
système linéaire, 57, 68
compatible, 58
définition, 58
homogène, 59
incompatible, 58

théorème du rang, 30, 41
transposée, 49

variable
libre, 65
liée, 65
Vect, 22
vecteur, 16

échange de lignes, 59
élimination de Gauss, 57, 61
élimination de Gauss–Jordan, 63
équations normales, 68

Chapitre 5

Déterminants

Introduction

Le *déterminant* est l'un des invariants les plus fondamentaux de l'algèbre linéaire. Il associe à toute matrice carrée un scalaire qui encode des informations essentielles : inversibilité, volume, orientation.

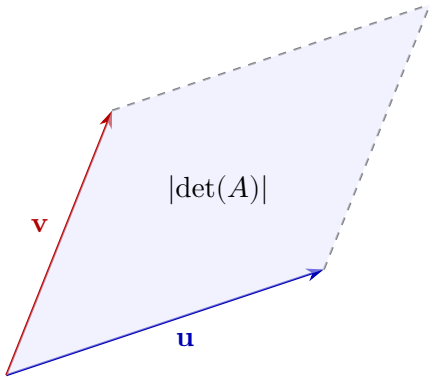
Avant d'en donner la construction rigoureuse, observons les cas les plus simples. Pour une matrice 2×2 :

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \qquad \det(A) = ad - bc.$$

Ce nombre apparaît naturellement en géométrie : si $\mathbf{u} = (a, c)$ et $\mathbf{v} = (b, d)$ sont les colonnes de A, alors $|\det(A)|$ est l'aire du parallélogramme engendré par $\mathbf{u}$ et $\mathbf{v}$, et le *signe* de $\det(A)$ détermine l'orientation du couple $(\mathbf{u}, \mathbf{v})$.

De plus, la matrice A est inversible si et seulement si $\det(A) \neq 0$. Cette caractérisation se généralise à toute dimension : le déterminant d'une matrice $n \times n$ est un scalaire qui s'annule exactement lorsque la matrice est singulière.

La construction du déterminant en dimension quelconque repose sur le *groupe symétrique* $\mathfrak{S}_n$ et la notion de *signature* d'une permutation. C'est pourquoi nous commençons ce chapitre par une étude des permutations.



Aire du parallélogramme $= |ad - bc|$

Fig. 5.1 : L'aire du parallélogramme engendré par les colonnes d'une matrice 2×2 est $|\det(A)|$.

5.1 Permutations et groupe symétrique

5.1.1 Définitions fondamentales

Définition 1 (Permutation)

Soit $n \in \mathbb{N}^*$. Une *permutation* de l'ensemble $\{1, 2, \dots, n\}$ est une bijection $\sigma : \{1, 2, \dots, n\} \to \{1, 2, \dots, n\}$. L'ensemble de toutes les permutations de $\{1, 2, \dots, n\}$, muni de la composition, forme un groupe appelé le *groupe symétrique* et noté $\mathfrak{S}_n$.

Le cardinal de $\mathfrak{S}_n$ est $n!$ (factorielle de n). On note une permutation σ par la notation en deux lignes :

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{pmatrix}.$$

Exemple 2 (Permutations de $\mathfrak{S}_3$)

Le groupe $\mathfrak{S}_3$ contient $3! = 6$ éléments :

$$\mathrm{Id} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}.$$

5.1.2 Transpositions

Définition 3 (Transposition)

Une *transposition* est une permutation $\tau \in \mathfrak{S}_n$ qui échange exactement deux éléments et laisse tous les autres fixes. On la note $\tau = (i\ j)$ avec $i \neq j$.

Proposition 4 (Décomposition en transpositions)

Toute permutation $\sigma \in \mathfrak{S}_n$ peut s'écrire comme composée d'un nombre fini de transpositions :

$$\sigma = \tau_1 \circ \tau_2 \circ \cdots \circ \tau_k.$$

La décomposition n'est pas unique, mais la *parité* du nombre de transpositions est toujours la même.

Démonstration. Il suffit de vérifier qu'un cycle $(a_1\ a_2\ \cdots\ a_r)$ se décompose en transpositions :

$$(a_1\ a_2\ \cdots\ a_r) = (a_1\ a_r) \circ (a_1\ a_{r-1}) \circ \cdots \circ (a_1\ a_2).$$

Comme toute permutation est un produit de cycles à supports disjoints, on obtient la décomposition voulue.

Pour la parité, considérons l'application $P(\sigma) := \prod_{1 \leq i < j \leq n} \frac{\sigma(j) - \sigma(i)}{j - i}$. Chaque transposition change le signe de P, donc $P(\sigma) = (-1)^k$ pour toute décomposition $\sigma = \tau_1 \circ \cdots \circ \tau_k$. La parité de k est ainsi indépendante de la décomposition choisie. □

5.1.3 Signature

Définition 5 (Signature d'une permutation)

La *signature* d'une permutation $\sigma \in \mathfrak{S}_n$ est définie par :

$$\operatorname{sgn}(\sigma) := (-1)^{N(\sigma)} = \prod_{1 \le i < j \le n} \frac{\sigma(j) - \sigma(i)}{j - i},$$

où $N(\sigma)$ est le nombre d'*inversions* de σ, c'est-à-dire le nombre de couples (i, j) avec $i < j$ et $\sigma(i) > \sigma(j)$.
Si $\operatorname{sgn}(\sigma) = +1$, la permutation est dite *paire* ; si $\operatorname{sgn}(\sigma) = -1$, elle est dite *impaire*.

Proposition 6 (Propriétés de la signature)

La signature est un morphisme de groupes $\operatorname{sgn} \colon \mathfrak{S}_n \to \{+1, -1\}$:

(i) $\operatorname{sgn}(\mathrm{Id}) = +1$.

(ii) $\operatorname{sgn}(\sigma \circ \tau) = \operatorname{sgn}(\sigma) \cdot \operatorname{sgn}(\tau)$ pour tous $\sigma, \tau \in \mathfrak{S}_n$.

(iii) $\operatorname{sgn}(\tau) = -1$ pour toute transposition τ.

(iv) $\operatorname{sgn}(\sigma^{-1}) = \operatorname{sgn}(\sigma)$ pour tout $\sigma \in \mathfrak{S}_n$.

Démonstration. *(i)* L'identité n'a aucune inversion, donc $\operatorname{sgn}(\mathrm{Id}) = (-1)^0 = 1$.

(ii) Pour $\sigma, \tau \in \mathfrak{S}_n$, on a :

$$\prod_{i<j} \frac{(\sigma \circ \tau)(j) - (\sigma \circ \tau)(i)}{j - i} = \prod_{i<j} \frac{\sigma(\tau(j)) - \sigma(\tau(i))}{\tau(j) - \tau(i)} \cdot \prod_{i<j} \frac{\tau(j) - \tau(i)}{j - i} = \operatorname{sgn}(\sigma) \cdot \operatorname{sgn}(\tau).$$

(iii) Soit $\tau = (a\ b)$ avec $a < b$. Parmi les couples (i, j) avec $i < j$, le couple (a, b) est inversé, et pour chaque élément c avec $a < c < b$, les deux couples (a, c) et (c, b) changent simultanément de statut. Le nombre total d'inversions créées est donc impair (il vaut $2(b - a - 1) + 1$), d'où $\operatorname{sgn}(\tau) = -1$.

(iv) Comme $\sigma \circ \sigma^{-1} = \mathrm{Id}$, on a $\operatorname{sgn}(\sigma) \cdot \operatorname{sgn}(\sigma^{-1}) = 1$, ce qui donne $\operatorname{sgn}(\sigma^{-1}) = \operatorname{sgn}(\sigma)$. □

Exemple 7 (Calcul de signatures)

Considérons $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 1 & 3 \end{pmatrix} \in \mathfrak{S}_4$. Les inversions sont : $(1, 3)$ car $\sigma(1) = 2 > \sigma(3) = 1$; $(2, 3)$ car $\sigma(2) = 4 > \sigma(3) = 1$; $(2, 4)$ car $\sigma(2) = 4 > \sigma(4) = 3$. Donc $N(\sigma) = 3$ et $\operatorname{sgn}(\sigma) = (-1)^3 = -1$: la permutation est impaire.

On vérifie : $\sigma = (1\ 2\ 4\ 3) = (1\ 3) \circ (1\ 4) \circ (1\ 2)$, produit de 3 transpositions, ce qui confirme $\operatorname{sgn}(\sigma) = -1$.

5.2 Définition du déterminant

5.2.1 Formule de Leibniz

Définition 8 (Déterminant (formule de Leibniz))

Soit $A = (a_{ij}) \in \mathcal{M}_n(\mathbb{K})$. Le *déterminant* de A est le scalaire :

$$\det(A) := \sum_{\sigma \in \mathfrak{S}_n} \operatorname{sgn}(\sigma)\, a_{1,\sigma(1)}\, a_{2,\sigma(2)} \cdots a_{n,\sigma(n)} = \sum_{\sigma \in \mathfrak{S}_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} a_{i,\sigma(i)}.$$

Remarque 9 (Nombre de termes)

La somme porte sur les $n!$ permutations de $\mathfrak{S}_n$. Pour $n = 2$, cela donne $2! = 2$ termes ; pour $n = 3$, $3! = 6$ termes ; pour $n = 10$, plus de 3,6 millions de termes. La formule de Leibniz est donc utile pour les démonstrations théoriques, mais peu efficace pour le calcul pratique en grande dimension.

Exemple 10 (Déterminant 2×2)

Pour $n = 2$, $\mathfrak{S}_2 = \{\mathrm{Id}, (1\ 2)\}$. La formule donne :

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \operatorname{sgn}(\mathrm{Id})\, a \cdot d + \operatorname{sgn}((1\ 2))\, b \cdot c = ad - bc.$$

Exemple 11 (Déterminant 3×3)

Pour $n = 3$, la formule de Leibniz donne (règle de Sarrus) :

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

5.2.2 Caractérisation par les formes multilinéaires alternées

On peut caractériser le déterminant de manière axiomatique, sans passer par les permutations. Notons $C_1, \dots, C_n \in \mathbb{K}^n$ les colonnes d'une matrice $A \in \mathcal{M}_n(\mathbb{K})$, de sorte que $A = (C_1 \mid \cdots \mid C_n)$.

Définition 12 (Forme multilinéaire alternée)

Une application $\varphi : (\mathbb{K}^n)^n \to \mathbb{K}$ est dite :

(i) *multilinéaire* si elle est linéaire en chacune de ses n variables (les autres étant fixées) ;

(ii) *alternée* si elle s'annule dès que deux de ses arguments sont égaux : $\varphi(\dots, C, \dots, C, \dots) = 0$.

Remarque 13 (Alternée implique antisymétrique)

Si φ est multilinéaire et alternée, alors échanger deux arguments change le signe :

$$\varphi(\dots, C_i, \dots, C_j, \dots) = -\varphi(\dots, C_j, \dots, C_i, \dots).$$

En effet, $0 = \varphi(\dots, C_i + C_j, \dots, C_i + C_j, \dots) = \varphi(\dots, C_i, \dots, C_j, \dots) + \varphi(\dots, C_j, \dots, C_i, \dots)$ par multilinéarité et le fait que φ s'annule sur les termes avec deux arguments identiques.

Théorème 14 (Caractérisation axiomatique du déterminant)

Il existe une unique application $\det : \mathcal{M}_n(\mathbb{K}) \to \mathbb{K}$ qui est :

(i) multilinéaire par rapport aux colonnes ;

(ii) alternée ;

(iii) normalisée : $\det(I_n) = 1$.

Cette application coïncide avec la formule de Leibniz.

Démonstration. Existence. La formule de Leibniz définit une application det qui vérifie les trois propriétés. La multilinéarité résulte du fait que chaque terme $\prod_{i=1}^{n} a_{i,\sigma(i)}$ est linéaire en chaque colonne. Le caractère alternée provient du fait que si deux colonnes $C_j = C_k$ $(j \neq k)$, on peut apparier les permutations σ et $\sigma \circ (j\ k)$ qui contribuent des termes opposés. La normalisation $\det(I_n) = 1$ se vérifie puisque seule l'identité contribue un terme non nul.

Unicité. Soit φ vérifiant (i)–(iii). En notant $e_1, \dots, e_n$ la base canonique de $\mathbb{K}^n$ et en écrivant $C_j = \sum_{i=1}^{n} a_{ij}\, e_i$, la multilinéarité donne :

$$\varphi(C_1, \dots, C_n) = \sum_{i_1, \dots, i_n = 1}^{n} a_{i_1,1} \cdots a_{i_n,n}\, \varphi(e_{i_1}, \dots, e_{i_n}).$$

Le caractère alternée impose $\varphi(e_{i_1}, \dots, e_{i_n}) = 0$ dès que deux indices coïncident. Donc seuls les n-uplets $(i_1, \dots, i_n)$ qui sont des permutations de $(1, \dots, n)$ subsistent. Pour une permutation σ, on a $\varphi(e_{\sigma(1)}, \dots, e_{\sigma(n)}) = \operatorname{sgn}(\sigma)\, \varphi(e_1, \dots, e_n) = \operatorname{sgn}(\sigma)$ par (iii). On retrouve la formule de Leibniz. □

5.3 Propriétés fondamentales

5.3.1 Multilinéarité par rapport aux lignes et colonnes

Le déterminant est multilinéaire par rapport aux colonnes par construction. Nous verrons dans la section 5.5.1 qu'il est aussi multilinéaire par rapport aux lignes.

Proposition 15 (Effet des opérations élémentaires)

Soit $A \in \mathcal{M}_n(\mathbb{K})$.

(i) **Échange de deux colonnes (ou lignes) :** $\det(\dots, C_j, \dots, C_k, \dots) = -\det(\dots, C_k, \dots, C_j, \dots)$.

(ii) **Multiplication d'une colonne (ou ligne) par $\lambda \in \mathbb{K}$:** $\det(\dots, \lambda C_j, \dots) = \lambda \det(\dots, C_j, \dots)$.

(iii) **Ajout d'un multiple d'une colonne (ou ligne) à une autre :** $\det(\dots, C_j +$

$\lambda C_k, \dots, C_k, \dots) = \det(\dots, C_j, \dots, C_k, \dots)$.

Démonstration. *(i)* C'est la propriété d'antisymétrie (voir remarque 13).

(ii) C'est l'homogénéité de la multilinéarité en la colonne j.

(iii) Par multilinéarité : $\det(\dots, C_j+\lambda C_k, \dots, C_k, \dots) = \det(\dots, C_j, \dots, C_k, \dots)+\lambda \det(\dots, C_k, \dots, C_k, \dots)$. Le second terme est nul car deux colonnes sont égales. □

Corollaire 16 (Colonne nulle ou colonnes proportionnelles)

(i) Si A possède une colonne (ou ligne) nulle, alors $\det(A) = 0$.

(ii) Si A possède deux colonnes (ou lignes) proportionnelles, alors $\det(A) = 0$.

5.3.2 Déterminant d'un produit

Théorème 17 (Multiplicativité du déterminant)

Pour toutes matrices $A, B \in \mathcal{M}_n(\mathbb{K})$:

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

Démonstration. Si $\det(B) = 0$, alors B n'est pas inversible, donc $\operatorname{rg}(AB) \leq \operatorname{rg}(B) < n$, ce qui implique que AB n'est pas inversible et $\det(AB) = 0 = \det(A) \cdot \det(B)$.

Si $\det(B) \neq 0$, considérons l'application $\varphi \colon \mathcal{M}_n(\mathbb{K}) \to \mathbb{K}$ définie par $\varphi(A) := \frac{\det(AB)}{\det(B)}$. Alors φ est multilinéaire alternée par rapport aux colonnes de A (car la multiplication à droite par B opère linéairement sur les colonnes) et $\varphi(I_n) = \frac{\det(B)}{\det(B)} = 1$. Par l'unicité du théorème 14, $\varphi(A) = \det(A)$, d'où $\det(AB) = \det(A) \cdot \det(B)$. □

Corollaire 18 (Déterminant de l'inverse)

Si $A \in \mathrm{GL}_n(\mathbb{K})$, alors :

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

Démonstration. On a $A \cdot A^{-1} = I_n$, donc $\det(A) \cdot \det(A^{-1}) = \det(I_n) = 1$. □

Corollaire 19 (Déterminant de matrices semblables)

Si A et B sont semblables (i.e. $B = P^{-1}AP$ pour $P \in \mathrm{GL}_n(\mathbb{K})$), alors $\det(A) = \det(B)$.

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

5.4 Développement selon une ligne ou une colonne

5.4.1 Mineur et cofacteur

Définition 20 (Mineur et cofacteur)

Soit $A = (a_{ij}) \in \mathcal{M}_n(\mathbb{K})$ avec $n \geq 2$.

(i) Le *mineur* M_{ij} est le déterminant de la matrice $(n-1) \times (n-1)$ obtenue en supprimant la i-ème ligne et la j-ème colonne de A.

(ii) Le *cofacteur* (ou *complément algébrique*) est $A_{ij} := (-1)^{i+j} M_{ij}$.

5.4.2 Développement de Laplace

Théorème 21 (Développement selon une ligne ou une colonne (Laplace))

Soit $A = (a_{ij}) \in \mathcal{M}_n(\mathbb{K})$ avec $n \geq 2$.

(i) **Développement selon la i-ème ligne :**

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

(ii) **Développement selon la j-ème colonne :**

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

Le résultat est indépendant du choix de la ligne ou de la colonne.

Démonstration. Montrons le développement selon la i-ème ligne. D'après la formule de Leibniz :

$$\det(A) = \sum_{\sigma \in \mathfrak{S}_n} \operatorname{sgn}(\sigma) \prod_{k=1}^{n} a_{k,\sigma(k)}.$$

On regroupe les termes selon la valeur $j = \sigma(i)$:

$$\det(A) = \sum_{j=1}^{n} a_{ij} \sum_{\substack{\sigma \in \mathfrak{S}_n \\ \sigma(i)=j}} \operatorname{sgn}(\sigma) \prod_{\substack{k=1 \\ k \neq i}}^{n} a_{k,\sigma(k)}.$$

Il faut montrer que la somme intérieure vaut $(-1)^{i+j} M_{ij}$.

Fixons j et considérons une permutation σ avec $\sigma(i) = j$. Pour ramener la i-ème ligne en première position et la j-ème colonne en première position, on effectue $i-1$ échanges de lignes adjacentes et $j-1$ échanges de colonnes adjacentes, ce qui contribue un facteur $(-1)^{i-1+j-1} = (-1)^{i+j}$.

Plus précisément, à chaque $\sigma \in \mathfrak{S}_n$ avec $\sigma(i) = j$, on associe bijectivement une permutation σ' de $\mathfrak{S}_{n-1}$ sur les indices restants $\{1, \dots, n\} \setminus \{i\} \to \{1, \dots, n\} \setminus \{j\}$, avec $\operatorname{sgn}(\sigma) = (-1)^{i+j} \operatorname{sgn}(\sigma')$. La somme intérieure devient :

$$\sum_{\sigma' \in \mathfrak{S}_{n-1}} \operatorname{sgn}(\sigma') \prod_{k'} a_{k',\sigma'(k')} = (-1)^{i+j}\, M_{ij}.$$

D'où le résultat. Le développement selon une colonne se démontre de manière analogue (ou en appliquant le résultat à A^T, cf. proposition 24). □

Exemple 22 (Développement d'un déterminant 3×3)

Calculons le déterminant suivant en développant selon la première ligne :

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

On calcule : $2(1-8) - 1(0-20) + 3(0+5) = 2(-7) - 1(-20) + 3(5) = -14 + 20 + 15 = 21$.

Remarque 23 (Stratégie de calcul)

En pratique, on développe selon la ligne ou la colonne qui contient le plus de zéros, afin de minimiser le nombre de termes à calculer. On peut aussi effectuer des opérations élémentaires sur les lignes et colonnes pour créer des zéros avant le développement.

5.5 Propriétés complémentaires

5.5.1 Déterminant de la transposée

Proposition 24 (Déterminant de la transposée)

Pour toute matrice $A \in \mathcal{M}_n(\mathbb{K})$:

$$\det(A^\mathsf{T}) = \det(A).$$

Démonstration. Notons $B = A^\mathsf{T}$, de sorte que $b_{ij} = a_{ji}$. Alors :

$$\det(B) = \sum_{\sigma \in \mathfrak{S}_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} b_{i,\sigma(i)} = \sum_{\sigma \in \mathfrak{S}_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} a_{\sigma(i),i}.$$

En posant $\tau = \sigma^{-1}$ (et donc $\operatorname{sgn}(\tau) = \operatorname{sgn}(\sigma)$), et en réindexant le produit par $j = \sigma(i)$, c'est-à-dire $i = \tau(j)$:

$$\det(B) = \sum_{\tau \in \mathfrak{S}_n} \operatorname{sgn}(\tau) \prod_{j=1}^{n} a_{j,\tau(j)} = \det(A).$$ □

Remarque 25 (Conséquence importante)

Cette propriété signifie que tout résultat établi pour les colonnes vaut aussi pour les lignes (et réciproquement). En particulier, le déterminant est aussi multilinéaire alternée par rapport aux lignes.

5.5.2 Matrices triangulaires

Proposition 26 (Déterminant d'une matrice triangulaire)

Le déterminant d'une matrice triangulaire (supérieure ou inférieure) est le produit de ses coefficients diagonaux :

$$\det(A) = \prod_{i=1}^{n} a_{ii}.$$

En particulier, $\det(I_n) = 1$ et $\det(\lambda I_n) = \lambda^n$.

Démonstration. Si A est triangulaire supérieure ($a_{ij} = 0$ pour $i > j$), alors dans la formule de Leibniz, un terme $\prod_{i=1}^{n} a_{i,\sigma(i)}$ est non nul seulement si $\sigma(i) \geq i$ pour tout i. La seule permutation de $\{1, \dots, n\}$ vérifiant $\sigma(i) \geq i$ pour tout i est l'identité (car $\sum_{i=1}^{n} \sigma(i) = \sum_{i=1}^{n} i$ impose $\sigma(i) = i$). Donc $\det(A) = \operatorname{sgn}(\mathrm{Id}) \prod_{i=1}^{n} a_{ii} = \prod_{i=1}^{n} a_{ii}$. □

5.5.3 Matrices par blocs

Proposition 27 (Déterminant d'une matrice triangulaire par blocs)

Soient $A \in \mathcal{M}_p(\mathbb{K})$, $B \in \mathcal{M}_q(\mathbb{K})$ et $C \in \mathcal{M}_{p,q}(\mathbb{K})$. Alors :

$$\det \begin{pmatrix} A & C \\ 0 & B \end{pmatrix} = \det(A) \cdot \det(B).$$

De même :

$$\det \begin{pmatrix} A & 0 \\ C & B \end{pmatrix} = \det(A) \cdot \det(B).$$

Démonstration. Notons $M = \begin{pmatrix} A & C \\ 0 & B \end{pmatrix} \in \mathcal{M}_{p+q}(\mathbb{K})$. Dans la formule de Leibniz, un terme $\prod_{i=1}^{p+q} m_{i,\sigma(i)}$ est non nul seulement si $\sigma(i) \leq p$ pour $i \leq p$ (car $m_{i,j} = 0$ pour $i > p$ et $j \leq p$ seulement si le bloc est nul, mais ici le bloc inférieur gauche est nul). Plus précisément, pour $i > p$, on a $m_{i,j} = 0$ si $j \leq p$, donc $\sigma(i) > p$. Ainsi σ doit préserver $\{1, \dots, p\}$ et $\{p+1, \dots, p+q\}$, ce qui implique que σ se décompose en $\sigma_1 \in \mathfrak{S}_p$ et $\sigma_2 \in \mathfrak{S}_q$ avec $\operatorname{sgn}(\sigma) = \operatorname{sgn}(\sigma_1) \cdot \operatorname{sgn}(\sigma_2)$. Le produit se factorise alors en $\prod_{i=1}^{p} a_{i,\sigma_1(i)} \cdot \prod_{j=1}^{q} b_{j,\sigma_2(j)}$, et la somme donne $\det(A) \cdot \det(B)$. □

5.6 Comatrice et règle de Cramer

5.6.1 Comatrice (matrice adjointe)

Définition 28 (Comatrice)

Soit $A = (a_{ij}) \in \mathcal{M}_n(\mathbb{K})$. La *comatrice* (ou *matrice des cofacteurs*) de A est la matrice $\operatorname{Com}(A) \in \mathcal{M}_n(\mathbb{K})$ dont le coefficient (i,j) est le cofacteur $A_{ij} = (-1)^{i+j} M_{ij}$.
La *matrice adjointe* (adjugate en anglais) est :

$$\widetilde{A} := (\operatorname{Com}(A))^{\mathsf{T}}.$$

Son coefficient (i,j) est donc $A_{ji} = (-1)^{i+j} M_{ji}$.

Théorème 29 (Formule de la comatrice)

Pour toute matrice $A \in \mathcal{M}_n(\mathbb{K})$:

$$A \cdot \widetilde{A} = \widetilde{A} \cdot A = \det(A) \cdot I_n.$$

Démonstration. Calculons le coefficient (i,k) du produit $A \cdot \widetilde{A}$:

$$(A \cdot \widetilde{A})_{ik} = \sum_{j=1}^{n} a_{ij} \, (\widetilde{A})_{jk} = \sum_{j=1}^{n} a_{ij} \, A_{kj} = \sum_{j=1}^{n} a_{ij} \, (-1)^{k+j} M_{kj}.$$

Cas $i = k$: C'est exactement le développement de $\det(A)$ selon la i-ème ligne (théorème 21).

Cas $i \neq k$: La somme $\sum_{j=1}^{n} a_{ij}(-1)^{k+j} M_{kj}$ est le développement selon la k-ème ligne du déterminant de la matrice A' obtenue en remplaçant la k-ème ligne de A par la i-ème ligne. Puisque A' possède deux lignes identiques (les lignes i et k), $\det(A') = 0$.

Donc $(A \cdot \widetilde{A})_{ik} = \delta_{ik} \det(A)$, c'est-à-dire $A \cdot \widetilde{A} = \det(A) \cdot I_n$. La preuve de $\widetilde{A} \cdot A = \det(A) \cdot I_n$ est analogue (on développe selon les colonnes). □

Corollaire 30 (Formule de l'inverse par la comatrice)

Si $A \in \mathrm{GL}_n(\mathbb{K})$ (i.e. $\det(A) \neq 0$), alors :

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

Exemple 31 (Inverse d'une matrice 2×2)

Pour $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ avec $ad - bc \neq 0$:

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

5.6.2 Règle de Cramer

Théorème 32 (Règle de Cramer)

Soit $A \in \mathrm{GL}_n(\mathbb{K})$ et $b \in \mathbb{K}^n$. L'unique solution du système linéaire $Ax = b$ est donnée par :

$$x_j = \frac{\det(A_j)}{\det(A)}, \qquad j = 1, \dots, n,$$

où A_j est la matrice obtenue en remplaçant la j-ème colonne de A par le vecteur b :

$$A_j = (C_1 \mid \cdots \mid C_{j-1} \mid b \mid C_{j+1} \mid \cdots \mid C_n).$$

Démonstration. Puisque A est inversible, le système admet une unique solution $x = A^{-1}b$. D'après le corollaire 30 :

$$x_j = (A^{-1}b)_j = \frac{1}{\det(A)} \sum_{i=1}^{n} A_{ij} \, b_i = \frac{1}{\det(A)} \sum_{i=1}^{n} (-1)^{i+j} M_{ij} \, b_i.$$

Or $\sum_{i=1}^{n}(-1)^{i+j}M_{ij}\,b_i$ est le développement de $\det(A_j)$ selon la j-ème colonne (dont les coefficients sont $b_1,\dots,b_n$), d'où $x_j = \frac{\det(A_j)}{\det(A)}$. □

Exemple 33 (Système 2×2 par Cramer)

Résolvons $\begin{cases} 2x+3y=5 \\ x-y=1 \end{cases}$ par Cramer.

$\det(A) = \det\begin{pmatrix} 2 & 3 \\ 1 & -1 \end{pmatrix} = -2-3 = -5.$

$x = \dfrac{\det\begin{pmatrix} 5 & 3 \\ 1 & -1 \end{pmatrix}}{-5} = \frac{-5-3}{-5} = \frac{-8}{-5} = \frac{8}{5}.$

$y = \dfrac{\det\begin{pmatrix} 2 & 5 \\ 1 & 1 \end{pmatrix}}{-5} = \frac{2-5}{-5} = \frac{-3}{-5} = \frac{3}{5}.$

Remarque 34 (Complexité de Cramer)

La règle de Cramer nécessite le calcul de $n+1$ déterminants $n\times n$. Pour de grands systèmes, l'élimination de Gauss est bien plus efficace ($O(n^3)$ contre $O(n!\cdot n)$ pour Cramer via la formule de Leibniz). La règle de Cramer est néanmoins utile pour les petites dimensions et pour les démonstrations théoriques.

5.7 Déterminant et inversibilité

Théorème 35 (Critère d'inversibilité)

Soit $A\in\mathcal{M}_n(\mathbb{K})$. Les assertions suivantes sont équivalentes :

(i) A est inversible ;

(ii) $\det(A)\neq 0$;

(iii) les colonnes de A forment une base de $\mathbb{K}^n$;

(iv) les lignes de A forment une base de $\mathbb{K}^n$;

(v) $\operatorname{rg}(A) = n$.

Démonstration. Les équivalences (i)⇔(iii)⇔(iv)⇔(v) ont été établies au chapitre 3.

(i)⇒(ii) : Si A est inversible, alors $1=\det(I_n)=\det(A\cdot A^{-1})=\det(A)\cdot\det(A^{-1})$, donc $\det(A)\neq 0$.

(ii)⇒(i) : Si $\det(A)\neq 0$, alors $A^{-1}=\frac{1}{\det(A)}\widetilde{A}$ d'après le corollaire 30, donc A est inversible. □

Corollaire 36 (Noyau et déterminant)

Soit $A\in\mathcal{M}_n(\mathbb{K})$. Le système homogène $Ax=0$ admet une solution non triviale si et seulement si $\det(A)=0$.

5.8 Déterminant d'un endomorphisme

Définition 37 (Déterminant d'un endomorphisme)

Soit E un $\mathbb{K}$-espace vectoriel de dimension finie n et $f \in \mathrm{End}(E)$. Le *déterminant de* f est défini par :

$$\det(f) := \det(\mathcal{M}_{\mathcal{B}}(f)),$$

où $\mathcal{B}$ est une base quelconque de E.

Proposition 38 (Indépendance de la base)

Le déterminant d'un endomorphisme est bien défini : il ne dépend pas du choix de la base.

Démonstration. Soient $\mathcal{B}$ et $\mathcal{B}'$ deux bases de E, et P la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$. Alors $\mathcal{M}_{\mathcal{B}'}(f) = P^{-1}\,\mathcal{M}_{\mathcal{B}}(f)\,P$. Par le corollaire 19 :

$$\det(\mathcal{M}_{\mathcal{B}'}(f)) = \det(P^{-1}\,\mathcal{M}_{\mathcal{B}}(f)\,P) = \det(\mathcal{M}_{\mathcal{B}}(f)).$$ □

Proposition 39 (Propriétés du déterminant d'un endomorphisme)

Soient $f, g \in \mathrm{End}(E)$ avec $\dim E = n$.

(i) $\det(g \circ f) = \det(g) \cdot \det(f)$.

(ii) $\det(\mathrm{Id}_E) = 1$.

(iii) f est un automorphisme si et seulement si $\det(f) \neq 0$.

(iv) $\det(\lambda f) = \lambda^n \det(f)$ pour tout $\lambda \in \mathbb{K}$.

5.9 Interprétation géométrique

5.9.1 Volume signé

Le déterminant possède une interprétation géométrique fondamentale en tant que *volume signé*.

Proposition 40 (Déterminant et volume)

Soient $\mathbf{v}_1, \dots, \mathbf{v}_n \in \mathbb{R}^n$ et $A = (\mathbf{v}_1 \mid \cdots \mid \mathbf{v}_n)$ la matrice dont les colonnes sont ces vecteurs. Alors :

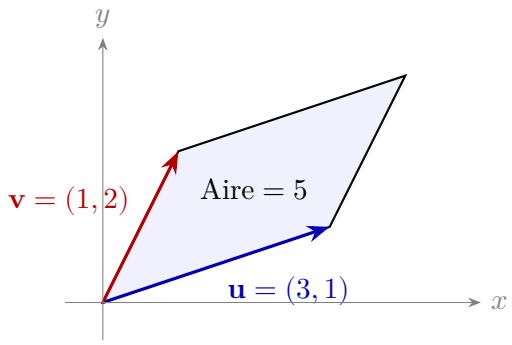
(i) $|\det(A)|$ est le *volume* du parallélotope engendré par $\mathbf{v}_1, \dots, \mathbf{v}_n$.

(ii) Le signe de $\det(A)$ indique l'*orientation* : $\det(A) > 0$ si la famille $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ est orientée positivement (même orientation que la base canonique), $\det(A) < 0$ sinon.

5.9.2 Cas du plan : aire du parallélogramme

Dans $\mathbb{R}^2$, si $\mathbf{u} = \begin{pmatrix} a \\ c \end{pmatrix}$ et $\mathbf{v} = \begin{pmatrix} b \\ d \end{pmatrix}$, l'aire du parallélogramme engendré par $\mathbf{u}$ et $\mathbf{v}$ est :

$$\text{Aire} = \left|\det \begin{pmatrix} a & b \\ c & d \end{pmatrix}\right| = |ad - bc|.$$



$$\det \begin{pmatrix} 3 & 1 \\ 1 & 2 \end{pmatrix} = 6 - 1 = 5$$

Fig. 5.2 : Aire du parallélogramme : $|\det(A)| = 5$.

5.9.3 Cas de l'espace : volume du parallélépipède

Dans $\mathbb{R}^3$, le volume du parallélépipède engendré par trois vecteurs $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$ est :

$$\text{Volume} = |\det(\mathbf{u} \mid \mathbf{v} \mid \mathbf{w})|.$$

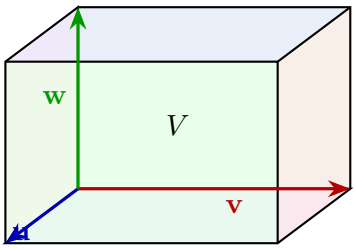


Fig. 5.3 : Volume du parallélépipède : $V = |\det(\mathbf{u} \mid \mathbf{v} \mid \mathbf{w})|$.

5.9.4 Orientation

Définition 41 (Orientation d'une base)

Soit $\mathcal{B} = (\mathbf{v}_1, \dots, \mathbf{v}_n)$ une base de $\mathbb{R}^n$. On dit que $\mathcal{B}$ est *orientée positivement* (ou *directe*) si $\det(\mathbf{v}_1 \mid \cdots \mid \mathbf{v}_n) > 0$, et *orientée négativement* (ou *indirecte*) sinon.

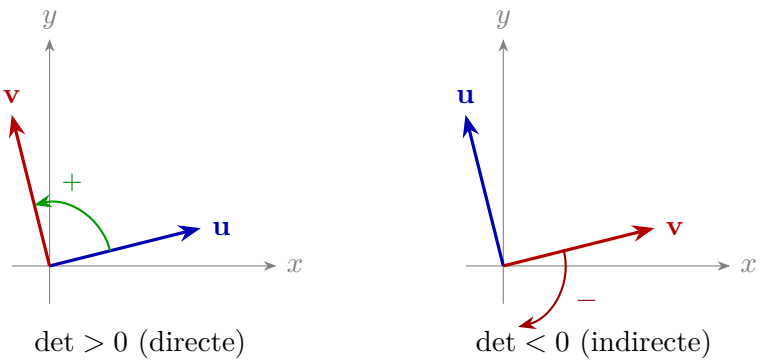


Fig. 5.4 : Orientation du plan : le signe du déterminant distingue les bases directes des bases indirectes.

5.10 Applications

5.10.1 Produit vectoriel dans $\mathbb{R}^3$

Définition 42 (Produit vectoriel)

Soient $\mathbf{u} = (u_1, u_2, u_3)$ et $\mathbf{v} = (v_1, v_2, v_3)$ dans $\mathbb{R}^3$. Le *produit vectoriel* $\mathbf{u} \times \mathbf{v}$ est le vecteur :

$$\mathbf{u} \times \mathbf{v} := \begin{pmatrix} u_2 v_3 - u_3 v_2 \\ u_3 v_1 - u_1 v_3 \\ u_1 v_2 - u_2 v_1 \end{pmatrix}.$$

On peut l'écrire symboliquement comme le développement selon la première ligne du déterminant formel :

$$\mathbf{u} \times \mathbf{v} = \det \begin{pmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{pmatrix}.$$

Proposition 43 (Propriétés du produit vectoriel)

Pour $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^3$ et $\lambda \in \mathbb{R}$:

(i) $\mathbf{u} \times \mathbf{v} = -(\mathbf{v} \times \mathbf{u})$ (antisymétrie) ;

(ii) $\mathbf{u} \times \mathbf{v}$ est orthogonal à $\mathbf{u}$ et à $\mathbf{v}$;

(iii) $\|\mathbf{u} \times \mathbf{v}\| = \|\mathbf{u}\| \, \|\mathbf{v}\| \, |\sin \theta|$ où θ est l'angle entre $\mathbf{u}$ et $\mathbf{v}$;

(iv) $\mathbf{u} \times \mathbf{v} = \mathbf{0}$ si et seulement si $\mathbf{u}$ et $\mathbf{v}$ sont colinéaires.

Remarque 44 (Produit mixte)

Le *produit mixte* $[\mathbf{u}, \mathbf{v}, \mathbf{w}] := \langle \mathbf{u} \times \mathbf{v}, \mathbf{w} \rangle$ est égal à $\det(\mathbf{u} \mid \mathbf{v} \mid \mathbf{w})$. Son module est le volume du parallélépipède construit sur $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$.

5.10.2 Formule de l'aire d'un triangle

Proposition 45 (Aire d'un triangle)

L'aire du triangle de sommets $P_1 = (x_1, y_1)$, $P_2 = (x_2, y_2)$, $P_3 = (x_3, y_3)$ dans $\mathbb{R}^2$ est :

$$\text{Aire} = \frac{1}{2} \left| \det \begin{pmatrix} x_2 - x_1 & x_3 - x_1 \\ y_2 - y_1 & y_3 - y_1 \end{pmatrix} \right| = \frac{1}{2} |(x_2 - x_1)(y_3 - y_1) - (x_3 - x_1)(y_2 - y_1)|.$$

5.10.3 Déterminant de Vandermonde

Théorème 46 (Déterminant de Vandermonde)

Soient $x_1, x_2, \dots, x_n \in \mathbb{K}$. Le *déterminant de Vandermonde* est :

$$V(x_1, \dots, x_n) := \det \begin{pmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \\ x_1^2 & x_2^2 & \cdots & x_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{n-1} & x_2^{n-1} & \cdots & x_n^{n-1} \end{pmatrix} = \prod_{1 \le i < j \le n} (x_j - x_i).$$

En particulier, $V(x_1, \dots, x_n) \neq 0$ si et seulement si les x_i sont deux à deux distincts.

Démonstration. Procédons par récurrence sur n.

Cas $n = 1$ *:* $V(x_1) = 1$ et le produit vide vaut 1.

Cas $n = 2$ *:* $V(x_1, x_2) = \det \begin{pmatrix} 1 & 1 \\ x_1 & x_2 \end{pmatrix} = x_2 - x_1$.

Hérédité : Supposons le résultat pour $n-1$. On effectue les opérations $C_j \leftarrow C_j - x_1\, C_{j-1}$ pour $j = n, n-1, \dots, 2$ (en partant de la dernière colonne). Cela ne change pas le déterminant et crée des zéros dans la première ligne sauf en position $(1,1)$. Plus précisément, après ces opérations la première ligne devient $(1, 0, 0, \dots, 0)$ et le coefficient en position (i, j) pour $i \ge 2$, $j \ge 2$ devient $x_j^{i-1} - x_1 \cdot x_j^{i-2} = x_j^{i-2}(x_j - x_1)$.

En développant selon la première ligne :

$$V(x_1, \dots, x_n) = \det\big((x_j - x_1) \cdot x_j^{i-2}\big)_{2 \le i \le n,\, 2 \le j \le n}.$$

En factorisant $(x_j - x_1)$ dans chaque colonne $j \ge 2$:

$$V(x_1, \dots, x_n) = \prod_{j=2}^{n} (x_j - x_1) \cdot V(x_2, \dots, x_n).$$

Par hypothèse de récurrence, $V(x_2, \dots, x_n) = \prod_{2 \le i < j \le n} (x_j - x_i)$, d'où :

$$V(x_1, \dots, x_n) = \prod_{j=2}^{n} (x_j - x_1) \cdot \prod_{2 \le i < j \le n} (x_j - x_i) = \prod_{1 \le i < j \le n} (x_j - x_i). \qquad \square$$

Remarque 47 (Application du déterminant de Vandermonde)

Le déterminant de Vandermonde intervient notamment dans :

- l'interpolation de Lagrange (existence et unicité du polynôme interpolateur) ;
- la théorie des polynômes symétriques ;
- la preuve que des scalaires distincts donnent des vecteurs propres linéairement indépendants.

5.11 Exercices

Exercice 48 (Calcul de déterminants 2×2 et 3×3)

Calculer les déterminants suivants :

(a) $\det \begin{pmatrix} 3 & 7 \\ 2 & 5 \end{pmatrix}$.

(b) $\det \begin{pmatrix} -1 & 4 \\ 6 & -2 \end{pmatrix}$.

(c) $\det \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$.

(d) $\det \begin{pmatrix} 2 & -1 & 0 \\ 3 & 1 & -2 \\ 1 & 4 & 3 \end{pmatrix}$.

Exercice 49 (Signature de permutations)

(a) Déterminer la signature de $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 4 & 2 \end{pmatrix}$.

(b) Décomposer σ en produit de transpositions et vérifier.

(c) Combien y a-t-il de permutations paires dans $\mathfrak{S}_4$?

Exercice 50 (Matrices triangulaires et diagonales)

(a) Calculer $\det \begin{pmatrix} 2 & 3 & -1 \\ 0 & 5 & 4 \\ 0 & 0 & -3 \end{pmatrix}$.

(b) Montrer que $\det(\lambda A) = \lambda^n \det(A)$ pour $A \in \mathcal{M}_n(\mathbb{K})$.

(c) Calculer $\det(A)$ pour $A = \operatorname{diag}(d_1, d_2, \dots, d_n)$.

Exercice 51 (Opérations élémentaires)

Soit $A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 7 \\ 3 & 7 & 11 \end{pmatrix}$.

(a) En effectuant des opérations élémentaires sur les lignes, réduire A sous forme triangulaire et en déduire $\det(A)$.

(b) Vérifier le résultat par un développement selon la première colonne.

Exercice 52 (Déterminant d'un produit)

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

(a) Calculer $\det(A)$, $\det(B)$ et $\det(AB)$ séparément.

(b) Vérifier que $\det(AB) = \det(A) \cdot \det(B)$.

(c) A-t-on $\det(A + B) = \det(A) + \det(B)$? Justifier.

Exercice 53 (Comatrice et inverse)

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

(a) Calculer tous les cofacteurs A_{ij}.

(b) En déduire la comatrice $\mathrm{Com}(A)$ et la matrice adjointe $\widetilde{A}$.

(c) Vérifier que $A \cdot \widetilde{A} = \det(A) \cdot I_3$.

(d) Calculer A^{-1}.

Exercice 54 (Règle de Cramer)

Résoudre les systèmes suivants par la règle de Cramer :

(a) $\begin{cases} 3x + 2y = 7 \\ x - y = 1 \end{cases}$.

(b) $\begin{cases} x + y + z = 6 \\ 2x - y + z = 3 \\ x + 2y - z = 2 \end{cases}$.

Exercice 55 (Développement selon une ligne ou colonne)

Calculer le déterminant de la matrice suivante par développement selon la colonne la plus avantageuse :

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

Exercice 56 (Déterminant et paramètre)

Soit $A(\lambda) = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 1 & 0 & \lambda \end{pmatrix}$.

(a) Calculer $\det(A(\lambda))$ en fonction de λ.

(b) Pour quelles valeurs de λ la matrice $A(\lambda)$ est-elle inversible ?

(c) Pour $\lambda = 2$, calculer $A(2)^{-1}$ par la méthode de la comatrice.

Exercice 57 (Déterminant de Vandermonde)

(a) Calculer $V(1,2,3) = \det \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 4 & 9 \end{pmatrix}$ directement, puis vérifier avec la formule $\prod_{i<j}(x_j - x_i)$.

(b) Montrer que si $P \in \mathbb{K}_n[X]$ admet $n+1$ racines distinctes, alors $P = 0$.
Indication : utiliser le déterminant de Vandermonde.

Exercice 58 (Formule de Leibniz et antisymétrie)

(a) Montrer directement à partir de la formule de Leibniz que si A a deux colonnes identiques, alors $\det(A) = 0$.
Indication : apparier les permutations σ et $\sigma \circ (j\ k)$.

(b) En déduire que le déterminant est une forme alternée.

(c) Montrer que si $\mathrm{car}(\mathbb{K}) \neq 2$, une forme multilinéaire antisymétrique est alternée. Donner un contre-exemple en caractéristique 2.

Exercice 59 (Déterminant par blocs)

(a) Soit $M = \begin{pmatrix} A & B \\ 0 & D \end{pmatrix}$ avec $A \in \mathcal{M}_p(\mathbb{K})$ et $D \in \mathcal{M}_q(\mathbb{K})$. Montrer que $\det(M) = \det(A) \cdot \det(D)$.

(b) La formule $\det \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det(A)\cdot\det(D) - \det(B)\cdot\det(C)$ est-elle vraie en général ? Donner un contre-exemple.

(c) Montrer que si A est inversible, alors $\det \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det(A) \cdot \det(D - CA^{-1}B)$.
Indication : multiplier par $\begin{pmatrix} I & 0 \\ -CA^{-1} & I \end{pmatrix}$ à gauche.

Exercice 60 (Déterminant d'un endomorphisme)

Soit $E = \mathbb{R}_2[X]$ (polynômes de degré ≤ 2) et $f : E \to E$ défini par $f(P) = P(X+1) - P(X)$.

(a) Montrer que f est un endomorphisme de E.

(b) Écrire la matrice de f dans la base $\mathcal{B} = (1, X, X^2)$.

(c) Calculer $\det(f)$. L'endomorphisme f est-il un automorphisme ?

(d) Déterminer $\mathrm{Ker}(f)$ et $\mathrm{Im}(f)$.

Exercice 61 (Déterminant de matrices particulières)

(a) Calculer le déterminant de la matrice $n \times n$ dont tous les coefficients valent 1 (matrice de uns).

Indication : soustraire la première ligne aux autres.

(b) Calculer le déterminant de la matrice $n \times n$: $A = (a_{ij})$ avec $a_{ij} = \min(i, j)$.

Indication : effectuer $L_i \leftarrow L_i - L_{i-1}$ pour $i = n, n-1, \ldots, 2$.

(c) Soit J_n la matrice $n \times n$ avec $a_{ij} = 1$ pour tout i, j, et $A = \alpha I_n + \beta J_n$. Calculer $\det(A)$ en fonction de α, β et n.

Indication : utiliser $J_n^2 = nJ_n$.

Résumé du chapitre

Ce qu'il faut retenir.

1. **Permutations.** Le groupe symétrique $\mathfrak{S}_n$ contient $n!$ éléments. Chaque permutation est un produit de transpositions, et sa signature $\mathrm{sgn}(\sigma) \in \{+1, -1\}$ est bien définie.
2. **Formule de Leibniz.** $\det(A) = \sum_{\sigma \in \mathfrak{S}_n} \mathrm{sgn}(\sigma) \prod_{i=1}^{n} a_{i,\sigma(i)}$.
3. **Caractérisation axiomatique.** Le déterminant est l'unique forme multilinéaire alternée normalisée ($\det(I_n) = 1$).
4. **Propriétés clés.**
   - $\det(A^\mathsf{T}) = \det(A)$.
   - $\det(AB) = \det(A) \cdot \det(B)$.
   - $\det(\lambda A) = \lambda^n \det(A)$.
   - $\det(A^{-1}) = (\det(A))^{-1}$.
5. **Opérations élémentaires.** Échange de lignes : facteur -1. Multiplication d'une ligne par λ : facteur λ. Ajout d'un multiple : déterminant inchangé.
6. **Développement de Laplace.** Développement selon la i-ème ligne : $\det(A) = \sum_j (-1)^{i+j} a_{ij} M_{ij}$.
7. **Comatrice et inverse.** $A \cdot \widetilde{A} = \det(A) \cdot I_n$, d'où $A^{-1} = \frac{1}{\det(A)} \widetilde{A}$ si $\det(A) \neq 0$.
8. **Règle de Cramer.** $x_j = \frac{\det(A_j)}{\det(A)}$.
9. **Inversibilité.** A inversible $\Leftrightarrow \det(A) \neq 0$.
10. **Déterminant d'un endomorphisme.** $\det(f)$ est indépendant de la base ; f est un automorphisme ssi $\det(f) \neq 0$.
11. **Interprétation géométrique.** $|\det(A)|$ est le volume du parallélotope ; le signe donne l'orientation.

12. Applications. Produit vectoriel, aire du triangle, déterminant de Vandermonde $V(x_1, \dots, x_n) = \prod_{i<j}(x_j - x_i)$.

Chapitre 6

Valeurs propres, vecteurs propres et diagonalisation

Introduction

L'un des objectifs fondamentaux de l'algèbre linéaire est de comprendre la structure d'un endomorphisme en le ramenant, si possible, à une forme aussi simple que possible. La *diagonalisation* consiste à trouver une base dans laquelle la matrice de l'endomorphisme est diagonale : l'action se réduit alors à de simples multiplications scalaires le long de directions privilégiées.

Cette idée est au cœur de nombreuses applications :

- **Équations différentielles linéaires :** la résolution du système $X'(t) = AX(t)$ se ramène, lorsque A est diagonalisable, à n équations scalaires indépendantes $x_i'(t) = \lambda_i x_i(t)$.
- **Chaînes de Markov :** le comportement asymptotique d'une chaîne de Markov est gouverné par les valeurs propres de sa matrice de transition.
- **Analyse vibratoire :** les modes propres d'un système mécanique correspondent aux vecteurs propres de la matrice de raideur.
- **Analyse en composantes principales (ACP) :** la réduction de dimension repose sur les vecteurs propres de la matrice de covariance.
- **Calcul de puissances de matrices :** si $A = P\,\mathrm{diag}(\lambda_1, \dots, \lambda_n)P^{-1}$, alors $A^k = P\,\mathrm{diag}(\lambda_1^k, \dots, \lambda_n^k)P^{-1}$, ce qui permet de résoudre des récurrences linéaires.

Dans tout ce chapitre, $\mathbb{K}$ désigne un corps commutatif ($\mathbb{R}$ ou $\mathbb{C}$ en pratique), E un $\mathbb{K}$-espace vectoriel de dimension finie $n \geq 1$, et $f \in \mathrm{End}(E)$.

6.1 Valeurs propres et vecteurs propres

Définition 1 (Valeur propre, vecteur propre)

Soit $f \in \mathrm{End}(E)$.

(i) Un scalaire $\lambda \in \mathbb{K}$ est une *valeur propre* de f s'il existe un vecteur $\mathbf{v} \in E \setminus \{\mathbf{0}\}$ tel que

$$f(\mathbf{v}) = \lambda \mathbf{v}.$$

Le vecteur $\mathbf{v}$ est alors appelé *vecteur propre* de f associé à la valeur propre λ.

(ii) De manière équivalente, λ est valeur propre de f si et seulement si $f - \lambda\,\mathrm{Id}$ n'est pas injective, c'est-à-dire si et seulement si $\mathrm{Ker}(f - \lambda\,\mathrm{Id}) \neq \{\mathbf{0}\}$.

Définition 2 (Valeur propre et vecteur propre d'une matrice)

Soit $A \in \mathcal{M}_n(\mathbb{K})$. Un scalaire $\lambda \in \mathbb{K}$ est une *valeur propre* de A s'il existe un vecteur colonne $X \in \mathbb{K}^n \setminus \{\mathbf{0}\}$ tel que $AX = \lambda X$, c'est-à-dire $(A - \lambda I_n)X = 0$. Ceci revient à dire que $\det(A - \lambda I_n) = 0$.

Remarque 3 (Lien endomorphisme–matrice)

Si $\mathcal{B}$ est une base de E et $A = \mathcal{M}_{\mathcal{B}}(f)$, alors λ est valeur propre de f si et seulement si λ est valeur propre de A. De plus, si $\mathbf{v}$ est vecteur propre de f pour λ, alors le vecteur de coordonnées de $\mathbf{v}$ dans $\mathcal{B}$ est vecteur propre de A pour λ.

Exemple 4 (Valeurs propres d'une matrice 2×2)

Soit $A = \begin{pmatrix} 3 & 1 \\ 0 & 2 \end{pmatrix}$. Cherchons les valeurs propres de A.
On résout $\det(A - \lambda I_2) = 0$:

$$\det \begin{pmatrix} 3-\lambda & 1 \\ 0 & 2-\lambda \end{pmatrix} = (3-\lambda)(2-\lambda) = 0.$$

Les valeurs propres sont $\lambda_1 = 3$ et $\lambda_2 = 2$.
Pour $\lambda_1 = 3$: $(A - 3I_2)X = 0$ donne $\begin{pmatrix} 0 & 1 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, soit $y = 0$. Les vecteurs propres associés sont de la forme $\begin{pmatrix} x \\ 0 \end{pmatrix}$, $x \neq 0$.
Pour $\lambda_2 = 2$: $(A - 2I_2)X = 0$ donne $\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, soit $x = -y$. Les vecteurs propres associés sont de la forme $\begin{pmatrix} -y \\ y \end{pmatrix}$, $y \neq 0$.

Proposition 5 (Indépendance linéaire des vecteurs propres)

Soient $\lambda_1, \dots, \lambda_r$ des valeurs propres *distinctes* de f, et soient $\mathbf{v}_1, \dots, \mathbf{v}_r$ des vecteurs propres associés respectivement. Alors la famille $(\mathbf{v}_1, \dots, \mathbf{v}_r)$ est libre.

Démonstration. Par récurrence sur r.
Initialisation : pour $r = 1$, $\mathbf{v}_1 \neq \mathbf{0}$ donc $(\mathbf{v}_1)$ est libre.
Hérédité : supposons le résultat vrai pour $r-1$ vecteurs propres associés à des valeurs propres distinctes. Supposons que

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_r\mathbf{v}_r = \mathbf{0}. \tag{6.1}$$

Appliquons f à cette relation :

$$\alpha_1\lambda_1\mathbf{v}_1 + \alpha_2\lambda_2\mathbf{v}_2 + \cdots + \alpha_r\lambda_r\mathbf{v}_r = \mathbf{0}.$$

En multipliant (6.1) par λ_r et en soustrayant :

$$\alpha_1(\lambda_1 - \lambda_r)\mathbf{v}_1 + \cdots + \alpha_{r-1}(\lambda_{r-1} - \lambda_r)\mathbf{v}_{r-1} = \mathbf{0}.$$

Par hypothèse de récurrence, $(\mathbf{v}_1, \dots, \mathbf{v}_{r-1})$ est libre. Comme les $\lambda_i - \lambda_r \neq 0$ pour $i < r$, on obtient $\alpha_i = 0$ pour $i = 1, \dots, r-1$. En reportant dans (6.1) : $\alpha_r \mathbf{v}_r = \mathbf{0}$, et comme $\mathbf{v}_r \neq \mathbf{0}$, $\alpha_r = 0$. □

Corollaire 6 (Nombre de valeurs propres)

Un endomorphisme d'un espace de dimension n possède au plus n valeurs propres distinctes.

6.2 Sous-espaces propres

Définition 7 (Sous-espace propre)

Soit λ une valeur propre de $f \in \mathrm{End}(E)$. Le *sous-espace propre* associé à λ est

$$E_\lambda(f) := \mathrm{Ker}(f - \lambda \,\mathrm{Id}) = \{\, \mathbf{v} \in E \mid f(\mathbf{v}) = \lambda \mathbf{v} \,\}.$$

Proposition 8 ($E_\lambda(f)$ est un sous-espace vectoriel)

Pour toute valeur propre λ de f, $E_\lambda(f)$ est un sous-espace vectoriel de E, de dimension ≥ 1.

Démonstration. $E_\lambda(f) = \mathrm{Ker}(f - \lambda \,\mathrm{Id})$ est le noyau d'un endomorphisme, donc un sous-espace vectoriel de E. De plus, λ étant valeur propre, il existe $\mathbf{v} \neq \mathbf{0}$ tel que $f(\mathbf{v}) = \lambda \mathbf{v}$, donc $E_\lambda(f) \neq \{\mathbf{0}\}$ et $\dim E_\lambda(f) \geq 1$. □

Proposition 9 (Somme directe des sous-espaces propres)

Soient $\lambda_1, \dots, \lambda_r$ les valeurs propres distinctes de f. Alors la somme $E_{\lambda_1}(f) + E_{\lambda_2}(f) + \cdots + E_{\lambda_r}(f)$ est directe. Autrement dit :

$$\sum_{i=1}^{r} E_{\lambda_i}(f) = \bigoplus_{i=1}^{r} E_{\lambda_i}(f).$$

Démonstration. Il suffit de montrer que si $\mathbf{v}_1 + \cdots + \mathbf{v}_r = \mathbf{0}$ avec $\mathbf{v}_i \in E_{\lambda_i}(f)$, alors $\mathbf{v}_i = \mathbf{0}$ pour tout i. Parmi les $\mathbf{v}_i$ non nuls, chacun est vecteur propre pour λ_i. Par la proposition 5, ces vecteurs propres (associés à des valeurs propres distinctes) forment une famille libre. La relation $\mathbf{v}_1 + \cdots + \mathbf{v}_r = \mathbf{0}$ impose alors que tous les coefficients (ici tous égaux à 1) soient nuls, ce qui n'est possible que si tous les $\mathbf{v}_i$ non nuls n'existent pas, c'est-à-dire $\mathbf{v}_i = \mathbf{0}$ pour tout i. □

Exemple 10 (Sous-espaces propres d'une matrice 3×3)

Soit $A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$. Les valeurs propres sont $\lambda_1 = 2$ et $\lambda_2 = 3$.

Pour $\lambda_1 = 2$: $A - 2I_3 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, donc $E_2(A) = \mathrm{Vect}\left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right)$. On a $\dim E_2(A) = 1$.

Pour $\lambda_2 = 3$: $A - 3I_3 = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$, donc $E_3(A) = \text{Vect}\left(\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right)$. On a $\dim E_3(A) = 1$.
Ici $\dim E_2(A) + \dim E_3(A) = 2 < 3 = \dim E$: la matrice A n'est *pas* diagonalisable, comme nous le verrons plus loin.

6.3 Polynôme caractéristique

Définition 11 (Polynôme caractéristique)

Soit $A \in \mathcal{M}_n(\mathbb{K})$. Le *polynôme caractéristique* de A est le polynôme

$$\chi_A(\lambda) := \det(A - \lambda I_n) \in \mathbb{K}[\lambda].$$

Si $f \in \text{End}(E)$ et $A = \mathcal{M}_{\mathcal{B}}(f)$ pour une base $\mathcal{B}$ de E, on pose $\chi_f := \chi_A$. Ce polynôme ne dépend pas du choix de la base $\mathcal{B}$.

Indépendance du choix de la base. Si $\mathcal{B}'$ est une autre base et P la matrice de passage, alors $\mathcal{M}_{\mathcal{B}'}(f) = P^{-1}AP$, et

$$\det(P^{-1}AP - \lambda I_n) = \det(P^{-1}(A - \lambda I_n)P) = \det(P^{-1})\det(A - \lambda I_n)\det(P) = \det(A - \lambda I_n). \quad \square$$

Proposition 12 (Propriétés du polynôme caractéristique)

Soit $A \in \mathcal{M}_n(\mathbb{K})$. Alors :

(i) χ_A est un polynôme de degré n en λ, de coefficient dominant $(-1)^n$:

$$\chi_A(\lambda) = (-1)^n\lambda^n + (-1)^{n-1}\,\text{tr}(A)\lambda^{n-1} + \cdots + \det(A).$$

(ii) $\chi_A(0) = \det(A)$.

(iii) Le coefficient de $(-1)^{n-1}\lambda^{n-1}$ est $\text{tr}(A)$.

(iv) λ_0 est valeur propre de A si et seulement si $\chi_A(\lambda_0) = 0$.

(v) Deux matrices semblables ont le même polynôme caractéristique.

Démonstration. *(i)* On développe $\det(A - \lambda I_n)$. Le terme de plus haut degré provient du produit diagonal $(a_{11} - \lambda)(a_{22} - \lambda)\cdots(a_{nn} - \lambda)$, ce qui donne $(-\lambda)^n + \text{tr}(A)(-\lambda)^{n-1} + \cdots$, soit $(-1)^n\lambda^n + (-1)^{n-1}\,\text{tr}(A)\lambda^{n-1} + \cdots$.
(ii) $\chi_A(0) = \det(A - 0 \cdot I_n) = \det(A)$.
(iii) Découle de (i).
(iv) λ_0 est valeur propre de A ssi $\text{Ker}(A - \lambda_0 I_n) \neq \{\mathbf{0}\}$ ssi $A - \lambda_0 I_n$ n'est pas inversible ssi $\det(A - \lambda_0 I_n) = 0$ ssi $\chi_A(\lambda_0) = 0$.
(v) Déjà prouvé dans la démonstration de l'indépendance du choix de base. $\square$

Exemple 13 (Polynôme caractéristique — cas 2×2)

Pour $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$:

$$\chi_A(\lambda) = \det \begin{pmatrix} a-\lambda & b \\ c & d-\lambda \end{pmatrix} = (a-\lambda)(d-\lambda) - bc = \lambda^2 - (a+d)\lambda + (ad-bc) = \lambda^2 - \mathrm{tr}(A)\lambda + \det(A).$$

Exemple 14 (Polynôme caractéristique — cas 3×3)

Soit $A = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 3 & 1 \\ 0 & 0 & 2 \end{pmatrix}$ (matrice triangulaire supérieure). Alors :

$$\chi_A(\lambda) = (1-\lambda)(3-\lambda)(2-\lambda) = -(\lambda-1)(\lambda-2)(\lambda-3).$$

Les valeurs propres sont 1, 2 et 3. On retrouve que le polynôme caractéristique d'une matrice triangulaire se lit directement sur la diagonale.

Proposition 15 (Cas d'une matrice triangulaire)

Si A est triangulaire (supérieure ou inférieure) avec des coefficients diagonaux $a_{11}, \dots, a_{nn}$, alors :

$$\chi_A(\lambda) = \prod_{i=1}^{n} (a_{ii} - \lambda).$$

En particulier, les valeurs propres de A sont exactement les coefficients diagonaux.

6.4 Théorème de Cayley-Hamilton

Théorème 16 (Cayley-Hamilton)

Soit $A \in \mathcal{M}_n(\mathbb{K})$ de polynôme caractéristique $\chi_A(\lambda) = (-1)^n \lambda^n + c_{n-1}\lambda^{n-1} + \cdots + c_1\lambda + c_0$. Alors :

$$\chi_A(A) = (-1)^n A^n + c_{n-1}A^{n-1} + \cdots + c_1 A + c_0 I_n = 0.$$

Autrement dit, toute matrice carrée est racine de son propre polynôme caractéristique.

Démonstration. Soit $B(\lambda) := {}^{\mathrm{t}}\mathrm{com}(A - \lambda I_n)$ la matrice adjointe (comatrice transposée) de $A - \lambda I_n$. Par la formule classique de l'inverse (ou de la comatrice) :

$$(A - \lambda I_n)\, B(\lambda) = \det(A - \lambda I_n) \cdot I_n = \chi_A(\lambda) \cdot I_n.$$

Chaque coefficient de $B(\lambda)$ est un cofacteur de $A - \lambda I_n$, donc un polynôme en λ de degré au plus $n-1$. On peut donc écrire :

$$B(\lambda) = B_0 + B_1\lambda + B_2\lambda^2 + \cdots + B_{n-1}\lambda^{n-1},$$

où $B_0, B_1, \dots, B_{n-1} \in \mathcal{M}_n(\mathbb{K})$ sont des matrices constantes (ne dépendant pas de λ).
D'autre part, $\chi_A(\lambda) = c_0 + c_1\lambda + \cdots + c_{n-1}\lambda^{n-1} + (-1)^n\lambda^n$.
En développant $(A - \lambda I_n)\, B(\lambda)$:

$$(A - \lambda I_n)(B_0 + B_1\lambda + \cdots + B_{n-1}\lambda^{n-1}) = \chi_A(\lambda) \cdot I_n.$$

On développe le membre de gauche et on identifie les coefficients de chaque puissance de λ :

$$\begin{aligned}
\lambda^0 : &\quad AB_0 = c_0 I_n, \\
\lambda^1 : &\quad AB_1 - B_0 = c_1 I_n, \\
\lambda^2 : &\quad AB_2 - B_1 = c_2 I_n, \\
&\vdots \\
\lambda^k : &\quad AB_k - B_{k-1} = c_k I_n, \\
&\vdots \\
\lambda^{n-1} : &\quad AB_{n-1} - B_{n-2} = c_{n-1} I_n, \\
\lambda^n : &\quad -B_{n-1} = (-1)^n I_n.
\end{aligned}$$

Multiplions la k-ième relation ($k = 0, 1, \dots, n$) par A^k *à gauche* :

$$\begin{aligned}
A^0 \cdot AB_0 &= c_0 I_n, \\
A^1(AB_1 - B_0) &= c_1 A, \\
A^2(AB_2 - B_1) &= c_2 A^2, \\
&\vdots \\
A^{n-1}(AB_{n-1} - B_{n-2}) &= c_{n-1} A^{n-1}, \\
A^n(-B_{n-1}) &= (-1)^n A^n.
\end{aligned}$$

En sommant toutes ces relations, le membre de gauche se téléscope : chaque terme $A^k B_{k-1}$ apparaît une fois avec le signe $+$ (dans la ligne k) et une fois avec le signe $-$ (dans la ligne $k-1$ sous la forme $A^{k-1} \cdot AB_{k-1} = A^k B_{k-1}$). Il reste :

$$0 = c_0 I_n + c_1 A + c_2 A^2 + \cdots + c_{n-1} A^{n-1} + (-1)^n A^n = \chi_A(A). \qquad \square$$

Exemple 17 (Vérification de Cayley-Hamilton)

Soit $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$. On a $\chi_A(\lambda) = \lambda^2 - 5\lambda - 2$. Vérifions :

$$A^2 - 5A - 2I_2 = \begin{pmatrix} 7 & 10 \\ 15 & 22 \end{pmatrix} - \begin{pmatrix} 5 & 10 \\ 15 & 20 \end{pmatrix} - \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}. \quad \checkmark$$

6.5 Spectre d'un endomorphisme

Définition 18 (Spectre)

Le *spectre* de $f \in \mathrm{End}(E)$ est l'ensemble de ses valeurs propres :

$$\mathrm{Spec}(f) := \{\, \lambda \in \mathbb{K} \mid \chi_f(\lambda) = 0 \,\}.$$

De même, le spectre d'une matrice $A \in \mathcal{M}_n(\mathbb{K})$ est $\mathrm{Spec}(A) := \{\, \lambda \in \mathbb{K} \mid \chi_A(\lambda) = 0 \,\}$.

Remarque 19 (Dépendance en le corps)

Le spectre dépend du corps $\mathbb{K}$. Par exemple, la matrice de rotation $R = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ a pour polynôme caractéristique $\chi_R(\lambda) = \lambda^2 + 1$. Si $\mathbb{K} = \mathbb{R}$, alors $\mathrm{Spec}(R) = \emptyset$; si $\mathbb{K} = \mathbb{C}$, alors

$\mathrm{Spec}(R) = \{i, -i\}$.

6.6 Multiplicité algébrique et multiplicité géométrique

Définition 20 (Multiplicité algébrique)

Soit λ une valeur propre de f. La *multiplicité algébrique* de λ, notée $m_a(\lambda)$, est l'ordre de multiplicité de λ comme racine de χ_f. Autrement dit, $m_a(\lambda)$ est le plus grand entier k tel que $(\lambda_0 - \lambda)^k$ divise $\chi_f(\lambda)$ (où λ_0 est la variable du polynôme), c'est-à-dire :

$$\chi_f(\lambda_0) = (\lambda_0 - \lambda)^{m_a(\lambda)}\, Q(\lambda_0), \quad Q(\lambda) \neq 0.$$

Définition 21 (Multiplicité géométrique)

La *multiplicité géométrique* de la valeur propre λ est la dimension du sous-espace propre associé :

$$m_g(\lambda) := \dim E_\lambda(f) = \dim \mathrm{Ker}(f - \lambda\,\mathrm{Id}).$$

Théorème 22 (Encadrement des multiplicités)

Pour toute valeur propre λ de f :

$$1 \leq m_g(\lambda) \leq m_a(\lambda).$$

Démonstration. L'inégalité $m_g(\lambda) \geq 1$ résulte de la définition : λ étant valeur propre, $E_\lambda(f) \neq \{\mathbf{0}\}$ donc $\dim E_\lambda(f) \geq 1$.

Montrons que $m_g(\lambda) \leq m_a(\lambda)$. Posons $p = m_g(\lambda)$ et choisissons une base $(\mathbf{v}_1, \dots, \mathbf{v}_p)$ de $E_\lambda(f)$. Complétons-la en une base $\mathcal{B} = (\mathbf{v}_1, \dots, \mathbf{v}_p, \mathbf{w}_{p+1}, \dots, \mathbf{w}_n)$ de E. Dans cette base, la matrice de f s'écrit :

$$\mathcal{M}_{\mathcal{B}}(f) = \begin{pmatrix} \lambda I_p & C \\ 0 & D \end{pmatrix}$$

où $C \in \mathcal{M}_{p,n-p}(\mathbb{K})$ et $D \in \mathcal{M}_{n-p}(\mathbb{K})$. Alors :

$$\chi_f(\mu) = \det \begin{pmatrix} (\lambda - \mu) I_p & C \\ 0 & D - \mu I_{n-p} \end{pmatrix} = (\lambda - \mu)^p \det(D - \mu I_{n-p}).$$

Donc $(\lambda - \mu)^p$ divise $\chi_f(\mu)$, ce qui donne $m_a(\lambda) \geq p = m_g(\lambda)$. □

Exemple 23 (Multiplicités distinctes)

Pour $A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$ (cf. exemple 10), on a $\chi_A(\lambda) = -(2-\lambda)^2(3-\lambda)$.

- Pour $\lambda = 2$: $m_a(2) = 2$ et $m_g(2) = \dim E_2(A) = 1$. Ici $m_g < m_a$.
- Pour $\lambda = 3$: $m_a(3) = 1$ et $m_g(3) = \dim E_3(A) = 1$. Ici $m_g = m_a$.

6.7 Diagonalisabilité

Définition 24 (Endomorphisme diagonalisable)

Un endomorphisme $f \in \mathrm{End}(E)$ est dit *diagonalisable* s'il existe une base de E formée de vecteurs propres de f. De manière équivalente, f est diagonalisable s'il existe une base $\mathcal{B}$ de E telle que $\mathcal{M}_{\mathcal{B}}(f)$ soit une matrice diagonale.

Définition 25 (Matrice diagonalisable)

Une matrice $A \in \mathcal{M}_n(\mathbb{K})$ est dite *diagonalisable* s'il existe $P \in \mathrm{GL}_n(\mathbb{K})$ telle que $P^{-1}AP$ soit diagonale. Autrement dit, A est semblable à une matrice diagonale.

Remarque 26 (Diagonalisation concrète)

Si f est diagonalisable et si $\mathcal{B} = (\mathbf{v}_1, \dots, \mathbf{v}_n)$ est une base de vecteurs propres avec $f(\mathbf{v}_i) = \lambda_i \mathbf{v}_i$, alors :

$$\mathcal{M}_{\mathcal{B}}(f) = \mathrm{diag}(\lambda_1, \dots, \lambda_n) = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

Si $A = \mathcal{M}_{\mathcal{E}}(f)$ dans la base canonique et P est la matrice dont les colonnes sont les coordonnées des $\mathbf{v}_i$, alors $A = P\,\mathrm{diag}(\lambda_1, \dots, \lambda_n)P^{-1}$.

Théorème 27 (Critère de diagonalisabilité)

Soit $f \in \mathrm{End}(E)$ avec $\dim E = n$. Notons $\lambda_1, \dots, \lambda_r$ les valeurs propres distinctes de f. Les assertions suivantes sont équivalentes :

(i) f est diagonalisable.

(ii) $E = \bigoplus_{i=1}^{r} E_{\lambda_i}(f)$.

(iii) $\sum_{i=1}^{r} \dim E_{\lambda_i}(f) = n$.

(iv) Le polynôme caractéristique χ_f est scindé sur $\mathbb{K}$ *et* $m_g(\lambda_i) = m_a(\lambda_i)$ pour tout $i \in \{1, \dots, r\}$.

Démonstration. *(i) ⇒ (ii) :* Si f est diagonalisable, il existe une base $\mathcal{B} = (\mathbf{v}_1, \dots, \mathbf{v}_n)$ de vecteurs propres. Chaque $\mathbf{v}_j$ appartient à un sous-espace propre $E_{\lambda_{i(j)}}(f)$. Comme $\mathcal{B}$ engendre E, on a $E = E_{\lambda_1}(f) + \cdots + E_{\lambda_r}(f)$. La somme est directe par la proposition 9.

(ii) ⇒ (iii) : Immédiat : si la somme directe vaut E, la somme des dimensions vaut n.

(iii) ⇒ (iv) : On a $\sum_i m_g(\lambda_i) = n$. Or $\sum_i m_a(\lambda_i) \leq n$ (car $\deg \chi_f = n$) et $m_g(\lambda_i) \leq m_a(\lambda_i)$ par le théorème 22. Donc :

$$n = \sum_{i=1}^{r} m_g(\lambda_i) \leq \sum_{i=1}^{r} m_a(\lambda_i) \leq n.$$

Toutes les inégalités sont des égalités. En particulier, $\sum_i m_a(\lambda_i) = n$, ce qui signifie que χ_f est scindé sur $\mathbb{K}$ (toutes les racines sont dans $\mathbb{K}$ et leurs multiplicités totalisent $n = \deg \chi_f$). De plus, $m_g(\lambda_i) = m_a(\lambda_i)$ pour tout i.

(iv) ⇒ (i) : Si χ_f est scindé et $m_g(\lambda_i) = m_a(\lambda_i)$ pour tout i, alors

$$\sum_{i=1}^{r} \dim E_{\lambda_i}(f) = \sum_{i=1}^{r} m_g(\lambda_i) = \sum_{i=1}^{r} m_a(\lambda_i) = n.$$

Par la proposition 9, la somme des sous-espaces propres est directe et de dimension n, donc $E = \bigoplus_i E_{\lambda_i}(f)$. En prenant la réunion de bases de chaque $E_{\lambda_i}(f)$, on obtient une base de E formée de vecteurs propres. □

Corollaire 28 (n valeurs propres distinctes)

Si f possède $n = \dim E$ valeurs propres distinctes, alors f est diagonalisable.

Démonstration. Si les n valeurs propres sont distinctes, chacune a multiplicité algébrique 1, donc multiplicité géométrique 1 (par l'encadrement). Le critère (iv) du théorème 27 est vérifié. □

6.8 Algorithme de diagonalisation

Voici la procédure à suivre pour diagonaliser un endomorphisme f (ou une matrice A) :

Étape 1. Calculer le polynôme caractéristique $\chi_f(\lambda) = \det(A - \lambda I_n)$.

Étape 2. Trouver les racines de χ_f dans $\mathbb{K}$. Ce sont les valeurs propres $\lambda_1, \dots, \lambda_r$ avec leurs multiplicités algébriques $m_a(\lambda_i)$.

Étape 3. Vérifier que χ_f est scindé : la somme des multiplicités algébriques doit valoir n. Sinon, f n'est pas diagonalisable (sur $\mathbb{K}$).

Étape 4. Calculer les sous-espaces propres $E_{\lambda_i} = \mathrm{Ker}(A - \lambda_i I_n)$ en résolvant les systèmes linéaires.

Étape 5. Vérifier que $\dim E_{\lambda_i} = m_a(\lambda_i)$ pour tout i. Si c'est le cas, f est diagonalisable ; sinon, non.

Étape 6. Former la matrice de passage P en juxtaposant les bases des sous-espaces propres. On obtient $P^{-1}AP = \mathrm{diag}(\lambda_1, \dots, \lambda_1, \lambda_2, \dots, \lambda_r)$.

Exemple 29 (Diagonalisation complète)

Diagonalisons la matrice $A = \begin{pmatrix} 4 & 1 \\ 2 & 3 \end{pmatrix}$.

Étape 1. $\chi_A(\lambda) = \det \begin{pmatrix} 4-\lambda & 1 \\ 2 & 3-\lambda \end{pmatrix} = (4-\lambda)(3-\lambda) - 2 = \lambda^2 - 7\lambda + 10 = (\lambda - 5)(\lambda - 2)$.

Étape 2. Les valeurs propres sont $\lambda_1 = 5$ et $\lambda_2 = 2$, chacune de multiplicité algébrique 1.

Étape 3. χ_A est scindé sur $\mathbb{R}$.

Étape 4.

- E_5 : $(A - 5I_2)X = 0$, soit $\begin{pmatrix} -1 & 1 \\ 2 & -2 \end{pmatrix} X = 0$, d'où $x = y$. Donc $E_5 = \mathrm{Vect}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$.
- E_2 : $(A - 2I_2)X = 0$, soit $\begin{pmatrix} 2 & 1 \\ 2 & 1 \end{pmatrix} X = 0$, d'où $2x + y = 0$. Donc $E_2 = \mathrm{Vect}\begin{pmatrix} 1 \\ -2 \end{pmatrix}$.

Étape 5. $\dim E_5 = 1 = m_a(5)$ et $\dim E_2 = 1 = m_a(2)$. La matrice est diagonalisable.

Étape 6. $P = \begin{pmatrix} 1 & 1 \\ 1 & -2 \end{pmatrix}$, $P^{-1} = \frac{1}{-3}\begin{pmatrix} -2 & -1 \\ -1 & 1 \end{pmatrix}$, et

$$P^{-1}AP = \begin{pmatrix} 5 & 0 \\ 0 & 2 \end{pmatrix}.$$

Exemple 30 (Diagonalisation d'une matrice 3×3)

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

Étape 1. $\chi_A(\lambda) = \det \begin{pmatrix} 1-\lambda & 1 & 0 \\ 0 & 2-\lambda & 0 \\ 0 & 1 & 3-\lambda \end{pmatrix} = (1-\lambda)(2-\lambda)(3-\lambda)$.

Étape 2. Les valeurs propres sont $\lambda_1 = 1$, $\lambda_2 = 2$, $\lambda_3 = 3$, chacune simple.

Étape 3. χ_A est scindé sur $\mathbb{R}$ et les trois valeurs propres sont distinctes : A est diagonalisable (corollaire 28).

Étape 4.

- E_1 : $(A - I_3)X = 0$ avec $A - I_3 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 2 \end{pmatrix}$, d'où $y = 0$, $z = 0$. Donc $E_1 = \text{Vect}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$.
- E_2 : $(A - 2I_3)X = 0$ avec $A - 2I_3 = \begin{pmatrix} -1 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix}$, d'où $x = y$ et $y = -z$. Donc $E_2 = \text{Vect}\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$.
- E_3 : $(A - 3I_3)X = 0$ avec $A - 3I_3 = \begin{pmatrix} -2 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 1 & 0 \end{pmatrix}$, d'où $y = 0$ et $x = 0$. Donc $E_3 = \text{Vect}\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$.

  Vérifions : $A\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 3 \end{pmatrix} = 3\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$.

Étape 6. $P = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix}$ et $P^{-1}AP = \text{diag}(1, 2, 3)$.

6.9 Trigonalisation

Lorsqu'un endomorphisme n'est pas diagonalisable, on peut chercher la meilleure forme simple suivante : une matrice triangulaire.

Définition 31 (Endomorphisme trigonalisable)

Un endomorphisme $f \in \mathrm{End}(E)$ est dit *trigonalisable* s'il existe une base de E dans laquelle la matrice de f est triangulaire supérieure. Une matrice A est trigonalisable si elle est semblable à une matrice triangulaire supérieure.

Théorème 32 (Critère de trigonalisation)

Un endomorphisme $f \in \mathrm{End}(E)$ est trigonalisable si et seulement si son polynôme caractéristique χ_f est scindé sur $\mathbb{K}$.
En particulier, sur $\mathbb{K} = \mathbb{C}$, **toute** matrice est trigonalisable (car tout polynôme est scindé sur $\mathbb{C}$ d'après le théorème fondamental de l'algèbre).

Démonstration. ($\Rightarrow$) Si f est trigonalisable, il existe une base dans laquelle sa matrice est triangulaire supérieure avec les coefficients diagonaux $t_{11}, \dots, t_{nn}$. Par la proposition 15, $\chi_f(\lambda) = \prod_{i=1}^{n}(t_{ii} - \lambda)$, qui est scindé sur $\mathbb{K}$.

($\Leftarrow$) Par récurrence sur $n = \dim E$.

Initialisation : si $n = 1$, toute matrice 1×1 est triangulaire.

Hérédité : soit $n \geq 2$. Comme χ_f est scindé, f admet au moins une valeur propre $\lambda_1 \in \mathbb{K}$. Soit $\mathbf{v}_1 \neq \mathbf{0}$ un vecteur propre associé. Soit H un supplémentaire de $\mathrm{Vect}(\mathbf{v}_1)$ dans E (un sous-espace tel que $E = \mathrm{Vect}(\mathbf{v}_1) \oplus H$, de dimension $n-1$). Posons $\mathcal{B}_1 = (\mathbf{v}_1, \mathbf{h}_2, \dots, \mathbf{h}_n)$ une base de E adaptée à cette décomposition. Dans cette base :

$$\mathcal{M}_{\mathcal{B}_1}(f) = \begin{pmatrix} \lambda_1 & * \\ 0 & A' \end{pmatrix}$$

avec $A' \in \mathcal{M}_{n-1}(\mathbb{K})$. Alors $\chi_f(\lambda) = (\lambda_1 - \lambda)\chi_{A'}(\lambda)$, et comme χ_f est scindé, $\chi_{A'}$ l'est aussi. Par hypothèse de récurrence, A' est trigonalisable : il existe $Q \in \mathrm{GL}_{n-1}(\mathbb{K})$ telle que $Q^{-1}A'Q$ soit triangulaire supérieure. En posant $\tilde{P} = \begin{pmatrix} 1 & 0 \\ 0 & Q \end{pmatrix}$, on obtient une matrice triangulaire supérieure pour f dans la base correspondante. □

Corollaire 33 (Trigonalisation sur $\mathbb{C}$)

Toute matrice $A \in \mathcal{M}_n(\mathbb{C})$ est trigonalisable : il existe $P \in \mathrm{GL}_n(\mathbb{C})$ et une matrice triangulaire supérieure T telles que $A = PTP^{-1}$. Les coefficients diagonaux de T sont les valeurs propres de A (comptées avec multiplicité).

6.10 Polynôme minimal

Définition 34 (Polynôme minimal)

Soit $f \in \mathrm{End}(E)$. Le *polynôme minimal* de f, noté μ_f, est le polynôme unitaire de plus petit degré tel que $\mu_f(f) = 0$ (l'endomorphisme nul).
De même, le polynôme minimal de $A \in \mathcal{M}_n(\mathbb{K})$ est le polynôme unitaire μ_A de plus petit degré tel que $\mu_A(A) = 0$.

Proposition 35 (Existence et unicité)

Le polynôme minimal existe et est unique. De plus, un polynôme $P \in \mathbb{K}[X]$ vérifie $P(f) = 0$ si et seulement si μ_f divise P.

Démonstration. L'espace $\mathrm{End}(E)$ est de dimension n^2, donc la famille $(\mathrm{Id}, f, f^2, \dots, f^{n^2})$ de n^2+1 éléments est liée : il existe une relation non triviale $\sum_{k=0}^{n^2} a_k f^k = 0$. Ceci montre l'existence d'un polynôme annulateur non nul (et le théorème de Cayley-Hamilton en fournit un de degré n). Parmi tous les polynômes annulateurs non nuls, prenons-en un de degré minimal et rendons-le unitaire : c'est μ_f.

Unicité : si μ' est un autre polynôme unitaire annulateur de degré minimal, alors $\mu_f - \mu'$ est annulateur de degré strictement inférieur à $\deg \mu_f$, donc $\mu_f - \mu' = 0$.

Divisibilité : si $P(f) = 0$, effectuons la division euclidienne $P = Q\mu_f + R$ avec $\deg R < \deg \mu_f$. Alors $R(f) = P(f) - Q(f)\mu_f(f) = 0 - 0 = 0$. Par minimalité de $\deg \mu_f$, on a $R = 0$, donc $\mu_f \mid P$. □

Théorème 36 (Relations entre μ_f et χ_f)

(i) μ_f divise χ_f (conséquence du théorème de Cayley-Hamilton).

(ii) μ_f et χ_f ont les mêmes racines (dans une clôture algébrique de $\mathbb{K}$).

(iii) $\deg \mu_f \leq n = \deg \chi_f$.

Démonstration. *(i)* Par le théorème 16, $\chi_f(f) = 0$, donc $\mu_f \mid \chi_f$ par la proposition 35.

(ii) Toute racine de μ_f est racine de χ_f (car $\mu_f \mid \chi_f$). Réciproquement, soit λ une racine de χ_f : c'est une valeur propre de f. Soit $\mathbf{v} \neq \mathbf{0}$ un vecteur propre associé. Alors $\mu_f(f)\mathbf{v} = \mu_f(\lambda)\mathbf{v} = \mathbf{0}$. Comme $\mathbf{v} \neq \mathbf{0}$, on a $\mu_f(\lambda) = 0$.

(iii) Conséquence immédiate de (i). □

Théorème 37 (Polynôme minimal et diagonalisabilité)

L'endomorphisme f est diagonalisable si et seulement si μ_f est scindé à racines simples sur $\mathbb{K}$, c'est-à-dire :

$$\mu_f(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_r)$$

où $\lambda_1, \dots, \lambda_r$ sont les valeurs propres distinctes de f.

Démonstration. ($\Rightarrow$) Supposons f diagonalisable. Alors il existe une base dans laquelle $\mathcal{M}(f) = \mathrm{diag}(\lambda_1, \dots, \lambda_1, \lambda_2, \dots, \lambda_r)$. Posons $P(\lambda) = (\lambda - \lambda_1) \cdots (\lambda - \lambda_r)$ (racines simples). On a $P(f) = (f - \lambda_1 \,\mathrm{Id}) \circ \cdots \circ (f - \lambda_r \,\mathrm{Id})$. Dans la base de diagonalisation, chaque vecteur de base $\mathbf{v}_j$ est vecteur propre pour une certaine valeur propre λ_{i_j}, donc $(f - \lambda_{i_j} \,\mathrm{Id})\mathbf{v}_j = \mathbf{0}$, et par conséquent $P(f)\mathbf{v}_j = \mathbf{0}$. Comme ceci vaut pour tous les vecteurs de la base, $P(f) = 0$. Donc $\mu_f \mid P$. Comme μ_f et χ_f ont les mêmes racines et que P a les mêmes racines que μ_f, le polynôme μ_f est à racines simples. Comme $\mu_f \mid P$ et P est à racines simples, μ_f est aussi à racines simples.

($\Leftarrow$) Supposons $\mu_f = (\lambda - \lambda_1) \cdots (\lambda - \lambda_r)$ scindé à racines simples. Alors $\mu_f(f) = (f - \lambda_1 \,\mathrm{Id}) \circ \cdots \circ (f - \lambda_r \,\mathrm{Id}) = 0$. Par le lemme des noyaux (appliqué aux polynômes deux à deux premiers entre eux $\lambda - \lambda_i$) :

$$E = \bigoplus_{i=1}^{r} \mathrm{Ker}(f - \lambda_i \,\mathrm{Id}) = \bigoplus_{i=1}^{r} E_{\lambda_i}(f).$$

Donc f est diagonalisable. □

Exemple 38 (Polynôme minimal)

Soit $A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$. On a $\chi_A(\lambda) = -(\lambda-2)^2(\lambda-3)$.

- $E_2(A) = \text{Vect}\left(\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}\right)$, de dimension 2.
- $E_3(A) = \text{Vect}\left(\begin{pmatrix}0\\0\\1\end{pmatrix}\right)$.

La matrice A est diagonalisable (c'est déjà une matrice diagonale). Le polynôme minimal est $\mu_A(\lambda) = (\lambda-2)(\lambda-3)$ (racines simples), de degré $2 < 3 = \deg \chi_A$.

6.11 Applications

6.11.1 Calcul de A^k

Proposition 39 (Puissances d'une matrice diagonalisable)

Si $A = P\ \text{diag}(\lambda_1, \dots, \lambda_n)\, P^{-1}$, alors pour tout $k \in \mathbb{N}$:

$$A^k = P\ \text{diag}(\lambda_1^k, \dots, \lambda_n^k)\, P^{-1}.$$

Démonstration. Par récurrence immédiate : $A^k = (PDP^{-1})^k = PD^kP^{-1} = P\,\text{diag}(\lambda_1^k, \dots, \lambda_n^k)P^{-1}$. □

6.11.2 Récurrences linéaires : la suite de Fibonacci

La suite de Fibonacci est définie par $F_0 = 0$, $F_1 = 1$ et $F_{n+2} = F_{n+1} + F_n$. Posons $X_n = \begin{pmatrix} F_{n+1} \\ F_n \end{pmatrix}$. Alors :

$$X_n = AX_{n-1} \quad \text{avec} \quad A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}.$$

Donc $X_n = A^n X_0$ avec $X_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$.

Diagonalisons A : $\chi_A(\lambda) = \lambda^2 - \lambda - 1$, de racines $\varphi = \frac{1+\sqrt{5}}{2}$ (nombre d'or) et $\hat{\varphi} = \frac{1-\sqrt{5}}{2}$.

Les vecteurs propres sont $\mathbf{v}_1 = \begin{pmatrix} \varphi \\ 1 \end{pmatrix}$ et $\mathbf{v}_2 = \begin{pmatrix} \hat{\varphi} \\ 1 \end{pmatrix}$. Avec $P = \begin{pmatrix} \varphi & \hat{\varphi} \\ 1 & 1 \end{pmatrix}$, on a $A = P\,\text{diag}(\varphi, \hat{\varphi})P^{-1}$, d'où :

$$A^n = P \begin{pmatrix} \varphi^n & 0 \\ 0 & \hat{\varphi}^n \end{pmatrix} P^{-1}.$$

En calculant la deuxième composante de $A^n X_0$, on obtient la **formule de Binet** :

$$\boxed{F_n = \frac{\varphi^n - \hat{\varphi}^n}{\sqrt{5}} = \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right].}$$

6.11.3 Chaînes de Markov : état stationnaire

Considérons une chaîne de Markov à deux états avec matrice de transition

$$M = \begin{pmatrix} 0{,}7 & 0{,}4 \\ 0{,}3 & 0{,}6 \end{pmatrix}.$$

(Les colonnes somment à 1 : matrice stochastique.)

On a $\chi_M(\lambda) = (\lambda - 1)(\lambda - 0{,}3)$, de valeurs propres $\lambda_1 = 1$ et $\lambda_2 = 0{,}3$.

Le vecteur propre associé à $\lambda_1 = 1$ (vecteur de probabilités stationnaires) vérifie $M\pi = \pi$, soit $\pi = \frac{1}{7}\begin{pmatrix} 4 \\ 3 \end{pmatrix}$ (normalisé pour que la somme des composantes vaille 1).

Comme $|\lambda_2| = 0{,}3 < 1$, pour toute distribution initiale X_0, la suite $X_n = M^n X_0$ converge vers π : l'état stationnaire est atteint.

6.11.4 Systèmes d'équations différentielles linéaires

Considérons le système $X'(t) = AX(t)$ avec $A = \begin{pmatrix} -1 & 1 \\ 0 & -2 \end{pmatrix}$.

On a $\chi_A(\lambda) = (\lambda + 1)(\lambda + 2)$, de valeurs propres $\lambda_1 = -1$ et $\lambda_2 = -2$.

- E_{-1} : vecteur propre $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$.
- E_{-2} : $(A + 2I)X = 0$ donne $\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} X = 0$, soit $\mathbf{v}_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$.

Dans la base $(\mathbf{v}_1, \mathbf{v}_2)$, le système devient $Y'(t) = \mathrm{diag}(-1, -2)Y(t)$, de solution $y_1(t) = c_1 e^{-t}$, $y_2(t) = c_2 e^{-2t}$. En revenant à la base canonique :

$$X(t) = c_1 e^{-t}\begin{pmatrix} 1 \\ 0 \end{pmatrix} + c_2 e^{-2t}\begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} c_1 e^{-t} - c_2 e^{-2t} \\ c_2 e^{-2t} \end{pmatrix}.$$

6.12 Illustrations

6.13 Exercices

Exercice 40 (Valeurs propres d'une matrice 2×2)

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

(a) Calculer le polynôme caractéristique de A.

(b) Déterminer les valeurs propres et les sous-espaces propres.

(c) La matrice A est-elle diagonalisable ? Si oui, la diagonaliser.

Exercice 41 (Valeurs propres d'une matrice 3×3)

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

(a) Calculer $\chi_A(\lambda)$.

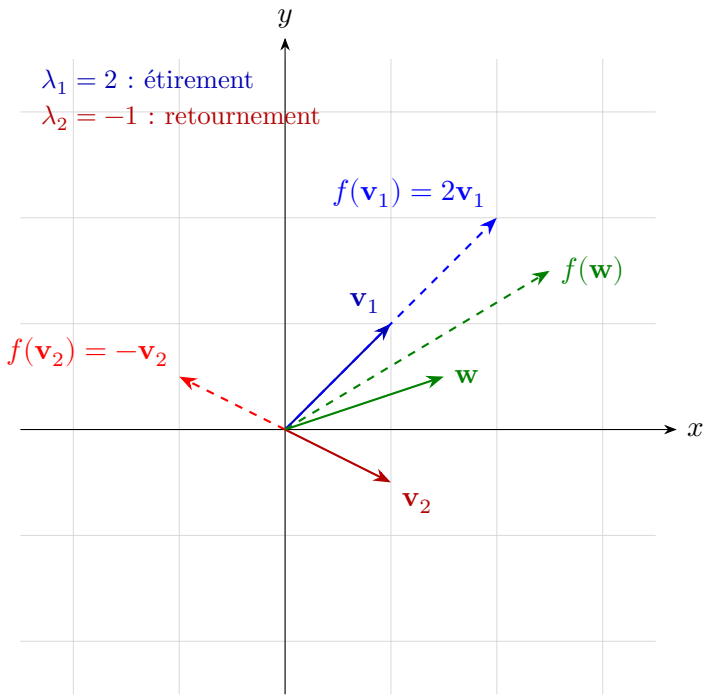


Fig. 6.1 : Action d'un endomorphisme sur ses vecteurs propres : l'endomorphisme étire le long de $\mathbf{v}_1$ (valeur propre 2) et retourne le long de $\mathbf{v}_2$ (valeur propre -1).

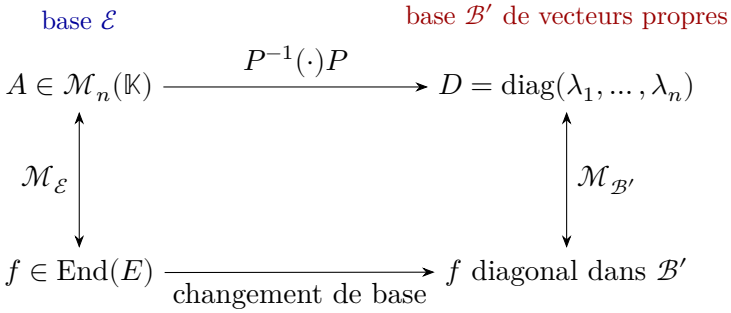


Fig. 6.2 : Diagramme de diagonalisation : le passage de la matrice A (dans la base canonique $\mathcal{E}$) à la matrice diagonale D (dans la base de vecteurs propres $\mathcal{B}'$) via la matrice de passage P.

(b) Déterminer $\mathrm{Spec}(A)$ et les sous-espaces propres.

(c) Diagonaliser A.

Exercice 42 (Polynôme caractéristique et trace/déterminant)

Soit $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M}_2(\mathbb{R})$.

(a) Montrer que $\chi_A(\lambda) = \lambda^2 - \mathrm{tr}(A)\lambda + \det(A)$.

(b) En déduire que si λ_1, λ_2 sont les valeurs propres de A (dans $\mathbb{C}$), alors $\lambda_1 + \lambda_2 = \mathrm{tr}(A)$ et $\lambda_1\lambda_2 = \det(A)$.

(c) Généraliser au cas $n \times n$: montrer que la somme des valeurs propres (comptées avec multiplicité) vaut $\mathrm{tr}(A)$ et leur produit vaut $\det(A)$.

Exercice 43 (Vérification de Cayley-Hamilton)

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

(a) Calculer $\chi_A(\lambda)$.

(b) Calculer A^2, puis vérifier que $\chi_A(A) = 0$.

(c) En déduire une expression de A^{-1} comme polynôme en A.

Exercice 44 (Matrice non diagonalisable)

Soit $A = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$.

(a) Déterminer $\chi_A(\lambda)$.

(b) Trouver les valeurs propres et les sous-espaces propres.

(c) Comparer multiplicités algébrique et géométrique. Conclure sur la diagonalisabilité de A.

(d) Calculer néanmoins A^n pour tout $n \in \mathbb{N}$ en écrivant $A = 2I_2 + N$ avec $N = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ et en utilisant la formule du binôme.

Exercice 45 (Puissance d'une matrice)

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

(a) Diagonaliser A.

(b) En déduire A^n pour tout $n \in \mathbb{N}$.

(c) Calculer A^{10}.

Exercice 46 (Récurrence linéaire)

On considère la suite définie par $u_0 = 1$, $u_1 = 2$ et $u_{n+2} = 5u_{n+1} - 6u_n$.

(a) Écrire cette récurrence sous forme matricielle : $X_{n+1} = AX_n$ avec $X_n = \begin{pmatrix} u_{n+1} \\ u_n \end{pmatrix}$.

(b) Diagonaliser la matrice A.

(c) En déduire l'expression de u_n en fonction de n.

Exercice 47 (Polynôme minimal)

Pour chacune des matrices suivantes, déterminer le polynôme minimal et dire si la matrice est diagonalisable :

(a) $A = \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix}$.

(b) $B = \begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix}$.

(c) $C = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

(d) $D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$.

Exercice 48 (Diagonalisation et système différentiel)

Résoudre le système $\begin{cases} x'(t) = 3x(t) - y(t) \\ y'(t) = -x(t) + 3y(t) \end{cases}$ en diagonalisant la matrice du système.

Exercice 49 (Chaîne de Markov)

Dans une ville, chaque jour, 80% des usagers du bus reprennent le bus et 20% passent au vélo ; 40% des cyclistes reprennent le vélo et 60% passent au bus.

(a) Écrire la matrice de transition M.

(b) Diagonaliser M.

(c) Déterminer le vecteur de probabilités stationnaires.

(d) Si initialement tout le monde prend le bus, quelle est la répartition après n jours ? Trouver la limite quand $n \to +\infty$.

Exercice 50 (Diagonalisation simultanée)

Soient $A, B \in \mathcal{M}_n(\mathbb{K})$ deux matrices diagonalisables telles que $AB = BA$.

(a) Montrer que pour toute valeur propre λ de A, le sous-espace propre $E_\lambda(A)$ est stable par B.

(b) En déduire que A et B sont *simultanément diagonalisables* : il existe $P \in \mathrm{GL}_n(\mathbb{K})$ telle que $P^{-1}AP$ et $P^{-1}BP$ soient toutes deux diagonales.

(c) Application : diagonaliser simultanément $A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ et $B = \begin{pmatrix} 3 & -2 \\ -2 & 3 \end{pmatrix}$.

Exercice 51 (Endomorphisme et polynôme annulateur)

Soit $f \in \mathrm{End}(E)$ avec $\dim E = n$. On suppose que $f^2 = f$ (i.e., f est un projecteur).

(a) Montrer que les seules valeurs propres possibles de f sont 0 et 1.
Indication : si $f(\mathbf{v}) = \lambda \mathbf{v}$, calculer $f^2(\mathbf{v})$.

(b) Montrer que μ_f divise $X^2 - X = X(X-1)$.

(c) En déduire que f est diagonalisable.

(d) Montrer que $E = \operatorname{Ker} f \oplus \operatorname{Im} f$ et interpréter les sous-espaces propres.

Exercice 52 (Cayley-Hamilton et calcul d'inverse)

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

(a) Calculer $\chi_A(\lambda)$.

(b) En utilisant le théorème de Cayley-Hamilton, exprimer A^{-1} comme un polynôme en A (de degré ≤ 2).

(c) Calculer A^{-1} par cette méthode et vérifier le résultat.

Exercice 53 (Trigonalisation)

Soit $A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$.

(a) Montrer que A n'est pas diagonalisable.

(b) Déterminer le polynôme minimal de A.

(c) Calculer A^n en écrivant $A = 2I_3 + N$ où N est nilpotente. Déterminer l'indice de nilpotence de N et utiliser la formule du binôme $(2I_3 + N)^n = \sum_{k=0}^{2} \binom{n}{k} 2^{n-k} N^k$.

Exercice 54 (Spectre et trace)

Soit $A \in \mathcal{M}_n(\mathbb{K})$ de valeurs propres $\lambda_1, \ldots, \lambda_n$ (comptées avec multiplicité, dans une clôture algébrique).

(a) Montrer que $\operatorname{tr}(A^k) = \sum_{i=1}^{n} \lambda_i^k$ pour tout $k \geq 1$.
Indication : trigonaliser A sur la clôture algébrique.

(b) En déduire que si $A^2 = A$, alors $\operatorname{tr}(A) = \operatorname{rg}(A)$.

(c) Montrer que si A est nilpotente, alors $\operatorname{tr}(A^k) = 0$ pour tout $k \geq 1$.

Résumé du chapitre

Ce qu'il faut retenir.

1. **Valeur propre / vecteur propre.** $\lambda \in \mathbb{K}$ est valeur propre de f si $\exists\, \mathbf{v} \neq \mathbf{0}$ tel que $f(\mathbf{v}) = \lambda \mathbf{v}$, i.e. $\operatorname{Ker}(f - \lambda \operatorname{Id}) \neq \{\mathbf{0}\}$.
2. **Sous-espace propre.** $E_\lambda(f) = \operatorname{Ker}(f - \lambda \operatorname{Id})$ est un sous-espace vectoriel. Les sous-espaces propres associés à des valeurs propres distinctes sont en somme directe.
3. **Polynôme caractéristique.** $\chi_f(\lambda) = \det(f - \lambda \operatorname{Id})$, de degré n, avec $\chi_f(0) = \det(f)$ et coefficient de $(-1)^{n-1}\lambda^{n-1}$ égal à $\operatorname{tr}(f)$. Les valeurs propres sont les racines de χ_f.

4. **Cayley-Hamilton.** $\chi_f(f) = 0$ — toute matrice est racine de son polynôme caractéristique.

5. **Multiplicités.** Pour toute valeur propre λ : $1 \leq m_g(\lambda) \leq m_a(\lambda)$.

6. **Diagonalisabilité.** f est diagonalisable ssi χ_f est scindé et $m_g = m_a$ pour chaque valeur propre, ssi μ_f est scindé à racines simples.

7. **Trigonalisation.** f est trigonalisable ssi χ_f est scindé. Sur $\mathbb{C}$, tout endomorphisme est trigonalisable.

8. **Polynôme minimal.** μ_f est le polynôme unitaire annulateur de plus petit degré. Il divise χ_f et a les mêmes racines.

9. **Applications.** Si $A = PDP^{-1}$:
   - $A^k = PD^kP^{-1}$ (puissances).
   - Résolution de récurrences linéaires ($u_{n+k} = \sum a_i u_{n+i}$).
   - Chaînes de Markov : l'état stationnaire est le vecteur propre pour $\lambda = 1$.
   - Systèmes différentiels : $X' = AX$ se résout par diagonalisation.

10. **Formules clés.**
    - $\sum \lambda_i = \mathrm{tr}(A)$, $\prod \lambda_i = \det(A)$.
    - A inversible $\Leftrightarrow 0 \notin \mathrm{Spec}(A)$.
    - Matrices semblables $\Rightarrow$ même χ, même μ, même spectre, même trace, même déterminant.

Chapitre 7

Espaces préhilbertiens et orthogonalité

Introduction

Les chapitres précédents ont construit la théorie des espaces vectoriels, des applications linéaires, du déterminant et de la diagonalisation. Cependant, cette théorie purement algébrique ne permet pas de définir des notions fondamentales de la géométrie : *longueur* d'un vecteur, *angle* entre deux vecteurs, *distance* entre deux points, ou encore *perpendicularité*.

Pour introduire ces notions géométriques dans un espace vectoriel, il faut le munir d'une structure supplémentaire : un *produit scalaire*. C'est une forme bilinéaire (ou sesquilinéaire dans le cas complexe) qui généralise le produit scalaire usuel de $\mathbb{R}^n$. L'espace vectoriel muni d'un produit scalaire est appelé *espace préhilbertien*, ou *espace euclidien* lorsqu'il est réel et de dimension finie, et *espace hermitien* lorsqu'il est complexe et de dimension finie.

Ce chapitre est au cœur de nombreuses applications :

- **Projection orthogonale :** trouver le vecteur d'un sous-espace le plus proche d'un vecteur donné (approximation au sens des moindres carrés).
- **Procédé de Gram-Schmidt :** construire une base orthonormée à partir d'une base quelconque, ce qui simplifie considérablement les calculs.
- **Analyse de Fourier :** décomposer une fonction en somme de fonctions trigonométriques orthogonales.
- **Factorisation QR :** décomposer une matrice en produit d'une matrice orthogonale et d'une matrice triangulaire, outil fondamental du calcul numérique.
- **Théorème spectral :** les matrices symétriques réelles (ou hermitiennes) sont diagonalisables en base orthonormée, résultat essentiel pour l'analyse en composantes principales.

Dans tout ce chapitre, $\mathbb{K}$ désigne $\mathbb{R}$ ou $\mathbb{C}$, E un $\mathbb{K}$-espace vectoriel, et $n = \dim E$ lorsque E est de dimension finie.

7.1 Formes bilinéaires

Définition 1 (Forme bilinéaire)

Soit E un $\mathbb{R}$-espace vectoriel. Une *forme bilinéaire* sur E est une application $\varphi : E \times E \to \mathbb{R}$ qui est linéaire en chacune de ses variables :

(i) Pour tout $\mathbf{v} \in E$, l'application $\mathbf{u} \mapsto \varphi(\mathbf{u}, \mathbf{v})$ est linéaire.

(ii) Pour tout $\mathbf{u} \in E$, l'application $\mathbf{v} \mapsto \varphi(\mathbf{u}, \mathbf{v})$ est linéaire.

Autrement dit, pour tous $\mathbf{u}, \mathbf{u}', \mathbf{v} \in E$ et tout $\lambda \in \mathbb{R}$:

$$\varphi(\mathbf{u} + \mathbf{u}', \mathbf{v}) = \varphi(\mathbf{u}, \mathbf{v}) + \varphi(\mathbf{u}', \mathbf{v}), \qquad \varphi(\lambda \mathbf{u}, \mathbf{v}) = \lambda\, \varphi(\mathbf{u}, \mathbf{v}),$$

et de même en la seconde variable.

Définition 2 (Symétrie, antisymétrie)

Soit φ une forme bilinéaire sur E.

(i) φ est *symétrique* si $\varphi(\mathbf{u}, \mathbf{v}) = \varphi(\mathbf{v}, \mathbf{u})$ pour tous $\mathbf{u}, \mathbf{v} \in E$.

(ii) φ est *antisymétrique* si $\varphi(\mathbf{u}, \mathbf{v}) = -\varphi(\mathbf{v}, \mathbf{u})$ pour tous $\mathbf{u}, \mathbf{v} \in E$.

Définition 3 (Matrice d'une forme bilinéaire)

Soit $\mathcal{B} = (\mathbf{e}_1, \dots, \mathbf{e}_n)$ une base de E et φ une forme bilinéaire sur E. La *matrice de φ dans la base $\mathcal{B}$* est la matrice $A = (a_{ij}) \in \mathcal{M}_n(\mathbb{R})$ définie par :

$$a_{ij} := \varphi(\mathbf{e}_i, \mathbf{e}_j), \qquad 1 \le i, j \le n.$$

Pour tous $\mathbf{u}, \mathbf{v} \in E$ de coordonnées $X, Y \in \mathbb{R}^n$ dans $\mathcal{B}$, on a :

$$\varphi(\mathbf{u}, \mathbf{v}) = X^{\mathsf{T}} A\, Y = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}\, x_i\, y_j.$$

Proposition 4 (Changement de base pour les formes bilinéaires)

Si P est la matrice de passage de la base $\mathcal{B}$ à la base $\mathcal{B}'$, et si A (resp. A') est la matrice de φ dans $\mathcal{B}$ (resp. $\mathcal{B}'$), alors :

$$A' = P^{\mathsf{T}} A\, P.$$

On dit que A et A' sont *congruentes*.

Démonstration. Soient $\mathbf{u}, \mathbf{v} \in E$ de coordonnées X, Y dans $\mathcal{B}$ et X', Y' dans $\mathcal{B}'$. On a $X = PX'$ et $Y = PY'$, donc :

$$\varphi(\mathbf{u}, \mathbf{v}) = X^{\mathsf{T}} A\, Y = (PX')^{\mathsf{T}}\, A\, (PY') = X'^{\mathsf{T}}\, (P^{\mathsf{T}} A\, P)\, Y'.$$

Par unicité de la matrice représentant φ dans $\mathcal{B}'$, on obtient $A' = P^{\mathsf{T}} A\, P$. □

Remarque 5 (Forme bilinéaire symétrique et matrice symétrique)

φ est symétrique si et seulement si $A = A^{\mathsf{T}}$. En effet, $\varphi(\mathbf{u},\mathbf{v}) = X^{\mathsf{T}}AY$ et $\varphi(\mathbf{v},\mathbf{u}) = Y^{\mathsf{T}}AX = X^{\mathsf{T}}A^{\mathsf{T}}Y$. L'égalité pour tous X, Y entraîne $A = A^{\mathsf{T}}$.

7.2 Formes quadratiques

Définition 6 (Forme quadratique)

Soit E un $\mathbb{R}$-espace vectoriel. Une *forme quadratique* sur E est une application $q\colon E \to \mathbb{R}$ telle qu'il existe une forme bilinéaire symétrique φ sur E vérifiant :

$$q(\mathbf{v}) = \varphi(\mathbf{v},\mathbf{v}) \quad \text{pour tout } \mathbf{v} \in E.$$

On dit que φ est la *forme polaire* de q.

Proposition 7 (Identité de polarisation)

La forme bilinéaire symétrique φ est entièrement déterminée par la forme quadratique q, via la *formule de polarisation* :

$$\varphi(\mathbf{u},\mathbf{v}) = \frac{1}{2}[q(\mathbf{u}+\mathbf{v}) - q(\mathbf{u}) - q(\mathbf{v})] = \frac{1}{4}[q(\mathbf{u}+\mathbf{v}) - q(\mathbf{u}-\mathbf{v})].$$

Démonstration. Par bilinéarité et symétrie de φ :

$$\begin{aligned} q(\mathbf{u}+\mathbf{v}) &= \varphi(\mathbf{u}+\mathbf{v},\mathbf{u}+\mathbf{v}) \\ &= \varphi(\mathbf{u},\mathbf{u}) + \varphi(\mathbf{u},\mathbf{v}) + \varphi(\mathbf{v},\mathbf{u}) + \varphi(\mathbf{v},\mathbf{v}) \\ &= q(\mathbf{u}) + 2\,\varphi(\mathbf{u},\mathbf{v}) + q(\mathbf{v}). \end{aligned}$$

D'où $\varphi(\mathbf{u},\mathbf{v}) = \frac{1}{2}[q(\mathbf{u}+\mathbf{v}) - q(\mathbf{u}) - q(\mathbf{v})]$. De même, $q(\mathbf{u}-\mathbf{v}) = q(\mathbf{u}) - 2\,\varphi(\mathbf{u},\mathbf{v}) + q(\mathbf{v})$, ce qui donne la seconde formule par soustraction. □

Théorème 8 (Loi d'inertie de Sylvester)

Soit q une forme quadratique sur un $\mathbb{R}$-espace vectoriel E de dimension n. Il existe une base $\mathcal{B}$ de E dans laquelle la matrice de la forme polaire φ est diagonale :

$$\mathcal{M}_{\mathcal{B}}(\varphi) = \operatorname{diag}(\underbrace{1,\dots,1}_{p},\ \underbrace{-1,\dots,-1}_{q},\ \underbrace{0,\dots,0}_{n-p-q}).$$

De plus, les entiers p (indice de positivité) et q (indice de négativité) ne dépendent pas du choix de la base : ils sont des invariants de q. Le couple (p,q) s'appelle la *signature* de q, et $p+q$ est le *rang* de q.

Remarque 9 (Interprétation de la signature)

La signature (p,q) mesure le nombre de directions dans lesquelles q est strictement positive (p) ou strictement négative (q). Une forme quadratique est :

- *définie positive* si $(p,q) = (n,0)$, i.e. $q(\mathbf{v}) > 0$ pour tout $\mathbf{v} \neq \mathbf{0}$.
- *définie négative* si $(p,q) = (0,n)$.

- *positive* (ou semi-définie positive) si $q = 0$, i.e. $q(\mathbf{v}) \geq 0$ pour tout $\mathbf{v}$.
- *non dégénérée* si $p + q = n$.

7.3 Produit scalaire

7.3.1 Cas réel

Définition 10 (Produit scalaire réel)

Soit E un $\mathbb{R}$-espace vectoriel. Un *produit scalaire* sur E est une forme bilinéaire $\langle \cdot, \cdot \rangle : E \times E \to \mathbb{R}$ qui est :

(i) **symétrique :** $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$ pour tous $\mathbf{u}, \mathbf{v} \in E$.

(ii) **définie positive :** $\langle \mathbf{v}, \mathbf{v} \rangle \geq 0$ pour tout $\mathbf{v} \in E$, et $\langle \mathbf{v}, \mathbf{v} \rangle = 0$ si et seulement si $\mathbf{v} = \mathbf{0}$.

Un $\mathbb{R}$-espace vectoriel muni d'un produit scalaire est un *espace préhilbertien réel*. S'il est de dimension finie, on l'appelle *espace euclidien*.

7.3.2 Cas complexe

Définition 11 (Forme sesquilinéaire)

Soit E un $\mathbb{C}$-espace vectoriel. Une *forme sesquilinéaire* sur E est une application $\varphi : E \times E \to \mathbb{C}$ qui est :

(i) linéaire en la seconde variable : $\varphi(\mathbf{u}, \lambda \mathbf{v} + \mathbf{v}') = \lambda\, \varphi(\mathbf{u}, \mathbf{v}) + \varphi(\mathbf{u}, \mathbf{v}')$.

(ii) semi-linéaire (antilinéaire) en la première variable : $\varphi(\lambda \mathbf{u} + \mathbf{u}', \mathbf{v}) = \bar{\lambda}\, \varphi(\mathbf{u}, \mathbf{v}) + \varphi(\mathbf{u}', \mathbf{v})$.

(Convention physique : la conjugaison porte sur le premier argument.)

Définition 12 (Produit scalaire hermitien)

Soit E un $\mathbb{C}$-espace vectoriel. Un *produit scalaire hermitien* (ou *produit hermitien*) sur E est une forme sesquilinéaire $\langle \cdot, \cdot \rangle : E \times E \to \mathbb{C}$ qui est :

(i) **hermitienne :** $\langle \mathbf{u}, \mathbf{v} \rangle = \overline{\langle \mathbf{v}, \mathbf{u} \rangle}$ pour tous $\mathbf{u}, \mathbf{v} \in E$.

(ii) **définie positive :** $\langle \mathbf{v}, \mathbf{v} \rangle \in \mathbb{R}$, $\langle \mathbf{v}, \mathbf{v} \rangle \geq 0$ pour tout $\mathbf{v} \in E$, et $\langle \mathbf{v}, \mathbf{v} \rangle = 0$ si et seulement si $\mathbf{v} = \mathbf{0}$.

Un $\mathbb{C}$-espace vectoriel de dimension finie muni d'un produit hermitien est un *espace hermitien*.

Remarque 13 (Convention)

La condition hermitienne entraîne $\langle \mathbf{v}, \mathbf{v} \rangle = \overline{\langle \mathbf{v}, \mathbf{v} \rangle}$, donc $\langle \mathbf{v}, \mathbf{v} \rangle \in \mathbb{R}$. De plus, dans le cas réel, la condition hermitienne se réduit à la symétrie. On peut donc énoncer les résultats

pour $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ de manière unifiée.

7.4 Exemples fondamentaux

Exemple 14 (Produit scalaire canonique de $\mathbb{R}^n$)

Sur $E = \mathbb{R}^n$, le *produit scalaire canonique* est :

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^{\mathsf{T}} \mathbf{y} = \sum_{i=1}^{n} x_i \, y_i, \qquad \mathbf{x} = (x_1, \dots, x_n),\ \mathbf{y} = (y_1, \dots, y_n).$$

Il est bilinéaire, symétrique, et défini positif car $\langle \mathbf{x}, \mathbf{x} \rangle = \sum_{i=1}^{n} x_i^2 \geq 0$, avec égalité si et seulement si $\mathbf{x} = \mathbf{0}$.

Exemple 15 (Produit hermitien canonique de $\mathbb{C}^n$)

Sur $E = \mathbb{C}^n$, le *produit hermitien canonique* est :

$$\langle \mathbf{z}, \mathbf{w} \rangle = \mathbf{z}^{*} \, \mathbf{w} = \sum_{i=1}^{n} \bar{z}_i \, w_i.$$

On vérifie immédiatement la sesquilinéarité, la condition hermitienne, et la positivité : $\langle \mathbf{z}, \mathbf{z} \rangle = \sum_{i=1}^{n} |z_i|^2 \geq 0$.

Exemple 16 (Produit scalaire L^2 sur $\mathcal{C}([a,b], \mathbb{R})$)

Sur l'espace $E = \mathcal{C}([a,b], \mathbb{R})$ des fonctions continues de $[a,b]$ dans $\mathbb{R}$, on pose :

$$\langle f, g \rangle = \int_a^b f(t) \, g(t) \, dt.$$

C'est un produit scalaire : la bilinéarité et la symétrie sont immédiates. Pour le caractère défini positif : $\langle f, f \rangle = \int_a^b f(t)^2 \, dt \geq 0$, et si $\langle f, f \rangle = 0$, alors $f(t)^2 \equiv 0$ par continuité, donc $f = 0$.

Exemple 17 (Produit scalaire à poids)

Soit $w \colon [a,b] \to \mathbb{R}_{>0}$ une fonction continue strictement positive (fonction poids). Alors :

$$\langle f, g \rangle_w = \int_a^b w(t) \, f(t) \, g(t) \, dt$$

est un produit scalaire sur $\mathcal{C}([a,b], \mathbb{R})$. Ce type de produit scalaire apparaît naturellement dans la théorie des polynômes orthogonaux (Legendre, Tchebychev, Hermite, Laguerre).

7.5 Norme, distance et inégalité de Cauchy-Schwarz

Définition 18 (Norme associée à un produit scalaire)

Soit $(E, \langle \cdot, \cdot \rangle)$ un espace préhilbertien. La *norme* associée au produit scalaire est l'application $\|\cdot\| : E \to \mathbb{R}_+$ définie par :

$$\|\mathbf{v}\| := \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}.$$

La *distance* entre deux vecteurs est $d(\mathbf{u}, \mathbf{v}) := \|\mathbf{u} - \mathbf{v}\|$.

Théorème 19 (Inégalité de Cauchy-Schwarz)

Soit $(E, \langle \cdot, \cdot \rangle)$ un espace préhilbertien (réel ou complexe). Pour tous $\mathbf{u}, \mathbf{v} \in E$:

$$\boxed{|\langle \mathbf{u}, \mathbf{v} \rangle| \leq \|\mathbf{u}\| \cdot \|\mathbf{v}\|.}$$

De plus, l'égalité a lieu si et seulement si $\mathbf{u}$ et $\mathbf{v}$ sont liés (i.e. l'un est multiple scalaire de l'autre).

Démonstration. Si $\mathbf{v} = \mathbf{0}$, les deux membres valent 0 et l'inégalité est triviale. Supposons donc $\mathbf{v} \neq \mathbf{0}$.

Cas réel. Pour tout $t \in \mathbb{R}$, considérons :

$$p(t) := \|\mathbf{u} - t\mathbf{v}\|^2 = \langle \mathbf{u} - t\mathbf{v}, \mathbf{u} - t\mathbf{v} \rangle = \|\mathbf{u}\|^2 - 2t \langle \mathbf{u}, \mathbf{v} \rangle + t^2 \|\mathbf{v}\|^2.$$

C'est un polynôme du second degré en t à coefficients réels, avec coefficient dominant $\|\mathbf{v}\|^2 > 0$, et $p(t) \geq 0$ pour tout $t \in \mathbb{R}$ (car c'est le carré d'une norme). La condition de non-négativité d'un trinôme du second degré impose que le discriminant soit négatif ou nul :

$$\Delta = 4 \langle \mathbf{u}, \mathbf{v} \rangle^2 - 4 \|\mathbf{u}\|^2 \|\mathbf{v}\|^2 \leq 0,$$

d'où $\langle \mathbf{u}, \mathbf{v} \rangle^2 \leq \|\mathbf{u}\|^2 \|\mathbf{v}\|^2$, ce qui est exactement l'inégalité de Cauchy-Schwarz.

Cas d'égalité. L'égalité a lieu si et seulement si $\Delta = 0$, c'est-à-dire si et seulement si p admet une racine réelle t_0, auquel cas $p(t_0) = \|\mathbf{u} - t_0\mathbf{v}\|^2 = 0$, donc $\mathbf{u} = t_0\mathbf{v}$: les vecteurs sont colinéaires.

Cas complexe. Pour tout $\lambda \in \mathbb{C}$, on a :

$$0 \leq \|\mathbf{u} - \lambda\mathbf{v}\|^2 = \|\mathbf{u}\|^2 - \lambda \langle \mathbf{u}, \mathbf{v} \rangle - \bar{\lambda} \overline{\langle \mathbf{u}, \mathbf{v} \rangle} + |\lambda|^2 \|\mathbf{v}\|^2.$$

Choisissons $\lambda = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{v}\|^2}$ (qui minimise l'expression). On obtient :

$$0 \leq \|\mathbf{u}\|^2 - \frac{|\langle \mathbf{u}, \mathbf{v} \rangle|^2}{\|\mathbf{v}\|^2},$$

d'où $|\langle \mathbf{u}, \mathbf{v} \rangle|^2 \leq \|\mathbf{u}\|^2 \cdot \|\mathbf{v}\|^2$, et en prenant la racine carrée :

$$|\langle \mathbf{u}, \mathbf{v} \rangle| \leq \|\mathbf{u}\| \cdot \|\mathbf{v}\|.$$

L'égalité a lieu si et seulement si $\|\mathbf{u} - \lambda\mathbf{v}\| = 0$, i.e. $\mathbf{u} = \lambda\mathbf{v}$. □

Corollaire 20 (Inégalité triangulaire)

Pour tous $\mathbf{u}, \mathbf{v} \in E$:

$$\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|.$$

Démonstration. On a :

$$\begin{aligned}\|\mathbf{u}+\mathbf{v}\|^2 &= \|\mathbf{u}\|^2 + 2\,\Re(\langle \mathbf{u},\mathbf{v}\rangle) + \|\mathbf{v}\|^2 \\ &\le \|\mathbf{u}\|^2 + 2\,|\langle \mathbf{u},\mathbf{v}\rangle| + \|\mathbf{v}\|^2 \\ &\le \|\mathbf{u}\|^2 + 2\,\|\mathbf{u}\|\|\mathbf{v}\| + \|\mathbf{v}\|^2 = \big(\|\mathbf{u}\| + \|\mathbf{v}\|\big)^2,\end{aligned}$$

en utilisant Cauchy-Schwarz à la dernière inégalité. On conclut en prenant la racine carrée. □

Proposition 21 (Propriétés de la norme)

La norme $\|\cdot\|$ vérifie les propriétés suivantes :

(i) **Séparation :** $\|\mathbf{v}\| = 0 \Leftrightarrow \mathbf{v} = \mathbf{0}$.

(ii) **Homogénéité :** $\|\lambda \mathbf{v}\| = |\lambda|\,\|\mathbf{v}\|$ pour tout $\lambda \in \mathbb{K}$.

(iii) **Inégalité triangulaire :** $\|\mathbf{u}+\mathbf{v}\| \le \|\mathbf{u}\| + \|\mathbf{v}\|$.

De plus, elle satisfait l'*identité du parallélogramme* :

$$\|\mathbf{u}+\mathbf{v}\|^2 + \|\mathbf{u}-\mathbf{v}\|^2 = 2(\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2).$$

Définition 22 (Angle entre deux vecteurs)

Soient $\mathbf{u}, \mathbf{v} \in E \setminus \{\mathbf{0}\}$ dans un espace euclidien. L'inégalité de Cauchy-Schwarz assure que :

$$-1 \le \frac{\langle \mathbf{u},\mathbf{v}\rangle}{\|\mathbf{u}\| \cdot \|\mathbf{v}\|} \le 1.$$

On définit l'*angle* $\theta \in [0,\pi]$ entre $\mathbf{u}$ et $\mathbf{v}$ par :

$$\cos\theta = \frac{\langle \mathbf{u},\mathbf{v}\rangle}{\|\mathbf{u}\| \cdot \|\mathbf{v}\|}.$$

7.6 Orthogonalité

Définition 23 (Vecteurs orthogonaux)

Soit $(E, \langle\cdot,\cdot\rangle)$ un espace préhilbertien.

(i) Deux vecteurs $\mathbf{u}, \mathbf{v} \in E$ sont *orthogonaux*, noté $\mathbf{u} \perp \mathbf{v}$, si $\langle \mathbf{u},\mathbf{v}\rangle = 0$.

(ii) Une famille $(\mathbf{v}_1, \dots, \mathbf{v}_p)$ est *orthogonale* si $\langle \mathbf{v}_i, \mathbf{v}_j\rangle = 0$ pour tout $i \neq j$.

(iii) Elle est *orthonormée* (ou *orthonormale*) si de plus $\|\mathbf{v}_i\| = 1$ pour tout i, i.e. $\langle \mathbf{v}_i, \mathbf{v}_j\rangle = \delta_{ij}$.

Proposition 24 (Famille orthogonale et liberté)

Toute famille orthogonale de vecteurs *non nuls* est libre.

Démonstration. Soit $(\mathbf{v}_1, \dots, \mathbf{v}_p)$ une famille orthogonale avec $\mathbf{v}_i \neq \mathbf{0}$ pour tout i. Supposons

que $\sum_{i=1}^{p} \alpha_i \mathbf{v}_i = \mathbf{0}$. En prenant le produit scalaire avec $\mathbf{v}_k$:

$$0 = \left\langle \sum_{i=1}^{p} \alpha_i \mathbf{v}_i, \mathbf{v}_k \right\rangle = \sum_{i=1}^{p} \alpha_i \langle \mathbf{v}_i, \mathbf{v}_k \rangle = \alpha_k \|\mathbf{v}_k\|^2.$$

Comme $\|\mathbf{v}_k\|^2 > 0$ (car $\mathbf{v}_k \neq \mathbf{0}$), on en déduit $\alpha_k = 0$ pour tout k. La famille est donc libre. □

Définition 25 (Orthogonal d'une partie)

Soit F une partie de E. L'*orthogonal* de F est :

$$F^\perp := \{ \mathbf{v} \in E \mid \langle \mathbf{v}, \mathbf{w} \rangle = 0 \text{ pour tout } \mathbf{w} \in F \}.$$

Proposition 26 (Propriétés de l'orthogonal)

Soient F, G des sous-espaces vectoriels d'un espace préhilbertien E de dimension finie. Alors :

(i) $F^\perp$ est un sous-espace vectoriel de E.

(ii) $E = F \oplus F^\perp$.

(iii) $\dim F^\perp = \dim E - \dim F$.

(iv) $(F^\perp)^\perp = F$.

(v) $F \subset G \implies G^\perp \subset F^\perp$.

(vi) $(F + G)^\perp = F^\perp \cap G^\perp$.

(vii) $(F \cap G)^\perp = F^\perp + G^\perp$.

Démonstration. *(i)* La linéarité du produit scalaire en chaque variable assure que $F^\perp$ est stable par combinaison linéaire.

(ii) Montrons d'abord que $F \cap F^\perp = \{\mathbf{0}\}$: si $\mathbf{v} \in F \cap F^\perp$, alors $\langle \mathbf{v}, \mathbf{v} \rangle = 0$ (car $\mathbf{v} \in F^\perp$ et $\mathbf{v} \in F$), donc $\mathbf{v} = \mathbf{0}$.

Soit $(\mathbf{e}_1, \dots, \mathbf{e}_p)$ une base orthonormée de F (dont l'existence sera établie par le procédé de Gram-Schmidt, théorème 31). Pour tout $\mathbf{v} \in E$, posons :

$$\mathbf{v}_F := \sum_{i=1}^{p} \langle \mathbf{e}_i, \mathbf{v} \rangle \, \mathbf{e}_i, \qquad \mathbf{v}_{F^\perp} := \mathbf{v} - \mathbf{v}_F.$$

On a $\mathbf{v}_F \in F$. Vérifions que $\mathbf{v}_{F^\perp} \in F^\perp$: pour tout $j \in \{1, \dots, p\}$,

$$\langle \mathbf{e}_j, \mathbf{v}_{F^\perp} \rangle = \langle \mathbf{e}_j, \mathbf{v} \rangle - \sum_{i=1}^{p} \langle \mathbf{e}_i, \mathbf{v} \rangle \langle \mathbf{e}_j, \mathbf{e}_i \rangle = \langle \mathbf{e}_j, \mathbf{v} \rangle - \langle \mathbf{e}_j, \mathbf{v} \rangle = 0.$$

Donc $\mathbf{v} = \mathbf{v}_F + \mathbf{v}_{F^\perp} \in F + F^\perp$. Avec $F \cap F^\perp = \{\mathbf{0}\}$, on obtient $E = F \oplus F^\perp$.

(iii) Conséquence immédiate de (ii) : $\dim F^\perp = \dim E - \dim F$.

(iv) L'inclusion $F \subset (F^\perp)^\perp$ est claire (si $\mathbf{v} \in F$, alors $\langle \mathbf{v}, \mathbf{w} \rangle = 0$ pour tout $\mathbf{w} \in F^\perp$). L'égalité des dimensions $\dim (F^\perp)^\perp = n - \dim F^\perp = n - (n - \dim F) = \dim F$ donne l'égalité.

Les propriétés (v)–(vii) se vérifient par double inclusion à partir des définitions. □

Théorème 27 (Théorème de Pythagore)

Soient $\mathbf{u}, \mathbf{v} \in E$ deux vecteurs orthogonaux. Alors :

$$\|\mathbf{u}+\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2.$$

Plus généralement, si $(\mathbf{v}_1, \dots, \mathbf{v}_p)$ est une famille orthogonale, alors :

$$\left\|\sum_{i=1}^{p} \mathbf{v}_i\right\|^2 = \sum_{i=1}^{p} \|\mathbf{v}_i\|^2.$$

Démonstration. $\|\mathbf{u}+\mathbf{v}\|^2 = \langle \mathbf{u}+\mathbf{v}, \mathbf{u}+\mathbf{v}\rangle = \|\mathbf{u}\|^2 + \langle \mathbf{u},\mathbf{v}\rangle + \langle \mathbf{v},\mathbf{u}\rangle + \|\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$, puisque $\langle \mathbf{u},\mathbf{v}\rangle = 0$ entraîne aussi $\langle \mathbf{v},\mathbf{u}\rangle = \overline{\langle \mathbf{u},\mathbf{v}\rangle} = 0$. La généralisation se fait par récurrence immédiate. □

7.7 Projection orthogonale

Définition 28 (Projection orthogonale)

Soit F un sous-espace vectoriel d'un espace préhilbertien E de dimension finie. D'après la décomposition $E = F \oplus F^\perp$, tout vecteur $\mathbf{v} \in E$ s'écrit de manière unique :

$$\mathbf{v} = \mathbf{v}_F + \mathbf{v}_{F^\perp}, \qquad \mathbf{v}_F \in F, \quad \mathbf{v}_{F^\perp} \in F^\perp.$$

L'application $\mathrm{proj}_F \colon E \to E$ définie par $\mathrm{proj}_F(\mathbf{v}) := \mathbf{v}_F$ est la *projection orthogonale* sur F parallèlement à $F^\perp$.

Théorème 29 (Caractérisation de la meilleure approximation)

Soit F un sous-espace vectoriel de dimension finie d'un espace préhilbertien E, et soit $\mathbf{v} \in E$. L'élément $\mathbf{v}_F = \mathrm{proj}_F(\mathbf{v})$ est l'unique élément de F qui minimise la distance de $\mathbf{v}$ à F :

$$\|\mathbf{v} - \mathbf{v}_F\| = \min_{\mathbf{w}\in F} \|\mathbf{v}-\mathbf{w}\| = d(\mathbf{v}, F).$$

De plus, $\mathbf{v}_F$ est caractérisé par les deux conditions :

(i) $\mathbf{v}_F \in F$.

(ii) $\mathbf{v} - \mathbf{v}_F \perp F$, i.e. $\langle \mathbf{v}-\mathbf{v}_F, \mathbf{w}\rangle = 0$ pour tout $\mathbf{w} \in F$.

Démonstration. Pour tout $\mathbf{w} \in F$, on a $\mathbf{v}-\mathbf{w} = (\mathbf{v}-\mathbf{v}_F) + (\mathbf{v}_F - \mathbf{w})$, avec $\mathbf{v}-\mathbf{v}_F \in F^\perp$ et $\mathbf{v}_F - \mathbf{w} \in F$. Ces deux vecteurs sont orthogonaux, donc par le théorème de Pythagore :

$$\|\mathbf{v}-\mathbf{w}\|^2 = \|\mathbf{v}-\mathbf{v}_F\|^2 + \|\mathbf{v}_F-\mathbf{w}\|^2 \geq \|\mathbf{v}-\mathbf{v}_F\|^2.$$

L'égalité a lieu si et seulement si $\|\mathbf{v}_F - \mathbf{w}\| = 0$, i.e. $\mathbf{w} = \mathbf{v}_F$. L'unicité est ainsi prouvée. □

Proposition 30 (Formule de la projection orthogonale)

Si $(\mathbf{e}_1, \dots, \mathbf{e}_p)$ est une base *orthonormée* de F, alors :

$$\mathrm{proj}_F(\mathbf{v}) = \sum_{i=1}^{p} \langle \mathbf{e}_i, \mathbf{v} \rangle \, \mathbf{e}_i.$$

Démonstration. Posons $\mathbf{v}_F = \sum_{i=1}^{p} \langle \mathbf{e}_i, \mathbf{v} \rangle \, \mathbf{e}_i \in F$. Pour tout j :

$$\langle \mathbf{e}_j, \mathbf{v} - \mathbf{v}_F \rangle = \langle \mathbf{e}_j, \mathbf{v} \rangle - \sum_{i=1}^{p} \langle \mathbf{e}_i, \mathbf{v} \rangle \, \delta_{ij} = 0.$$

Donc $\mathbf{v} - \mathbf{v}_F \in F^\perp$, et par unicité de la décomposition $E = F \oplus F^\perp$, $\mathbf{v}_F = \mathrm{proj}_F(\mathbf{v})$. □

7.8 Procédé de Gram-Schmidt

Théorème 31 (Procédé d'orthonormalisation de Gram-Schmidt)

Soit $(\mathbf{v}_1, \dots, \mathbf{v}_p)$ une famille libre dans un espace préhilbertien E. Il existe une famille orthonormée $(\mathbf{e}_1, \dots, \mathbf{e}_p)$ telle que :

$$\mathrm{Vect}(\mathbf{e}_1, \dots, \mathbf{e}_k) = \mathrm{Vect}(\mathbf{v}_1, \dots, \mathbf{v}_k) \quad \text{pour tout } k \in \{1, \dots, p\}.$$

De plus, on peut choisir les $\mathbf{e}_k$ de sorte que $\langle \mathbf{e}_k, \mathbf{v}_k \rangle > 0$.

Démonstration. On procède par récurrence sur k.

Initialisation ($k = 1$). On pose $\mathbf{e}_1 := \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|}$ (bien défini car $\mathbf{v}_1 \neq \mathbf{0}$ puisque la famille est libre). On a $\|\mathbf{e}_1\| = 1$ et $\mathrm{Vect}(\mathbf{e}_1) = \mathrm{Vect}(\mathbf{v}_1)$.

Hérédité. Supposons construits $\mathbf{e}_1, \dots, \mathbf{e}_{k-1}$ orthonormés avec $\mathrm{Vect}(\mathbf{e}_1, \dots, \mathbf{e}_j) = \mathrm{Vect}(\mathbf{v}_1, \dots, \mathbf{v}_j)$ pour $j \leq k-1$. Posons $F_{k-1} = \mathrm{Vect}(\mathbf{e}_1, \dots, \mathbf{e}_{k-1})$ et :

$$\mathbf{w}_k := \mathbf{v}_k - \sum_{i=1}^{k-1} \langle \mathbf{e}_i, \mathbf{v}_k \rangle \, \mathbf{e}_i = \mathbf{v}_k - \mathrm{proj}_{F_{k-1}}(\mathbf{v}_k).$$

Le vecteur $\mathbf{w}_k$ est la composante de $\mathbf{v}_k$ orthogonale à F_{k-1}. Comme $\mathbf{v}_k \notin F_{k-1} = \mathrm{Vect}(\mathbf{v}_1, \dots, \mathbf{v}_{k-1})$ (la famille $(\mathbf{v}_i)$ étant libre), on a $\mathbf{w}_k \neq \mathbf{0}$.

Vérifions l'orthogonalité : pour tout $j \in \{1, \dots, k-1\}$,

$$\langle \mathbf{e}_j, \mathbf{w}_k \rangle = \langle \mathbf{e}_j, \mathbf{v}_k \rangle - \sum_{i=1}^{k-1} \langle \mathbf{e}_i, \mathbf{v}_k \rangle \underbrace{\langle \mathbf{e}_j, \mathbf{e}_i \rangle}_{=\delta_{ij}} = \langle \mathbf{e}_j, \mathbf{v}_k \rangle - \langle \mathbf{e}_j, \mathbf{v}_k \rangle = 0.$$

On pose alors $\mathbf{e}_k := \frac{\mathbf{w}_k}{\|\mathbf{w}_k\|}$, qui vérifie $\|\mathbf{e}_k\| = 1$ et $\mathbf{e}_k \perp \mathbf{e}_j$ pour $j < k$. De plus :

$$\mathbf{v}_k = \mathbf{w}_k + \sum_{i=1}^{k-1} \langle \mathbf{e}_i, \mathbf{v}_k \rangle \, \mathbf{e}_i \in \mathrm{Vect}(\mathbf{e}_1, \dots, \mathbf{e}_k),$$

donc $\mathrm{Vect}(\mathbf{v}_1, \dots, \mathbf{v}_k) \subset \mathrm{Vect}(\mathbf{e}_1, \dots, \mathbf{e}_k)$. Comme $\mathbf{e}_k$ est combinaison linéaire de $\mathbf{v}_k$ et des $\mathbf{e}_i$ ($i < k$), et que $\mathrm{Vect}(\mathbf{e}_1, \dots, \mathbf{e}_{k-1}) = \mathrm{Vect}(\mathbf{v}_1, \dots, \mathbf{v}_{k-1})$ par hypothèse de récurrence, l'inclusion réciproque est vérifiée. D'où l'égalité des sous-espaces. □

Exemple 32 (Gram-Schmidt dans $\mathbb{R}^3$)

Orthonormalisons la base $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ de $\mathbb{R}^3$ avec le produit scalaire canonique, où :

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad \mathbf{v}_3 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}.$$

Étape 1. $\|\mathbf{v}_1\| = \sqrt{1+1+0} = \sqrt{2}$, donc $\mathbf{e}_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$.

Étape 2. $\langle \mathbf{e}_1, \mathbf{v}_2 \rangle = \frac{1}{\sqrt{2}}(1 \cdot 1 + 1 \cdot 0 + 0 \cdot 1) = \frac{1}{\sqrt{2}}$.

$$\mathbf{w}_2 = \mathbf{v}_2 - \langle \mathbf{e}_1, \mathbf{v}_2 \rangle \, \mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} - \frac{1}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} - \begin{pmatrix} 1/2 \\ 1/2 \\ 0 \end{pmatrix} = \begin{pmatrix} 1/2 \\ -1/2 \\ 1 \end{pmatrix}.$$

$\|\mathbf{w}_2\| = \sqrt{1/4 + 1/4 + 1} = \sqrt{3/2}$, donc $\mathbf{e}_2 = \sqrt{\frac{2}{3}} \begin{pmatrix} 1/2 \\ -1/2 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}$.

Étape 3. $\langle \mathbf{e}_1, \mathbf{v}_3 \rangle = \frac{1}{\sqrt{2}}(0 + 1 + 0) = \frac{1}{\sqrt{2}}$, $\quad \langle \mathbf{e}_2, \mathbf{v}_3 \rangle = \frac{1}{\sqrt{6}}(0 - 1 + 2) = \frac{1}{\sqrt{6}}$.

$$\mathbf{w}_3 = \mathbf{v}_3 - \langle \mathbf{e}_1, \mathbf{v}_3 \rangle \, \mathbf{e}_1 - \langle \mathbf{e}_2, \mathbf{v}_3 \rangle \, \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} - \frac{1}{6} \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} = \begin{pmatrix} -2/3 \\ 2/3 \\ 2/3 \end{pmatrix}.$$

$\|\mathbf{w}_3\| = \frac{2}{3}\sqrt{3}$, donc $\mathbf{e}_3 = \frac{1}{\sqrt{3}} \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$.

On peut vérifier que $\langle \mathbf{e}_i, \mathbf{e}_j \rangle = \delta_{ij}$.

7.9 Bases orthonormées

Corollaire 33 (Existence de bases orthonormées)

Tout espace préhilbertien de dimension finie admet une base orthonormée.

Démonstration. Appliquer le procédé de Gram-Schmidt à une base quelconque. □

Proposition 34 (Coordonnées dans une base orthonormée)

Soit $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ une base orthonormée de E. Pour tout $\mathbf{v} \in E$:

$$\mathbf{v} = \sum_{i=1}^{n} \langle \mathbf{e}_i, \mathbf{v} \rangle \, \mathbf{e}_i.$$

Les scalaires $\langle \mathbf{e}_i, \mathbf{v} \rangle$ sont les *coefficients de Fourier* de $\mathbf{v}$ dans la base $(\mathbf{e}_i)$.

Démonstration. En écrivant $\mathbf{v} = \sum_{i=1}^{n} \alpha_i \mathbf{e}_i$ et en prenant le produit scalaire avec $\mathbf{e}_j$, on obtient $\langle \mathbf{e}_j, \mathbf{v} \rangle = \alpha_j$. □

Théorème 35 (Identité de Parseval)

Soit $(\mathbf{e}_1, \dots, \mathbf{e}_n)$ une base orthonormée de E. Pour tous $\mathbf{u}, \mathbf{v} \in E$:

$$\langle \mathbf{u}, \mathbf{v} \rangle = \sum_{i=1}^{n} \overline{\langle \mathbf{e}_i, \mathbf{u} \rangle} \, \langle \mathbf{e}_i, \mathbf{v} \rangle.$$

En particulier :

$$\|\mathbf{v}\|^2 = \sum_{i=1}^{n} |\langle \mathbf{e}_i, \mathbf{v} \rangle|^2.$$

Démonstration. En décomposant $\mathbf{u} = \sum_i \alpha_i \mathbf{e}_i$ et $\mathbf{v} = \sum_j \beta_j \mathbf{e}_j$ avec $\alpha_i = \langle \mathbf{e}_i, \mathbf{u} \rangle$ et $\beta_j = \langle \mathbf{e}_j, \mathbf{v} \rangle$:

$$\langle \mathbf{u}, \mathbf{v} \rangle = \left\langle \sum_i \alpha_i \mathbf{e}_i, \sum_j \beta_j \mathbf{e}_j \right\rangle = \sum_i \sum_j \bar{\alpha}_i \beta_j \underbrace{\langle \mathbf{e}_i, \mathbf{e}_j \rangle}_{=\delta_{ij}} = \sum_i \bar{\alpha}_i \beta_i. \qquad \square$$

Théorème 36 (Inégalité de Bessel)

Soit $(\mathbf{e}_1, \dots, \mathbf{e}_p)$ une famille orthonormée (non nécessairement base) de E. Pour tout $\mathbf{v} \in E$:

$$\sum_{i=1}^{p} |\langle \mathbf{e}_i, \mathbf{v} \rangle|^2 \leq \|\mathbf{v}\|^2.$$

L'égalité a lieu si et seulement si $\mathbf{v} \in \mathrm{Vect}(\mathbf{e}_1, \dots, \mathbf{e}_p)$.

Démonstration. Posons $F = \mathrm{Vect}(\mathbf{e}_1, \dots, \mathbf{e}_p)$ et $\mathbf{v}_F = \mathrm{proj}_F(\mathbf{v}) = \sum_{i=1}^{p} \langle \mathbf{e}_i, \mathbf{v} \rangle \, \mathbf{e}_i$. Par le théorème de Pythagore :

$$\|\mathbf{v}\|^2 = \|\mathbf{v}_F\|^2 + \|\mathbf{v} - \mathbf{v}_F\|^2 = \sum_{i=1}^{p} |\langle \mathbf{e}_i, \mathbf{v} \rangle|^2 + \|\mathbf{v} - \mathbf{v}_F\|^2 \geq \sum_{i=1}^{p} |\langle \mathbf{e}_i, \mathbf{v} \rangle|^2.$$

L'égalité a lieu ssi $\|\mathbf{v} - \mathbf{v}_F\| = 0$, i.e. $\mathbf{v} = \mathbf{v}_F \in F$. $\square$

7.10 Matrices orthogonales et unitaires

Définition 37 (Matrice orthogonale)

Une matrice $Q \in \mathcal{M}_n(\mathbb{R})$ est *orthogonale* si :

$$Q^{\mathsf{T}} Q = I_n,$$

ou de manière équivalente, $Q^{-1} = Q^{\mathsf{T}}$. L'ensemble des matrices orthogonales de taille n forme un groupe pour la multiplication, noté $\mathrm{O}(n)$ et appelé *groupe orthogonal*.

Définition 38 (Matrice unitaire)

Une matrice $U \in \mathcal{M}_n(\mathbb{C})$ est *unitaire* si :

$$U^* U = I_n,$$

ou de manière équivalente, $U^{-1} = U^*$. L'ensemble des matrices unitaires de taille n forme le *groupe unitaire* $\mathrm{U}(n)$.

Proposition 39 (Caractérisations des matrices orthogonales)

Pour $Q \in \mathcal{M}_n(\mathbb{R})$, les conditions suivantes sont équivalentes :

(i) Q est orthogonale.

(ii) Les colonnes de Q forment une base orthonormée de $\mathbb{R}^n$.

(iii) Les lignes de Q forment une base orthonormée de $\mathbb{R}^n$.

(iv) Q conserve le produit scalaire : $\langle Q\mathbf{x}, Q\mathbf{y}\rangle = \langle \mathbf{x}, \mathbf{y}\rangle$ pour tous $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.

(v) Q est une isométrie : $\|Q\mathbf{x}\| = \|\mathbf{x}\|$ pour tout $\mathbf{x} \in \mathbb{R}^n$.

Démonstration. (i)⇔(ii) : $Q^{\mathsf{T}}Q = I_n$ signifie que le produit scalaire de la i-ème et de la j-ème colonne de Q vaut δ_{ij}.

(ii)⇔(iii) : Si $Q^{\mathsf{T}}Q = I_n$, alors Q est inversible et $QQ^{\mathsf{T}} = I_n$, ce qui signifie que les lignes sont orthonormées.

(i)⇒(iv) : $\langle Q\mathbf{x}, Q\mathbf{y}\rangle = (Q\mathbf{x})^{\mathsf{T}}Q\mathbf{y} = \mathbf{x}^{\mathsf{T}}Q^{\mathsf{T}}Q\mathbf{y} = \mathbf{x}^{\mathsf{T}}\mathbf{y} = \langle \mathbf{x}, \mathbf{y}\rangle$.

(iv)⇒(v) : prendre $\mathbf{y} = \mathbf{x}$.

(v)⇒(i) : par l'identité de polarisation, $\|Q\mathbf{x}\| = \|\mathbf{x}\|$ pour tout $\mathbf{x}$ entraîne $\langle Q\mathbf{x}, Q\mathbf{y}\rangle = \langle \mathbf{x}, \mathbf{y}\rangle$, d'où $\mathbf{x}^{\mathsf{T}}(Q^{\mathsf{T}}Q - I_n)\mathbf{y} = 0$ pour tous $\mathbf{x}, \mathbf{y}$, ce qui impose $Q^{\mathsf{T}}Q = I_n$. □

Proposition 40 (Déterminant d'une matrice orthogonale)

Si $Q \in \mathrm{O}(n)$, alors $\det(Q) = \pm 1$.

Démonstration. $\det(Q^{\mathsf{T}}Q) = \det(Q^{\mathsf{T}})\det(Q) = \det(Q)^2 = \det(I_n) = 1$. □

Définition 41 (Groupe spécial orthogonal)

Le *groupe spécial orthogonal* est :

$$\mathrm{SO}(n) := \{\, Q \in \mathrm{O}(n) \mid \det(Q) = 1 \,\}.$$

Les éléments de $\mathrm{SO}(n)$ sont les *rotations* de $\mathbb{R}^n$.

Exemple 42 (Matrices de SO(2))

Les éléments de $\mathrm{SO}(2)$ sont les matrices de rotation d'angle θ :

$$R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \qquad \theta \in \mathbb{R}.$$

Les éléments de $\mathrm{O}(2) \setminus \mathrm{SO}(2)$ (de déterminant -1) sont les réflexions : $S_\theta = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}$.

7.11 Opérateur adjoint

Définition 43 (Adjoint d'un endomorphisme)

Soit $(E, \langle \cdot, \cdot \rangle)$ un espace préhilbertien de dimension finie et $f \in \mathrm{End}(E)$. L'*adjoint* de f est l'unique endomorphisme $f^* \in \mathrm{End}(E)$ tel que :

$$\langle f(\mathbf{u}), \mathbf{v} \rangle = \langle \mathbf{u}, f^*(\mathbf{v}) \rangle \quad \text{pour tous } \mathbf{u}, \mathbf{v} \in E.$$

Théorème 44 (Existence et unicité de l'adjoint)

Pour tout $f \in \mathrm{End}(E)$, l'adjoint f^* existe et est unique.

Démonstration. Unicité. Si g_1 et g_2 satisfont la propriété, alors $\langle \mathbf{u}, g_1(\mathbf{v}) - g_2(\mathbf{v}) \rangle = 0$ pour tous $\mathbf{u}, \mathbf{v}$. En prenant $\mathbf{u} = g_1(\mathbf{v}) - g_2(\mathbf{v})$, on obtient $g_1(\mathbf{v}) = g_2(\mathbf{v})$ pour tout $\mathbf{v}$.

Existence. Soit $(\mathbf{e}_1, \dots, \mathbf{e}_n)$ une base orthonormée de E. Posons $A = \mathcal{M}_{\mathcal{B}}(f)$. Définissons g par $\mathcal{M}_{\mathcal{B}}(g) = A^*$ (ou A^{T} dans le cas réel). Alors :

$$\langle f(\mathbf{u}), \mathbf{v} \rangle = (AX)^* Y = X^* A^* Y = \langle \mathbf{u}, g(\mathbf{v}) \rangle,$$

où X, Y sont les vecteurs de coordonnées dans la base orthonormée. Donc $g = f^*$. □

Proposition 45 (Propriétés de l'adjoint)

Soient $f, g \in \mathrm{End}(E)$ et $\lambda \in \mathbb{K}$. Alors :

(i) $(f + g)^* = f^* + g^*$.

(ii) $(\lambda f)^* = \bar{\lambda}\, f^*$.

(iii) $(f \circ g)^* = g^* \circ f^*$.

(iv) $(f^*)^* = f$.

(v) $\mathrm{Ker}(f^*) = (\mathrm{Im}\, f)^\perp$ et $\mathrm{Im}(f^*) = (\mathrm{Ker}\, f)^\perp$.

Démonstration. Les propriétés (i)–(iv) se déduisent de la définition par unicité de l'adjoint.

Pour (v) : $\mathbf{v} \in \mathrm{Ker}(f^*)$ si et seulement si $f^*(\mathbf{v}) = \mathbf{0}$, si et seulement si $\langle \mathbf{u}, f^*(\mathbf{v}) \rangle = 0$ pour tout $\mathbf{u}$, si et seulement si $\langle f(\mathbf{u}), \mathbf{v} \rangle = 0$ pour tout $\mathbf{u}$, si et seulement si $\mathbf{v} \perp \mathrm{Im}(f)$, i.e. $\mathbf{v} \in (\mathrm{Im}\, f)^\perp$. La seconde égalité s'obtient en appliquant la première à f^* et en utilisant $(f^*)^* = f$. □

Remarque 46 (Matrice de l'adjoint)

Dans une base orthonormée $\mathcal{B}$, si $A = \mathcal{M}_{\mathcal{B}}(f)$, alors $\mathcal{M}_{\mathcal{B}}(f^*) = A^*$ (ou A^{T} si $\mathbb{K} = \mathbb{R}$). C'est pourquoi on écrit souvent A^* pour la matrice adjointe.

Définition 47 (Endomorphismes spéciaux)

Soit $f \in \mathrm{End}(E)$ sur un espace préhilbertien de dimension finie.

(i) f est *auto-adjoint* (ou *symétrique* si $\mathbb{K} = \mathbb{R}$, *hermitien* si $\mathbb{K} = \mathbb{C}$) si $f^* = f$. En base orthonormée, cela correspond à $A = A^*$.

(ii) f est *orthogonal* (si $\mathbb{K} = \mathbb{R}$) ou *unitaire* (si $\mathbb{K} = \mathbb{C}$) si $f^* \circ f = \mathrm{Id}$, i.e. f est une isométrie.

(iii) f est *normal* si $f^* \circ f = f \circ f^*$.

7.12 Applications

7.12.1 Approximation aux moindres carrés

Considérons un système linéaire $A\mathbf{x} = \mathbf{b}$ avec $A \in \mathcal{M}_{m\times n}(\mathbb{R})$ et $m > n$ (système surdéterminé, plus d'équations que d'inconnues). En général, ce système n'a pas de solution. On cherche alors $\mathbf{x}^* \in \mathbb{R}^n$ qui minimise $\|A\mathbf{x} - \mathbf{b}\|^2$, c'est-à-dire le vecteur $\mathbf{x}^*$ tel que $A\mathbf{x}^*$ soit la projection orthogonale de $\mathbf{b}$ sur $\mathrm{Im}(A)$.

Proposition 48 (Équations normales)

Le vecteur $\mathbf{x}^*$ minimisant $\|A\mathbf{x} - \mathbf{b}\|^2$ est l'unique solution (si A est de rang n) de :

$$\boxed{A^\mathsf{T} A\,\mathbf{x}^* = A^\mathsf{T}\,\mathbf{b}.}$$

Démonstration. La condition d'optimalité est $\mathbf{b} - A\mathbf{x}^* \perp \mathrm{Im}(A)$, i.e. $\langle A\mathbf{y}, \mathbf{b} - A\mathbf{x}^*\rangle = 0$ pour tout $\mathbf{y} \in \mathbb{R}^n$. Cela donne $\mathbf{y}^\mathsf{T} A^\mathsf{T}(\mathbf{b} - A\mathbf{x}^*) = 0$ pour tout $\mathbf{y}$, d'où $A^\mathsf{T}A\mathbf{x}^* = A^\mathsf{T}\mathbf{b}$. □

7.12.2 Aperçu de la factorisation QR

Le procédé de Gram-Schmidt fournit une factorisation matricielle importante. Si $A \in \mathcal{M}_{m\times n}(\mathbb{R})$ est de rang n (colonnes linéairement indépendantes), en appliquant Gram-Schmidt aux colonnes $\mathbf{a}_1, \dots, \mathbf{a}_n$ de A, on obtient une famille orthonormée $\mathbf{q}_1, \dots, \mathbf{q}_n$. Ceci conduit à :

$$A = QR,$$

où $Q \in \mathcal{M}_{m\times n}(\mathbb{R})$ a des colonnes orthonormées ($Q^\mathsf{T}Q = I_n$) et $R \in \mathcal{M}_n(\mathbb{R})$ est triangulaire supérieure avec des éléments diagonaux strictement positifs. La factorisation QR est un outil fondamental en calcul numérique pour résoudre les systèmes aux moindres carrés et calculer les valeurs propres.

7.12.3 Aperçu des séries de Fourier

Sur l'espace $\mathcal{C}([-\pi, \pi], \mathbb{R})$ muni du produit scalaire $\langle f, g\rangle = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)g(t)\,dt$, la famille :

$$\left(\frac{1}{\sqrt{2}},\ \cos(t),\ \sin(t),\ \cos(2t),\ \sin(2t),\ \dots\right)$$

est orthonormée. La projection orthogonale d'une fonction f sur le sous-espace engendré par les $2N+1$ premières fonctions est le polynôme trigonométrique :

$$S_N f(t) = \frac{a_0}{2} + \sum_{k=1}^{N} (a_k \cos(kt) + b_k \sin(kt)),$$

où $a_k = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\cos(kt)\,dt$ et $b_k = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\sin(kt)\,dt$ sont les *coefficients de Fourier* de f. L'inégalité de Bessel assure :

$$\frac{a_0^2}{2} + \sum_{k=1}^{N}(a_k^2 + b_k^2) \le \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)^2\,dt,$$

et l'identité de Parseval (pour les séries de Fourier) affirme que l'égalité a lieu à la limite $N \to +\infty$.

7.13 Illustrations

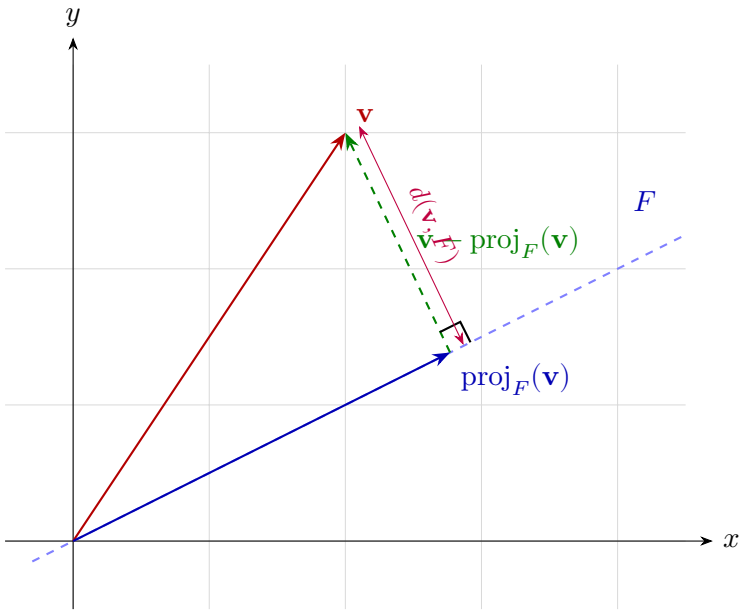


Fig. 7.1 : Projection orthogonale de $\mathbf{v}$ sur un sous-espace F : le projeté $\text{proj}_F(\mathbf{v})$ est le point de F le plus proche de $\mathbf{v}$, et le vecteur $\mathbf{v} - \text{proj}_F(\mathbf{v})$ est orthogonal à F.

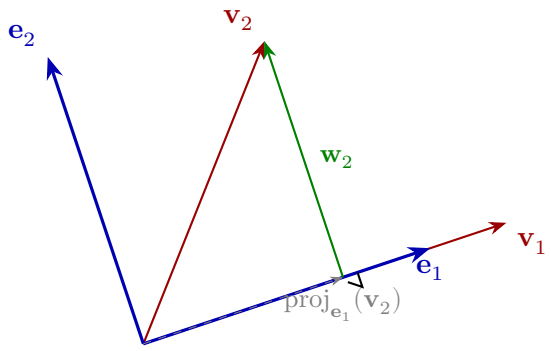


Fig. 7.2 : Procédé de Gram-Schmidt en dimension 2 : à partir de $\mathbf{v}_1$ et $\mathbf{v}_2$, on construit $\mathbf{e}_1$ (normalisation de $\mathbf{v}_1$) puis $\mathbf{w}_2 = \mathbf{v}_2 - \text{proj}_{\mathbf{e}_1}(\mathbf{v}_2)$, et enfin $\mathbf{e}_2$ (normalisation de $\mathbf{w}_2$).

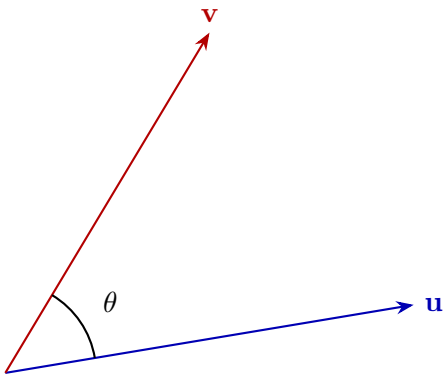
7.14 Exercices

Exercice 49 (Produit scalaire et norme sur $\mathbb{R}^2$)

Soit $E = \mathbb{R}^2$ muni du produit scalaire canonique.

(a) Calculer $\langle \mathbf{u}, \mathbf{v} \rangle$, $\|\mathbf{u}\|$, $\|\mathbf{v}\|$ et l'angle θ entre $\mathbf{u}$ et $\mathbf{v}$, pour $\mathbf{u} = (1, 2)$ et $\mathbf{v} = (3, -1)$.

(b) Trouver tous les vecteurs de $\mathbb{R}^2$ de norme 1 orthogonaux à $\mathbf{u} = (3, 4)$.



$$\cos\theta = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{u}\| \cdot \|\mathbf{v}\|}$$

Fig. 7.3 : Angle θ entre deux vecteurs : le cosinus de l'angle est donné par le rapport du produit scalaire au produit des normes.

Exercice 50 (Vérification des axiomes)

Déterminer si les applications suivantes définissent un produit scalaire sur $\mathbb{R}^2$:

(a) $\varphi(\mathbf{x}, \mathbf{y}) = x_1y_1 - x_1y_2 - x_2y_1 + 4x_2y_2$.

(b) $\psi(\mathbf{x}, \mathbf{y}) = x_1y_1 - x_1y_2 - x_2y_1 + x_2y_2$.

(c) $\chi(\mathbf{x}, \mathbf{y}) = 2x_1y_1 + x_1y_2 + x_2y_1 + 2x_2y_2$.

Indication : vérifier la bilinéarité, la symétrie et le caractère défini positif. Pour ce dernier, examiner la matrice associée.

Exercice 51 (Orthogonal d'un sous-espace)

Dans $\mathbb{R}^3$ muni du produit scalaire canonique, soit $F = \text{Vect}\left(\begin{pmatrix}1\\1\\0\end{pmatrix}, \begin{pmatrix}0\\1\\1\end{pmatrix}\right)$.

(a) Déterminer $F^\perp$.

(b) Vérifier que $\mathbb{R}^3 = F \oplus F^\perp$.

(c) Calculer la projection orthogonale de $\mathbf{v} = \begin{pmatrix}1\\2\\3\end{pmatrix}$ sur F.

Exercice 52 (Gram-Schmidt dans $\mathbb{R}^2$)

Appliquer le procédé de Gram-Schmidt à la base $\mathbf{v}_1 = (3, 4)$, $\mathbf{v}_2 = (1, 0)$ de $\mathbb{R}^2$ muni du produit scalaire canonique.

Exercice 53 (Inégalité de Cauchy-Schwarz)

(a) En appliquant l'inégalité de Cauchy-Schwarz dans $\mathbb{R}^n$ avec des vecteurs bien choisis,

montrer que pour tous réels $a_1, \dots, a_n$:

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

(b) En appliquant Cauchy-Schwarz sur $\mathcal{C}([0,1], \mathbb{R})$, montrer que pour toute $f \in \mathcal{C}([0,1], \mathbb{R})$:

$$\left(\int_0^1 f(t)\, dt\right)^2 \leq \int_0^1 f(t)^2\, dt.$$

Exercice 54 (Matrice de Gram et produit scalaire)

Soit $(\mathbf{v}_1, \dots, \mathbf{v}_p)$ une famille de vecteurs d'un espace préhilbertien E. La *matrice de Gram* est la matrice $G \in \mathcal{M}_p(\mathbb{K})$ définie par $G_{ij} = \langle \mathbf{v}_i, \mathbf{v}_j \rangle$.

(a) Montrer que G est hermitienne (symétrique si $\mathbb{K} = \mathbb{R}$) et semi-définie positive.

(b) Montrer que $\det(G) = 0$ si et seulement si la famille $(\mathbf{v}_i)$ est liée.

(c) Calculer la matrice de Gram de la famille $\mathbf{v}_1 = (1, 0, 1)$, $\mathbf{v}_2 = (1, 1, 0)$ dans $\mathbb{R}^3$.

Exercice 55 (Projection orthogonale et distance)

Dans $\mathbb{R}^4$ muni du produit scalaire canonique, soit $F = \text{Vect}\left(\begin{pmatrix}1\\0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\\1\end{pmatrix}\right)$.

(a) Construire une base orthonormée de F.

(b) Calculer la projection orthogonale de $\mathbf{v} = (1, 2, 3, 4)$ sur F.

(c) En déduire $d(\mathbf{v}, F)$.

Exercice 56 (Gram-Schmidt et factorisation QR)

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

(a) Appliquer le procédé de Gram-Schmidt aux colonnes de A pour obtenir une famille orthonormée $(\mathbf{q}_1, \mathbf{q}_2)$.

(b) Écrire la factorisation $A = QR$ où Q a des colonnes orthonormées et R est triangulaire supérieure.

(c) Utiliser cette factorisation pour résoudre le système aux moindres carrés $A\mathbf{x} \approx \mathbf{b}$ avec $\mathbf{b} = (1, 0, 0)$.

Exercice 57 (Matrices orthogonales)

(a) Montrer que le produit de deux matrices orthogonales est orthogonal.

(b) Montrer que l'inverse d'une matrice orthogonale est orthogonale.

(c) Déterminer toutes les matrices de $\mathrm{O}(2)$ et les classer en rotations et réflexions.

(d) La matrice $\frac{1}{3}\begin{pmatrix} 2 & -1 & 2 \\ 2 & 2 & -1 \\ -1 & 2 & 2 \end{pmatrix}$ est-elle orthogonale ? Si oui, est-elle dans $\mathrm{SO}(3)$?

Exercice 58 (Endomorphisme auto-adjoint)

Soit $E = \mathbb{R}^3$ muni du produit scalaire canonique et $f : E \to E$ défini par $f(x, y, z) = (2x + y,\, x + 2y,\, z)$.

(a) Écrire la matrice A de f dans la base canonique.

(b) Vérifier que f est auto-adjoint.

(c) Diagonaliser f en base orthonormée.

Exercice 59 (Polynômes orthogonaux de Legendre)

Sur $E = \mathbb{R}_2[X]$ (polynômes de degré ≤ 2), on définit le produit scalaire :

$$\langle P, Q \rangle = \int_{-1}^{1} P(t)\, Q(t)\, dt.$$

(a) Vérifier que c'est bien un produit scalaire.

(b) Appliquer le procédé de Gram-Schmidt à la base $(1, X, X^2)$ pour obtenir une base orthogonale (P_0, P_1, P_2).

(c) Normaliser pour obtenir une base orthonormée.

(d) Les polynômes P_0, P_1, P_2 obtenus sont (à un facteur multiplicatif près) les premiers polynômes de Legendre. Vérifier que $P_2(t)$ est proportionnel à $\frac{1}{2}(3t^2 - 1)$.

Exercice 60 (Moindres carrés : régression linéaire)

On dispose des données $(x_i, y_i)_{1 \leq i \leq 4}$ suivantes : $(1, 2)$, $(2, 3)$, $(3, 5)$, $(4, 4)$. On cherche la droite $y = a + bx$ qui minimise $\sum_{i=1}^{4} (y_i - a - bx_i)^2$.

(a) Formuler ce problème comme un système surdéterminé $A\mathbf{c} = \mathbf{y}$ avec $\mathbf{c} = (a, b)$.

(b) Écrire et résoudre les équations normales.

(c) Calculer l'erreur résiduelle $\|A\mathbf{c}^* - \mathbf{y}\|$.

Exercice 61 (Isométries du plan)

Soit $f : \mathbb{R}^2 \to \mathbb{R}^2$ un endomorphisme dont la matrice dans la base canonique est $A \in \mathcal{M}_2(\mathbb{R})$.

(a) Montrer que f est une isométrie si et seulement si $A \in \mathrm{O}(2)$.

(b) En déduire que toute isométrie de $\mathbb{R}^2$ est soit une rotation, soit une réflexion.

(c) Soit f la réflexion par rapport à la droite passant par l'origine et faisant un angle α avec l'axe des abscisses. Montrer que :

$$\mathcal{M}(f) = \begin{pmatrix} \cos 2\alpha & \sin 2\alpha \\ \sin 2\alpha & -\cos 2\alpha \end{pmatrix}.$$

(d) Montrer que la composée de deux réflexions par rapport à des droites passant par l'origine est une rotation, et déterminer son angle en fonction des angles des deux droites.

Exercice 62 (Adjoint et noyau)

Soit $f \in \mathrm{End}(E)$ sur un espace euclidien de dimension n.

(a) Montrer que $\mathrm{rg}(f) = \mathrm{rg}(f^*) = \mathrm{rg}(A^\mathsf{T} A)$ où $A = \mathcal{M}(f)$ dans une base orthonormée.

(b) Montrer que $\mathrm{Ker}(f^* \circ f) = \mathrm{Ker}(f)$.
Indication : considérer $\langle f^*(f(\mathbf{v})), \mathbf{v}\rangle = \|f(\mathbf{v})\|^2$.

(c) En déduire que $\mathrm{rg}(A^\mathsf{T} A) = \mathrm{rg}(A)$ pour toute matrice réelle A.

Exercice 63 (Approximation de Fourier)

Sur $\mathcal{C}([-\pi, \pi], \mathbb{R})$ muni du produit scalaire $\langle f, g\rangle = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)g(t)\,dt$, on considère $f(t) = t$.

(a) Calculer a_0, $a_k = \frac{1}{\pi}\int_{-\pi}^{\pi} t\cos(kt)\,dt$ et $b_k = \frac{1}{\pi}\int_{-\pi}^{\pi} t\sin(kt)\,dt$ pour $k \geq 1$.

(b) Écrire le polynôme trigonométrique $S_N f(t)$ de meilleure approximation de f à l'ordre N.

(c) En utilisant l'identité de Parseval (à la limite), montrer que $\sum_{k=1}^{+\infty} \frac{1}{k^2} = \frac{\pi^2}{6}$.

Résumé du chapitre

Ce qu'il faut retenir.

1. **Forme bilinéaire et forme quadratique.** Une forme bilinéaire symétrique φ est représentée par une matrice symétrique. La forme quadratique associée est $q(\mathbf{v}) = \varphi(\mathbf{v}, \mathbf{v})$, et on peut retrouver φ à partir de q par polarisation. La loi d'inertie de Sylvester assure que la signature (p, q) est un invariant.

2. **Produit scalaire.** Un produit scalaire est une forme bilinéaire symétrique (cas réel) ou

sesquilinéaire hermitienne (cas complexe), définie positive. Il permet de définir norme, distance et angle.

3. **Cauchy-Schwarz.** $|\langle \mathbf{u}, \mathbf{v}\rangle| \leq \|\mathbf{u}\| \cdot \|\mathbf{v}\|$, avec égalité ssi $\mathbf{u}$ et $\mathbf{v}$ sont colinéaires. Conséquence : inégalité triangulaire.

4. **Orthogonalité.** $\mathbf{u} \perp \mathbf{v}$ ssi $\langle \mathbf{u}, \mathbf{v}\rangle = 0$. Toute famille orthogonale de vecteurs non nuls est libre. En dimension finie, $E = F \oplus F^\perp$ et $\dim F^\perp = \dim E - \dim F$.

5. **Projection orthogonale.** $\mathrm{proj}_F(\mathbf{v})$ est l'unique élément de F le plus proche de $\mathbf{v}$. En base orthonormée de F : $\mathrm{proj}_F(\mathbf{v}) = \sum_i \langle \mathbf{e}_i, \mathbf{v}\rangle\, \mathbf{e}_i$.

6. **Gram-Schmidt.** Algorithme d'orthonormalisation : à partir d'une famille libre, on construit une famille orthonormée engendrant les mêmes sous-espaces. Tout espace préhilbertien de dimension finie admet une base orthonormée.

7. **Parseval et Bessel.** En base orthonormée complète : $\|\mathbf{v}\|^2 = \sum_i |\langle \mathbf{e}_i, \mathbf{v}\rangle|^2$ (Parseval). Pour une famille orthonormée partielle, on a l'inégalité de Bessel $\leq$.

8. **Matrices orthogonales et unitaires.** $Q \in \mathrm{O}(n) \Leftrightarrow Q^\mathsf{T} Q = I_n$: les colonnes forment une base orthonormée. $\det(Q) = \pm 1$; $\mathrm{SO}(n) = \{\, Q \in \mathrm{O}(n) \mid \det Q = 1 \,\}$.

9. **Adjoint.** f^* est l'unique endomorphisme tel que $\langle f(\mathbf{u}), \mathbf{v}\rangle = \langle \mathbf{u}, f^*(\mathbf{v})\rangle$. En base orthonormée, $\mathcal{M}(f^*) = \mathcal{M}(f)^*$. $\mathrm{Ker}(f^*) = (\mathrm{Im}\, f)^\perp$.

10. **Applications.**
    - Moindres carrés : $A^\mathsf{T} A \mathbf{x}^* = A^\mathsf{T} \mathbf{b}$.
    - Factorisation QR : $A = QR$ (Gram-Schmidt sur les colonnes).
    - Séries de Fourier : projection orthogonale sur les sous-espaces trigonométriques.

Chapitre 8

Théorème spectral et matrices symétriques

Introduction

Les matrices symétriques occupent une place privilégiée en algèbre linéaire : elles apparaissent naturellement dans de très nombreux contextes appliqués et jouissent de propriétés spectrales remarquables qui les rendent particulièrement maniables.

Voici quelques-unes de leurs occurrences les plus fréquentes :

- **Matrices de covariance :** en statistiques, la matrice de covariance $\Sigma = \mathbb{E}[(X-\mu)(X-\mu)^\mathsf{T}]$ d'un vecteur aléatoire est symétrique (et semi-définie positive). L'analyse en composantes principales (ACP) repose entièrement sur sa diagonalisation.

- **Hessiennes :** en optimisation, la matrice hessienne $H_f(x) = (\frac{\partial^2 f}{\partial x_i \partial x_j})$ d'une fonction de classe $\mathcal{C}^2$ est symétrique, et la nature de ses valeurs propres détermine la nature des points critiques.

- **Formes quadratiques :** toute forme quadratique $q(\mathbf{x}) = \mathbf{x}^\mathsf{T} A\mathbf{x}$ peut être associée à une matrice symétrique A, et sa classification (définie positive, semi-définie, indéfinie) est dictée par les valeurs propres de A.

- **Vibrations et modes propres :** en mécanique, l'équation $M\ddot{x} + Kx = 0$ conduit au problème aux valeurs propres généralisé $Kv = \omega^2 Mv$ où K et M sont symétriques.

- **Graphes et réseaux :** la matrice laplacienne d'un graphe est symétrique et semi-définie positive ; ses valeurs propres gouvernent la connectivité et les processus de diffusion sur le graphe.

Le résultat central de ce chapitre — le *théorème spectral* — affirme que toute matrice symétrique réelle (ou hermitienne complexe) est *orthogonalement diagonalisable* : il existe une base orthonormée de vecteurs propres. Ce théorème est l'un des résultats les plus importants et les plus utilisés de toute l'algèbre linéaire.

Dans tout ce chapitre, sauf mention contraire, E désigne un espace vectoriel de dimension finie $n \geq 1$ muni d'un produit scalaire (réel) ou hermitien (complexe). On note $\langle \cdot, \cdot \rangle$ ce produit et $\|\cdot\|$ la norme associée.

8.1 Opérateurs auto-adjoints

Définition 1 (Adjoint d'un endomorphisme)

Soit E un espace préhilbertien de dimension finie et $f \in \mathrm{End}(E)$. L'*adjoint* de f, noté f^*, est l'unique endomorphisme de E vérifiant

$$\forall \mathbf{u}, \mathbf{v} \in E, \qquad \langle f(\mathbf{u}), \mathbf{v} \rangle = \langle \mathbf{u}, f^*(\mathbf{v}) \rangle.$$

Remarque 2 (Existence et unicité de l'adjoint)

L'existence de f^* découle du théorème de représentation de Riesz appliqué à la forme linéaire $\mathbf{v} \mapsto \langle f(\mathbf{u}), \mathbf{v} \rangle$. L'unicité provient de la non-dégénérescence du produit scalaire : si $\langle \mathbf{u}, g_1(\mathbf{v}) \rangle = \langle \mathbf{u}, g_2(\mathbf{v}) \rangle$ pour tout $\mathbf{u}$, alors $g_1 = g_2$.

Proposition 3 (Matrice de l'adjoint)

Soit $\mathcal{B}$ une base *orthonormée* de E et $A = \mathcal{M}_{\mathcal{B}}(f)$.

(i) Sur $\mathbb{R}$: $\mathcal{M}_{\mathcal{B}}(f^*) = A^{\mathsf{T}}$.

(ii) Sur $\mathbb{C}$: $\mathcal{M}_{\mathcal{B}}(f^*) = A^* = \overline{A^{\mathsf{T}}}$.

Démonstration. Soient $(\mathbf{e}_1, \dots, \mathbf{e}_n)$ la base orthonormée $\mathcal{B}$, et $A = (a_{ij})$, $B = (b_{ij})$ les matrices de f et f^*. On a

$$b_{ij} = \langle f^*(\mathbf{e}_j), \mathbf{e}_i \rangle = \langle \mathbf{e}_j, f(\mathbf{e}_i) \rangle = \overline{\langle f(\mathbf{e}_i), \mathbf{e}_j \rangle} = \overline{a_{ji}}.$$

Donc $B = A^*$. Sur $\mathbb{R}$, $\overline{a_{ji}} = a_{ji}$, d'où $B = A^{\mathsf{T}}$. □

Définition 4 (Opérateur auto-adjoint (symétrique/hermitien))

Un endomorphisme $f \in \mathrm{End}(E)$ est dit *auto-adjoint* (ou *symétrique* si $\mathbb{K} = \mathbb{R}$, *hermitien* si $\mathbb{K} = \mathbb{C}$) si $f^* = f$, c'est-à-dire

$$\forall \mathbf{u}, \mathbf{v} \in E, \qquad \langle f(\mathbf{u}), \mathbf{v} \rangle = \langle \mathbf{u}, f(\mathbf{v}) \rangle.$$

En termes matriciels (dans une base orthonormée) :

- sur $\mathbb{R}$: $A^{\mathsf{T}} = A$ (matrice symétrique),
- sur $\mathbb{C}$: $A^* = A$ (matrice hermitienne).

Exemple 5 (Matrices symétriques et hermitiennes)

La matrice $A = \begin{pmatrix} 2 & -1 \\ -1 & 3 \end{pmatrix}$ est symétrique réelle : $A^{\mathsf{T}} = A$.

La matrice $B = \begin{pmatrix} 3 & 2-i \\ 2+i & 5 \end{pmatrix}$ est hermitienne : $B^* = B$. On note que les éléments diagonaux sont réels et que $b_{ji} = \overline{b_{ij}}$.

Proposition 6 (Propriétés élémentaires des opérateurs auto-adjoints)

Soit f un opérateur auto-adjoint sur E.

(i) $\langle f(\mathbf{v}), \mathbf{v}\rangle \in \mathbb{R}$ pour tout $\mathbf{v} \in E$ (même si $\mathbb{K} = \mathbb{C}$).

(ii) Si g est aussi auto-adjoint et $\alpha, \beta \in \mathbb{R}$, alors $\alpha f + \beta g$ est auto-adjoint.

(iii) Si f est inversible, alors f^{-1} est auto-adjoint.

(iv) $\mathrm{Ker}(f)$ et $\mathrm{Im}(f)$ sont des sous-espaces orthogonaux : $\mathrm{Ker}(f) = \mathrm{Im}(f)^{\perp}$.

Démonstration.

(i) $\langle f(\mathbf{v}), \mathbf{v}\rangle = \langle \mathbf{v}, f(\mathbf{v})\rangle = \overline{\langle f(\mathbf{v}), \mathbf{v}\rangle}$, donc $\langle f(\mathbf{v}), \mathbf{v}\rangle \in \mathbb{R}$.

(ii) $(\alpha f + \beta g)^* = \alpha f^* + \beta g^* = \alpha f + \beta g$ (car $\alpha, \beta \in \mathbb{R}$).

(iii) $\langle f^{-1}(\mathbf{u}), \mathbf{v}\rangle = \langle f^{-1}(\mathbf{u}), f(f^{-1}(\mathbf{v}))\rangle = \langle f(f^{-1}(\mathbf{u})), f^{-1}(\mathbf{v})\rangle = \langle \mathbf{u}, f^{-1}(\mathbf{v})\rangle$.

(iv) $\mathbf{v} \in \mathrm{Ker}(f) \iff f(\mathbf{v}) = \mathbf{0} \iff \langle f(\mathbf{v}), \mathbf{u}\rangle = 0\ \forall \mathbf{u} \iff \langle \mathbf{v}, f(\mathbf{u})\rangle = 0\ \forall \mathbf{u} \iff \mathbf{v} \in \mathrm{Im}(f)^{\perp}$. □

8.2 Valeurs propres réelles

Théorème 7 (Les valeurs propres d'un opérateur auto-adjoint sont réelles)

Soit f un opérateur auto-adjoint sur un espace préhilbertien complexe (ou euclidien) de dimension finie. Alors toutes les valeurs propres de f sont réelles.

Démonstration. Soit $\lambda \in \mathbb{C}$ une valeur propre de f et $\mathbf{v} \neq \mathbf{0}$ un vecteur propre associé : $f(\mathbf{v}) = \lambda \mathbf{v}$.

Calculons $\langle f(\mathbf{v}), \mathbf{v}\rangle$ de deux manières.

D'une part :

$$\langle f(\mathbf{v}), \mathbf{v}\rangle = \langle \lambda \mathbf{v}, \mathbf{v}\rangle = \lambda \langle \mathbf{v}, \mathbf{v}\rangle = \lambda \|\mathbf{v}\|^2.$$

D'autre part, comme f est auto-adjoint :

$$\langle f(\mathbf{v}), \mathbf{v}\rangle = \langle \mathbf{v}, f(\mathbf{v})\rangle = \langle \mathbf{v}, \lambda \mathbf{v}\rangle = \overline{\lambda} \langle \mathbf{v}, \mathbf{v}\rangle = \overline{\lambda} \|\mathbf{v}\|^2.$$

Puisque $\mathbf{v} \neq \mathbf{0}$, $\|\mathbf{v}\|^2 > 0$, et en égalant :

$$\lambda \|\mathbf{v}\|^2 = \overline{\lambda} \|\mathbf{v}\|^2 \implies \lambda = \overline{\lambda} \implies \lambda \in \mathbb{R}.$$ □

Remarque 8 (Nécessité du produit scalaire)

Ce résultat est faux sans hypothèse sur le produit scalaire : la matrice $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ vérifie $A^{\mathsf{T}} = -A$ (antisymétrique) et possède les valeurs propres $\pm i$. En revanche, pour une matrice symétrique réelle, la preuve ci-dessus s'applique en travaillant dans $\mathbb{C}^n$ muni du produit hermitien standard.

8.3 Orthogonalité des vecteurs propres

Théorème 9 (Orthogonalité des sous-espaces propres)

Soit f un opérateur auto-adjoint. Si λ et μ sont deux valeurs propres *distinctes* de f, alors les sous-espaces propres $E_\lambda(f)$ et $E_\mu(f)$ sont orthogonaux :

$$\lambda \neq \mu \implies E_\lambda(f) \perp E_\mu(f).$$

Démonstration. Soient $\mathbf{u} \in E_\lambda(f)$ et $\mathbf{v} \in E_\mu(f)$. Alors $f(\mathbf{u}) = \lambda\mathbf{u}$ et $f(\mathbf{v}) = \mu\mathbf{v}$.

D'une part :

$$\langle f(\mathbf{u}), \mathbf{v}\rangle = \langle \lambda\mathbf{u}, \mathbf{v}\rangle = \lambda\langle \mathbf{u}, \mathbf{v}\rangle.$$

(On utilise que $\lambda \in \mathbb{R}$ d'après le théorème 7.)

D'autre part, par auto-adjonction :

$$\langle f(\mathbf{u}), \mathbf{v}\rangle = \langle \mathbf{u}, f(\mathbf{v})\rangle = \langle \mathbf{u}, \mu\mathbf{v}\rangle = \overline{\mu}\langle \mathbf{u}, \mathbf{v}\rangle = \mu\langle \mathbf{u}, \mathbf{v}\rangle.$$

En soustrayant :

$$(\lambda - \mu)\langle \mathbf{u}, \mathbf{v}\rangle = 0.$$

Comme $\lambda \neq \mu$, on conclut $\langle \mathbf{u}, \mathbf{v}\rangle = 0$. □

8.4 Théorème spectral pour les matrices symétriques réelles

Théorème 10 (Théorème spectral réel)

Soit $A \in \mathcal{M}_n(\mathbb{R})$ une matrice symétrique ($A^\mathsf{T} = A$). Alors :

(i) Toutes les valeurs propres de A sont réelles.

(ii) Il existe une matrice orthogonale $P \in \mathrm{O}(n)$ ($P^\mathsf{T}P = I_n$) et une matrice diagonale $D = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$ telles que
$$A = PDP^\mathsf{T}.$$

(iii) Les colonnes de P forment une base orthonormée de $\mathbb{R}^n$ constituée de vecteurs propres de A.

De manière intrinsèque : tout opérateur auto-adjoint sur un espace euclidien de dimension finie admet une base orthonormée de vecteurs propres.

Démonstration. La preuve procède par récurrence forte sur la dimension n.

Initialisation : pour $n = 1$, toute matrice $A = (a)$ est triviallement symétrique et diagonale ; on prend $P = (1)$ et $D = (a)$.

Hérédité : supposons le résultat vrai pour tout espace euclidien de dimension $< n$, et soit f un opérateur auto-adjoint sur un espace euclidien E de dimension n.

Étape 1 : Existence d'une valeur propre réelle.

Considérons la matrice $A = \mathcal{M}_{\mathcal{B}}(f)$ dans une base orthonormée $\mathcal{B}$. Comme $A^\mathsf{T} = A$, on peut voir A comme une matrice hermitienne dans $\mathbb{C}^n$. Le polynôme caractéristique $\chi_A(\lambda) = \det(A - \lambda I_n)$ est de degré n à coefficients réels. Par le théorème fondamental de l'algèbre, il possède au moins une racine $\lambda_1 \in \mathbb{C}$. Par le théorème 7, $\lambda_1 \in \mathbb{R}$.

Étape 2 : Choix d'un vecteur propre unitaire.

Soit $\mathbf{v}_1 \in E$ un vecteur propre associé à λ_1 avec $\|\mathbf{v}_1\| = 1$.

Étape 3 : Invariance du complément orthogonal.

Posons $W = (\mathrm{Vect}(\mathbf{v}_1))^\perp = \{\mathbf{w} \in E : \langle \mathbf{w}, \mathbf{v}_1 \rangle = 0\}$. Montrons que W est stable par f.

Soit $\mathbf{w} \in W$. Alors :

$$\langle f(\mathbf{w}), \mathbf{v}_1 \rangle = \langle \mathbf{w}, f(\mathbf{v}_1) \rangle = \langle \mathbf{w}, \lambda_1 \mathbf{v}_1 \rangle = \lambda_1 \langle \mathbf{w}, \mathbf{v}_1 \rangle = 0.$$

Donc $f(\mathbf{w}) \in W$: le sous-espace W est invariant par f.

Étape 4 : Application de l'hypothèse de récurrence.

La restriction $g = f|_W \in \mathrm{End}(W)$ est auto-adjointe sur l'espace euclidien W (le produit scalaire est hérité de E) de dimension $n-1$. Par hypothèse de récurrence, il existe une base orthonormée $(\mathbf{v}_2, \dots, \mathbf{v}_n)$ de W constituée de vecteurs propres de g (donc de f), associés à des valeurs propres réelles $\lambda_2, \dots, \lambda_n$.

Étape 5 : Conclusion.

La famille $(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n)$ est une base orthonormée de E (car $\mathbf{v}_1 \perp W$ et $(\mathbf{v}_2, \dots, \mathbf{v}_n)$ est orthonormée dans W). Chaque $\mathbf{v}_i$ est vecteur propre de f pour λ_i.

En posant $P = (\mathbf{v}_1 \mid \mathbf{v}_2 \mid \cdots \mid \mathbf{v}_n)$, la matrice dont les colonnes sont les $\mathbf{v}_i$, on obtient $P \in \mathrm{O}(n)$ et $P^\mathsf{T} A P = D = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$, soit $A = PDP^\mathsf{T}$. □

Corollaire 11 (Existence d'une base orthonormée de vecteurs propres)

Soit $A \in \mathrm{Sym}_n(\mathbb{R})$. Alors $\mathbb{R}^n$ possède une base orthonormée formée de vecteurs propres de A, et le polynôme caractéristique de A est scindé sur $\mathbb{R}$.

Remarque 12 (Interprétation géométrique)

L'écriture $A = PDP^\mathsf{T}$ signifie que l'application linéaire associée à A agit par :

(1) rotation (passage en coordonnées dans la base propre via P^T),

(2) dilatation le long des axes propres (multiplication par les valeurs propres via D),

(3) rotation inverse (retour à la base de départ via P).

8.5 Théorème spectral pour les matrices hermitiennes

Théorème 13 (Théorème spectral hermitien)

Soit $A \in \mathcal{M}_n(\mathbb{C})$ une matrice hermitienne ($A^* = A$). Alors :

(i) Toutes les valeurs propres de A sont réelles.

(ii) Il existe une matrice unitaire $U \in \mathcal{M}_n(\mathbb{C})$ ($U^*U = I_n$) et une matrice diagonale réelle $D = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$ telles que

$$A = UDU^*.$$

(iii) Les colonnes de U forment une base orthonormée de $\mathbb{C}^n$ (pour le produit hermitien standard) constituée de vecteurs propres de A.

Démonstration. La preuve est identique à celle du théorème 10, en remplaçant « orthogonal » par « unitaire », « transposée » par « transconjuguée » et le produit scalaire euclidien par le produit hermitien. L'existence d'une valeur propre est immédiate car $\mathbb{C}$ est algébriquement

clos, et la réalité des valeurs propres est assurée par le théorème 7. L'argument d'invariance du complément orthogonal reste valide : si $f(\mathbf{v}_1) = \lambda_1 \mathbf{v}_1$ avec $\lambda_1 \in \mathbb{R}$ et $\mathbf{w} \perp \mathbf{v}_1$, alors $\langle f(\mathbf{w}), \mathbf{v}_1 \rangle = \langle \mathbf{w}, f(\mathbf{v}_1) \rangle = \overline{\lambda_1} \langle \mathbf{w}, \mathbf{v}_1 \rangle = 0$. □

8.6 Diagonalisation orthogonale : procédure et exemples

La procédure de *diagonalisation orthogonale* d'une matrice symétrique réelle $A \in \mathrm{Sym}_n(\mathbb{R})$ se déroule en quatre étapes :

Étape 1. **Polynôme caractéristique :** calculer $\chi_A(\lambda) = \det(A - \lambda I_n)$ et trouver ses racines $\lambda_1, \dots, \lambda_r$ (avec multiplicités).

Étape 2. **Sous-espaces propres :** pour chaque λ_i, déterminer $E_{\lambda_i} = \mathrm{Ker}(A - \lambda_i I_n)$ par résolution du système homogène.

Étape 3. **Orthonormalisation :** si $\dim E_{\lambda_i} > 1$, appliquer le procédé de Gram–Schmidt pour obtenir une base orthonormée de chaque sous-espace propre. (Les vecteurs propres associés à des valeurs propres *distinctes* sont déjà orthogonaux par le théorème 9.)

Étape 4. **Assemblage :** former P en juxtaposant les vecteurs propres orthonormés, et $D = P^{\mathsf{T}} A P$.

Exemple 14 (Diagonalisation orthogonale d'une matrice 2×2)

Soit $A = \begin{pmatrix} 5 & 2 \\ 2 & 2 \end{pmatrix}$.

Étape 1. $\chi_A(\lambda) = (5-\lambda)(2-\lambda) - 4 = \lambda^2 - 7\lambda + 6 = (\lambda - 1)(\lambda - 6)$. Valeurs propres : $\lambda_1 = 1$, $\lambda_2 = 6$.

Étape 2.

- $E_1 = \mathrm{Ker}(A - I_2) = \mathrm{Ker} \begin{pmatrix} 4 & 2 \\ 2 & 1 \end{pmatrix} = \mathrm{Vect} \begin{pmatrix} 1 \\ -2 \end{pmatrix}$.
- $E_6 = \mathrm{Ker}(A - 6I_2) = \mathrm{Ker} \begin{pmatrix} -1 & 2 \\ 2 & -4 \end{pmatrix} = \mathrm{Vect} \begin{pmatrix} 2 \\ 1 \end{pmatrix}$.

Étape 3. On normalise : $\mathbf{v}_1 = \frac{1}{\sqrt{5}} \begin{pmatrix} 1 \\ -2 \end{pmatrix}$, $\mathbf{v}_2 = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 \\ 1 \end{pmatrix}$. On vérifie : $\langle \mathbf{v}_1, \mathbf{v}_2 \rangle = \frac{1}{5}(2 - 2) = 0$.

Étape 4. $P = \frac{1}{\sqrt{5}} \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix}$, $D = \begin{pmatrix} 1 & 0 \\ 0 & 6 \end{pmatrix}$, et $A = PDP^{\mathsf{T}}$.

Exemple 15 (Diagonalisation orthogonale d'une matrice 3×3)

Soit $A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$.

Étape 1. On calcule :

$$\chi_A(\lambda) = -(\lambda - 4)(\lambda - 1)^2.$$

Valeurs propres : $\lambda_1 = 4$ (multiplicité 1), $\lambda_2 = 1$ (multiplicité 2).

Étape 2.

- $E_4 = \mathrm{Ker}(A - 4I_3) = \mathrm{Vect} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.

- $E_1 = \mathrm{Ker}(A - I_3) = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \,\middle|\, x + y + z = 0 \right\} = \mathrm{Vect}\left(\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \right)$.

Étape 3. Pour E_4 : $\mathbf{v}_1 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.

Pour E_1, on applique Gram–Schmidt. Posons $\mathbf{w}_1 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$, $\mathbf{w}_2 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$.

$\mathbf{u}_1 = \mathbf{w}_1$, $\|\mathbf{u}_1\| = \sqrt{2}$.

$\mathbf{u}_2 = \mathbf{w}_2 - \frac{\langle \mathbf{w}_2, \mathbf{u}_1 \rangle}{\|\mathbf{u}_1\|^2} \mathbf{u}_1 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -1/2 \\ -1/2 \\ 1 \end{pmatrix}$. On a $\|\mathbf{u}_2\| = \sqrt{3/2}$.

Après normalisation : $\mathbf{v}_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$, $\mathbf{v}_3 = \frac{1}{\sqrt{6}} \begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix}$.

Étape 4.

$$P = \begin{pmatrix} 1/\sqrt{3} & -1/\sqrt{2} & -1/\sqrt{6} \\ 1/\sqrt{3} & 1/\sqrt{2} & -1/\sqrt{6} \\ 1/\sqrt{3} & 0 & 2/\sqrt{6} \end{pmatrix}, \qquad D = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

8.7 Opérateurs normaux

Définition 16 (Opérateur normal)

Un endomorphisme $f \in \mathrm{End}(E)$ est dit *normal* si f commute avec son adjoint :

$$f \circ f^* = f^* \circ f.$$

En termes matriciels (dans une base orthonormée), A est normale si $A^* A = A A^*$.

Remarque 17 (Exemples d'opérateurs normaux)

Les classes suivantes sont toutes constituées d'opérateurs normaux :

- opérateurs auto-adjoints ($f^* = f$),
- opérateurs anti-auto-adjoints ($f^* = -f$),
- opérateurs unitaires/orthogonaux ($f^* = f^{-1}$).

Lemme 18 (Caractérisation par la norme)

f est normal si et seulement si $\|f(\mathbf{v})\| = \|f^*(\mathbf{v})\|$ pour tout $\mathbf{v} \in E$.

Démonstration. On a $\|f(\mathbf{v})\|^2 = \langle f(\mathbf{v}), f(\mathbf{v}) \rangle = \langle f^* f(\mathbf{v}), \mathbf{v} \rangle$ et de même $\|f^*(\mathbf{v})\|^2 = \langle f f^*(\mathbf{v}), \mathbf{v} \rangle$. Donc $\|f(\mathbf{v})\|^2 - \|f^*(\mathbf{v})\|^2 = \langle (f^* f - f f^*)(\mathbf{v}), \mathbf{v} \rangle$.

Si f est normal, $f^* f = f f^*$ et l'expression est nulle. Réciproquement, si $\langle g(\mathbf{v}), \mathbf{v} \rangle = 0$ pour tout $\mathbf{v}$ avec $g = f^* f - f f^*$ auto-adjoint, alors $g = 0$ (par polarisation), donc $f^* f = f f^*$. □

Théorème 19 (Théorème spectral pour les opérateurs normaux sur $\mathbb{C}$)

Soit E un espace hermitien de dimension finie et $f \in \mathrm{End}(E)$. Les assertions suivantes sont équivalentes :

(i) f est normal : $f \circ f^* = f^* \circ f$.

(ii) Il existe une base orthonormée de E constituée de vecteurs propres de f.

En termes matriciels : $A \in \mathcal{M}_n(\mathbb{C})$ est normale si et seulement s'il existe une matrice unitaire U telle que $A = U \operatorname{diag}(\lambda_1, \dots, \lambda_n) U^*$.

Démonstration.
(ii) $\Rightarrow$ (i) : Si $A = UDU^*$ avec D diagonale, alors $A^* = U\overline{D}U^*$, et $AA^* = UD\overline{D}U^* = U\overline{D}DU^* = A^*A$ car $D\overline{D} = \overline{D}D$ (matrices diagonales commutent).

(i) $\Rightarrow$ (ii) : Par récurrence sur $n = \dim E$.

Pour $n = 1$, c'est trivial. Supposons le résultat pour $n-1$ et soit f normal sur E de dimension n.

Comme $\mathbb{C}$ est algébriquement clos, f possède une valeur propre λ_1 et un vecteur propre unitaire $\mathbf{v}_1$: $f(\mathbf{v}_1) = \lambda_1 \mathbf{v}_1$.

Fait clé : $\mathbf{v}_1$ est aussi vecteur propre de f^* pour la valeur propre $\overline{\lambda_1}$. En effet, $\|(f^* - \overline{\lambda_1}\,\mathrm{Id})(\mathbf{v}_1)\| = \|(f - \lambda_1 \,\mathrm{Id})(\mathbf{v}_1)\| = 0$, en utilisant le lemme 18 appliqué à l'opérateur normal $f - \lambda_1 \,\mathrm{Id}$.

Posons $W = (\mathrm{Vect}(\mathbf{v}_1))^\perp$. Pour $\mathbf{w} \in W$: $\langle f(\mathbf{w}), \mathbf{v}_1 \rangle = \langle \mathbf{w}, f^*(\mathbf{v}_1) \rangle = \langle \mathbf{w}, \overline{\lambda_1}\mathbf{v}_1 \rangle = \lambda_1 \langle \mathbf{w}, \mathbf{v}_1 \rangle = 0$. Donc W est stable par f.

De même, $\langle f^*(\mathbf{w}), \mathbf{v}_1 \rangle = \langle \mathbf{w}, f(\mathbf{v}_1) \rangle = \overline{\lambda_1} \langle \mathbf{w}, \mathbf{v}_1 \rangle = 0$, donc W est stable par f^*. Ainsi $g = f|_W$ est normal sur W (de dimension $n-1$). Par hypothèse de récurrence, W possède une base orthonormée de vecteurs propres de g, et on conclut en ajoutant $\mathbf{v}_1$. □

8.8 Matrices définies positives et semi-définies positives

Définition 20 (Matrice définie positive, semi-définie positive)

Soit $A \in \mathrm{Sym}_n(\mathbb{R})$.

(i) A est *définie positive* si $\mathbf{x}^\mathsf{T} A \mathbf{x} > 0$ pour tout $\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$.

(ii) A est *semi-définie positive* si $\mathbf{x}^\mathsf{T} A \mathbf{x} \geq 0$ pour tout $\mathbf{x} \in \mathbb{R}^n$.

(iii) Les notions de *définie négative* et *semi-définie négative* s'obtiennent en inversant les inégalités.

(iv) A est *indéfinie* si elle n'entre dans aucune des catégories précédentes.

On note $A \succ 0$ (définie positive), $A \succeq 0$ (semi-définie positive).

Théorème 21 (Caractérisations des matrices définies positives)

Pour $A \in \mathrm{Sym}_n(\mathbb{R})$, les assertions suivantes sont équivalentes :

(i) A est définie positive : $\mathbf{x}^\mathsf{T} A \mathbf{x} > 0$ pour tout $\mathbf{x} \neq \mathbf{0}$.

(ii) Toutes les valeurs propres de A sont strictement positives : $\lambda_i > 0$ pour $i = 1, \dots, n$.

(iii) Tous les mineurs principaux dominants de A sont strictement positifs (critère de Sylvester) : $\Delta_k = \det(A_{1..k,1..k}) > 0$ pour $k = 1, \dots, n$.

(iv) Il existe une matrice triangulaire inférieure L à coefficients diagonaux strictement positifs telle que $A = LL^\mathsf{T}$ (décomposition de Cholesky).

(v) Il existe une matrice inversible B telle que $A = B^\mathsf{T}B$.

Démonstration. Nous démontrons les implications principales.

(i) $\Rightarrow$ (ii) : Soit λ une valeur propre de A et $\mathbf{v}$ un vecteur propre associé ($\|\mathbf{v}\| = 1$). Alors $0 < \mathbf{v}^\mathsf{T}A\mathbf{v} = \lambda\mathbf{v}^\mathsf{T}\mathbf{v} = \lambda$, donc $\lambda > 0$.

(ii) $\Rightarrow$ (i) : Par le théorème spectral, $A = PDP^\mathsf{T}$ avec $D = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$, $\lambda_i > 0$, et P orthogonale. Pour $\mathbf{x} \neq \mathbf{0}$, posons $\mathbf{y} = P^\mathsf{T}\mathbf{x} \neq \mathbf{0}$:

$$\mathbf{x}^\mathsf{T}A\mathbf{x} = \mathbf{y}^\mathsf{T}D\mathbf{y} = \sum_{i=1}^{n} \lambda_i y_i^2 > 0.$$

(i) $\Rightarrow$ (iii) : Si $A \succ 0$, alors pour tout $k \leq n$, la sous-matrice $A_k = A_{1..k,1..k}$ est définie positive (restreindre à $\mathbf{x} = (x_1, \dots, x_k, 0, \dots, 0)^\mathsf{T}$). Donc ses valeurs propres sont positives, et $\Delta_k = \det(A_k) > 0$.

(ii) $\Rightarrow$ (iv) : Si $\lambda_i > 0$, on pose $A = PDP^\mathsf{T} = PD^{1/2}(PD^{1/2})^\mathsf{T} = (D^{1/2}P^\mathsf{T})^\mathsf{T}(D^{1/2}P^\mathsf{T})$ avec $D^{1/2} = \mathrm{diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n})$. La factorisation de Cholesky proprement dite s'obtient ensuite par un algorithme constructif (élimination de Gauss adaptée).

(iv) $\Rightarrow$ (v) : prendre $B = L$.

(v) $\Rightarrow$ (i) : $\mathbf{x}^\mathsf{T}A\mathbf{x} = \mathbf{x}^\mathsf{T}B^\mathsf{T}B\mathbf{x} = \|B\mathbf{x}\|^2 \geq 0$, avec égalité si et seulement si $B\mathbf{x} = \mathbf{0}$, c'est-à-dire $\mathbf{x} = \mathbf{0}$ car B est inversible. □

Proposition 22 (Critère de Sylvester)

Soit $A \in \mathrm{Sym}_n(\mathbb{R})$.

(i) $A \succ 0 \iff \Delta_k > 0$ pour tout $k = 1, \dots, n$.

(ii) $A \prec 0 \iff (-1)^k\Delta_k > 0$ pour tout $k = 1, \dots, n$.

Pour la semi-définie positivité, il faut vérifier que *tous* les mineurs principaux (pas seulement les dominants) sont ≥ 0.

Exemple 23 (Critère de Sylvester appliqué)

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

$\Delta_1 = 2 > 0$, $\Delta_2 = \det\begin{pmatrix} 2 & -1 \\ -1 & 3 \end{pmatrix} = 6 - 1 = 5 > 0$, $\Delta_3 = \det(A) = 2(6-1) + 1(-2-0) = 10 - 2 = 8 > 0$.
Par le critère de Sylvester, A est définie positive.

8.9 Décomposition en valeurs singulières (SVD)

Théorème 24 (Décomposition en valeurs singulières)

Soit $A \in \mathcal{M}_{m\times n}(\mathbb{R})$ une matrice quelconque. Il existe :

- une matrice orthogonale $U \in \mathrm{O}(m)$,
- une matrice orthogonale $V \in \mathrm{O}(n)$,
- une matrice $\Sigma \in \mathcal{M}_{m\times n}(\mathbb{R})$ de la forme $\Sigma = \mathrm{diag}(\sigma_1, \dots, \sigma_p, 0, \dots, 0)$ avec $p = \min(m, n)$ et $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_p \geq 0$,

telles que $A = U\Sigma V^{\mathsf{T}}$.
Les $\sigma_i > 0$ sont les *valeurs singulières* de A : ce sont les racines carrées des valeurs propres non nulles de $A^{\mathsf{T}}A$ (ou de AA^{T}). Le nombre de valeurs singulières non nulles est égal au rang de A.

Idée de la preuve. La matrice $A^{\mathsf{T}}A \in \mathrm{Sym}_n(\mathbb{R})$ est semi-définie positive. Par le théorème spectral, $A^{\mathsf{T}}A = VDV^{\mathsf{T}}$ avec V orthogonale et $D = \mathrm{diag}(\sigma_1^2, \dots, \sigma_r^2, 0, \dots, 0)$, $\sigma_i > 0$.

Pour $i = 1, \dots, r$, posons $\mathbf{u}_i = \frac{1}{\sigma_i} A\mathbf{v}_i$ où les $\mathbf{v}_i$ sont les colonnes de V. On vérifie que les $\mathbf{u}_i$ sont orthonormés :

$$\langle \mathbf{u}_i, \mathbf{u}_j \rangle = \frac{1}{\sigma_i \sigma_j} \mathbf{v}_i{}^{\mathsf{T}} A^{\mathsf{T}} A \mathbf{v}_j = \frac{\sigma_j^2}{\sigma_i \sigma_j} \delta_{ij} = \delta_{ij}.$$

On complète $(\mathbf{u}_1, \dots, \mathbf{u}_r)$ en une base orthonormée $(\mathbf{u}_1, \dots, \mathbf{u}_m)$ de $\mathbb{R}^m$ et on pose $U = (\mathbf{u}_1 \mid \cdots \mid \mathbf{u}_m)$. Alors $A\mathbf{v}_i = \sigma_i \mathbf{u}_i$ pour $i \leq r$ et $A\mathbf{v}_i = \mathbf{0}$ pour $i > r$, ce qui donne $AV = U\Sigma$, soit $A = U\Sigma V^{\mathsf{T}}$. □

Remarque 25 (Interprétation géométrique de la SVD)

L'écriture $A = U\Sigma V^{\mathsf{T}}$ décompose l'application linéaire $\mathbf{x} \mapsto A\mathbf{x}$ en trois étapes :

(1) rotation/réflexion dans $\mathbb{R}^n$ (via V^{T}),

(2) dilatation le long des axes et éventuellement changement de dimension (via Σ),

(3) rotation/réflexion dans $\mathbb{R}^m$ (via U).

Géométriquement, l'image de la sphère unité de $\mathbb{R}^n$ par A est un ellipsoïde dans $\mathbb{R}^m$ dont les demi-axes ont pour longueurs $\sigma_1, \dots, \sigma_r$.

Exemple 26 (SVD d'une matrice 2×2)

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

$A^{\mathsf{T}}A = \begin{pmatrix} 9 & 0 \\ 0 & 1 \end{pmatrix}$: valeurs propres $9, 1$, donc $\sigma_1 = 3$, $\sigma_2 = 1$. Ici $V = I_2$, $U = I_2$, $\Sigma = A$, et la SVD est triviale.

Pour un exemple non trivial, soit $B = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$.

$B^{\mathsf{T}}B = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$: valeurs propres $2, 0$, donc $\sigma_1 = \sqrt{2}$, $\mathrm{rg}(B) = 1$.

Vecteur propre de $B^\mathsf{T}B$ pour $\lambda = 2$: $\mathbf{v}_1 = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}$, et pour $\lambda = 0$: $\mathbf{v}_2 = \frac{1}{\sqrt{2}}\begin{pmatrix}-1\\1\end{pmatrix}$.

$\mathbf{u}_1 = \frac{1}{\sqrt{2}}B\mathbf{v}_1 = \frac{1}{\sqrt{2}}\cdot\frac{1}{\sqrt{2}}\begin{pmatrix}2\\0\end{pmatrix} = \begin{pmatrix}1\\0\end{pmatrix}$.

On complète : $\mathbf{u}_2 = \begin{pmatrix}0\\1\end{pmatrix}$.

Donc : $B = \begin{pmatrix}1 & 0\\0 & 1\end{pmatrix}\begin{pmatrix}\sqrt{2} & 0\\0 & 0\end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix}1 & -1\\1 & 1\end{pmatrix}^\mathsf{T}$.

8.10 Applications

8.10.1 Analyse en composantes principales (ACP)

L'*analyse en composantes principales* est la technique fondamentale de réduction de dimension en statistiques.

Soit $X \in \mathcal{M}_{n\times p}(\mathbb{R})$ une matrice de données centrées (n observations, p variables). La matrice de covariance empirique est

$$C = \frac{1}{n-1}X^\mathsf{T}X \in \mathrm{Sym}_p(\mathbb{R}).$$

Cette matrice est semi-définie positive. Par le théorème spectral, $C = PDP^\mathsf{T}$ avec $D = \mathrm{diag}(\lambda_1, \dots, \lambda_p)$, $\lambda_1 \geq \cdots \geq \lambda_p \geq 0$.

Les colonnes de P sont les *directions principales* : la première composante principale est la direction de variance maximale (λ_1), la deuxième est la direction de variance maximale orthogonale à la première (λ_2), et ainsi de suite.

La proportion de variance expliquée par les k premières composantes est

$$\frac{\sum_{i=1}^{k}\lambda_i}{\sum_{i=1}^{p}\lambda_i} = \frac{\sum_{i=1}^{k}\lambda_i}{\mathrm{tr}(C)}.$$

8.10.2 Classification des formes quadratiques

Une *forme quadratique* $q\colon \mathbb{R}^n \to \mathbb{R}$ est définie par $q(\mathbf{x}) = \mathbf{x}^\mathsf{T}A\mathbf{x}$ pour une matrice symétrique A unique. Par le théorème spectral, dans la base propre orthonormée de A :

$$q(\mathbf{x}) = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2$$

où $\mathbf{y} = P^\mathsf{T}\mathbf{x}$. La classification est alors immédiate :

- *Définie positive* : tous les $\lambda_i > 0$.
- *Semi-définie positive* : tous les $\lambda_i \geq 0$.
- *Définie négative* : tous les $\lambda_i < 0$.
- *Semi-définie négative* : tous les $\lambda_i \leq 0$.
- *Indéfinie* : $\lambda_i > 0$ et $\lambda_j < 0$ pour certains i, j.

La *signature* (p, q) de la forme (nombre de valeurs propres positives, négatives) est un invariant (loi d'inertie de Sylvester).

8.10.3 Optimisation : min/max des formes quadratiques

Théorème 27 (Principe du min-max (Courant–Fischer))

Soit $A \in \mathrm{Sym}_n(\mathbb{R})$ avec valeurs propres $\lambda_1 \le \lambda_2 \le \cdots \le \lambda_n$. Alors :

$$\lambda_k = \min_{\substack{V \le \mathbb{R}^n \\ \dim V = k}} \max_{\substack{\mathbf{x} \in V \\ \|\mathbf{x}\|=1}} \mathbf{x}^{\mathsf{T}} A \mathbf{x} = \max_{\substack{V \le \mathbb{R}^n \\ \dim V = n-k+1}} \min_{\substack{\mathbf{x} \in V \\ \|\mathbf{x}\|=1}} \mathbf{x}^{\mathsf{T}} A \mathbf{x}.$$

En particulier :

$$\lambda_1 = \min_{\|\mathbf{x}\|=1} \mathbf{x}^{\mathsf{T}} A \mathbf{x}, \qquad \lambda_n = \max_{\|\mathbf{x}\|=1} \mathbf{x}^{\mathsf{T}} A \mathbf{x}.$$

Démonstration. Démontrons le cas particulier $\lambda_n = \max_{\|\mathbf{x}\|=1} \mathbf{x}^{\mathsf{T}} A \mathbf{x}$.

Par le théorème spectral, $A = PDP^{\mathsf{T}}$ avec $D = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$. Posons $\mathbf{y} = P^{\mathsf{T}}\mathbf{x}$; $\|\mathbf{x}\| = 1$ équivaut à $\|\mathbf{y}\| = 1$, et

$$\mathbf{x}^{\mathsf{T}} A \mathbf{x} = \sum_{i=1}^{n} \lambda_i y_i^2 \le \lambda_n \sum_{i=1}^{n} y_i^2 = \lambda_n.$$

Le maximum est atteint pour $\mathbf{y} = \mathbf{e}_n$, c'est-à-dire $\mathbf{x} = P\mathbf{e}_n$ (le vecteur propre associé à λ_n). Le cas de λ_1 est analogue. □

8.10.4 Vibrations et modes propres

Considérons un système mécanique linéarisé au voisinage d'un équilibre :

$$M\ddot{\mathbf{x}} + K\mathbf{x} = \mathbf{0},$$

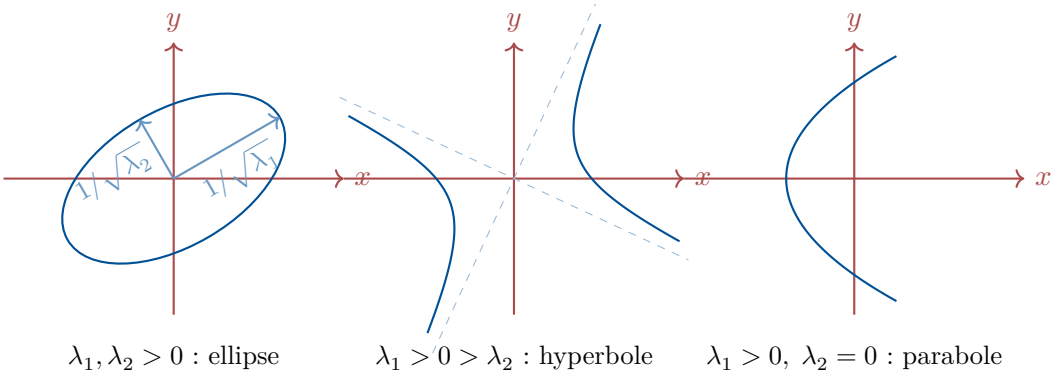
où $M \succ 0$ (matrice de masse) et $K \succeq 0$ (matrice de raideur) sont symétriques. On cherche des solutions harmoniques $\mathbf{x}(t) = \mathbf{v}\cos(\omega t)$, ce qui conduit au problème aux valeurs propres généralisé :

$$K\mathbf{v} = \omega^2 M \mathbf{v}.$$

En posant $\tilde{K} = M^{-1/2} K M^{-1/2}$ (qui est symétrique et semi-définie positive), on se ramène au problème standard $\tilde{K}\mathbf{w} = \omega^2 \mathbf{w}$, $\mathbf{w} = M^{1/2}\mathbf{v}$. Par le théorème spectral, il existe une base orthonormée de *modes propres*, chacun vibrant à sa propre fréquence $\omega_i = \sqrt{\lambda_i}$.

8.11 Illustrations

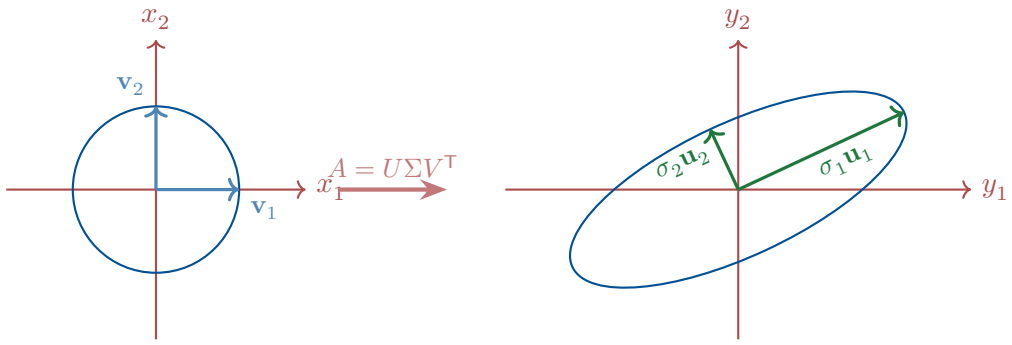
8.11.1 Classification des coniques par les valeurs propres



$\lambda_1, \lambda_2 > 0$: ellipse $\quad$ $\lambda_1 > 0 > \lambda_2$: hyperbole $\quad$ $\lambda_1 > 0,\ \lambda_2 = 0$: parabole

Fig. 8.1 : Classification des coniques $\mathbf{x}^{\mathsf{T}} A \mathbf{x} + \mathbf{b}^{\mathsf{T}} \mathbf{x} + c = 0$ selon les valeurs propres de la partie quadratique A.

8.11.2 Géométrie de la SVD



Cercle unité dans $\mathbb{R}^2$ — Image : ellipse de demi-axes σ_1, σ_2

Fig. 8.2 : La SVD transforme le cercle unité en un ellipsoïde. Les directions principales sont données par les vecteurs singuliers gauches $\mathbf{u}_i$, et les longueurs des demi-axes sont les valeurs singulières σ_i.

8.12 Exercices

Exercice 28 (Diagonalisation orthogonale 2×2)

Diagonaliser orthogonalement la matrice $A = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$. Vérifier que $A = PDP^\top$.

Exercice 29 (Diagonalisation orthogonale 3×3)

Diagonaliser orthogonalement la matrice $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$.

Exercice 30 (Matrice hermitienne)

Diagonaliser unitairement la matrice hermitienne $H = \begin{pmatrix} 2 & 1+i \\ 1-i & 3 \end{pmatrix}$. Vérifier que les valeurs propres sont réelles et que les vecteurs propres sont orthogonaux.

Exercice 31 (Critère de Sylvester)

En utilisant le critère de Sylvester, déterminer la nature (définie positive, définie négative, indéfinie) de :

$$A = \begin{pmatrix} 4 & 2 \\ 2 & 3 \end{pmatrix}, \qquad B = \begin{pmatrix} -3 & 1 \\ 1 & -5 \end{pmatrix}, \qquad C = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}.$$

Exercice 32 (Formes quadratiques et coniques)

Classifier la conique $5x^2 + 4xy + 2y^2 = 6$ en déterminant les valeurs propres de la matrice associée à la forme quadratique. Tracer la conique dans la base propre.

Exercice 33 (Calcul de SVD)

Calculer la décomposition en valeurs singulières de la matrice $A = \begin{pmatrix} 1 & 1 \\ 1 & -1 \\ 0 & 0 \end{pmatrix}$.

Exercice 34 (Décomposition de Cholesky)

Calculer la décomposition de Cholesky de la matrice $A = \begin{pmatrix} 4 & 2 & -2 \\ 2 & 5 & -4 \\ -2 & -4 & 14 \end{pmatrix}$. Vérifier que $A = LL^{\mathsf{T}}$.

Exercice 35 (Puissances d'une matrice symétrique)

Soit $A = \begin{pmatrix} 5 & 4 \\ 4 & 5 \end{pmatrix}$.

(a) Diagonaliser orthogonalement A.

(b) Calculer A^{10}.

(c) Exprimer $e^{tA} = \sum_{k=0}^{\infty} \frac{t^k}{k!} A^k$ sous forme explicite.

Exercice 36 (Opérateur normal non auto-adjoint)

Soit $A = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \in \mathcal{M}_2(\mathbb{R})$.

(a) Montrer que A n'est pas symétrique.

(b) Vérifier que A est normale : $A^{\mathsf{T}}A = AA^{\mathsf{T}}$.

(c) Montrer que A n'est pas diagonalisable sur $\mathbb{R}$ mais l'est sur $\mathbb{C}$. Trouver une matrice unitaire U telle que $A = UDU^{*}$.

Exercice 37 (Approximation de rang faible et SVD)

Soit $A \in \mathcal{M}_{m\times n}(\mathbb{R})$ de SVD $A = U\Sigma V^{\mathsf{T}}$.

(a) Montrer que pour tout $k \leq \mathrm{rg}(A)$, la matrice $A_k = \sum_{i=1}^{k} \sigma_i \mathbf{u}_i \mathbf{v}_i^{\mathsf{T}}$ est de rang k.

(b) Admettre le théorème d'Eckart–Young : A_k est la meilleure approximation de rang k de A au sens de la norme de Frobenius : $\|A - A_k\|_F = \min_{\mathrm{rg}(B)\leq k} \|A - B\|_F$.

(c) Calculer $\|A - A_k\|_F^2$ en fonction des valeurs singulières.

(d) Application : si $A = \begin{pmatrix} 3 & 0 \\ 0 & 2 \\ 0 & 0 \end{pmatrix}$, quelle est la meilleure approximation de rang 1 ?

Exercice 38 (Rayleigh et convergence)

Soit $A \in \mathrm{Sym}_n(\mathbb{R})$ avec valeurs propres $\lambda_1 < \lambda_2 \leq \cdots \leq \lambda_n$. Le *quotient de Rayleigh* est $R(\mathbf{x}) = \frac{\mathbf{x}^\mathsf{T} A \mathbf{x}}{\mathbf{x}^\mathsf{T}\mathbf{x}}$ pour $\mathbf{x} \neq \mathbf{0}$.

(a) Montrer que $\lambda_1 \leq R(\mathbf{x}) \leq \lambda_n$.

(b) Montrer que $R(\mathbf{x}) = \lambda_k$ si et seulement si $\mathbf{x}$ est vecteur propre de A pour λ_k.

(c) Calculer le gradient $\nabla R(\mathbf{x})$ et montrer que ses points critiques sont les vecteurs propres de A.

Exercice 39 (Preuve de la décomposition de Cholesky)

Soit $A \in \mathrm{Sym}_n(\mathbb{R})$ définie positive.

(a) Montrer que $a_{11} > 0$.

(b) Écrire A sous la forme par blocs $A = \begin{pmatrix} a_{11} & \mathbf{c}^\mathsf{T} \\ \mathbf{c} & A' \end{pmatrix}$ et montrer que $S = A' - \frac{1}{a_{11}}\mathbf{c}\mathbf{c}^\mathsf{T}$ (le complément de Schur) est définie positive de taille $(n-1)\times(n-1)$.

(c) En déduire par récurrence l'existence de la décomposition $A = LL^\mathsf{T}$ avec L triangulaire inférieure à diagonale strictement positive.

(d) Montrer l'unicité de cette décomposition.

Résumé du chapitre

Résumé — Chapitre 8

Opérateurs auto-adjoints. f est auto-adjoint si $f = f^*$, i.e. $\langle f(\mathbf{u}), \mathbf{v}\rangle = \langle \mathbf{u}, f(\mathbf{v})\rangle$ pour tous $\mathbf{u}, \mathbf{v}$. En base orthonormée : matrice symétrique ($\mathbb{R}$) ou hermitienne ($\mathbb{C}$).

Propriétés spectrales fondamentales.

- Les valeurs propres d'un opérateur auto-adjoint sont *réelles*.
- Les sous-espaces propres associés à des valeurs propres distinctes sont *orthogonaux*.

Théorème spectral (réel). Toute matrice symétrique réelle $A \in \mathrm{Sym}_n(\mathbb{R})$ est *orthogonalement diagonalisable* : $A = PDP^\mathsf{T}$, $P \in \mathrm{O}(n)$, D diagonale réelle.

Théorème spectral (hermitien). Toute matrice hermitienne $A = A^*$ est *unitairement diagonalisable* : $A = UDU^*$, U unitaire, D diagonale réelle.

Opérateurs normaux. f est normal si $ff^* = f^*f$. Sur $\mathbb{C}$, f est normal $\iff$ f admet une base orthonormée de vecteurs propres.

Matrices définies positives. $A \succ 0 \iff \lambda_i > 0\ \forall i \iff \Delta_k > 0\ \forall k$ (Sylvester) $\iff$ $A = LL^\mathsf{T}$ (Cholesky) $\iff$ $A = B^\mathsf{T}B$, B inversible.

SVD. Toute matrice $A \in \mathcal{M}_{m\times n}(\mathbb{R})$ s'écrit $A = U\Sigma V^\mathsf{T}$ avec $U \in \mathrm{O}(m)$, $V \in \mathrm{O}(n)$, $\Sigma = \mathrm{diag}(\sigma_1, \ldots, \sigma_p)$, $\sigma_i \geq 0$. Les σ_i sont les racines carrées des valeurs propres de $A^\mathsf{T}A$.

Applications. ACP (directions de variance maximale), classification des formes quadratiques et coniques, principe du min-max (Courant–Fischer), approximation de rang faible (Eckart–Young), vibrations et modes propres.

Chapitre 9

Forme Normale de Jordan

Introduction

Au chapitre 6, nous avons vu qu'un endomorphisme f d'un $\mathbb{K}$-espace vectoriel E de dimension finie est diagonalisable si et seulement si ses sous-espaces propres engendrent E. Malheureusement, de nombreuses matrices ne sont pas diagonalisables : par exemple, la matrice

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

possède la valeur propre double $\lambda = 2$, mais son sous-espace propre $E_2(A)$ est de dimension 1. Il est donc impossible de trouver une base de $\mathbb{R}^2$ formée de vecteurs propres de A.

Se pose alors une question naturelle : *quelle est la forme matricielle la plus simple que l'on puisse atteindre par changement de base pour un endomorphisme quelconque ?* La réponse est donnée par la *forme normale de Jordan*, qui constitue la forme canonique la plus fine pour les endomorphismes sur un corps algébriquement clos (typiquement $\mathbb{C}$).

La forme de Jordan généralise la diagonalisation : une matrice diagonalisable a une forme de Jordan diagonale, tandis qu'une matrice non diagonalisable acquiert des 1 sur la sur-diagonale, regroupés en *blocs de Jordan*. Cette structure encode très précisément le défaut de diagonalisabilité.

Les applications sont considérables :

- **Systèmes différentiels :** le calcul de e^{tA} pour une matrice non diagonalisable repose sur la forme de Jordan.
- **Puissances de matrices :** A^k se calcule explicitement via les blocs de Jordan.
- **Stabilité :** la classification des points d'équilibre de systèmes dynamiques linéaires fait intervenir la forme de Jordan.
- **Classification des endomorphismes :** deux matrices sont semblables si et seulement si elles ont la même forme de Jordan.

Dans tout ce chapitre, sauf mention contraire, E désigne un $\mathbb{C}$-espace vectoriel de dimension finie $n \geq 1$, et $f \in \mathrm{End}(E)$.

9.1 Endomorphismes nilpotents

La théorie de Jordan repose de manière essentielle sur la compréhension des endomorphismes nilpotents. Ce sont les endomorphismes dont une certaine puissance s'annule.

Définition 1 (Endomorphisme nilpotent)

Un endomorphisme $f \in \mathrm{End}(E)$ est dit *nilpotent* s'il existe un entier $p \geq 1$ tel que $f^p = 0$ (l'endomorphisme nul). Une matrice $A \in \mathcal{M}_n(\mathbb{C})$ est dite *nilpotente* s'il existe $p \geq 1$ tel que $A^p = 0$.

Définition 2 (Indice de nilpotence)

Soit $f \in \mathrm{End}(E)$ un endomorphisme nilpotent. Le plus petit entier $\nu \geq 1$ tel que $f^\nu = 0$ est appelé l'*indice de nilpotence* de f. On écrit $\nu(f) = \nu$.

Proposition 3 (Propriétés des endomorphismes nilpotents)

Soit $f \in \mathrm{End}(E)$ un endomorphisme nilpotent d'indice ν, avec $\dim E = n$. Alors :

(i) $\nu \leq n$.

(ii) La seule valeur propre de f est 0, c'est-à-dire $\mathrm{Spec}(f) = \{0\}$.

(iii) $\det(f) = 0$ et $\mathrm{tr}(f) = 0$.

(iv) Le polynôme caractéristique de f est $\chi_f(X) = X^n$.

(v) Le polynôme minimal de f est $\mu_f(X) = X^\nu$.

Démonstration. (i) Considérons la suite de sous-espaces emboîtés :

$$\{\mathbf{0}\} = \mathrm{Ker}(f^0) \subseteq \mathrm{Ker}(f) \subseteq \mathrm{Ker}(f^2) \subseteq \cdots \subseteq \mathrm{Ker}(f^\nu) = E.$$

Chaque inclusion $\mathrm{Ker}(f^k) \subseteq \mathrm{Ker}(f^{k+1})$ est large, et l'inclusion est stricte pour $k = 0, 1, \ldots, \nu - 1$ (sinon $f^k = 0$ contredirait la minimalité de ν). Donc $\dim \mathrm{Ker}(f^{k+1}) \geq \dim \mathrm{Ker}(f^k) + 1$, ce qui donne $n = \dim \mathrm{Ker}(f^\nu) \geq \nu$.

(ii) Si λ est valeur propre de f, il existe $\mathbf{v} \neq \mathbf{0}$ tel que $f(\mathbf{v}) = \lambda \mathbf{v}$. Alors $f^\nu(\mathbf{v}) = \lambda^\nu \mathbf{v} = \mathbf{0}$, donc $\lambda^\nu = 0$ et $\lambda = 0$.

(iii) Le déterminant est le produit des valeurs propres (comptées avec multiplicité) sur $\mathbb{C}$, et la trace est leur somme. Puisque la seule valeur propre est 0, on obtient $\det(f) = 0$ et $\mathrm{tr}(f) = 0$.

(iv) Le polynôme caractéristique est de degré n et ses racines (dans $\mathbb{C}$) sont les valeurs propres. Puisque 0 est la seule valeur propre, de multiplicité algébrique n, on a $\chi_f(X) = X^n$.

(v) Le polynôme minimal divise le polynôme caractéristique X^n, donc $\mu_f(X) = X^m$ pour un certain $m \leq n$. Par définition, $f^m = 0$ et $f^{m-1} \neq 0$, donc $m = \nu$. □

Exemple 4 (Matrice nilpotente)

Soit $N = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$. Calculons :

$$N^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad N^3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = 0.$$

Donc N est nilpotente d'indice $\nu(N) = 3$. On vérifie : $\chi_N(X) = X^3$ et $\mu_N(X) = X^3$.

Proposition 5 (Suite des noyaux itérés)

Soit $f \in \mathrm{End}(E)$ nilpotent d'indice ν. On a la chaîne strictement croissante :

$$\{\mathbf{0}\} = \mathrm{Ker}(f^0) \subsetneq \mathrm{Ker}(f) \subsetneq \mathrm{Ker}(f^2) \subsetneq \cdots \subsetneq \mathrm{Ker}(f^\nu) = E.$$

De plus, pour tout $k \geq \nu$, $\mathrm{Ker}(f^k) = E$.

Démonstration. La croissance $\mathrm{Ker}(f^k) \subseteq \mathrm{Ker}(f^{k+1})$ est immédiate : si $f^k(\mathbf{v}) = \mathbf{0}$, alors $f^{k+1}(\mathbf{v}) = f(f^k(\mathbf{v})) = f(\mathbf{0}) = \mathbf{0}$.

Montrons que l'inclusion est stricte pour $k < \nu$. Par définition de l'indice de nilpotence, $f^k \neq 0$ pour $k < \nu$, donc il existe $\mathbf{w} \in E$ tel que $f^k(\mathbf{w}) \neq \mathbf{0}$, c'est-à-dire $\mathbf{w} \notin \mathrm{Ker}(f^k)$. Montrons que si $\mathrm{Ker}(f^k) = \mathrm{Ker}(f^{k+1})$, alors $\mathrm{Ker}(f^j) = \mathrm{Ker}(f^k)$ pour tout $j \geq k$, par récurrence sur j. Si $\mathbf{v} \in \mathrm{Ker}(f^{k+2})$, alors $f(\mathbf{v}) \in \mathrm{Ker}(f^{k+1}) = \mathrm{Ker}(f^k)$, donc $f^{k+1}(\mathbf{v}) = f^k(f(\mathbf{v})) = \mathbf{0}$, d'où $\mathbf{v} \in \mathrm{Ker}(f^{k+1}) = \mathrm{Ker}(f^k)$. On obtient $\mathrm{Ker}(f^k) = E$ pour tout $k < \nu$, ce qui contredit $f^k \neq 0$.

Pour $k \geq \nu$, $f^\nu = 0$ implique $f^k = f^{k-\nu} \circ f^\nu = 0$, donc $\mathrm{Ker}(f^k) = E$. □

9.2 Sous-espaces invariants

Définition 6 (Sous-espace invariant)

Soit $f \in \mathrm{End}(E)$. Un sous-espace vectoriel F de E est dit *f-invariant* (ou *stable par f*) si $f(F) \subseteq F$, c'est-à-dire si pour tout $\mathbf{v} \in F$, on a $f(\mathbf{v}) \in F$.

Exemple 7 (Exemples de sous-espaces invariants)

(i) $\{\mathbf{0}\}$ et E sont toujours f-invariants.

(ii) $\mathrm{Ker}(f)$ et $\mathrm{Im}(f)$ sont f-invariants.

(iii) Tout sous-espace propre $E_\lambda(f) = \mathrm{Ker}(f - \lambda\,\mathrm{Id})$ est f-invariant.

(iv) Plus généralement, $\mathrm{Ker}(P(f))$ est f-invariant pour tout polynôme P.

Preuve de (iv). Soit $\mathbf{v} \in \mathrm{Ker}(P(f))$, c'est-à-dire $P(f)(\mathbf{v}) = \mathbf{0}$. Comme f commute avec $P(f)$ (car $P(f)$ est un polynôme en f), on a :

$$P(f)(f(\mathbf{v})) = f(P(f)(\mathbf{v})) = f(\mathbf{0}) = \mathbf{0},$$

donc $f(\mathbf{v}) \in \mathrm{Ker}(P(f))$. □

Proposition 8 (Restriction à un sous-espace invariant)

Si F est un sous-espace f-invariant, alors la restriction $f|_F \colon F \to F$ est un endomorphisme de F. De plus :

(i) Le polynôme minimal de $f|_F$ divise le polynôme minimal de f.

(ii) Le polynôme caractéristique de $f|_F$ divise le polynôme caractéristique de f.

Démonstration. (i) Si $\mu_f(f) = 0$, alors pour tout $\mathbf{v} \in F$, $\mu_f(f|_F)(\mathbf{v}) = \mu_f(f)(\mathbf{v}) = \mathbf{0}$. Donc μ_f annule $f|_F$, et le polynôme minimal de $f|_F$ divise μ_f.

(ii) Complétons une base $(e_1, \dots, e_p)$ de F en une base $(e_1, \dots, e_p, e_{p+1}, \dots, e_n)$ de E. Dans cette base, la matrice de f a la forme

$$\mathcal{M}(f) = \begin{pmatrix} A & B \\ 0 & C \end{pmatrix},$$

où $A = \mathcal{M}(f|_F) \in \mathcal{M}_p(\mathbb{C})$. Alors $\chi_f(X) = \det(XI_n - \mathcal{M}(f)) = \det(XI_p - A) \cdot \det(XI_{n-p} - C) = \chi_{f|_F}(X) \cdot \chi_C(X)$. □

9.3 Sous-espaces caractéristiques

Les sous-espaces propres ne suffisent pas en général à décomposer E. Il faut les « épaissir » en remplaçant $\operatorname{Ker}(f - \lambda \operatorname{Id})$ par $\operatorname{Ker}((f - \lambda \operatorname{Id})^k)$ pour k assez grand.

Définition 9 (Sous-espace caractéristique)

Soit $f \in \operatorname{End}(E)$ et $\lambda \in \operatorname{Spec}(f)$. Le *sous-espace caractéristique* (ou *sous-espace propre généralisé*) de f associé à λ est

$$F_\lambda(f) := \operatorname{Ker}((f - \lambda \operatorname{Id})^n) = \bigcup_{k=0}^{n} \operatorname{Ker}((f - \lambda \operatorname{Id})^k),$$

où $n = \dim E$.

Définition 10 (Vecteur propre généralisé)

Un vecteur $\mathbf{v} \in E \setminus \{\mathbf{0}\}$ est un *vecteur propre généralisé* de f associé à la valeur propre λ s'il existe un entier $k \geq 1$ tel que $(f - \lambda \operatorname{Id})^k(\mathbf{v}) = \mathbf{0}$.

Proposition 11 (Propriétés des sous-espaces caractéristiques)

Soit $f \in \operatorname{End}(E)$ avec $\operatorname{Spec}(f) = \{\lambda_1, \dots, \lambda_r\}$ (valeurs propres distinctes sur $\mathbb{C}$), et notons m_i la multiplicité algébrique de λ_i. Alors :

(i) Chaque $F_{\lambda_i}(f)$ est un sous-espace f-invariant.

(ii) $E_{\lambda_i}(f) \subseteq F_{\lambda_i}(f)$.

(iii) $\dim F_{\lambda_i}(f) = m_i$.

(iv) $E = F_{\lambda_1}(f) \oplus F_{\lambda_2}(f) \oplus \cdots \oplus F_{\lambda_r}(f)$.

(v) La restriction $(f - \lambda_i \operatorname{Id})|_{F_{\lambda_i}(f)}$ est nilpotente.

Démonstration. (i) $F_{\lambda_i}(f) = \operatorname{Ker}((f - \lambda_i \operatorname{Id})^n)$ est le noyau d'un polynôme en f, donc f-invariant (cf. exemple 7).

(ii) Si $\mathbf{v} \in E_{\lambda_i}(f)$, alors $(f - \lambda_i \operatorname{Id})(\mathbf{v}) = \mathbf{0}$, donc $(f - \lambda_i \operatorname{Id})^n(\mathbf{v}) = \mathbf{0}$, c'est-à-dire $\mathbf{v} \in F_{\lambda_i}(f)$.

(iii) et (iv) : écrivons le polynôme caractéristique sous la forme

$$\chi_f(X) = \prod_{i=1}^{r} (X - \lambda_i)^{m_i}, \qquad \sum_{i=1}^{r} m_i = n.$$

Posons $P_i(X) = \prod_{j \neq i} (X - \lambda_j)^{m_j}$. Les polynômes $P_1, \dots, P_r$ sont premiers entre eux dans leur ensemble (car les λ_i sont distincts), donc il existe des polynômes $U_1, \dots, U_r$ tels que $\sum_{i=1}^{r} U_i(X) P_i(X) = 1$.

Par le théorème de Cayley-Hamilton, $\chi_f(f) = 0$, donc $(f - \lambda_i \operatorname{Id})^{m_i} \circ P_i(f) = 0$ pour tout i, ce qui entraîne $\operatorname{Im}(P_i(f)) \subseteq \operatorname{Ker}((f - \lambda_i \operatorname{Id})^{m_i}) \subseteq F_{\lambda_i}(f)$.

De la relation $\sum_i U_i(f) \circ P_i(f) = \operatorname{Id}$, on déduit que pour tout $\mathbf{v} \in E$:

$$\mathbf{v} = \sum_{i=1}^{r} U_i(f)(P_i(f)(\mathbf{v})),$$

et chaque terme $U_i(f)(P_i(f)(\mathbf{v})) \in F_{\lambda_i}(f)$ (car $F_{\lambda_i}(f)$ est f-invariant et contient $\operatorname{Im}(P_i(f))$). Ainsi $E = \sum_i F_{\lambda_i}(f)$.

Montrons que la somme est directe. Si $\mathbf{v} \in F_{\lambda_i}(f) \cap \sum_{j \neq i} F_{\lambda_j}(f)$, alors $(f - \lambda_i \operatorname{Id})^{m_i}(\mathbf{v}) = \mathbf{0}$ car $\mathbf{v} \in F_{\lambda_i}(f) \subseteq \operatorname{Ker}((f - \lambda_i \operatorname{Id})^{m_i})$. Par ailleurs, $\mathbf{v} = \sum_{j \neq i} \mathbf{v}_j$ avec $\mathbf{v}_j \in F_{\lambda_j}(f)$, et sur chaque $F_{\lambda_j}(f)$ $(j \neq i)$, l'endomorphisme $f - \lambda_i \operatorname{Id}$ est inversible (car $\lambda_j \neq \lambda_i$ implique que λ_i n'est pas valeur propre de $f|_{F_{\lambda_j}(f)}$). Donc $(f - \lambda_i \operatorname{Id})^{m_i}$ est inversible sur $\sum_{j \neq i} F_{\lambda_j}(f)$, ce qui force $\mathbf{v} = \mathbf{0}$.

La somme directe $E = \bigoplus_i F_{\lambda_i}(f)$ donne $n = \sum_i \dim F_{\lambda_i}(f)$. De plus, $F_{\lambda_i}(f) \subseteq \operatorname{Ker}((f - \lambda_i \operatorname{Id})^{m_i})$ (par le raisonnement ci-dessus) et l'on sait que $\dim \operatorname{Ker}((f - \lambda_i \operatorname{Id})^{m_i}) \leq m_i$ (car le polynôme caractéristique de $f|_{F_{\lambda_i}(f)}$ est de degré $\dim F_{\lambda_i}(f)$, et $(X - \lambda_i)^{m_i}$ divise χ_f). L'égalité $\sum_i \dim F_{\lambda_i}(f) = \sum_i m_i = n$ force $\dim F_{\lambda_i}(f) = m_i$.

(v) Sur $F_{\lambda_i}(f)$, tout vecteur $\mathbf{v}$ vérifie $(f - \lambda_i \operatorname{Id})^{m_i}(\mathbf{v}) = \mathbf{0}$, donc $(f - \lambda_i \operatorname{Id})|_{F_{\lambda_i}(f)}$ est nilpotent (d'indice $\leq m_i$). □

9.4 Blocs de Jordan et matrices de Jordan

Définition 12 (Bloc de Jordan)

Soit $\lambda \in \mathbb{C}$ et $k \geq 1$ un entier. Le *bloc de Jordan* de taille k associé à λ est la matrice $J_k(\lambda) \in \mathcal{M}_k(\mathbb{C})$ définie par

$$J_k(\lambda) := \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda & 1 \\ 0 & \cdots & \cdots & 0 & \lambda \end{pmatrix} = \lambda I_k + N_k,$$

où N_k est la matrice nilpotente ayant des 1 sur la sur-diagonale et des 0 partout ailleurs.

Remarque 13 (Cas particuliers)

- Pour $k = 1$: $J_1(\lambda) = (\lambda)$, c'est un scalaire.
- Un bloc de Jordan $J_k(\lambda)$ avec $\lambda \neq 0$ n'est pas nilpotent, mais $J_k(\lambda) - \lambda I_k = N_k$ l'est.
- Un bloc $J_k(0) = N_k$ est nilpotent d'indice k.

Définition 14 (Matrice de Jordan)

Une *matrice de Jordan* est une matrice diagonale par blocs dont chaque bloc diagonal est un bloc de Jordan :

$$J = \operatorname{diag}(J_{k_1}(\lambda_1),\, J_{k_2}(\lambda_2),\, \dots,\, J_{k_s}(\lambda_s)) = \begin{pmatrix} J_{k_1}(\lambda_1) & & & 0 \\ & J_{k_2}(\lambda_2) & & \\ & & \ddots & \\ 0 & & & J_{k_s}(\lambda_s) \end{pmatrix}.$$

Les λ_i ne sont pas nécessairement distincts, et $k_1 + k_2 + \cdots + k_s = n$.

Proposition 15 (Propriétés d'un bloc de Jordan)

Soit $J_k(\lambda)$ un bloc de Jordan. Alors :

(i) $\chi_{J_k(\lambda)}(X) = (X - \lambda)^k$.

(ii) $\mu_{J_k(\lambda)}(X) = (X - \lambda)^k$.

(iii) $\dim E_\lambda(J_k(\lambda)) = 1$.

(iv) $J_k(\lambda)$ est diagonalisable si et seulement si $k = 1$.

Démonstration. (i) $J_k(\lambda) - XI_k$ est triangulaire supérieure avec $\lambda - X$ sur la diagonale, donc $\chi_{J_k(\lambda)}(X) = (\lambda - X)^k = (X - \lambda)^k$ (à signe près, selon la convention ; nous utilisons $\chi(X) = \det(XI - A)$).

(ii) On a $(J_k(\lambda) - \lambda I_k) = N_k$ et $N_k^k = 0$, $N_k^{k-1} \neq 0$. Donc $\mu_{J_k(\lambda)}(X) = (X - \lambda)^k$.

(iii) $\operatorname{Ker}(J_k(\lambda) - \lambda I_k) = \operatorname{Ker}(N_k)$ est engendré par le premier vecteur de la base canonique, donc de dimension 1.

(iv) Si $k \geq 2$, la multiplicité algébrique de λ est k mais la multiplicité géométrique est 1, donc $J_k(\lambda)$ n'est pas diagonalisable. Si $k = 1$, $J_1(\lambda) = (\lambda)$ est déjà diagonale. □

9.5 Théorème de la forme normale de Jordan

Théorème 16 (Forme normale de Jordan)

Soit E un $\mathbb{C}$-espace vectoriel de dimension $n \geq 1$ et $f \in \operatorname{End}(E)$. Alors il existe une base $\mathcal{B}$ de E dans laquelle la matrice de f est une matrice de Jordan :

$$\mathcal{M}_{\mathcal{B}}(f) = \operatorname{diag}(J_{k_1}(\lambda_1),\, \dots,\, J_{k_s}(\lambda_s)),$$

où $\lambda_1, \dots, \lambda_s \in \mathbb{C}$ (non nécessairement distincts) et $k_1 + \cdots + k_s = n$.
De plus, cette forme de Jordan est unique à l'ordre des blocs près : les tailles des blocs et les valeurs propres associées sont uniquement déterminées par f.

De manière équivalente, pour toute matrice $A \in \mathcal{M}_n(\mathbb{C})$, il existe une matrice inversible $P \in \mathrm{GL}_n(\mathbb{C})$ telle que $P^{-1}AP = J$ est une matrice de Jordan.

Preuve du théorème 16. La preuve se déroule en plusieurs étapes.

Étape 1 : Décomposition en sous-espaces caractéristiques.

Par la proposition 11, on dispose de la décomposition

$$E = F_{\lambda_1}(f) \oplus \cdots \oplus F_{\lambda_r}(f),$$

où $\lambda_1, \dots, \lambda_r$ sont les valeurs propres distinctes de f, et $\dim F_{\lambda_i}(f) = m_i$ (multiplicité algébrique). Chaque $F_{\lambda_i}(f)$ est f-invariant, et la restriction $g_i = (f - \lambda_i \operatorname{Id})|_{F_{\lambda_i}(f)}$ est nilpotente.

Il suffit donc de montrer qu'un endomorphisme nilpotent d'un espace de dimension finie admet une forme de Jordan (avec $\lambda = 0$). En effet, si dans une base de $F_{\lambda_i}(f)$, la matrice de g_i est $\operatorname{diag}(N_{k_1}, \dots, N_{k_t})$ (blocs nilpotents), alors la matrice de $f|_{F_{\lambda_i}(f)}$ dans cette même base est $\operatorname{diag}(J_{k_1}(\lambda_i), \dots, J_{k_t}(\lambda_i))$.

Étape 2 : Forme de Jordan d'un endomorphisme nilpotent.

Soit $g\colon V \to V$ un endomorphisme nilpotent d'un espace vectoriel V de dimension m, d'indice de nilpotence ν. On procède par récurrence forte sur $m = \dim V$.

Cas de base : si $m = 1$, alors $g = 0$ et la matrice de g est $(0) = J_1(0)$.

Étape de récurrence : supposons le résultat vrai pour tout espace de dimension $< m$. Puisque g est nilpotent et $m \geq 2$, g n'est pas inversible, donc $\operatorname{Ker}(g) \neq \{\mathbf{0}\}$ et $W = \operatorname{Im}(g) \subsetneq V$ est un sous-espace propre de V (de dimension $< m$), stable par g.

Par hypothèse de récurrence, il existe une base de W dans laquelle $g|_W$ est de la forme

$$\operatorname{diag}(N_{p_1}, \dots, N_{p_t})$$

avec $p_1 \geq p_2 \geq \cdots \geq p_t \geq 1$.

Pour chaque bloc N_{p_j}, notons $(w_1^{(j)}, \dots, w_{p_j}^{(j)})$ les vecteurs de base correspondants dans W, de sorte que

$$g(w_i^{(j)}) = w_{i+1}^{(j)} \quad \text{pour } i < p_j, \qquad g(w_{p_j}^{(j)}) = \mathbf{0}.$$

Le vecteur $w_1^{(j)} \in \operatorname{Im}(g)$, donc il existe $v_0^{(j)} \in V$ tel que $g(v_0^{(j)}) = w_1^{(j)}$. Alors la famille $(v_0^{(j)}, w_1^{(j)}, \dots, w_{p_j}^{(j)})$ vérifie

$$g(v_0^{(j)}) = w_1^{(j)}, \quad g(w_1^{(j)}) = w_2^{(j)}, \quad \dots, \quad g(w_{p_j}^{(j)}) = \mathbf{0},$$

ce qui donne un bloc de Jordan N_{p_j+1}... mais il faut vérifier que les $v_0^{(j)}$ peuvent être choisis de manière à former avec les vecteurs de base de W et un complément de $\operatorname{Ker}(g) \cap (\operatorname{Im}(g))^{\perp}$ une base de V.

Plus précisément, construisons la base de Jordan directement. Choisissons $\mathbf{v} \in V$ tel que $g^{\nu-1}(\mathbf{v}) \neq \mathbf{0}$ (un tel vecteur existe par définition de ν). Alors les vecteurs

$$\mathbf{v},\ g(\mathbf{v}),\ g^2(\mathbf{v}),\ \dots,\ g^{\nu-1}(\mathbf{v})$$

sont linéairement indépendants (car si $\sum_{i=0}^{\nu-1} \alpha_i g^i(\mathbf{v}) = \mathbf{0}$, en appliquant $g^{\nu-1}$ on obtient $\alpha_0 g^{\nu-1}(\mathbf{v}) = \mathbf{0}$, d'où $\alpha_0 = 0$, puis par récurrence $\alpha_i = 0$ pour tout i).

Posons $V_1 = \operatorname{Vect}(\mathbf{v}, g(\mathbf{v}), \dots, g^{\nu-1}(\mathbf{v}))$. C'est un sous-espace de dimension ν, g-invariant, et dans la base $(g^{\nu-1}(\mathbf{v}), \dots, g(\mathbf{v}), \mathbf{v})$, la matrice de $g|_{V_1}$ est N_ν.

Si $V_1 = V$, on a terminé. Sinon, il faut trouver un supplémentaire g-invariant de V_1 dans V. Cela se fait en choisissant un supplémentaire V_2 de V_1 dans V tel que $g(V_2) \subseteq V_2$ (un tel supplémentaire existe, voir ci-dessous), et on applique l'hypothèse de récurrence à $g|_{V_2}$.

Existence d'un supplémentaire g-invariant. Nous l'établissons par le lemme suivant. □

Lemme 17 (Supplémentaire invariant pour un nilpotent)

Soit $g\colon V \to V$ nilpotent et soit V_1 un sous-espace g-invariant de V ayant une base cyclique $(g^{\nu-1}(\mathbf{v}), \dots, g(\mathbf{v}), \mathbf{v})$ pour un certain $\mathbf{v} \in V$ avec $g^{\nu}(\mathbf{v}) = \mathbf{0}$ et $g^{\nu-1}(\mathbf{v}) \neq \mathbf{0}$. Alors il existe

un sous-espace V_2 de V, g-invariant, tel que $V = V_1 \oplus V_2$.

Démonstration. Nous procédons par récurrence sur $\dim V$. Si $\dim V = \nu$, alors $V_1 = V$ et $V_2 = \{\mathbf{0}\}$ convient. Supposons $\dim V > \nu$.

Comme $g^{\nu-1}(\mathbf{v}) \in \mathrm{Ker}(g) \setminus \{\mathbf{0}\}$, choisissons un sous-espace W de $\mathrm{Ker}(g)$ tel que $\mathrm{Ker}(g) = \mathrm{Vect}(g^{\nu-1}(\mathbf{v})) \oplus W$.

Cas $\nu = 1$. Alors $g(\mathbf{v}) = \mathbf{0}$, $V_1 = \mathrm{Vect}(\mathbf{v})$ et $V_1 \subseteq \mathrm{Ker}(g)$. Choisissons un supplémentaire W' de V_1 dans $\mathrm{Ker}(g)$, et un supplémentaire V_2 de $\mathrm{Ker}(g)$ dans V. Alors $V = V_1 \oplus (W' \oplus V_2)$. On peut vérifier que $g(W' \oplus V_2) \subseteq \mathrm{Ker}(g)$... La construction nécessite plus de soin. On peut étendre W en un supplémentaire de V_1 et vérifier la g-invariance par récurrence.

Cas $\nu \geq 2$. Posons $V' = g^{-1}(W + V_1)$. Considérons $\bar{V} = V/V_1$ et l'endomorphisme induit $\bar{g}\colon \bar{V} \to \bar{V}$ (bien défini car V_1 est g-invariant). Alors $\bar{g}$ est nilpotent sur $\bar{V}$ et $\dim \bar{V} < \dim V$. Par un argument de relèvement, on construit le supplémentaire voulu.

La construction complète peut se faire par récurrence sur ν : on pose $\mathbf{v}' = g(\mathbf{v})$, $V_1' = \mathrm{Vect}(g^{\nu-2}(\mathbf{v}'), \dots, g(\mathbf{v}'), \mathbf{v}')$, et $\bar{V} = \mathrm{Im}(g)$. Alors $g|_{\bar{V}}$ est nilpotent et $V_1' \subseteq \bar{V}$ est un sous-espace cyclique de dimension $\nu - 1$. Par hypothèse de récurrence, il existe $V_2' \subseteq \bar{V}$ tel que $\bar{V} = V_1' \oplus V_2'$, avec V_2' g-stable.

Pour chaque vecteur de base w_j de V_2', choisissons un antécédent $u_j \in V$ tel que $g(u_j) = w_j$. Les vecteurs u_j, complétés par les chaînes cycliques qu'ils engendrent et les vecteurs de $\mathrm{Ker}(g) \cap V_2'$, forment le supplémentaire V_2 cherché (on vérifie que V_2 est g-stable et que $V = V_1 \oplus V_2$ par un argument de dimension). □

Fin de la preuve du théorème 16 — Unicité. **Étape 3 : Unicité.**

Soit λ une valeur propre de f, et notons $k_1 \geq k_2 \geq \cdots \geq k_t$ les tailles des blocs de Jordan associés à λ. Posons $g = (f - \lambda \mathrm{Id})|_{F_\lambda(f)}$, qui est nilpotent. Pour tout $j \geq 0$, calculons :

$$r_j := \mathrm{rg}(g^j) = \dim F_\lambda(f) - \dim \mathrm{Ker}(g^j).$$

Or, pour un bloc N_k de taille k, on a $\dim \mathrm{Ker}(N_k^j) = \min(j, k)$. Donc

$$\dim \mathrm{Ker}(g^j) = \sum_{i=1}^{t} \min(j, k_i).$$

Les quantités $\dim \mathrm{Ker}(g^j)$ sont des invariants de f (elles ne dépendent pas de la base choisie). À partir de la suite $(\dim \mathrm{Ker}(g^j))_{j \geq 0}$, on retrouve les tailles des blocs : le nombre de blocs de taille $\geq j$ est $\dim \mathrm{Ker}(g^j) - \dim \mathrm{Ker}(g^{j-1})$, et le nombre de blocs de taille exactement j est

$$(\dim \mathrm{Ker}(g^j) - \dim \mathrm{Ker}(g^{j-1})) - (\dim \mathrm{Ker}(g^{j+1}) - \dim \mathrm{Ker}(g^j)) = 2\dim \mathrm{Ker}(g^j) - \dim \mathrm{Ker}(g^{j-1}) - \dim \mathrm{Ker}(g^{j+1}).$$

Les tailles de blocs sont donc uniquement déterminées. □

9.6 Calcul de la forme de Jordan

Pour calculer la forme de Jordan d'une matrice $A \in \mathcal{M}_n(\mathbb{C})$, on suit l'algorithme suivant.

1. **Trouver les valeurs propres :** calculer le polynôme caractéristique $\chi_A(X) = \det(XI_n - A)$ et le factoriser sur $\mathbb{C}$:
$$\chi_A(X) = \prod_{i=1}^{r} (X - \lambda_i)^{m_i}.$$
La multiplicité algébrique de λ_i est m_i.

2. **Pour chaque valeur propre λ_i, déterminer les tailles des blocs :** calculer successivement :
$$d_j^{(i)} = \dim \mathrm{Ker}((A - \lambda_i I_n)^j), \qquad j = 1, 2, 3, ...$$
jusqu'à atteindre $d_j^{(i)} = m_i$. On pose $d_0^{(i)} = 0$.

3. **Déterminer le nombre de blocs de chaque taille :** le nombre de blocs de Jordan de taille j associés à λ_i est
$$b_j^{(i)} = d_j^{(i)} - 2d_{j-1}^{(i)} + d_{j-2}^{(i)},$$
où $d_{-1}^{(i)} = 0$ par convention. De manière équivalente, le nombre total de blocs associés à λ_i est $d_1^{(i)}$ (la multiplicité géométrique), et le nombre de blocs de taille $\geq j$ est $d_j^{(i)} - d_{j-1}^{(i)}$.

4. **Construire la matrice de Jordan :** assembler les blocs $J_k(\lambda_i)$ selon les tailles déterminées.

Remarque 18 (Vérifications)

On vérifie toujours que :

- la somme des tailles de tous les blocs vaut n,
- pour chaque λ_i, la somme des tailles des blocs associés vaut m_i,
- le nombre de blocs de Jordan associés à λ_i égale la multiplicité géométrique $\dim E_{\lambda_i}(A)$.

9.7 Exemples détaillés

Exemple 19 (Matrice 3×3)

Soit $A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$.

Étape 1 : A est triangulaire, donc ses valeurs propres sont les éléments diagonaux : $\lambda = 2$ de multiplicité algébrique $m = 3$.

Étape 2 : Posons $B = A - 2I_3 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$.

$$d_1 = \dim \mathrm{Ker}(B) = 1,$$
$$d_2 = \dim \mathrm{Ker}(B^2) = 2, \quad \text{car } B^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$
$$d_3 = \dim \mathrm{Ker}(B^3) = 3, \quad \text{car } B^3 = 0.$$

Étape 3 : Nombre de blocs de taille j :

- $j = 1$: $b_1 = d_1 - 2d_0 + d_{-1} = 1 - 0 + 0 = 1$... Utilisons plutôt la formule des différences successives.

Nombre de blocs de taille $\geq j$:

$$\Delta_1 = d_1 - d_0 = 1 - 0 = 1,$$
$$\Delta_2 = d_2 - d_1 = 2 - 1 = 1,$$
$$\Delta_3 = d_3 - d_2 = 3 - 2 = 1.$$

Nombre de blocs de taille exactement j : $\Delta_j - \Delta_{j+1}$:

- taille 1 : $\Delta_1 - \Delta_2 = 1 - 1 = 0$,
- taille 2 : $\Delta_2 - \Delta_3 = 1 - 1 = 0$,
- taille 3 : $\Delta_3 - 0 = 1$.

Il y a un seul bloc de taille 3.
Conclusion : La forme de Jordan de A est

$$J = J_3(2) = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix} = A.$$

La matrice A est déjà sous forme de Jordan.

Exemple 20 (Matrice 4×4)

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

Étape 1 : $\chi_A(X) = (X-3)^4$, donc $\lambda = 3$ de multiplicité 4.

Étape 2 : $B = A - 3I_4 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$.

$$d_1 = \dim \mathrm{Ker}(B) = 2 \quad \text{(les colonnes 1 et 3 sont nulles)},$$
$$d_2 = \dim \mathrm{Ker}(B^2) = 4 \quad \text{(car } B^2 = 0\text{)}.$$

Étape 3 : $\Delta_1 = d_1 - d_0 = 2$, $\Delta_2 = d_2 - d_1 = 2$. Blocs de taille exactement 1 : $\Delta_1 - \Delta_2 = 0$.
Blocs de taille exactement 2 : $\Delta_2 = 2$.
Conclusion : La forme de Jordan est

$$J = \mathrm{diag}(J_2(3), J_2(3)) = \begin{pmatrix} 3 & 1 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 3 \end{pmatrix} = A.$$

Exemple 21 (Matrice 4×4 avec deux valeurs propres)

Soit $A = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix}$.
Le polynôme caractéristique est $\chi_A(X) = (X-1)^2(X-2)^2$.
Pour $\lambda_1 = 1$ *:* $B_1 = A - I_4$. On trouve $d_1 = \dim \mathrm{Ker}(B_1) = 1$, $d_2 = \dim \mathrm{Ker}(B_1^2) = 2$. Donc un seul bloc $J_2(1)$.
Pour $\lambda_2 = 2$ *:* $B_2 = A - 2I_4$. On trouve $d_1 = \dim \mathrm{Ker}(B_2) = 1$, $d_2 = \dim \mathrm{Ker}(B_2^2) = 2$. Donc un seul bloc $J_2(2)$.
La forme de Jordan est :

$$J = \mathrm{diag}(J_2(1), J_2(2)) = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix}.$$

Exemple 22 (Matrice non triviale)

Soit $A = \begin{pmatrix} 5 & 4 & 2 & 1 \\ 0 & 1 & -1 & -1 \\ -1 & -1 & 3 & 0 \\ 1 & 1 & -1 & 2 \end{pmatrix}$.
On calcule $\chi_A(X) = (X-1)(X-2)(X-4)^2$. Les valeurs propres sont $\lambda_1 = 1$, $\lambda_2 = 2$, $\lambda_3 = 4$.
Pour $\lambda_1 = 1$ et $\lambda_2 = 2$, la multiplicité algébrique est 1, donc le bloc associé est $J_1(1) = (1)$ et $J_1(2) = (2)$.
Pour $\lambda_3 = 4$ de multiplicité 2 : on calcule $d_1 = \dim \mathrm{Ker}(A - 4I_4)$. Si $d_1 = 2$, il y a deux blocs $J_1(4)$ (diagonalisable pour cette valeur propre). Si $d_1 = 1$, il y a un bloc $J_2(4)$.
On calcule $A - 4I_4 = \begin{pmatrix} 1 & 4 & 2 & 1 \\ 0 & -3 & -1 & -1 \\ -1 & -1 & -1 & 0 \\ 1 & 1 & -1 & -2 \end{pmatrix}$. Par réduction, on trouve $\mathrm{rg}(A - 4I_4) = 3$, donc $d_1 = 1$ et la forme de Jordan est

$$J = \mathrm{diag}(J_1(1),\, J_1(2),\, J_2(4)) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 4 & 1 \\ 0 & 0 & 0 & 4 \end{pmatrix}.$$

9.8 Forme de Jordan et polynôme minimal

Théorème 23 (Polynôme minimal et blocs de Jordan)

Soit $f \in \mathrm{End}(E)$ de forme de Jordan $J = \mathrm{diag}(J_{k_1}(\lambda_1), \dots, J_{k_s}(\lambda_s))$, et soient $\mu_1, \dots, \mu_r$ les valeurs propres distinctes de f. Pour chaque μ_i, notons ν_i la taille du plus grand bloc

de Jordan associé à μ_i. Alors le polynôme minimal de f est

$$\mu_f(X) = \prod_{i=1}^{r} (X - \mu_i)^{\nu_i}.$$

Démonstration. Posons $Q(X) = \prod_{i=1}^{r} (X - \mu_i)^{\nu_i}$. Montrons que $Q(f) = 0$ et que μ_f divise Q et Q divise μ_f.

Pour chaque bloc $J_k(\mu_i)$, on a $(J_k(\mu_i) - \mu_i I_k)^{\nu_i} = N_k^{\nu_i} = 0$ car $k \leq \nu_i$. De plus, les facteurs $(X - \mu_j)^{\nu_j}$ pour $j \neq i$ donnent des matrices inversibles sur le bloc $J_k(\mu_i)$. Donc $Q(J_k(\mu_i)) = 0$ pour chaque bloc, et $Q(J) = 0$, d'où $Q(f) = 0$. Ainsi $\mu_f \mid Q$.

Inversement, $\mu_f(f) = 0$ implique $\mu_f(J_k(\mu_i)) = 0$ pour chaque bloc. En écrivant $\mu_f(X) = \prod_i (X - \mu_i)^{a_i}$, la condition $\mu_f(J_k(\mu_i)) = 0$ impose $N_k^{a_i} \cdot C = 0$ où C est inversible (produit des $(J_k(\mu_i) - \mu_j I_k)^{a_j}$ pour $j \neq i$). Donc $N_k^{a_i} = 0$, ce qui requiert $a_i \geq k$. Prenant le plus grand bloc ($k = \nu_i$), on obtient $a_i \geq \nu_i$, d'où $Q \mid \mu_f$.

Finalement $\mu_f = Q$. □

Corollaire 24 (Diagonalisabilité et blocs)

f est diagonalisable si et seulement si tous les blocs de Jordan sont de taille 1, c'est-à-dire si et seulement si le polynôme minimal est à racines simples.

Démonstration. f est diagonalisable si et seulement si μ_f est scindé à racines simples, c'est-à-dire $\nu_i = 1$ pour tout i, ce qui signifie que tous les blocs de Jordan sont de taille 1. □

9.9 Cayley-Hamilton via la forme de Jordan

Théorème 25 (Théorème de Cayley-Hamilton (nouvelle preuve))

Soit $A \in \mathcal{M}_n(\mathbb{C})$. Alors $\chi_A(A) = 0$.

Démonstration. Par le théorème 16, il existe $P \in \mathrm{GL}_n(\mathbb{C})$ telle que $P^{-1}AP = J = \mathrm{diag}(J_{k_1}(\lambda_1), \ldots, J_{k_s}(\lambda_s))$.

Comme $\chi_A = \chi_J$, on a :

$$\chi_A(A) = P\chi_J(J)P^{-1}.$$

Il suffit de montrer $\chi_J(J) = 0$. Puisque $\chi_J(X) = \prod_{j=1}^{s} (X - \lambda_j)^{k_j}$ et que J est diagonale par blocs :

$$\chi_J(J) = \mathrm{diag}(\chi_J(J_{k_1}(\lambda_1)), \ldots, \chi_J(J_{k_s}(\lambda_s))).$$

Pour chaque bloc $J_{k_j}(\lambda_j)$:

$$\chi_J(J_{k_j}(\lambda_j)) = \prod_{i=1}^{s} (J_{k_j}(\lambda_j) - \lambda_i I_{k_j})^{k_i}.$$

Le facteur correspondant à $i = j$ donne $(J_{k_j}(\lambda_j) - \lambda_j I_{k_j})^{k_j} = N_{k_j}^{k_j} = 0$. Donc l'ensemble du produit est nul, et $\chi_J(J_{k_j}(\lambda_j)) = 0$.

Ceci étant vrai pour chaque bloc, $\chi_J(J) = 0$ et $\chi_A(A) = 0$. □

Remarque 26 (Avantage de cette preuve)

Cette preuve est conceptuellement très claire : elle réduit le théorème de Cayley-Hamilton au cas des blocs de Jordan, pour lesquels la vérification est immédiate. Elle requiert toutefois le théorème de Jordan, qui est plus profond que Cayley-Hamilton.

9.10 Exponentielle de matrice

Définition 27 (Exponentielle de matrice)

Pour toute matrice $A \in \mathcal{M}_n(\mathbb{C})$, l'*exponentielle de matrice* est définie par la série :

$$e^A := \exp(A) := \sum_{k=0}^{\infty} \frac{A^k}{k!} = I_n + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \cdots$$

Cette série converge absolument pour toute matrice A.

Proposition 28 (Propriétés de l'exponentielle)

(i) $e^0 = I_n$.

(ii) Si $AB = BA$, alors $e^{A+B} = e^A e^B$.

(iii) e^A est toujours inversible, et $(e^A)^{-1} = e^{-A}$.

(iv) Si $A = P^{-1}BP$, alors $e^A = P^{-1}e^B P$.

(v) $\det(e^A) = e^{\mathrm{tr}(A)}$.

Démonstration. (i) Immédiat : $e^0 = \sum_{k\geq 0} \frac{0^k}{k!} = I_n$.

(ii) Lorsque $AB = BA$, le binôme de Newton s'applique : $(A+B)^k = \sum_{j=0}^{k} \binom{k}{j} A^j B^{k-j}$. On vérifie alors que le produit de Cauchy des séries de e^A et e^B donne e^{A+B}.

(iii) Comme A et $-A$ commutent, $e^A e^{-A} = e^{A-A} = e^0 = I_n$.

(iv) $e^A = \sum_k \frac{A^k}{k!} = \sum_k \frac{(P^{-1}BP)^k}{k!} = \sum_k \frac{P^{-1}B^kP}{k!} = P^{-1}(\sum_k \frac{B^k}{k!})P = P^{-1}e^B P$.

(v) Par la propriété (iv), il suffit de montrer le résultat pour les matrices triangulaires (ou de Jordan). Pour $J = \mathrm{diag}(J_{k_1}(\lambda_1), \ldots)$, $\det(e^J) = \prod_j e^{\lambda_j k_j}$ (à préciser via les blocs), et $\mathrm{tr}(J) = \sum_j \lambda_j k_j$, d'où le résultat. Plus rigoureusement, $e^{J_k(\lambda)}$ est triangulaire supérieure avec e^λ sur la diagonale, donc $\det(e^{J_k(\lambda)}) = e^{k\lambda}$ et $\mathrm{tr}(J_k(\lambda)) = k\lambda$. □

Proposition 29 (Exponentielle d'un bloc de Jordan)

L'exponentielle du bloc $tJ_k(\lambda)$ (où $t \in \mathbb{C}$) est :

$$e^{tJ_k(\lambda)} = e^{\lambda t} \begin{pmatrix} 1 & t & \frac{t^2}{2!} & \cdots & \frac{t^{k-1}}{(k-1)!} \\ 0 & 1 & t & \cdots & \frac{t^{k-2}}{(k-2)!} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & t \\ 0 & \cdots & \cdots & 0 & 1 \end{pmatrix}.$$

Démonstration. On écrit $tJ_k(\lambda) = \lambda t I_k + tN_k$, où $\lambda t I_k$ et tN_k commutent. Donc :

$$e^{tJ_k(\lambda)} = e^{\lambda t I_k} \cdot e^{tN_k} = e^{\lambda t} I_k \cdot e^{tN_k}.$$

Or N_k est nilpotent d'indice k, donc la série exponentielle est finie :

$$e^{tN_k} = \sum_{j=0}^{k-1} \frac{t^j}{j!} N_k^j.$$

La matrice N_k^j a des 1 sur la j-ième sur-diagonale, ce qui donne le résultat annoncé. □

Exemple 30 (Exponentielle d'une matrice 2×2)

Soit $A = \begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix} = J_2(3)$. Alors :

$$e^{tA} = e^{3t} \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} e^{3t} & te^{3t} \\ 0 & e^{3t} \end{pmatrix}.$$

Exemple 31 (Exponentielle d'une matrice 3×3)

Soit $A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix} = J_3(2)$. Alors :

$$e^{tA} = e^{2t} \begin{pmatrix} 1 & t & \frac{t^2}{2} \\ 0 & 1 & t \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} e^{2t} & te^{2t} & \frac{t^2}{2}e^{2t} \\ 0 & e^{2t} & te^{2t} \\ 0 & 0 & e^{2t} \end{pmatrix}.$$

9.11 Applications

9.11.1 Systèmes d'équations différentielles linéaires

Considérons le système différentiel linéaire à coefficients constants

$$X'(t) = AX(t), \qquad X(0) = X_0, \tag{9.1}$$

où $A \in \mathcal{M}_n(\mathbb{C})$ et $X(t) \in \mathbb{C}^n$.

Théorème 32 (Solution du système différentiel)

L'unique solution du problème de Cauchy (9.1) est

$$X(t) = e^{tA} X_0.$$

Démonstration. Existence : posons $X(t) = e^{tA}X_0$. Alors $X'(t) = Ae^{tA}X_0 = AX(t)$ (car $\frac{d}{dt}e^{tA} = Ae^{tA}$, ce qui se vérifie terme à terme sur la série) et $X(0) = e^0 X_0 = X_0$.

Unicité : si $Y(t)$ est une autre solution, posons $Z(t) = e^{-tA}Y(t)$. Alors $Z'(t) = -Ae^{-tA}Y(t) + e^{-tA}Y'(t) = -Ae^{-tA}Y(t) + e^{-tA}AY(t) = 0$. Donc Z est constante : $Z(t) = Z(0) = X_0$, d'où $Y(t) = e^{tA}X_0$. □

Exemple 33 (Système 2×2 non diagonalisable)

Résoudre $X' = AX$ avec $A = \begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix}$ et $X(0) = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$.

On a $e^{tA} = \begin{pmatrix} e^{3t} & te^{3t} \\ 0 & e^{3t} \end{pmatrix}$, donc :

$$X(t) = \begin{pmatrix} e^{3t} & te^{3t} \\ 0 & e^{3t} \end{pmatrix} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \begin{pmatrix} (x_0 + y_0 t)e^{3t} \\ y_0 e^{3t} \end{pmatrix}.$$

On retrouve le phénomène caractéristique : la présence du terme te^{3t} (croissance polynomiale multipliée par l'exponentielle) est la signature d'un bloc de Jordan non trivial.

Exemple 34 (Système 3×3)

Résoudre $X' = AX$ avec $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$.

La matrice A est déjà sous forme de Jordan : $A = \operatorname{diag}(J_1(1), J_2(2))$. Donc :

$$e^{tA} = \operatorname{diag}(e^t,\, e^{2t} \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}) = \begin{pmatrix} e^t & 0 & 0 \\ 0 & e^{2t} & te^{2t} \\ 0 & 0 & e^{2t} \end{pmatrix}.$$

La solution générale est :

$$X(t) = \begin{pmatrix} c_1 e^t \\ c_2 e^{2t} + c_3 t e^{2t} \\ c_3 e^{2t} \end{pmatrix},$$

où c_1, c_2, c_3 sont déterminés par la condition initiale.

9.11.2 Fonctions de matrices

La forme de Jordan permet de définir et calculer des fonctions de matrices.

Définition 35 (Fonction d'une matrice)

Soit f une fonction analytique au voisinage de chaque valeur propre de A. Si $A = PJP^{-1}$ avec $J = \operatorname{diag}(J_{k_1}(\lambda_1), \dots, J_{k_s}(\lambda_s))$, on définit :

$$f(A) := P \operatorname{diag}(f(J_{k_1}(\lambda_1)), \dots, f(J_{k_s}(\lambda_s)))\, P^{-1},$$

où, pour un bloc de Jordan :

$$f(J_k(\lambda)) = \begin{pmatrix} f(\lambda) & f'(\lambda) & \frac{f''(\lambda)}{2!} & \cdots & \frac{f^{(k-1)}(\lambda)}{(k-1)!} \\ 0 & f(\lambda) & f'(\lambda) & \cdots & \frac{f^{(k-2)}(\lambda)}{(k-2)!} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & f(\lambda) & f'(\lambda) \\ 0 & \cdots & \cdots & 0 & f(\lambda) \end{pmatrix}.$$

Exemple 36 (Racine carrée d'une matrice)

Soit $A = \begin{pmatrix} 4 & 1 \\ 0 & 4 \end{pmatrix} = J_2(4)$. Avec $f(x) = \sqrt{x}$, $f'(x) = \frac{1}{2\sqrt{x}}$:

$$\sqrt{A} = \begin{pmatrix} \sqrt{4} & \frac{1}{2\sqrt{4}} \\ 0 & \sqrt{4} \end{pmatrix} = \begin{pmatrix} 2 & \frac{1}{4} \\ 0 & 2 \end{pmatrix}.$$

Vérifions : $\left(\sqrt{A}\right)^2 = \begin{pmatrix} 4 & 1 \\ 0 & 4 \end{pmatrix} = A$.

9.11.3 Puissances d'une matrice

Proposition 37 (Puissances d'un bloc de Jordan)

Pour $m \geq 0$:

$$\left(J_k(\lambda)\right)^m = \begin{pmatrix} \lambda^m & \binom{m}{1}\lambda^{m-1} & \binom{m}{2}\lambda^{m-2} & \cdots & \binom{m}{k-1}\lambda^{m-k+1} \\ 0 & \lambda^m & \binom{m}{1}\lambda^{m-1} & \cdots & \binom{m}{k-2}\lambda^{m-k+2} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \lambda^m & \binom{m}{1}\lambda^{m-1} \\ 0 & \cdots & \cdots & 0 & \lambda^m \end{pmatrix},$$

où $\binom{m}{j} = \frac{m!}{j!(m-j)!}$ pour $j \leq m$ et $\binom{m}{j} = 0$ pour $j > m$.

Démonstration. On écrit $J_k(\lambda) = \lambda I_k + N_k$. Puisque λI_k et N_k commutent, le binôme de Newton donne :

$$J_k(\lambda)^m = \sum_{j=0}^{k-1} \binom{m}{j} \lambda^{m-j} N_k^j.$$

Le coefficient en position $(i, i+j)$ de N_k^j est 1 (pour $1 \leq i \leq k-j$), ce qui donne la formule annoncée. □

9.12 Exercices

Exercice 38 (Endomorphismes nilpotents : propriétés de base)

Soit $f \in \mathrm{End}(E)$ nilpotent.

(a) Montrer que $\mathrm{Id} + f$ est inversible. *Indication : utiliser la série géométrique finie* $(\mathrm{Id}+f)^{-1} = \sum_{k=0}^{\nu-1} (-1)^k f^k$.

(b) Montrer que si g commute avec f et est nilpotent, alors $f + g$ est nilpotent.

(c) Montrer que $\mathrm{tr}(f^k) = 0$ pour tout $k \geq 1$.

Exercice 39 (Classification des nilpotents en dimension 3)

Déterminer toutes les formes de Jordan possibles pour un endomorphisme nilpotent d'un espace de dimension 3. Pour chaque cas, préciser l'indice de nilpotence, le polynôme minimal et la dimension du noyau.

Exercice 40 (Calcul de forme de Jordan)

Déterminer la forme de Jordan de chacune des matrices suivantes :

(a) $A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$.

(b) $B = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$.

(c) $C = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}$.

Exercice 41 (Forme de Jordan d'une matrice 4×4)

Soit $A = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}$.

(a) Calculer le polynôme caractéristique de A.

(b) Calculer $\dim \mathrm{Ker}(A - 2I_4)$, $\dim \mathrm{Ker}(A - 2I_4)^2$, $\dim \mathrm{Ker}(A - 2I_4)^3$.

(c) En déduire la forme de Jordan de A.

(d) Déterminer le polynôme minimal de A.

Exercice 42 (Exponentielle de matrice)

Calculer e^{tA} pour :

(a) $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$.

(b) $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$.

(c) $A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$.

(d) $A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$.

Exercice 43 (Système différentiel)

Résoudre le système différentiel $X' = AX$, $X(0) = X_0$, où

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}, \qquad X_0 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}.$$

Exercice 44 (Polynôme minimal et forme de Jordan)

Pour chacun des cas suivants, donner toutes les formes de Jordan possibles (à l'ordre des blocs près) pour une matrice $A \in \mathcal{M}_5(\mathbb{C})$:

(a) $\chi_A(X) = (X-1)^3(X-2)^2$ et $\mu_A(X) = (X-1)^2(X-2)$.

(b) $\chi_A(X) = (X-3)^5$ et $\mu_A(X) = (X-3)^3$.

(c) $\chi_A(X) = (X-1)^5$ et $\mu_A(X) = (X-1)^2$.

Exercice 45 (Matrice semblable à son bloc de Jordan)

Soit $A = \begin{pmatrix} 5 & -4 \\ 1 & 1 \end{pmatrix}$.

(a) Montrer que A n'est pas diagonalisable sur $\mathbb{C}$.

(b) Déterminer sa forme de Jordan J.

(c) Trouver une matrice P telle que $A = PJP^{-1}$.

(d) Calculer A^m pour tout $m \geq 0$.

(e) Calculer e^{tA}.

Exercice 46 (Commutant d'un bloc de Jordan)

Soit $J = J_n(\lambda)$ un bloc de Jordan de taille n.

(a) Montrer qu'une matrice $B \in \mathcal{M}_n(\mathbb{C})$ commute avec J si et seulement si B est un polynôme (de degré $\leq n-1$) en la matrice nilpotente $N_n = J - \lambda I_n$.

(b) En déduire que le commutant de J est de dimension n.

(c) Montrer que deux blocs de Jordan $J_n(\lambda)$ et $J_m(\mu)$ (avec $(n,\lambda) \neq (m,\mu)$) ne sont pas semblables.

Exercice 47 (Logarithme de matrice)

Soit $A \in \mathcal{M}_n(\mathbb{C})$ dont toutes les valeurs propres ont une partie réelle strictement positive.

(a) Montrer qu'il existe une matrice $B \in \mathcal{M}_n(\mathbb{C})$ telle que $e^B = A$. *Indication : utiliser la forme de Jordan et le logarithme complexe.*

(b) Calculer explicitement un tel B pour $A = \begin{pmatrix} e & 1 \\ 0 & e \end{pmatrix}$.

Exercice 48 (Décomposition de Dunford)

Soit $f \in \mathrm{End}(E)$ avec E un $\mathbb{C}$-espace vectoriel de dimension finie.

(a) Montrer qu'il existe un unique couple $(d, n) \in \mathrm{End}(E)^2$ tel que :

- $f = d + n$,
- d est diagonalisable,
- n est nilpotent,
- $d \circ n = n \circ d$.

Indication : utiliser la forme de Jordan pour construire d et n.

(b) Montrer que d et n sont des polynômes en f.

(c) Calculer la décomposition de Dunford de $A = \begin{pmatrix} 3 & 1 & 0 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{pmatrix}$.

Exercice 49 (Résolution d'une récurrence linéaire)

Considérons la récurrence linéaire

$$u_{n+2} - 4u_{n+1} + 4u_n = 0, \qquad u_0 = 1, \quad u_1 = 3.$$

(a) Écrire cette récurrence sous forme matricielle $X_{n+1} = AX_n$ avec $X_n = \begin{pmatrix} u_{n+1} \\ u_n \end{pmatrix}$.

(b) Déterminer la forme de Jordan de A.

(c) Calculer A^n et en déduire u_n.

Résumé du chapitre

- Un endomorphisme f est **nilpotent** si $f^\nu = 0$ pour un certain $\nu \geq 1$. L'indice de nilpotence est le plus petit tel ν, et il vérifie $\nu \leq \dim E$. Le polynôme caractéristique est X^n et le polynôme minimal est X^ν.
- Un **sous-espace caractéristique** (ou espace propre généralisé) est $F_\lambda(f) = \mathrm{Ker}((f - \lambda\,\mathrm{Id})^n)$. Sa dimension égale la multiplicité algébrique de λ.
- L'espace E se décompose en somme directe de sous-espaces caractéristiques : $E = \bigoplus_{\lambda \in \mathrm{Spec}(f)} F_\lambda(f)$.
- Un **bloc de Jordan** $J_k(\lambda)$ est une matrice $k \times k$ avec λ sur la diagonale et des 1 sur la sur-diagonale.
- **Théorème de Jordan :** toute matrice sur $\mathbb{C}$ est semblable à une matrice de Jordan, unique à l'ordre des blocs près.
- Le **polynôme minimal** est $\mu_f(X) = \prod_i (X - \lambda_i)^{\nu_i}$ où ν_i est la taille du plus grand bloc de Jordan associé à λ_i.

- Une matrice est **diagonalisable** si et seulement si tous ses blocs de Jordan sont de taille 1.
- Le **calcul des tailles des blocs** se fait via les dimensions $\dim \mathrm{Ker}((A - \lambda I)^k)$ pour $k = 1, 2, ...$
- L'**exponentielle de matrice** e^{tA} se calcule bloc par bloc via la forme de Jordan, ce qui permet de résoudre le système $X' = AX$.
- La **décomposition de Dunford** $f = d + n$ (diagonalisable + nilpotent, commutants) est une conséquence directe de la forme de Jordan.

Index

$\mathbb{K}^n$, 17
$\oplus$, 21

anneau, 5
    commutatif, 6
application, 2
application linéaire, 37, 38
automorphisme, 44

base, 25
base canonique, 25
base duale, 51
bidual, 51
bijection, 3

caractéristique, 9
changement de base, 47
combinaison de lignes, 59
combinaison linéaire, 22
composition, 3
conservation du flux, 67
coordonnées, 25
coordonnées homogènes, 51
corps, 6
corps fini, 8
critère de base, 28

dimension, 27
    invariance, 26
dimension finie, 27
droite, 66
dual, 50
dérivation, 39
déterminant, 50

endomorphisme, 38
ensemble, 1
ensemble des solutions, 58
espace de fonctions, 17
espace des matrices, 17
espace des polynômes, 17
espace des solutions, 30
espace des suites, 18
espace dual, 50
espace vectoriel, 16

famille génératrice, 23
famille libre, 24
famille liée, 24
forme linéaire, 50
forme échelonnée, 60
forme échelonnée réduite, 60
formule de Grassmann, 28

$\mathrm{GL}_n(\mathbb{K})$, 49
$\mathrm{GL}(E)$, 44
Grassmann, formule de, 28
groupe, 5
    abélien, 5
groupe linéaire, 50

homomorphisme, 38
homothétie, 37

image, 40
indépendance linéaire, 24
infographie, 51
injection, 2
injectivité
    application linéaire, 43
intersection de sous-espaces, 20
inégalité de Sylvester, 54
isomorphe, 43
isomorphisme, 43

Kirchhoff (lois de), 57
Kronecker–Capelli (théorème de), 64

$\mathcal{L}(E, F)$, 38
lemme d'échange, 26
lemme de Steinitz, 26
Leontief (modèle de), 57
loi de composition interne, 4

matrice

augmentée, 58
d'une application linéaire, 45
des coefficients, 58
inversible, 59
matrice stochastique, 52
matrice de passage, 47
matrice de transition, 52
matrice inversible, 49
matrices semblables, 47
moindres carrés, 68
multiplication d'une ligne, 59
multiplication scalaire, 16

nilpotent, 54
nombres complexes, 8
noyau, 40

opérations élémentaires, 59
sur les lignes, 59

pivot, 60
pivot de Gauss, 61
plan, 66
polynôme impair, 31
polynôme pair, 31
produit matriciel, 48
projecteur, 39
projection, 37

rang, 41, 63
colonne, 63
d'une famille de vecteurs, 29
ligne, 63
relation binaire, 3
relation d'ordre, 4
relation d'équivalence, 3
rotation, 37
Rouché–Capelli (théorème de), 64
RREF, 60

réflexion, 37

scalaire, 15
solution
triviale, 59
solution d'un système, 58
somme de sous-espaces, 20
somme directe, 21
sous-espace
trivial, 20
sous-espace affine, 65
sous-espace engendré, 22
sous-espace vectoriel, 18, 59
sous-espaces supplémentaires, 21
substitution arrière, 61
supplémentaire, 21
surjection, 2
surjectivité
application linéaire, 43
système homogène, 18
système linéaire, 57, 68
compatible, 58
définition, 58
homogène, 59
incompatible, 58

théorème du rang, 30, 41
transposée, 49

variable
libre, 65
liée, 65
Vect, 22
vecteur, 16

échange de lignes, 59
élimination de Gauss, 57, 61
élimination de Gauss–Jordan, 63
équations normales, 68