

# Calcul matriciel

Fabien Priziac

Licence 3 de Mathématiques, année universitaire 2020-2021



# Table des matières

<b>1</b>	<b>Dualité linéaire</b>	<b>7</b>
1.1	Introduction . . . . .	7
1.2	Formes linéaires sur un espace vectoriel et espace dual . . . . .	7
1.3	Base duale . . . . .	9
1.4	Aspects matriciels . . . . .	13
1.5	Annulateur d'un sous-espace vectoriel et correspondance duale . . . . .	15
1.6	Application transposée . . . . .	18
1.7	Bidual . . . . .	20
<b>2</b>	<b>Espaces euclidiens</b>	<b>23</b>
2.1	Introduction . . . . .	23
2.2	Produit scalaire sur un espace vectoriel réel . . . . .	23
2.3	Orthogonalité dans les espaces euclidiens . . . . .	27
2.4	Orthogonalité et dualité dans les espaces euclidiens . . . . .	30
2.5	Bases orthogonales et bases orthonormales . . . . .	31
2.6	Représentation matricielle du produit scalaire . . . . .	35
2.7	Endomorphisme adjoint . . . . .	37
2.8	Endomorphismes orthogonaux et matrices orthogonales . . . . .	41
2.9	Décomposition $QR$ d'une matrice inversible . . . . .	45
<b>3</b>	<b>Réduction des endomorphismes</b>	<b>49</b>
3.1	Introduction . . . . .	49
3.2	Valeurs propres et espaces propres . . . . .	49
3.3	Polynôme caractéristique . . . . .	50
3.4	Diagonalisabilité et diagonalisation . . . . .	51
3.5	Polynômes d'endomorphismes, polynômes annulateurs et diagonalisabilité . . . . .	54
3.6	Polynôme minimal . . . . .	55
3.7	Triangularisabilité et triangularisation . . . . .	56
3.8	Réduction de Jordan . . . . .	60
3.8.1	Etape 1 . . . . .	61
3.8.2	Etape 2 . . . . .	62
3.8.3	Description matricielle de la méthode de réduction à la forme de Jordan . . . . .	65

<b>4</b>	<b>Exponentielle de matrices</b>	<b>77</b>
4.1	Introduction . . . . .	77
4.2	Norme de matrices . . . . .	77
4.3	Définition et propriétés de base . . . . .	79
4.4	Calcul via la réduction de Jordan . . . . .	83
4.5	Résolution des systèmes différentiels linéaires . . . . .	86
<b>5</b>	<b>Orthogonalité et réduction</b>	<b>91</b>
5.1	Introduction . . . . .	91
5.2	Diagonalisabilité des endomorphismes auto-adjoints . . . . .	91
5.3	Matrices symétriques positives . . . . .	95
5.4	Décomposition polaire . . . . .	102
5.5	Réduction des endomorphismes et matrices orthogonaux . . . . .	105
<b>6</b>	<b>Normes subordonnées et rayon spectral</b>	<b>111</b>
6.1	Introduction . . . . .	111
6.2	Normes matricielles subordonnées . . . . .	111
6.3	Rayon spectral . . . . .	117
6.4	Conditionnement . . . . .	122
<b>7</b>	<b>Matrices stochastiques et théorème de Perron</b>	<b>127</b>
7.1	Introduction . . . . .	127
7.2	Matrices stochastiques et vecteurs stochastiques . . . . .	130
7.3	Matrices positives, strictement positives, primitives, irréductibles . . . . .	132
7.4	Les théorèmes de Perron-Frobenius . . . . .	133
7.5	Le cas des matrices primitives stochastiques . . . . .	136
<b>8</b>	<b>Résolution de systèmes linéaires</b>	<b>143</b>
8.1	Introduction . . . . .	143
8.2	Méthode du pivot de Gauss . . . . .	147
8.3	La décomposition $LU$ . . . . .	151
8.4	La décomposition $PLU$ . . . . .	157
8.5	La décomposition de Cholesky . . . . .	163

# Index des notations

Ci-dessous,  $\mathbb{K}$  désigne un corps commutatif quelconque.  $E$  et  $F$  désignent deux espaces vectoriels sur  $\mathbb{K}$ .  $f$  désigne une application linéaire de  $E$  dans  $F$  et  $g$  désigne un endomorphisme de  $E$ .  $v, v_1, \dots, v_k$  désignent des vecteurs de  $E$ .  $\mathcal{B}, \mathcal{B}'$  désignent deux bases de  $E$ .  $A$  désigne une matrice à coefficients dans  $\mathbb{K}$ .  $\mathcal{C}$  désigne une base de  $F$ .  $n$  et  $p$  désignent des entiers naturels non nuls.  $i$  et  $j$  désignent deux nombres de l'ensemble  $\{1, \dots, n\}$ .  $M$  désigne une matrice carrée de taille  $n$ .  $P$  désigne un polynôme de  $\mathbb{K}[X]$ .  $S$  désigne un sous-ensemble quelconque de  $E$ .  $R$  désigne un anneau commutatif et  $x_1, \dots, x_k$  des éléments de  $R$ .

- $\mathcal{L}(E, F)$  : espace vectoriel sur  $\mathbb{K}$  des applications linéaires de  $E$  dans  $F$ .
- $\text{Ker } f$  : noyau de  $f$  (il s'agit d'un sous-espace vectoriel de  $E$ ).
- $\text{Im } f$  : image de  $f$  (il s'agit d'un sous-espace vectoriel de  $F$ ).
- $\dim(E)$  : dimension de  $E$  sur  $\mathbb{K}$  si  $E$  est de dimension finie.
- $\text{Vect}\{v_1, \dots, v_k\}$  : sous-espace vectoriel de  $E$  engendré par  $v_1, \dots, v_k$ .
- $\text{Mat}_{\mathcal{B}}(v)$  : vecteur colonne (matrice colonne) des coordonnées du vecteur  $v$  dans la base  $\mathcal{B}$ .
- ${}^tA$  : transposée de la matrice  $A$ .
- $\text{Id}_E$  : application identité de  $E$ .
- $\text{Mat}_{\mathcal{B}, \mathcal{C}}(f)$  : matrice représentative de  $f$  dans les bases  $\mathcal{B}$  de  $E$  et  $\mathcal{C}$  de  $F$ .
- $P_{\mathcal{B} \rightarrow \mathcal{B}'}$  : matrice de passage de la base  $\mathcal{B}$  à la base  $\mathcal{B}'$  (il s'agit de la matrice  $\text{Mat}_{\mathcal{B}', \mathcal{B}}(\text{Id}_E)$ ).
- $M_n(\mathbb{K})$  : espace vectoriel des matrices carrées à  $n$  lignes et  $n$  colonnes et à coefficients dans  $\mathbb{K}$ .
- $E_{i,j}$  : matrice de  $M_n(\mathbb{K})$  avec un coefficient 1 sur la ligne  $i$  et la colonne  $j$ , et des coefficients nuls partout ailleurs.
- $\mathbb{K}_n[X]$  : espace vectoriel sur  $\mathbb{K}$  des polynômes en une indéterminée à coefficients dans  $\mathbb{K}$  et de degré au plus  $n$ .

- $\text{Tr}(M)$  : trace de la matrice  $M$  i.e. la somme des coefficients diagonaux de  $M$ .
- $I_n$  : matrice identité de taille  $n$ .
- $0_{n,p}$  : matrice à  $n$  lignes et  $p$  colonnes avec uniquement des coefficients nuls.
- $\text{Vect}(S)$  : sous-espace vectoriel de  $E$  engendré par les vecteurs de  $E$  contenus dans  $S$ .
- $\mathcal{L}(E)$  : espace vectoriel sur  $\mathbb{K}$  des endomorphismes de  $E$  ( $\mathcal{L}(E) = \mathcal{L}(E, E)$ ).
- $\text{rg}(A)$  : rang de la matrice  $A$ .
- $\text{rg}(f)$  : rang de l'application linéaire  $f$ .
- $\det(M)$  : déterminant de la matrice  $M$ .
- $\det(g)$  : déterminant de l'endomorphisme  $g$ .
- $M_{n,p}(\mathbb{K})$  : espace vectoriel des matrices à  $n$  lignes et  $p$  colonnes et à coefficients dans  $\mathbb{K}$ .
- $\text{Ker } A$  : noyau de la matrice  $A$ .
- $\deg(P)$  : degré du polynôme  $P$ .
- $(x_1, \dots, x_k)$  : idéal de  $A$  engendré par les éléments  $x_1, \dots, x_k$ .

# Chapitre 1

## Dualité linéaire

### 1.1 Introduction

La dualité linéaire est la théorie des formes linéaires sur un espace vectoriel, c'est-à-dire, pour  $\mathbb{K}$  un corps commutatif quelconque, la théorie des applications linéaires  $E \rightarrow \mathbb{K}$  où  $E$  est un espace vectoriel sur  $\mathbb{K}$ . On peut également voir la dualité linéaire comme la théorie des équations linéaires sur un espace vectoriel. En particulier, cette théorie nous donne une correspondance explicite entre les sous-espaces vectoriels d'un espace vectoriel  $E$  et les systèmes d'équations linéaires sur  $E$ . La dualité linéaire nous fournit également une interprétation vectorielle de l'opération de transposition sur les matrices.

Tout au long de ce chapitre,  $\mathbb{K}$  désigne un corps commutatif quelconque.

### 1.2 Formes linéaires sur un espace vectoriel et espace dual

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

**Définition 1.2.1.** On appelle *forme linéaire sur  $E$*  toute application linéaire de  $E$  dans  $\mathbb{K}$ . L'ensemble des formes linéaires sur  $E$  est appelé espace dual de  $E$  et noté  $E^*$ .

*Exemple 1.2.2* (exemple "fil rouge"). L'application  $\varphi : \begin{array}{ccc} \mathbb{R}^3 & \rightarrow & \mathbb{R} \\ (x, y, z) & \mapsto & 2x + 3y - 5z \end{array}$  est une forme linéaire sur  $\mathbb{R}^3$ .

*Remarque 1.2.3.* •  $E^* = \mathcal{L}(E, \mathbb{K})$  est un espace vectoriel sur  $\mathbb{K}$ .

- Si  $E$  est de dimension finie  $n \in \mathbb{N} \setminus \{0\}$  et si  $\mathcal{B} = \{e_1, \dots, e_n\}$  est une base de  $E$ , alors, pour tout vecteur  $v$  de  $E$  de coordonnées  $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$  dans la base  $\mathcal{B}$  et toute forme linéaire  $\varphi$  de

$E^*$ , on a

$$\begin{aligned}\varphi(v) &= \varphi(x_1 e_1 + \cdots + x_n e_n) \\ &= x_1 \underbrace{\varphi(e_1)}_{\in \mathbb{K}} + \cdots + x_n \underbrace{\varphi(e_n)}_{\in \mathbb{K}} \\ &= (\varphi(e_1) \cdots \varphi(e_n)) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}\end{aligned}$$

*Exemple 1.2.4* (suite de l'exemple "fil rouge"). Pour tout vecteur  $(x, y, z)$  de  $\mathbb{R}^3$ , on a

$$\varphi(x, y, z) = (2 \ 3 \ -5) \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

**Définition 1.2.5.** Soit  $\varphi \in E^*$  une forme linéaire sur  $E$  non identiquement nulle. On appelle hyperplan de  $E$  déterminé par  $\varphi$  le sous-espace vectoriel  $\text{Ker } \varphi$  de  $E$ .

*Exemple 1.2.6* (suite de l'exemple "fil rouge"). L'hyperplan de  $\mathbb{R}^3$  déterminé par  $\varphi$  est le sous-espace vectoriel  $\{(x, y, z) \in \mathbb{R}^3 \mid 2x + 3y - 5z = 0\}$  de  $\mathbb{R}^3$ .

L'appellation "hyperplan" est justifiée par le fait que, si  $E$  est de dimension finie  $n \in \mathbb{N} \setminus \{0\}$ , l'hyperplan déterminé par une forme linéaire sur  $E$  non identiquement nulle est effectivement un sous-espace vectoriel de  $E$  de dimension  $n - 1$ . Nous allons montrer ce fait ci-dessous, ainsi que sa réciproque :

**Proposition 1.2.7.** Supposons que  $E$  est de dimension finie  $n \in \mathbb{N} \setminus \{0\}$ .

Les hyperplans de  $E$  déterminés par les formes linéaires non nulles de  $E^*$  sont exactement les hyperplans linéaires de  $E$  i.e. ("id est" : c'est-à-dire) les sous-espaces vectoriels de  $E$  de dimension  $n - 1$ .

*Démonstration.* Soit  $\varphi \in E^* \setminus \{0\}$ . L'image  $\text{Im } \varphi$  de  $\varphi$  étant un sous-espace vectoriel de  $\mathbb{K}$ , la dimension de  $\text{Im } \varphi$  (sur  $\mathbb{K}$ ) est inférieure ou égale à 1 (la dimension de  $\mathbb{K}$  sur  $\mathbb{K}$  est 1). Comme  $\varphi$  est non identiquement nulle, la dimension de  $\text{Im } \varphi$  ne peut être 0. La dimension de  $\text{Im } \varphi$  est donc 1 et, par le théorème du rang,

$$\dim(\text{Ker } \varphi) = \dim(E) - \dim(\text{Im } \varphi) = n - 1,$$

i.e. l'hyperplan de  $E$  déterminé par  $\varphi$  est de dimension  $n - 1$ .

Réciproquement, soit  $H$  un sous-espace vectoriel de  $E$  de dimension  $n - 1$ . Soit  $v_0$  un vecteur de  $E$  n'appartenant pas à  $H$ . Alors les sous-espaces vectoriels  $H$  et  $\text{Vect}\{v_0\}$  de  $E$  sont en somme directe et, par un argument de dimension,  $E = H \oplus \text{Vect}\{v_0\}$ . Ainsi, tout vecteur  $v$  de  $E$  se décompose de façon unique en une somme  $v = u_v + \lambda_v v_0$  avec  $u_v \in H$  et  $\lambda_v \in \mathbb{K}$ . Si  $\varphi$  désigne alors la forme linéaire  $\begin{matrix} E & \rightarrow & \mathbb{K} \\ v & \mapsto & \lambda_v \end{matrix}$ , on a  $H = \text{Ker } \varphi$ , i.e.  $H$  est l'hyperplan de  $E$  déterminé par la forme linéaire  $\varphi$ . □



*Remarque 1.2.8.* Reprenons les notations de la remarque 1.2.3. Si  $\varphi \in E^*$ , l'hyperplan de  $E$  déterminé par  $\varphi$  est le sous-espace vectoriel de  $E$  caractérisé par l'équation linéaire

$$\underbrace{\varphi(e_1)}_{\in \mathbb{K}} x_1 + \cdots + \underbrace{\varphi(e_n)}_{\in \mathbb{K}} x_n = 0$$

en les coordonnées  $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$  dans la base  $\mathcal{B}$ .

Réciproquement, à tout sous-espace vectoriel  $H$  de  $E$  caractérisé par une équation linéaire

$$a_1 x_1 + \cdots + a_n x_n = 0$$

avec  $(a_1, \dots, a_n) \in \mathbb{K}^n \setminus \{(0, \dots, 0)\}$ , on peut associer la forme linéaire

$$\varphi : \begin{array}{ccc} E & \rightarrow & \mathbb{K} \\ x_1 e_1 + \cdots + x_n e_n & \mapsto & a_1 x_1 + \cdots + a_n x_n \end{array},$$

et alors  $H = \text{Ker } \varphi$ .

On obtient ainsi une "correspondance" entre l'espace dual  $E^*$  de  $E$  et les équations linéaires en les coordonnées dans la base  $\mathcal{B}$ . A noter que l'espace des solutions d'un système d'équations linéaires peut être vu comme une intersection d'hyperplans.

### 1.3 Base duale

Soit  $E$  un espace vectoriel sur  $\mathbb{K}$  de dimension finie  $n \in \mathbb{N} \setminus \{0\}$ .

En utilisant le fait que  $\dim(\mathcal{L}(E, \mathbb{K})) = \dim(E) \times \dim(\mathbb{K}) = \dim(E)$ , on peut montrer directement que  $\dim(E^*) = \dim(E)$ . On peut également le montrer en associant à toute base de  $E$  une base de  $E^*$  :

**Théorème et Définition 1.3.1.** *Soit  $\mathcal{B} = \{e_1, \dots, e_n\}$  une base de  $E$ .*

*Pour  $i \in \{1, \dots, n\}$ , on note  $e_i^*$  la forme linéaire sur  $E$  définie par*

$$\text{pour tout } j \in \{1, \dots, n\}, e_i^*(e_j) = \delta_{ij} = \begin{cases} 1 & \text{si } i = j, \\ 0 & \text{si } i \neq j. \end{cases}$$

*La famille  $\{e_1^*, \dots, e_n^*\}$  de  $E^*$  est une base de  $E^*$ , appelée base duale de  $\mathcal{B}$ . On la note  $\mathcal{B}^*$ .*

*Démonstration.* Soit  $i \in \{1, \dots, n\}$ . Commençons par remarquer que, par définition, si  $v$  est un

vecteur de  $E$  de coordonnées  $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$  dans la base  $\mathcal{B}$ ,

$$e_i^*(v) = e^*(x_1 e_1 + \cdots + x_n e_n) = x_1 e_i^*(e_1) + \cdots + x_n e_i^*(e_n) = x_i,$$

autrement dit  $e_i^*$  associe à tout vecteur  $v$  de  $E$  sa  $i^{\text{ème}}$  coordonnée dans la base  $\mathcal{B}$ .

A présent, montrons que la famille  $\{e_1^*, \dots, e_n^*\}$  de  $E^*$  est libre : soient  $\lambda_1, \dots, \lambda_n \in \mathbb{K}$  tels que  $\lambda_1 e_1^* + \dots + \lambda_n e_n^*$  soit la forme linéaire nulle, i.e., pour tout vecteur  $v$  de  $E$ ,  $\lambda_1 e_1^*(v) + \dots + \lambda_n e_n^*(v) = 0$ . En particulier, pour tout  $j \in \{1, \dots, n\}$ ,  $0 = \lambda_1 e_1^*(e_j) + \dots + \lambda_n e_n^*(e_j) = \lambda_j$  et la famille  $\{e_1^*, \dots, e_n^*\}$  de  $E^*$  est donc libre.

Montrons ensuite que la famille  $\{e_1^*, \dots, e_n^*\}$  engendre  $E^*$ . Soit donc  $\varphi \in E^*$ , et soit  $v$  un vecteur de coordonnées  $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$  dans la base  $\mathcal{B}$ . On a alors

$$\varphi(v) = \varphi(x_1 e_1 + \dots + x_n e_n) = \underbrace{\varphi(e_1)}_{\in \mathbb{K}} x_1 + \dots + \underbrace{\varphi(e_n)}_{\in \mathbb{K}} x_n = \varphi(e_1) e_1^*(v) + \dots + \varphi(e_n) e_n^*(v).$$

Ainsi,  $\varphi = \varphi(e_1) e_1^* + \dots + \varphi(e_n) e_n^* \in \text{Vect}\{e_1^*, \dots, e_n^*\}$  et la famille  $\{e_1^*, \dots, e_n^*\}$  est donc génératrice de  $E^*$ .  $\square$

*Remarque 1.3.2.* • On aurait pu se contenter de montrer le caractère libre ou le caractère générateur de la famille  $\{e_1^*, \dots, e_n^*\}$  puis d'utiliser le fait, établi précédemment, que  $\dim(E^*) = \dim(E) = n$  pour montrer que la famille  $\{e_1^*, \dots, e_n^*\}$  est une base de  $E^*$ . On a cependant fait le choix de la démonstration “complète” ci-dessus pour son intérêt didactique.

- $E$  et  $E^*$  étant deux espaces vectoriels de même dimension finie, ils sont isomorphes. Cependant, en général, ils ne le sont pas de façon “canonique” : un isomorphisme entre ces deux espaces vectoriels dépend, en général, d'un choix de bases pour  $E$  et  $E^*$ .

*Exemple 1.3.3.* Si  $\mathcal{B} = \{e_1, \dots, e_n\}$  est la base canonique de  $\mathbb{K}^n$ , pour tout  $i \in \{1, \dots, n\}$ ,  $e_i^*$  est la forme linéaire

$$e_i^* : \begin{array}{ccc} \mathbb{K}^n & \rightarrow & \mathbb{K} \\ (x_1, \dots, x_n) & \mapsto & x_i \end{array}$$

*Exemple 1.3.4.* On considère la base  $\mathcal{B} = \{e_1, e_2, e_3\}$  de  $\mathbb{R}^3$  formé par les vecteurs  $e_1 := (1, 1, 1)$ ,  $e_2 := (1, 0, -1)$  et  $e_3 := (0, 1, 1)$ . Déterminons la base duale  $\mathcal{B}^*$  de  $\mathcal{B}$  : précisément, nous allons déterminer les expressions des formes linéaires  $e_1^*$ ,  $e_2^*$  et  $e_3^*$  sur  $\mathbb{R}^3$ .

On cherche  $a, b, c \in \mathbb{R}$  tels que, pour tout  $(x_1, x_2, x_3) \in \mathbb{R}^3$ ,  $e_1^*(x_1, x_2, x_3) = ax_1 + bx_2 + cx_3$ . Or

$$\begin{cases} e_1^*(e_1) = 1 \\ e_1^*(e_2) = 0 \\ e_1^*(e_3) = 0 \end{cases} \Leftrightarrow \begin{cases} a + b + c = 1 \\ a - c = 0 \\ b + c = 0 \end{cases} \Leftrightarrow \begin{cases} a = 1 \\ b = -1 \\ c = 1 \end{cases}$$

Ainsi,  $e_1^*$  est l'application

$$e_1^* : \begin{array}{ccc} \mathbb{R}^3 & \rightarrow & \mathbb{R} \\ (x_1, x_2, x_3) & \mapsto & x_1 - x_2 + x_3 \end{array}$$

On cherche à présent  $a, b, c \in \mathbb{R}$  tels que, pour tout  $(x_1, x_2, x_3) \in \mathbb{R}^3$ ,  $e_2^*(x_1, x_2, x_3) = ax_1 + bx_2 + cx_3$ . Or

$$\begin{cases} e_2^*(e_1) = 0 \\ e_2^*(e_2) = 1 \\ e_2^*(e_3) = 0 \end{cases} \Leftrightarrow \begin{cases} a + b + c = 0 \\ a & -c = 1 \\ & b + c = 0 \end{cases} \Leftrightarrow \begin{cases} a = 0 \\ b = 1 \\ c = -1 \end{cases}$$

Ainsi,  $e_2^*$  est l'application

$$e_2^* : \begin{array}{ccc} \mathbb{R}^3 & \rightarrow & \mathbb{R} \\ (x_1, x_2, x_3) & \mapsto & x_2 - x_3 \end{array}$$

Enfin, on cherche  $a, b, c \in \mathbb{R}$  tels que, pour tout  $(x_1, x_2, x_3) \in \mathbb{R}^3$ ,  $e_3^*(x_1, x_2, x_3) = ax_1 + bx_2 + cx_3$ . Or

$$\begin{cases} e_3^*(e_1) = 0 \\ e_3^*(e_2) = 0 \\ e_3^*(e_3) = 1 \end{cases} \Leftrightarrow \begin{cases} a + b + c = 0 \\ a & -c = 0 \\ & b + c = 1 \end{cases} \Leftrightarrow \begin{cases} a = -1 \\ b = 2 \\ c = -1 \end{cases}$$

Ainsi,  $e_3^*$  est l'application

$$e_3^* : \begin{array}{ccc} \mathbb{R}^3 & \rightarrow & \mathbb{R} \\ (x_1, x_2, x_3) & \mapsto & -x_1 + 2x_2 - x_3 \end{array}$$

*Remarque 1.3.5.* Attention : parler de “dual d'un vecteur” n'a pas de sens. Si  $\mathcal{B}_1$  et  $\mathcal{B}_2$  sont deux bases de  $E$  et  $v$  est un vecteur de  $E$  appartenant à chacune de ces deux bases, les vecteurs “ $v^*$ ” dans  $\mathcal{B}_1^*$  et “ $v^*$ ” dans  $\mathcal{B}_2^*$  sont a priori différents (on devrait écrire  $v^{*\mathcal{B}_1}$ , respectivement  $v^{*\mathcal{B}_2}$ ).

Reprenons les vecteurs  $e_1 = (1, 1, 1)$  et  $e_2 = (1, 0, -1)$  de  $\mathbb{R}^3$  de l'exemple précédent mais posons cette fois  $v_3 := (1, 0, 0)$ . La famille  $\mathcal{B}' := \{e_1, e_2, v_3\}$  est également une base de  $\mathbb{R}^3$ . Déterminons les formes linéaires de la base duale  $\mathcal{B}'^*$  de  $\mathcal{B}'$  : on cherche  $a, b, c \in \mathbb{R}$  tels que, pour tout  $(x_1, x_2, x_3) \in \mathbb{R}^3$ ,  $e_1^*(x_1, x_2, x_3) = ax_1 + bx_2 + cx_3$ . Or

$$\begin{cases} e_1^*(e_1) = 1 \\ e_1^*(e_2) = 0 \\ e_1^*(v_3) = 0 \end{cases} \Leftrightarrow \begin{cases} a + b + c = 1 \\ a & -c = 0 \\ a & = 0 \end{cases} \Leftrightarrow \begin{cases} a = 0 \\ b = 1 \\ c = 0 \end{cases}$$

Ainsi,  $e_1^*$  est l'application

$$e_1^* : \begin{array}{ccc} \mathbb{R}^3 & \rightarrow & \mathbb{R} \\ (x_1, x_2, x_3) & \mapsto & x_2 \end{array}$$

On cherche ensuite  $a, b, c \in \mathbb{R}$  tels que, pour tout  $(x_1, x_2, x_3) \in \mathbb{R}^3$ ,  $e_2^*(x_1, x_2, x_3) = ax_1 + bx_2 + cx_3$ . Or

$$\begin{cases} e_2^*(e_1) = 0 \\ e_2^*(e_2) = 1 \\ e_2^*(v_3) = 0 \end{cases} \Leftrightarrow \begin{cases} a + b + c = 0 \\ a & -c = 1 \\ a & = 0 \end{cases} \Leftrightarrow \begin{cases} a = 0 \\ b = 1 \\ c = -1 \end{cases}$$

Ainsi,  $e_2^*$  est l'application

$$e_2^* : \begin{array}{ccc} \mathbb{R}^3 & \rightarrow & \mathbb{R} \\ (x_1, x_2, x_3) & \mapsto & x_2 - x_3 \end{array}$$

Enfin, on cherche  $a, b, c \in \mathbb{R}$  tels que, pour tout  $(x_1, x_2, x_3) \in \mathbb{R}^3$ ,  $v_3^*(x_1, x_2, x_3) = ax_1 + bx_2 + cx_3$ . Or

$$\begin{cases} v_3^*(e_1) = 0 \\ v_3^*(e_2) = 0 \\ v_3^*(e_3) = 1 \end{cases} \Leftrightarrow \begin{cases} a + b + c = 0 \\ a - c = 0 \\ a = 1 \end{cases} \Leftrightarrow \begin{cases} a = 1 \\ b = -2 \\ c = 1 \end{cases}$$

Ainsi,  $v_3^*$  est l'application

$$v_3^* : \begin{array}{ccc} \mathbb{R}^3 & \rightarrow & \mathbb{R} \\ (x_1, x_2, x_3) & \mapsto & x_1 - 2x_2 + x_3 \end{array}$$

On remarque ainsi que  $e_2^{*\mathcal{B}'} = e_2^{*\mathcal{B}}$  mais que  $e_1^{*\mathcal{B}'} \neq e_1^{*\mathcal{B}}$ . A noter également que, même si  $v_3$  est le troisième vecteur de la base canonique de  $\mathbb{R}^3$ ,  $v_3^{*\mathcal{B}'}$  n'est pas l'application  $\begin{array}{ccc} \mathbb{R}^3 & \rightarrow & \mathbb{R} \\ (x_1, x_2, x_3) & \mapsto & x_3 \end{array}$ .

Dans la suite,  $\mathcal{B} = \{e_1, \dots, e_n\}$  désignera une base de  $E$ .

**Proposition 1.3.6.** 1. Pour toute forme linéaire  $\varphi \in E^*$ ,  $\varphi = \sum_{i=1}^n \varphi(e_i)e_i^*$ , autrement dit  $\varphi$

a pour coordonnées  $\begin{pmatrix} \varphi(e_1) \\ \vdots \\ \varphi(e_n) \end{pmatrix}$  dans la base duale  $\mathcal{B}^*$  de  $\mathcal{B}$ .

2. Pour tout vecteur  $v \in E$ ,  $v = \sum_{j=1}^n e_j^*(v)e_j$ , autrement dit  $v$  a pour coordonnées  $\begin{pmatrix} e_1^*(v) \\ \vdots \\ e_n^*(v) \end{pmatrix}$  dans la base  $\mathcal{B}$ .

*Démonstration.* 1. Soit  $\varphi \in E^*$ . Comme  $\mathcal{B}^*$  est une base de  $E^*$ , il existe  $\lambda_1, \dots, \lambda_n \in \mathbb{K}$  (uniques) tels que  $\varphi = \sum_{i=1}^n \lambda_i e_i^*$ . Si  $j \in \{1, \dots, n\}$ , on a alors  $\varphi(e_j) = \sum_{i=1}^n \lambda_i e_i^*(e_j) = \lambda_j$ , et donc  $\varphi = \sum_{i=1}^n \varphi(e_i)e_i^*$ .

2. Soit  $v \in E$ . Comme  $\mathcal{B}$  est une base de  $E$ , il existe  $\mu_1, \dots, \mu_n \in \mathbb{K}$  (uniques) tels que  $v = \sum_{j=1}^n \mu_j e_j$ . Si  $i \in \{1, \dots, n\}$ , on a alors  $e_i^*(v) = \sum_{j=1}^n \mu_j e_i^*(e_j) = \mu_i$ , et donc  $v = \sum_{j=1}^n e_j^*(v)e_j$ .  $\square$

*Remarque 1.3.7.* Cela peut constituer un moyen "efficace" de déterminer les coordonnées d'une forme linéaire dans une base duale donnée, resp. ("respectivement") d'un vecteur dans une base donnée.

*Exemple 1.3.8.* • Reprenons la forme linéaire  $\varphi : \begin{array}{ccc} \mathbb{R}^3 & \rightarrow & \mathbb{R} \\ (x_1, x_2, x_3) & \mapsto & 2x_1 + 3x_2 - 5x_3 \end{array}$  de l'exemple fil rouge et déterminons ses coordonnées dans la base duale  $\mathcal{B}^*$  de l'exemple

1.3.4. On a  $\varphi(e_1) = 0$ ,  $\varphi(e_2) = 7$ ,  $\varphi(e_3) = -2$  et donc  $\varphi = 7e_2^* - 2e_3^*$ . Remarquons que l'on n'a pas besoin de l'expression des formes linéaires de la base duale pour déterminer les coordonnées de  $\varphi$  dans celle-ci (les expressions obtenues dans l'exemple 1.3.4 nous permettent cependant de vérifier que la décomposition précédente est bien correcte).

- Si l'on considère le vecteur  $v = (3, -4, 1)$  de  $\mathbb{R}^3$ , on obtient ses coordonnées dans la base  $\mathcal{B}$  de l'exemple 1.3.4 en calculant  $e_1^*(v) = 8$ ,  $e_2^*(v) = -5$  et  $e_3^*(v) = -12$ . On a donc  $v = 8e_1 - 5e_2 - 12e_3$ .

La proposition 1.3.6 nous permet également de montrer de l'opération qui à toute base de  $E$  associe sa base duale est injective. Nous montrerons sa surjectivité dans la section suivante.

**Corollaire 1.3.9.** *Soit  $\mathcal{B}' = \{f_1, \dots, f_n\}$  une base de  $E$ . Si  $\mathcal{B}^* = \mathcal{B}'^*$ , alors  $\mathcal{B} = \mathcal{B}'$ .*

*Démonstration.* Soit  $i \in \{1, \dots, n\}$ . D'après la proposition 1.3.6,

$$e_i = \sum_{j=1}^n f_j^*(e_i) f_j = \sum_{j=1}^n e_j^*(e_i) f_j = f_i.$$

Ainsi  $\mathcal{B} = \mathcal{B}'$ . □

*Remarque 1.3.10.* Comme, sur l'espace vectoriel de dimension finie  $E^*$ , on a accès à des bases, on peut utiliser les outils matriciels pour étudier les formes linéaires de  $E^*$ .

## 1.4 Aspects matriciels

Comme dans la section précédente, on considère un espace vectoriel  $E$  de dimension finie  $n \in \mathbb{N} \setminus \{0\}$  et une base  $\mathcal{B} = \{e_1, \dots, e_n\}$  de  $E$ .

Commençons par une première remarque : si  $\varphi$  est une forme linéaire de  $E^*$  et  $v$  est un vecteur de  $E$ , alors, d'après la remarque 1.2.3,

$$\varphi(v) = {}^t \text{Mat}_{\mathcal{B}^*}(\varphi) \text{Mat}_{\mathcal{B}}(v).$$

A présent, nous allons nous intéresser au changement de base pour les bases duales : précisément, soit  $\mathcal{B}' = \{f_1, \dots, f_n\}$  une autre base de  $E$ , on peut calculer la matrice de passage de la base duale  $\mathcal{B}^*$  à la base duale  $\mathcal{B}'^*$  à partir de la matrice de passage de la base  $\mathcal{B}$  à la base  $\mathcal{B}'$ .

**Proposition 1.4.1.** *On a*

$$P_{\mathcal{B}^* \rightarrow \mathcal{B}'^*} = {}^t P_{\mathcal{B} \rightarrow \mathcal{B}'}^{-1} = {}^t P_{\mathcal{B}' \rightarrow \mathcal{B}}$$

(rappel : la transposition et l'inversion des matrices inversibles commutent).

*Démonstration.* Pour simplifier les écritures, on note  $P = (p_{ij})_{1 \leq i, j \leq n} := P_{\mathcal{B} \rightarrow \mathcal{B}'}$  et  $Q = (q_{ij})_{1 \leq i, j \leq n} := P_{\mathcal{B}^* \rightarrow \mathcal{B}'^*}$ . Ainsi, pour  $i, j \in \{1, \dots, n\}$ ,  $f_i^* = \sum_{k=1}^n q_{ki} e_k^*$  (il s'agit de la décomposition de  $f_i^*$  dans la base  $\mathcal{B}^*$ ) et  $f_j = \sum_{l=1}^n p_{lj} e_l$  (il s'agit de la décomposition de  $f_j$  dans la base  $\mathcal{B}$ ). Nous allons montrer que  $I_n = {}^t Q P$  (et donc  $Q = {}^t P^{-1}$ ).

Soient  $i, j \in \{1, \dots, n\}$ . Le coefficient situé sur la ligne  $i$  et la colonne  $j$  de la matrice identité  $I_n$  est  $\delta_{i,j}$  et, par définition de la base duale  $\mathcal{B}'^* = \{f_1^*, \dots, f_n^*\}$ , on a

$$\begin{aligned} \delta_{i,j} = f_i^*(f_j) &= \left( \sum_{k=1}^n q_{k,i} e_k^* \right) \left( \sum_{l=1}^n p_{l,j} e_l \right) \\ &= \sum_{k=1}^n \sum_{l=1}^n q_{k,i} p_{l,j} e_k^*(e_l) \\ &= \sum_{k=1}^n \sum_{l=1}^n q_{k,i} p_{l,j} \delta_{k,l} \\ &= \sum_{k=1}^n q_{k,i} p_{k,j} \end{aligned}$$

Or  $\sum_{k=1}^n q_{k,i} p_{k,j}$  est justement le coefficient situé sur la ligne  $i$  et la colonne  $j$  de la matrice produit  ${}^tQP$ . Ainsi, on a bien  $I_n = {}^tQP$  i.e.  $Q = {}^tP^{-1}$  i.e.  $P_{\mathcal{B}'^* \rightarrow \mathcal{B}^*} = {}^tP_{\mathcal{B} \rightarrow \mathcal{B}'}^{-1}$ .  $\square$

A l'aide de cette propriété, on peut également montrer la surjectivité, et donc la bijectivité (voir corollaire 1.3.9 ci-dessus), de l'opération qui associe à toute base  $\mathcal{B}$  de  $E$  sa base duale  $\mathcal{B}^*$  de  $E^*$  :

**Corollaire et Définition 1.4.2.** *Pour toute base  $\mathcal{C}$  de  $E^*$ , il existe une et une seule base  $\mathcal{B}$  de  $E$  telle que  $\mathcal{C} = \mathcal{B}^*$ . On appelle  $\mathcal{B}$  la base antéduale de  $\mathcal{C}$ .*

*Démonstration.* Soit  $\mathcal{C}$  une base de  $E^*$ . Fixons maintenant  $\mathcal{B}_0$  une base quelconque de  $E$  et considérons la matrice  $Q := {}^tP_{\mathcal{B}_0^* \rightarrow \mathcal{C}}^{-1}$ . Comme il s'agit d'une matrice inversible de taille  $n$ ,  $Q$  peut être considérée comme la matrice de passage  $P_{\mathcal{B}_0 \rightarrow \mathcal{B}}$  de la base  $\mathcal{B}_0$  de  $E$  à une base  $\mathcal{B}$  (les coordonnées des vecteurs de  $\mathcal{B}$  dans la base  $\mathcal{B}_0$  sont données par les colonnes de  $Q$ ) et on a alors, d'après la proposition précédente,

$$P_{\mathcal{B}_0^* \rightarrow \mathcal{B}^*} = {}^tP_{\mathcal{B}_0 \rightarrow \mathcal{B}}^{-1} = {}^tQ^{-1} = P_{\mathcal{B}_0^* \rightarrow \mathcal{C}},$$

de sorte que les coordonnées des vecteurs de la base  $\mathcal{C}$  dans la base  $\mathcal{B}_0^*$  sont les mêmes que les coordonnées des vecteurs de la base  $\mathcal{B}^*$  dans la base  $\mathcal{B}_0^*$  et donc  $\mathcal{C} = \mathcal{B}^*$ .  $\square$

*Exemple 1.4.3.* On considère les formes linéaires  $\varphi_1 : \mathbb{R}^3 \rightarrow \mathbb{R}$ ,  $\varphi_2 : \mathbb{R}^3 \rightarrow \mathbb{R}$  :

$$\begin{aligned} \mathbb{R}^3 &\rightarrow \mathbb{R} & \text{et } \varphi_3 : \mathbb{R}^3 &\rightarrow \mathbb{R} \\ (x_1, x_2, x_3) &\rightarrow -x_1 + x_3 & (x_1, x_2, x_3) &\rightarrow x_2 + x_3 \end{aligned} \quad \text{sur } \mathbb{R}^3. \text{ La famille } \mathcal{C} = \{\varphi_1, \varphi_2, \varphi_3\}$$

est une base de  $(\mathbb{R}^3)^*$ . Pour le voir, on écrit les coordonnées de  $\varphi_1, \varphi_2$  et  $\varphi_3$  dans la base duale  $\mathcal{B}_0^* = \{e_1^*, e_2^*, e_3^*\}$  de la base canonique  $\mathcal{B}_0 = \{e_1, e_2, e_3\}$  de  $\mathbb{R}^3$  : on a  $\varphi_1 = e_1^* + e_2^* + e_3^*$ ,  $\varphi_2 = -e_1^* + e_3^*$  et  $\varphi_3 = e_2^* + e_3^*$ , et la matrice

$$P := \begin{pmatrix} 1 & -1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

dont les colonnes sont les coordonnées de  $\varphi_1, \varphi_2$  et  $\varphi_3$  dans la base  $\mathcal{B}^*$ , est inversible.

On cherche maintenant à déterminer la base antéduale  $\mathcal{B} = \{v_1, v_2, v_3\}$  de la base  $\mathcal{C}$ . On procède comme dans la démonstration précédente : la matrice  $P$  ci-dessus est la matrice de passage de  $\mathcal{B}_0^*$  à  $\mathcal{C}$  et la matrice de passage de la base  $\mathcal{B}_0$  à la base  $\mathcal{B}$  est alors la matrice  ${}^tP^{-1}$ . On obtient

$${}^tP^{-1} = \begin{pmatrix} 1 & 0 & -1 \\ -1 & -1 & 2 \\ 1 & 1 & -1 \end{pmatrix}$$

et on a donc  $v_1 = e_1 - e_2 + e_3 = (1, -1, 1)$ ,  $v_2 = -e_2 + e_3 = (0, -1, 1)$  et  $v_3 = -e_1 + 2e_2 - e_3 = (-1, 2, -1)$ .

## 1.5 Annulateur d'un sous-espace vectoriel et correspondance duale

Soit  $E$  un espace vectoriel de dimension finie  $n \in \mathbb{N} \setminus \{0\}$ .

Dans cette section, on va définir des outils de la dualité qui vont nous donner une correspondance explicite entre les sous-espaces vectoriels de  $E$  et les systèmes d'équations linéaires qui les caractérisent.

Soit  $F$  un sous-espace vectoriel de  $E$  et soit  $W$  un sous-espace vectoriel de  $E^*$ .

**Définition 1.5.1.** • L'ensemble, noté  $F^0$ , des formes linéaires de  $E^*$  qui s'annulent sur  $F$  est appelé annulateur de  $F$ .

• L'ensemble, noté  $W^0$ , des vecteurs de  $E$  qui sont annulés par toutes les formes linéaires de  $W$  est appelé annulateur de  $W$ .

L'ensemble  $F^0 = \{\varphi \in E^* \mid \text{pour tout } v \in F, \varphi(v) = 0\}$  est un sous-espace vectoriel de  $E^*$  : la forme linéaire identiquement nulle sur  $E$  appartient à  $F^0$  et, si  $\varphi, \psi \in E^*$  et  $\lambda, \mu \in \mathbb{K}$ , pour tout  $v \in E$ ,  $(\lambda\varphi + \mu\psi)(v) = \lambda\varphi(v) + \mu\psi(v) = 0$ .

De façon analogue,  $W^0 = \{v \in E \mid \text{pour tout } \varphi \in W, \varphi(v) = 0\}$  est un sous-espace vectoriel de  $E$  : le vecteur nul de  $E$  appartient à  $W^0$  et, si  $v, w \in E$  et  $\lambda, \mu \in \mathbb{K}$ , pour tout  $\varphi \in E^*$ ,  $\varphi(\lambda v + \mu w) = \lambda\varphi(v) + \mu\varphi(w) = 0$ .

**Proposition 1.5.2.** 1. Si  $\{v_1, \dots, v_p\}$  est une base de  $F$ , alors  $F^0 = \{\varphi \in E^* \mid \varphi(v_1) = 0, \dots, \varphi(v_p) = 0\}$ .

2. Si  $\{\varphi_1, \dots, \varphi_q\}$  est une base de  $W$ , alors  $W^0 = \{v \in E \mid \varphi_1(v) = 0, \dots, \varphi_q(v) = 0\}$ .

*Démonstration.* 1. Soit  $\varphi \in F^0$ , alors, comme  $v_1, \dots, v_p \in F$ ,  $\varphi(v_1) = 0, \dots, \varphi(v_p) = 0$ . Réciproquement, soit maintenant  $\varphi$  une forme linéaire sur  $E$  annulant les vecteurs  $v_1, \dots, v_p$  et soit  $v \in F$ . Comme  $\{v_1, \dots, v_p\}$  est une base de  $F$ , il existe  $\lambda_1, \dots, \lambda_p \in \mathbb{K}$  (uniques) tels que  $v = \lambda_1 v_1 + \dots + \lambda_p v_p$  et alors

$$\varphi(v) = \lambda_1 \varphi(v_1) + \dots + \lambda_p \varphi(v_p) = 0,$$

et  $\varphi$  appartient donc à  $F^0$ .

2. Soit  $v \in W^0$ , alors, comme  $\varphi_1, \dots, \varphi_q \in W$ ,  $\varphi_1(v) = 0, \dots, \varphi_q(v) = 0$ . Réciproquement, soit maintenant  $v$  un vecteur de  $E$  annulé par les formes linéaires  $\varphi_1, \dots, \varphi_q$  et soit  $\varphi \in W$ . Comme  $\{\varphi_1, \dots, \varphi_q\}$  est une base de  $W$ , il existe  $\mu_1, \dots, \mu_q \in \mathbb{K}$  (uniques) tels que  $\varphi = \mu_1\varphi_1 + \dots + \mu_q\varphi_q$  et alors

$$\varphi(v) = \mu_1\varphi_1(v) + \dots + \mu_q\varphi_q(v) = 0,$$

et  $v$  appartient donc à  $W^0$ . □

Remarquons que, si  $\{\varphi_1, \dots, \varphi_q\}$  est une base de  $W$ ,  $W^0 = \bigcap_{i=1}^q \text{Ker } \varphi_i$  et que, si  $\mathcal{B} = \{e_1, \dots, e_n\}$  est une base de  $E$ ,  $W^0$  est le sous-espace vectoriel de  $E$  caractérisé par le système linéaire

$$\begin{cases} \varphi_1(e_1)x_1 + \dots + \varphi_1(e_n)x_n = 0 \\ \vdots \\ \varphi_q(e_1)x_1 + \dots + \varphi_q(e_n)x_n = 0 \end{cases}$$

en les coordonnées  $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$  dans la base  $\mathcal{B}$ .

La proposition suivante affirme notamment que si  $W = F^0$ , alors  $F$  peut être décrit par le système linéaire ci-dessus, autrement dit que les vecteurs de  $F$  sont exactement les solutions de ce système.

**Proposition 1.5.3.** *On a*

1.  $\dim(E) = \dim(F) + \dim(F^0)$  et  $\dim(E^*) = \dim(W) + \dim(W^0)$ ,
2.  $(F^0)^0 = F$  et  $(W^0)^0 = W$ .

*Démonstration.* 1. Montrons tout d'abord que  $\dim(E) = \dim(F) + \dim(F^0)$ . Soit  $\{v_1, \dots, v_p\}$  une base de  $F$  que l'on complète en une base  $\mathcal{B} = \{v_1, \dots, v_p, v_{p+1}, \dots, v_n\}$  de  $E$ . Considérons la base duale  $\mathcal{B}^* = \{v_1^*, \dots, v_p^*, v_{p+1}^*, \dots, v_n^*\}$  et montrons que la famille  $\{v_{p+1}^*, \dots, v_n^*\}$  est une base de  $F^0$  :

- Soit  $i \in \{p+1, \dots, n\}$ , alors  $v_i^* \in F^0$  car, pour tout  $j \in \{1, \dots, p\}$ ,  $v_i^*(v_j) = 0$  ( $i \neq j$ ) et les vecteurs  $v_1, \dots, v_p$  engendrent  $F$ .
- La famille  $\{v_{p+1}^*, \dots, v_n^*\}$  de  $E^*$  est libre comme sous-famille de la base  $\mathcal{B}^*$ .
- De plus, elle engendre  $F^0$  : en effet, soit  $\varphi \in F^0$ , alors, d'après la proposition 1.3.6 i),

$$\begin{aligned} \varphi &= \varphi(v_1)v_1^* + \dots + \varphi(v_p)v_p^* + \varphi(v_{p+1})v_{p+1}^* + \dots + \varphi(v_n)v_n^* \\ &= \varphi(v_{p+1})v_{p+1}^* + \dots + \varphi(v_n)v_n^* \in \text{Vect} \{v_{p+1}^*, \dots, v_n^*\} \end{aligned}$$

$$(\varphi(v_1) = \dots = \varphi(v_p) = 0 \text{ car } v_1, \dots, v_p \in F \text{ et } \varphi \in F^0).$$



En particulier,  $\dim(F^0) = n - p = \dim(E) - \dim(F)$ .

L'égalité  $\dim(E^0) = \dim(W) + \dim(W^0)$  se démontre de façon tout à fait similaire, à l'aide de la notion de base antéduale (corollaire et définition 1.4.2) : soit  $\{\varphi_1, \dots, \varphi_q\}$  une base de  $W$  que l'on complète en une base  $\mathcal{C} = \{\varphi_1, \dots, \varphi_q, \varphi_{q+1}, \dots, \varphi_n\}$  de  $E^*$  et notons  $\mathcal{B} = \{v_1, \dots, v_q, v_{q+1}, \dots, v_n\}$  la base antéduale de  $\mathcal{C}$ . On montre que la famille  $\{v_{q+1}, \dots, v_n\}$  est une base de  $W^0$  :

- Soit  $j \in \{q+1, \dots, n\}$ , alors  $v_j \in W^0$  car, pour tout  $i \in \{1, \dots, q\}$ ,  $\varphi_i(v_j) = v_i^*(v_j) = 0$  ( $i \neq j$ ) et les vecteurs  $\varphi_1, \dots, \varphi_q$  engendrent  $W$ .
- La famille  $\{v_{q+1}, \dots, v_n\}$  de  $E$  est libre comme sous-famille de la base  $\mathcal{B}$ .
- De plus, elle engendre  $W^0$  : en effet, soit  $v \in W^0$ , alors, d'après la proposition 1.3.6 ii),

$$\begin{aligned} v &= v_1^*(v)v_1 + \dots + v_q^*(v)v_q + v_{q+1}^*(v)v_{q+1} + \dots + v_n^*(v)v_n \\ &= \varphi_1(v)v_1 + \dots + \varphi_q(v)v_q + \varphi_{q+1}(v)v_{q+1} + \dots + \varphi_n(v)v_n \\ &= \varphi_{q+1}(v)v_{q+1} + \dots + \varphi_n(v)v_n \in \text{Vect}\{v_{q+1}, \dots, v_n\} \end{aligned}$$

$$(\varphi_1(v) = \dots = \varphi_q(v) = 0 \text{ car } \varphi_1, \dots, \varphi_q \in W \text{ et } v \in W^0).$$

En particulier,  $\dim(W^0) = n - q = \dim(E^*) - \dim(W) = \dim(E) - \dim(W)$ .

2. Des deux égalités démontrées précédemment, on déduit la double inclusion  $(F^0)^0 = F$  : on a  $F \subset (F^0)^0$  (car si  $v \in F$  et  $\varphi \in F^0$  alors  $\varphi(v) = 0$ ) et

$$\dim((F^0)^0) = \dim(E) - \dim(F^0) = \dim(E) - (\dim(E) - \dim(F)) = \dim(F).$$

De même,  $(W^0)^0 = W$  car  $W \subset (W^0)^0$  (si  $\varphi \in W$  et  $v \in W^0$  alors  $\varphi(v) = 0$ ) et

$$\dim((W^0)^0) = \dim(E) - \dim(W^0) = \dim(E) - (\dim(E) - \dim(W)) = \dim(W).$$

□

Cette proposition et sa démonstration nous donnent en particulier une méthode pour, à partir d'une base de  $F$ , obtenir un système d'équations linéaires "linéairement indépendantes" (i.e. d'équations correspondant à des formes linéaires linéairement indépendantes) décrivant  $F$  :

*Exemple 1.5.4.* Soit  $F$  le sous-espace vectoriel de  $\mathbb{R}^3$  engendré par le vecteur  $v_1 = (1, 1, 1)$ . On note  $v_2$  le vecteur  $(1, 0, -1)$  et  $v_3$  le vecteur  $(0, 1, 1)$ , puis on complète la famille libre  $\{v_1\}$  de  $\mathbb{R}^3$  en la base  $\mathcal{B} = \{v_1, v_2, v_3\}$  (voir également exemple 1.3.4). On considère ensuite la base duale  $\mathcal{B}^* = \{v_1^*, v_2^*, v_3^*\}$  et, d'après ce que l'on a vu dans la démonstration précédente,  $F^0 = \text{Vect}\{v_2^*, v_3^*\}$ . L'expression de  $v_2^*$  sur  $\mathbb{R}^3$  est  $v_2^* : \begin{matrix} \mathbb{R}^3 & \rightarrow & \mathbb{R} \\ (x_1, x_2, x_3) & \mapsto & x_2 - x_3 \end{matrix}$  et l'expression

de  $v_3^*$  sur  $\mathbb{R}^3$  est  $v_3^*$  :  $\mathbb{R}^3 \rightarrow \mathbb{R}$ . Ainsi

$$(x_1, x_2, x_3) \mapsto -x_1 + 2x_2 - x_3$$

$$\begin{aligned} F = (F^0)^0 &= \{(x_1, x_2, x_3) \in \mathbb{R}^3 \mid v_2^*(x_1, x_2, x_3) = 0, v_3^*(x_1, x_2, x_3) = 0\} \\ &= \{(x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_2 - x_3 = 0, -x_1 + 2x_2 - x_3 = 0\}. \end{aligned}$$

Une méthode analogue permet d'obtenir, à partir d'une description de  $F$  comme ensemble des solutions d'un système d'équations linéaires linéairement indépendantes, une base de  $F$  :

*Exemple 1.5.5.* Notons  $\mathcal{B}_0 = \{e_1, e_2, e_3\}$  la base canonique de  $\mathbb{R}^3$  et considérons les formes linéaires  $\varphi_1 = e_1^* + e_2^* + e_3^*$  et  $\varphi_2 = -e_1^* + e_3^*$  sur  $\mathbb{R}^3$ . On note  $W := \text{Vect}\{\varphi_1, \varphi_2\} \subset \mathbb{R}^{3*}$  et on cherche à déterminer une base de  $W^0 = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_1 + x_2 + x_3 = 0, -x_1 + x_3 = 0\}$ .

Tout d'abord, remarquons que les formes linéaires  $\varphi_1$  et  $\varphi_2$  sont linéairement indépendantes : on peut par exemple constituer la matrice dont les colonnes sont les coordonnées de  $\varphi_1$  et  $\varphi_2$  dans la base  $\mathcal{B}_0^*$  et montrer qu'elle est bien de rang 2. On note ensuite  $\varphi_3 := e_2^* + e_3^*$  et on complète la famille libre  $\{\varphi_1, \varphi_2\}$  en la base  $\mathcal{C} := \{\varphi_1, \varphi_2\}$  de  $(\mathbb{R}^3)^*$ . D'après l'exemple 1.4.3, la base préduale de  $\mathcal{C}$  est la base  $\mathcal{B} = \{(1, -1, 1), (0, -1, 1), (-1, 2, -1)\}$  de  $\mathbb{R}^3$  et, d'après la démonstration de la proposition 1.5.3,  $W^0 = \text{Vect}\{(-1, 2, -1)\}$ .

*Remarque 1.5.6.* On a  $\{0_E\}^0 = E^*$ ,  $E^0 = \{0_{E^*}\}$ ,  $\{0_{E^*}\}^0 = E$  et  $(E^*)^0 = \{0_E\}$ .

## 1.6 Application transposée

La dualité linéaire va également nous permettre de donner une interprétation vectorielle à l'opération de transposition sur les matrices.

Soient  $E$  et  $F$  deux espaces vectoriels sur  $\mathbb{K}$ .

**Définition 1.6.1.** Soit  $f \in \mathcal{L}(E, F)$ . On appelle transposée de  $f$  l'application linéaire

$${}^t f : \begin{array}{ccc} F^* & \rightarrow & E^* \\ \varphi & \mapsto & \varphi \circ f \end{array}$$

*Remarque 1.6.2.* • Si  $\varphi \in F^* = \mathcal{L}(F, \mathbb{K})$ , on a bien  $\varphi \circ f \in E^* = \mathcal{L}(E, \mathbb{K})$ .

- L'application  ${}^t f$  est bien linéaire : si  $\varphi, \psi \in F^*$  et  $\lambda, \mu \in \mathbb{K}$ ,

$$\begin{aligned} {}^t f(\lambda\varphi + \mu\psi) &= (\lambda\varphi + \mu\psi) \circ f \\ &= \lambda\varphi \circ f + \mu\psi \circ f \\ &= \lambda {}^t f(\varphi) + \mu {}^t f(\psi) \end{aligned}$$

Les propriétés de base de la transposée d'une application linéaire sont réunies dans la proposition suivante :

**Proposition 1.6.3.** 1.  ${}^t(\text{Id}_E) = \text{Id}_{E^*}$ .

2. Si  $f, g \in \mathcal{L}(E, F)$  et  $\lambda, \mu \in \mathbb{K}$ ,  ${}^t(\lambda f + \mu g) = \lambda {}^t f + \mu {}^t g$ .

3. Si  $G$  est un espace vectoriel sur  $\mathbb{K}$ ,  $f \in \mathcal{L}(E, F)$  et  $g \in \mathcal{L}(F, G)$ ,  ${}^t(g \circ f) = {}^t f \circ {}^t g$ .
4. Si  $f$  est une application linéaire bijective de  $E$  dans  $F$ , alors  ${}^t f : F^* \rightarrow E^*$  est également bijective et  $({}^t f)^{-1} = {}^t(f^{-1})$ .

*Démonstration.* 1. Pour tout  $\varphi \in E^*$ , on a

$${}^t(\text{Id}_E)(\varphi) = \varphi \circ \text{Id}_E = \varphi = \text{Id}_{E^*}(\varphi).$$

2. Pour tout  $\varphi \in F^*$ , on a

$${}^t(\lambda f + \mu g)(\varphi) = \varphi \circ (\lambda f + \mu g) = \lambda \varphi \circ f + \mu \varphi \circ g = \lambda {}^t f(\varphi) + \mu {}^t g(\varphi) = ({}^t f + {}^t g)(\varphi).$$

3. Pour tout  $\varphi \in G^*$ , on a

$${}^t(g \circ f)(\varphi) = \varphi \circ (g \circ f) = (\varphi \circ g) \circ f = {}^t g(\varphi) \circ f = {}^t f({}^t g(\varphi)) = ({}^t f \circ {}^t g)(\varphi).$$

4. On a  ${}^t f \circ {}^t(f^{-1}) = {}^t(f^{-1} \circ f) = {}^t(\text{Id}_E) = \text{Id}_{E^*}$  et  ${}^t(f^{-1}) \circ {}^t f = {}^t(f \circ f^{-1}) = {}^t(\text{Id}_F) = \text{Id}_{F^*}$ .  $\square$

On en vient à la justification matricielle de l'appellation "transposée" et, "réciproquement", à une interprétation vectorielle, via la dualité, de la transposition matricielle :

**Proposition 1.6.4.** *On suppose que les espaces vectoriels  $E$  et  $F$  sont tous deux de dimension finie. Soient alors  $\mathcal{B} = \{e_1, \dots, e_n\}$  une base de  $E$  et  $\mathcal{C} = \{v_1, \dots, v_m\}$  une base de  $F$ , et soit  $f \in \mathcal{L}(E, F)$ . On a*

$$\text{Mat}_{\mathcal{C}^*, \mathcal{B}^*}({}^t f) = {}^t \text{Mat}_{\mathcal{B}, \mathcal{C}}(f),$$

autrement dit la matrice de la transposée  ${}^t f$  de  $f$  dans les bases duales  $\mathcal{C}^*$  et  $\mathcal{B}^*$  est la transposée de la matrice de  $f$  dans les bases  $\mathcal{B}$  et  $\mathcal{C}$ , ou encore, symétriquement, la transposée de la matrice de  $f$  dans les bases  $\mathcal{B}$  et  $\mathcal{C}$  est la matrice de la transposée  ${}^t f$  de  $f$  dans les bases duales  $\mathcal{C}^*$  et  $\mathcal{B}^*$ .

*Démonstration.* Soit  $j \in \{1, \dots, m\}$  et notons  $A = (a_{kl})_{1 \leq k \leq m, 1 \leq l \leq n} := \text{Mat}_{\mathcal{B}, \mathcal{C}}(f)$ . Alors

$$\begin{aligned} {}^t f(v_j^*) &= \underbrace{v_j^* \circ f}_{\in E^*} = \sum_{i=1}^n v_j^* \circ f(e_i) e_i^* \text{ par la proposition 1.3.6, 1.} \\ &= \sum_{i=1}^n v_j^*(f(e_i)) e_i^* \\ &= \sum_{i=1}^n v_j^* \left( \sum_{k=1}^m a_{ki} v_k \right) e_i^* \\ &= \sum_{i=1}^n \sum_{k=1}^m a_{ki} v_j^*(v_k) e_i^* \\ &= \sum_{i=1}^n a_{ji} e_i^* \end{aligned}$$

Ainsi la matrice  $\text{Mat}_{\mathcal{C}^*, \mathcal{B}^*}({}^t f)$  est exactement la transposée de la matrice  $A = \text{Mat}_{\mathcal{B}, \mathcal{C}}(f)$ .  $\square$

*Remarque 1.6.5.* Ce résultat est également à mettre en lien avec la propriété 1.4.1 concernant la matrice de passage d'une base duale à une autre : si  $\mathcal{B}$  et  $\mathcal{B}'$  sont deux bases d'un espace vectoriel de dimension finie  $E$ , la matrice de passage  $P_{\mathcal{B} \rightarrow \mathcal{B}'}$  est la matrice  $\text{Mat}_{\mathcal{B}', \mathcal{B}}(\text{Id}_E)$  de l'identité de  $E$  dans les bases  $\mathcal{B}'$  et  $\mathcal{B}$ , et la transposée de cette matrice est, d'après les propositions précédentes, la matrice  $\text{Mat}_{\mathcal{B}^*, \mathcal{B}'^*}(\text{Id}_{E^*})$ , i.e. la matrice de passage  $P_{\mathcal{B}'^* \rightarrow \mathcal{B}^*}$ . Autrement dit,  ${}^t P_{\mathcal{B} \rightarrow \mathcal{B}'} = P_{\mathcal{B}'^* \rightarrow \mathcal{B}^*}$  et donc, de façon équivalente,  $P_{\mathcal{B}^* \rightarrow \mathcal{B}'^*} = P_{\mathcal{B}'^* \rightarrow \mathcal{B}^*}^{-1} = {}^t P_{\mathcal{B} \rightarrow \mathcal{B}'}^{-1} = {}^t P_{\mathcal{B}' \rightarrow \mathcal{B}}$ .

Ce point de vue "vectoriel" sur la transposition matricielle permet également, à partir de la proposition 1.6.3 et de la correspondance entre produit de matrices et composition d'applications linéaires, de montrer sans calcul les propriétés matricielles " ${}^t(AB) = {}^tB {}^tA$ " et " $({}^tA)^{-1} = {}^t(A^{-1})$ ".

## 1.7 Bidual

Soit  $E$  un espace vectoriel sur  $\mathbb{K}$ . Son dual  $E^*$  est également un espace vectoriel sur  $\mathbb{K}$  : on peut donc aussi considérer son dual que l'on note  $E^{**}$ . On a alors, par définition,

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

**Définition 1.7.1.** On appelle  $E^{**}$  le bidual de  $E$ .

On a vu qu'un espace vectoriel de dimension finie est isomorphe à son dual et donc, comme le dual est isomorphe à son propre dual, à son bidual. Néanmoins, comme on l'a dit plus haut, on ne dispose pas, en général, d'isomorphisme "canonique" entre un espace vectoriel de dimension finie et son dual. Une propriété remarquable du bidual est que, en dimension finie, tout espace vectoriel est canoniquement isomorphe à son bidual :

**Proposition 1.7.2.** On suppose que  $E$  est de dimension finie. Alors  $E$  est canoniquement isomorphe à son bidual  $E^{**}$ , i.e. on peut construire un isomorphisme linéaire de  $E$  sur  $E^{**}$  sans faire appel à des choix de bases.

*Démonstration.* Pour  $v \in E$ , définissons tout d'abord l'application  $\Phi_v : \begin{array}{ccc} E^* & \rightarrow & \mathbb{K} \\ \varphi & \mapsto & \varphi(v) \end{array}$  (l'application d'"évaluation" des formes linéaires sur  $E$  en le vecteur  $v$ ). Pour tout  $v \in E$ , l'application  $\Phi_v$  est linéaire : si  $\varphi, \psi \in E^*$  et  $\lambda, \mu \in \mathbb{K}$ ,

$$\Phi_v(\lambda\varphi + \mu\psi) = (\lambda\varphi + \mu\psi)(v) = \lambda\varphi(v) + \mu\psi(v) = \lambda\Phi_v(\varphi) + \mu\Phi_v(\psi).$$

Ainsi, pour tout  $v \in E$ ,  $\Phi_v \in \mathcal{L}(E^*, \mathbb{K}) = E^{**}$ .

On considère alors l'application

$$\Phi : \begin{array}{ccc} E & \rightarrow & E^{**} \\ v & \rightarrow & \Phi_v \end{array}$$

$\Phi$  est une application linéaire : si  $v, w \in E$  et  $\lambda, \mu \in \mathbb{K}$  alors, pour tout  $\varphi \in E^*$ ,

$$\begin{aligned}
 \Phi(\lambda v + \mu w)(\varphi) &= \Phi_{\lambda v + \mu w}(\varphi) \\
 &= \varphi(\lambda v + \mu w) \\
 &= \lambda\varphi(v) + \mu\varphi(w) \\
 &= \lambda\Phi_v(\varphi) + \mu\Phi_w(\varphi) \\
 &= \lambda\Phi(v)(\varphi) + \mu\Phi(w)(\varphi) \\
 &= (\lambda\Phi(v) + \mu\Phi(w))(\varphi)
 \end{aligned}$$

i.e.  $\Phi(\lambda v + \mu w) = \lambda\Phi(v) + \mu\Phi(w)$ .

On montre enfin que l'application linéaire  $\Phi : E \rightarrow E^{**}$  est bijective : comme  $\dim(E^{**}) = \dim(E^*) = \dim(E)$ , on peut se contenter de montrer que  $\Phi$  est injective. Soit donc  $v \in E$  tel que  $\Phi(v) = \Phi_v = 0$ , i.e., pour tout  $\varphi \in E^*$ ,  $\varphi(v) = 0$ . En particulier, si l'on considère une base  $\mathcal{B} = \{e_1, \dots, e_n\}$  de  $E$  et sa base duale  $\mathcal{B}^* = \{e_1^*, \dots, e_n^*\}$ , on a, pour tout  $j \in \{1, \dots, n\}$ ,  $e_j^*(v) = 0$  et donc  $v = \sum_{j=1}^n e_j^*(v)e_j = 0$  :  $\Phi$  est donc bien injective.

Ainsi, l'application  $\Phi$  est bien un isomorphisme linéaire de  $E$  sur  $E^{**}$  (et sa définition ne dépend pas d'un choix de bases).  $\square$



# Chapitre 2

## Espaces euclidiens

### 2.1 Introduction

On introduit sur les  $\mathbb{R}$ -espaces vectoriels de dimension finie une structure supplémentaire : le produit scalaire, notion qui généralise le produit scalaire classique sur  $\mathbb{R}^2$  ou  $\mathbb{R}^3$ . Cette structure supplémentaire nous donne accès aux notions géométriques d'orthogonalité et de distance.

### 2.2 Produit scalaire sur un espace vectoriel réel

Soit  $E$  un espace vectoriel sur  $\mathbb{R}$ .

**Définition 2.2.1.** *Considérons une application  $\langle \cdot, \cdot \rangle : E \times E \rightarrow \mathbb{R}$  :  $(v, w) \mapsto \langle v, w \rangle$ . On dit que  $\langle \cdot, \cdot \rangle$  est un produit scalaire sur  $E$  si  $\langle \cdot, \cdot \rangle$  est*

1. *bilinéaire, i.e. pour tous  $v_1, v_2, w_1, w_2 \in E$  et tous  $\lambda, \mu \in \mathbb{R}$ ,  $\langle \lambda v_1 + \mu v_2, w \rangle = \lambda \langle v_1, w \rangle + \mu \langle v_2, w \rangle$  et  $\langle v, \lambda w_1 + \mu w_2 \rangle = \lambda \langle v, w_1 \rangle + \mu \langle v, w_2 \rangle$ ,*
2. *symétrique, i.e. pour tous  $v, w \in E$ ,  $\langle v, w \rangle = \langle w, v \rangle$ ,*
3. *définie positive, i.e. pour tout  $v \in E$ ,  $\langle v, v \rangle \geq 0$  et  $\langle v, v \rangle = 0$  si et seulement si  $v = 0_E$ .*

*Exemple 2.2.2.* 1. Soit  $n \in \mathbb{N} \setminus \{0\}$ , pour tous  $v = (x_1, \dots, x_n)$  et  $w = (y_1, \dots, y_n)$  dans  $\mathbb{R}^n$ , on définit

$$\langle v, w \rangle_{\text{can}} := x_1 y_1 + \dots + x_n y_n.$$

L'application  $\langle \cdot, \cdot \rangle_{\text{can}} : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$  :  $(v, w) \mapsto \langle v, w \rangle_{\text{can}}$  est alors un produit scalaire sur  $\mathbb{R}^n$ , appelé produit scalaire canonique sur  $\mathbb{R}^n$ . Montrons le caractère défini positif de l'appli-

cation  $\langle \cdot, \cdot \rangle_{\text{can}}$  : si  $v = (x_1, \dots, x_n) \in \mathbb{R}^n$ ,  $\langle v, v \rangle_{\text{can}} = \sum_i^n x_i^2 \geq 0$  et  $\langle v, v \rangle_{\text{can}} = \sum_i^n x_i^2 = 0$  ssi ("si et seulement si") pour tout  $i \in \{1, \dots, n\}$ ,  $x_i = 0$  ssi  $v = (0, \dots, 0)$ .

2. Supposons que  $E$  est un espace vectoriel de dimension finie  $n \in \mathbb{N} \setminus \{0\}$  et soit  $\mathcal{B} = \{e_1, \dots, e_n\}$  une base de  $E$ . Pour tous vecteurs  $v$  et  $w$  de  $E$ , de coordonnées respectives  $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$  et  $\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$  dans la base  $\mathcal{B}$ , on définit

$$\langle v, w \rangle_{\mathcal{B}} := x_1 y_1 + \dots + x_n y_n = {}^t \text{Mat}_{\mathcal{B}}(v) \text{Mat}_{\mathcal{B}}(w).$$

L'application  $\langle \cdot, \cdot \rangle_{\mathcal{B}} : E \times E \rightarrow \mathbb{R}$   
 $(v, w) \mapsto \langle v, w \rangle_{\mathcal{B}}$  est alors un produit scalaire sur  $E$ , appelé produit scalaire associé à la base  $\mathcal{B}$ .

3. Soit  $n \in \mathbb{N} \setminus \{0\}$ . Pour toutes matrices  $A = (a_{ij})_{1 \leq i, j \leq n}$  et  $B = (b_{ij})_{1 \leq i, j \leq n}$  dans  $M_n(\mathbb{R})$ , on définit

$$\langle A, B \rangle := \sum_{1 \leq i, j \leq n} a_{ij} b_{ij} = \text{Tr}({}^t A B).$$

L'application  $\langle \cdot, \cdot \rangle : M_n(\mathbb{R}) \times M_n(\mathbb{R}) \rightarrow \mathbb{R}$   
 $(A, B) \mapsto \langle A, B \rangle$  est alors un produit scalaire sur  $M_n(\mathbb{R})$  : il s'agit du produit scalaire associé à la base canonique  $\{E_{ij}\}_{1 \leq i, j \leq n}$  de  $M_n(\mathbb{R})$ .

4. Soit  $n \in \mathbb{N} \setminus \{0\}$ . Pour tous polynômes  $P$  et  $Q$  de  $\mathbb{R}_n[X]$ , on définit

$$\langle P, Q \rangle := \int_0^1 P(t)Q(t)dt.$$

L'application  $\langle \cdot, \cdot \rangle : \mathbb{R}_n[X] \times \mathbb{R}_n[X] \rightarrow \mathbb{R}$   
 $(P, Q) \mapsto \langle P, Q \rangle$  est alors un produit scalaire sur  $\mathbb{R}_n[X]$ .

La bilinéarité de l'application  $\langle \cdot, \cdot \rangle$  provient de la linéarité de l'intégrale. Montrons que  $\langle \cdot, \cdot \rangle$  est bien définie positive. Soit  $P \in \mathbb{R}_n[X]$ , on a  $\langle P, P \rangle = \int_0^1 (P(t))^2 dt \geq 0$  (l'intégrale sur un segment d'une fonction positive est positive). Si  $\langle P, P \rangle = \int_0^1 (P(t))^2 dt = 0$ , comme la fonction  $\mathbb{R} \rightarrow \mathbb{R} ; t \mapsto P(t)^2$  est continue et positive, on a, pour tout  $t \in [0, 1]$ ,  $(P(t))^2 = 0$  (l'intégrale d'une fonction continue et positive sur un segment est nulle si et seulement si la fonction est identiquement nulle sur ce segment) et donc, pour tout  $t \in [0, 1]$ ,  $P(t) = 0$ , donc (un polynôme ayant une infinité de racines étant nécessairement nul) le polynôme  $P$  est nul.

*Remarque 2.2.3.* Si  $\langle \cdot, \cdot \rangle$  est un produit scalaire sur  $E$  et si  $F$  est un sous-espace vectoriel de  $E$ , la restriction

$$\langle \cdot, \cdot \rangle|_{F \times F} : F \times F \rightarrow \mathbb{R}$$

$$(v, w) \mapsto \langle v, w \rangle$$

est un produit scalaire sur  $F$ .

**Définition 2.2.4.** Si  $E$  est un espace vectoriel de dimension finie muni d'un produit scalaire  $\langle \cdot, \cdot \rangle$ , le couple  $(E, \langle \cdot, \cdot \rangle)$  est appelé espace euclidien.



*Remarque 2.2.5.* D'après la remarque 2.2.3, si  $(E, \langle \cdot, \cdot \rangle)$  est un espace euclidien et si  $F$  est un sous-espace vectoriel de  $E$ , alors  $(F, \langle \cdot, \cdot \rangle|_{F \times F})$  est également un espace euclidien. On le notera simplement  $(F, \langle \cdot, \cdot \rangle)$ .

Supposons donc dans la suite de cette section que  $E$  est un espace vectoriel de dimension finie muni d'un produit scalaire  $\langle \cdot, \cdot \rangle$ . Pour tout vecteur  $v$  de  $E$ ,  $\langle v, v \rangle \geq 0$  et on définit alors  $\|v\| := \sqrt{\langle v, v \rangle}$ . Une première propriété importante des espaces euclidiens est l'inégalité de Cauchy-Schwarz ci-dessous. Cette inégalité permet en particulier de montrer que l'application qui à tout vecteur  $v$  de  $E$  associe  $\|v\| \in [0, +\infty[$  est une norme.

**Lemme 2.2.6** (Inégalité de Cauchy-Schwarz). *Pour tous vecteurs  $v$  et  $w$  de  $E$ , on a l'inégalité*

$$|\langle v, w \rangle| \leq \|v\| \|w\|,$$

et  $|\langle v, w \rangle| = \|v\| \|w\|$  si et seulement si les vecteurs  $v$  et  $w$  sont liés.

*Démonstration.* Soient  $v, w \in E$ .

Si  $w$  est le vecteur nul  $0_E$  de  $E$ , on a

$$\langle v, w \rangle = \langle v, 0_E \rangle = \langle v, 0 \cdot 0_E \rangle = 0 \cdot \langle v, 0_E \rangle = 0$$

et  $\|w\| = \sqrt{\langle w, w \rangle} = \sqrt{\langle 0_E, 0_E \rangle} = 0$ . L'inégalité ci-dessus est donc vérifiée : il s'agit d'une égalité et on a  $w = 0_E = 0 \cdot v$ .

On suppose maintenant que  $w \neq 0_E$ . Soit alors  $\lambda \in \mathbb{R}$ , on a  $\|v + \lambda w\|^2 \geq 0$ . Or

$$\begin{aligned} \|v + \lambda w\|^2 &= \langle v + \lambda w, v + \lambda w \rangle \\ &= \langle v, v \rangle + \lambda \langle v, w \rangle + \lambda \langle w, v \rangle + \lambda^2 \langle w, w \rangle \\ &= \|v\|^2 + 2\lambda \langle v, w \rangle + \lambda^2 \|w\|^2. \end{aligned}$$

Ainsi, pour tout  $\lambda \in \mathbb{R}$ ,  $\|w\|^2 \lambda^2 + 2\langle v, w \rangle \lambda + \|v\|^2 \geq 0$ , en d'autres termes, la fonction polynomiale du second degré (remarquons que  $\|w\|^2 \neq 0$  car  $\langle w, w \rangle \neq 0$  car  $w \neq 0_E$ )

$$\begin{array}{ccc} \mathbb{R} & \rightarrow & \mathbb{R} \\ \lambda & \mapsto & \|w\|^2 \lambda^2 + 2\langle v, w \rangle \lambda + \|v\|^2 \end{array}$$

est positive sur tout  $\mathbb{R}$ , ce qui est équivalent au fait que le discriminant associé  $4\langle v, w \rangle^2 - 4\|w\|^2 \|v\|^2$  soit négatif ou nul. Ainsi, on a  $\langle v, w \rangle^2 \leq \|v\|^2 \|w\|^2$  i.e.  $|\langle v, w \rangle| \leq \|v\| \|w\|$ .

De plus, si  $|\langle v, w \rangle| = \|v\| \|w\| \Leftrightarrow \langle v, w \rangle^2 = \|v\|^2 \|w\|^2$ , le discriminant associé à la fonction polynomiale du second degré ci-dessus est nul et donc le polynôme associé possède une racine (double)  $\lambda_0 \in \mathbb{R}$ . On a ainsi

$$\langle v + \lambda_0 w, v + \lambda_0 w \rangle = \|v + \lambda_0 w\|^2 = \|w\|^2 \lambda_0^2 + 2\langle v, w \rangle \lambda_0 + \|v\|^2 = 0$$

et donc  $v + \lambda_0 w = 0_E$ , en particulier les vecteurs  $v$  et  $w$  sont liés.

Réciproquement, supposons qu'il existe  $(\mu_1, \mu_2) \in \mathbb{R}^2 \setminus \{(0, 0)\}$  tel que  $\mu_1 v + \mu_2 w = 0$ . Si  $\mu_2 = 0$ , nécessairement  $v = 0_E$  et, comme ci-dessus,  $|\langle v, w \rangle| = 0 = \|v\| \|w\|$ . Si  $\mu_2 \neq 0$ , on a  $w = \frac{\mu_1}{\mu_2} v$  et alors

$$\begin{aligned} \langle v, w \rangle^2 &= \left\langle v, \frac{\mu_1}{\mu_2} v \right\rangle^2 \\ &= \left( \frac{\mu_1}{\mu_2} \right)^2 \langle v, v \rangle^2 \\ &= \langle v, v \rangle \left\langle \frac{\mu_1}{\mu_2} v, \frac{\mu_1}{\mu_2} v \right\rangle \\ &= \|v\|^2 \|w\|^2. \end{aligned}$$

□

**Corollaire et Définition 2.2.7.** L'application  $\| \cdot \| : \begin{array}{l} E \rightarrow [0, +\infty[ \\ v \mapsto \|v\| \end{array}$  est une norme, i.e.

1. pour tous  $v \in E$  et  $\lambda \in \mathbb{R}$ ,  $\|\lambda v\| = |\lambda| \|v\|$ ,
2. pour tout  $v \in E$ ,  $\|v\| = 0$  si et seulement si  $v = 0_E$ ,
3. pour tous  $v, w \in E$ ,  $\|v + w\| \leq \|v\| + \|w\|$ .

En conséquence, le couple  $(E, \| \cdot \|)$  est un espace vectoriel normé. La norme  $\| \cdot \|$  est appelée norme euclidienne associée au produit scalaire  $\langle \cdot, \cdot \rangle$ .

*Démonstration.* 1. Soient  $v \in E$  et  $\lambda \in \mathbb{R}$ , alors

$$\|\lambda v\| = \sqrt{\langle \lambda v, \lambda v \rangle} = \sqrt{\lambda^2} \sqrt{\langle v, v \rangle} = |\lambda| \|v\|.$$

2. Soit  $v \in E$ , alors  $\|v\| = \sqrt{\langle v, v \rangle} = 0$  ssi  $\langle v, v \rangle = 0$  ssi  $v = 0_E$ .

3. Soient  $v, w \in E$ . Alors

$$\begin{aligned} \|v + w\|^2 &= \langle v + w, v + w \rangle \\ &= \langle v, v \rangle + \langle v, w \rangle + \langle w, v \rangle + \langle w, w \rangle \\ &= \|v\|^2 + 2\langle v, w \rangle + \|w\|^2 \\ &\leq \|v\|^2 + 2|\langle v, w \rangle| + \|w\|^2 \\ &\leq \|v\|^2 + 2\|v\|^2 \|w\|^2 + \|w\|^2 \quad (\text{par l'inégalité de Cauchy-Schwarz}) \\ &= (\|v\| + \|w\|)^2 \end{aligned}$$

et donc  $\|v + w\| \leq \|v\| + \|w\|$ .

□

*Remarque 2.2.8.* D'après les premières égalités ci-dessus, on a, pour tous  $v, w \in E$ ,

$$\langle v, w \rangle = \frac{1}{2} (\|v + w\|^2 - \|v\|^2 - \|w\|^2).$$

## 2.3 Orthogonalité dans les espaces euclidiens

Soit  $(E, \langle \cdot, \cdot \rangle)$  un espace euclidien.

La structure supplémentaire qu'apporte le produit scalaire  $\langle \cdot, \cdot \rangle$  sur l'espace vectoriel de dimension finie  $E$  nous permet d'introduire une notion d'orthogonalité :

**Définition 2.3.1.** Soient  $v$  et  $w$  deux vecteurs de  $E$ . On dit que  $v$  et  $w$  sont orthogonaux si  $\langle v, w \rangle = 0$ . Dans ce cas, on dira également que  $v$  est orthogonal à  $w$  et que  $w$  est orthogonal à  $v$ .

Soit maintenant  $A$  un sous-ensemble non vide de  $E$ . On note  $A^\perp$  l'ensemble des vecteurs de  $E$  orthogonaux à tous les vecteurs contenus dans  $A$ , i.e.

$$A^\perp := \{v \in E \mid \text{pour tout } w \in A, \langle v, w \rangle = 0\}.$$

On appelle  $A^\perp$  l'orthogonal de  $A$  (par rapport à  $\langle \cdot, \cdot \rangle$ ).

*Remarque 2.3.2.* Une famille finie  $\{v_1, \dots, v_p\}$  de vecteurs non nuls de  $E$  deux à deux orthogonaux (i.e., pour tous  $i, j \in \{1, \dots, p\}$ , si  $i \neq j$  alors  $\langle v_i, v_j \rangle = 0$ ) est libre. En effet, soient  $\lambda_1, \dots, \lambda_p \in \mathbb{R}$  tels que  $\lambda_1 v_1 + \dots + \lambda_p v_p = 0_E$ , alors, pour  $i \in \{1, \dots, p\}$ ,

$$0 = \left\langle v_i, \sum_{j=1}^p \lambda_j v_j \right\rangle = \sum_{j=1}^p \lambda_j \langle v_i, v_j \rangle = \lambda_i \langle v_i, v_i \rangle$$

et donc, comme  $\langle v_i, v_i \rangle \neq 0$  (car  $v_i \neq 0_E$ ),  $\lambda_i = 0$ .

Une famille finie de vecteurs non nuls de  $E$  deux à deux orthogonaux sera appelée famille orthogonale. Etant libre, une famille orthogonale possède au plus  $\dim(E)$  éléments.

*Exemple 2.3.3.* Dans  $\mathbb{R}^3$  muni du produit scalaire canonique, les vecteurs  $(1, -1, 2)$  et  $(1, 3, 1)$  sont orthogonaux, et  $\{(-2, 5, 3)\}^\perp = \{(x, y, z) \in \mathbb{R}^3 \mid \langle (x, y, z), (-2, 5, 3) \rangle_{\text{can}} = -2x + 5y + 3z = 0\}$ .

Dans l'exemple ci-dessus, on peut remarquer que  $\{(-2, 5, 3)\}^\perp$  est un sous-espace vectoriel de  $\mathbb{R}^3$ . L'orthogonal d'un sous-ensemble est en fait toujours un sous-espace vectoriel :

**Lemme 2.3.4.** Soit  $A$  un sous-ensemble de  $E$ . Alors  $A^\perp$  est un sous-espace vectoriel de  $E$  (même si  $A$  ne l'est pas).

*Démonstration.*  $0_E \in A^\perp$  car, pour tout  $w \in A$ ,  $\langle 0_E, w \rangle = 0$ , et, si  $v_1, v_2 \in A^\perp$  et  $\lambda, \mu \in \mathbb{R}$ , on a, pour tout  $w \in A$ ,

$$\langle \lambda v_1 + \mu v_2, w \rangle = \lambda \langle v_1, w \rangle + \mu \langle v_2, w \rangle = 0.$$

Ainsi,  $A^\perp$  est bien un sous-espace vectoriel de  $E$ . □

*Remarque 2.3.5.* • Si  $A$  est un sous-ensemble de  $E$ , on a  $A^\perp = (\text{Vect}(A))^\perp$ .

- On a  $\{0_E\}^\perp = E$  (car, pour tout  $v \in E$ ,  $\langle v, 0_E \rangle = 0$ ) et  $E^\perp = \{0_E\}$  (en effet, si  $v \in E^\perp$ ,  $\langle v, v \rangle = 0$ , car  $v \in E^\perp$  et  $v \in E$ , et donc  $v = 0_E$ ).

Remarquons ensuite que le théorème de Pythagore peut être étendu à tout cadre euclidien général :

**Lemme 2.3.6** (Théorème de Pythagore). *Soient  $v$  et  $w$  deux vecteurs de  $E$ . Alors  $v$  et  $w$  sont orthogonaux si et seulement si  $\|v + w\|^2 = \|v\|^2 + \|w\|^2$ .*

*Démonstration.* Dans la section précédente, on a vu que  $\|v + w\|^2 = \|v\|^2 + 2\langle v, w \rangle + \|w\|^2$  et donc  $\|v + w\|^2 = \|v\|^2 + \|w\|^2$  ssi  $\langle v, w \rangle = 0$  ssi  $v$  et  $w$  sont orthogonaux.  $\square$

Revenons maintenant à l'orthogonal d'un sous-ensemble de  $E$ . Si  $F$  est un sous-espace vectoriel de  $E$ , on va pouvoir décomposer  $E$  en la somme directe de  $F$  et de son orthogonal  $F^\perp$ . On montrera l'égalité  $\dim(E) = \dim(F) + \dim(F^\perp)$  dans la section suivante, qui exhibera un lien important entre orthogonalité et dualité pour un espace euclidien.

**Proposition 2.3.7.** *Soit  $F$  un sous-espace vectoriel de  $E$ . On a :*

1.  $\dim(E) = \dim(F) + \dim(F^\perp)$ ,
2.  $E = F \oplus F^\perp$ ,
3.  $(F^\perp)^\perp = F$ .

*Démonstration.* 1. Il s'agit du corollaire 2.4.4 de la section suivante.

2. On a  $F \cap F^\perp = \{0_E\}$  car, si  $v \in F \cap F^\perp$ ,  $\langle v, v \rangle = 0$  (car  $v \in F^\perp$  et  $v \in F$ ) et donc  $v = 0_E$ . Ainsi,  $F$  et  $F^\perp$  sont en somme directe et, comme  $\dim(E) = \dim(F) + \dim(F^\perp)$  par 1.,  $E = F \oplus F^\perp$ .

3. On a  $F \subset (F^\perp)^\perp$  car, si  $v \in F$ , pour tout  $w \in F^\perp$ ,  $\langle v, w \rangle = \langle w, v \rangle = 0$ . De plus,

$$\dim\left((F^\perp)^\perp\right) = \dim(E) - \dim(F^\perp) = \dim(E) - (\dim(E) - \dim(F)) = \dim(F).$$

En conséquence,  $F = (F^\perp)^\perp$ .  $\square$

Soit  $F$  un sous-espace vectoriel de  $E$ . Comme  $E$  se décompose en la somme directe de  $F$  et son orthogonal  $F^\perp$ , on peut considérer la projection de  $E$  sur  $F$  parallèlement à  $F^\perp$  :

**Définition 2.3.8.** *On appelle projection orthogonale sur  $F$  la projection de  $E$  sur  $F$  parallèlement à  $F^\perp$ . Il s'agit de l'application linéaire surjective qui à tout vecteur  $v = w + u$  de  $E$  avec  $w \in F$  et  $u \in F^\perp$  associe sa composante  $w$  dans  $F$ . On note  $p_F$  cette application.*

*Remarque 2.3.9.* Si  $\{e_1, \dots, e_p\}$  est une base de  $F$  et  $\{e_{p+1}, \dots, e_n\}$  est une base de  $F^\perp$  alors  $\mathcal{B} := \{e_1, \dots, e_p, e_{p+1}, \dots, e_n\}$  est une base de  $E$  dans laquelle la matrice représentative de  $p_F$  est

$$\text{Mat}_{\mathcal{B}}(p_F) = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix} = \begin{pmatrix} I_p & 0_{p, n-p} \\ 0_{n-p, p} & 0_{n-p, n-p} \end{pmatrix}$$

La projection orthogonale permet notamment de calculer explicitement la distance euclidienne d'un vecteur de  $E$  à un sous-espace vectoriel de  $E$ . On donne tout d'abord la définition de cette notion de distance euclidienne :

**Définition 2.3.10.** *Si  $v$  et  $w$  sont deux vecteurs de  $E$ , on définit la distance euclidienne de  $v$  à  $w$  comme étant le réel positif ou nul  $d(v, w) := \|w - v\|$ .*

*Si  $v$  est un vecteur de  $E$  et  $A$  un sous-ensemble non vide de  $E$ , on définit également la distance euclidienne de  $v$  à  $A$  comme étant le réel positif ou nul  $d(v, A) := \inf_{w \in A} d(v, w) = \inf_{w \in A} \|w - v\|$ .*

*Remarque 2.3.11.* • La distance euclidienne d'un vecteur de  $E$  à un autre est bien une distance, dans le sens où

1. pour tous  $v, w \in E$ ,  $d(v, w) \geq 0$ ,
2. pour tous  $v, w \in E$ ,  $d(v, w) = d(w, v)$  (en effet  $\|w - v\| = \|v - w\|$ ),
3. pour tous  $v, w \in E$ ,  $d(v, w) = 0$  ssi  $v = w$  (en effet,  $\|w - v\| = 0$  ssi  $w - v = 0_E$ ),
4. pour tous  $v, w, u \in E$ ,  $d(v, u) \leq d(v, w) + d(w, u)$  (en effet,  $\|u - v\| = \|u - w + w - v\| \leq \|u - w\| + \|w - v\|$ ).

Plus généralement, toute norme sur un espace vectoriel induit, de cette façon, une distance.

- Pour  $v$  un vecteur de  $E$  et  $A$  un sous-ensemble non vide de  $E$ , la borne inférieure de l'ensemble  $\{d(v, w) \mid w \in A\}$  existe bien et est positive ou nulle : ce sous-ensemble de  $\mathbb{R}$  est non vide et minorée par 0.

**Proposition 2.3.12.** *Soit  $v \in E$ . On a*

$$d(v, F) = \|v - p_F(v)\|.$$

*Démonstration.* Soit  $w \in F$ . On a

$$\|v - w\|^2 = \|v - p_F(v) + p_F(v) - w\|^2 = \|v - p_F(v)\|^2 + \|p_F(v) - w\|^2$$

par le théorème de Pythagore (lemme 2.3.6) car  $v - p_F(v) \in F^\perp$  (par définition de la projection orthogonale : il existe un unique vecteur  $u \in F^\perp$  tel que  $v = p_F(v) + u$ ) et  $p_F(v) - w \in F$  (car  $F$  est un sous-espace vectoriel de  $E$ ).

Ainsi, pour tout  $w \in F$ ,  $\|v - w\|^2 \geq \|v - p_F(v)\|^2$  donc  $\|v - w\| \geq \|v - p_F(v)\|$  donc  $\inf \{\|v - w\| \mid w \in F\} \geq \|v - p_F(v)\|$ . De plus, comme  $p_F(v) \in F$ ,  $\inf \{\|v - w\| \mid w \in F\} \leq \|v - p_F(v)\|$ . Au total,  $\|v - p_F(v)\| = \inf \{\|v - w\| \mid w \in F\} = d(v, F)$ .  $\square$

On définit également la symétrie orthogonale par rapport à  $F$  :

**Définition 2.3.13.** *On appelle symétrie orthogonale par rapport à  $F$  la symétrie par rapport à  $F$  parallèlement à  $F^\perp$ . Il s'agit de l'involution linéaire de  $E$  qui à tout vecteur  $v = w + u$  de  $E$  avec  $w \in F$  et  $u \in F^\perp$  associe le vecteur  $w - u$ . On note  $s_F$  cette application.*

*Remarque 2.3.14.* • “Involution linéaire de  $E$ ” signifie que  $s_F$  est une application linéaire de  $E$  dans  $E$  telle que  $s_F \circ s_F = \text{Id}_E$ . Autrement dit,  $s_F$  est un automorphisme linéaire de  $E$  d’inverse lui-même.

- Comme pour toute symétrie vectorielle, on a  $s_F = 2p_F - \text{Id}_E$  (si  $v = w + u \in E$  avec  $w \in F$  et  $u \in F^\perp$ , on a  $(2p_F - \text{Id}_E)(v) = 2w - (w + u) = w - u = s_F(v)$ ).
- Si  $\{e_1, \dots, e_p\}$  est une base de  $F$  et  $\{e_{p+1}, \dots, e_n\}$  est une base de  $F^\perp$  alors  $\mathcal{B} := \{e_1, \dots, e_p, e_{p+1}, \dots, e_n\}$  est une base de  $E$  dans laquelle la matrice représentative de  $s_F$  est

$$\text{Mat}_{\mathcal{B}}(s_F) = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & -1 & & \\ & & & & \ddots & \\ & & & & & -1 \end{pmatrix} = \begin{pmatrix} I_p & 0_{p,n-p} \\ 0_{n-p,p} & -I_{n-p} \end{pmatrix}$$

## 2.4 Orthogonalité et dualité dans les espaces euclidiens

Effectuons un petit retour sur la dualité linéaire. Soit  $E$  un espace vectoriel de dimension finie sur  $\mathbb{R}$ . On a vu au chapitre précédent que, même si  $E$  et son dual  $E^*$  sont isomorphes, il n’existe en général pas d’isomorphisme canonique (i.e. ne dépendant pas d’un choix de bases) entre  $E$  et  $E^*$ . Mais si  $E$  est maintenant muni d’un produit scalaire  $\langle \cdot, \cdot \rangle$ , celui-ci permet de construire un tel isomorphisme canonique entre  $E$  et  $E^*$  :

**Théorème 2.4.1.** *Soit  $(E, \langle \cdot, \cdot \rangle)$  un espace euclidien. Pour tout  $v \in E$ , on définit l’application linéaire  $\Lambda_v : E \rightarrow \mathbb{R}$  ;  $w \mapsto \langle v, w \rangle$ . On définit ensuite l’application*

$$\Lambda : \begin{array}{l} E \rightarrow E^* \\ v \mapsto \Lambda_v \end{array}$$

L’application  $\Lambda$  est un isomorphisme linéaire.

*Démonstration.* Remarquons tout d’abord que, si  $v \in E$ ,  $\Lambda_v$  est bien une application linéaire de  $E$  dans  $\mathbb{R}$  (en d’autres termes  $\Lambda_v \in E^*$ ) car le produit scalaire  $\langle \cdot, \cdot \rangle$  est bilinéaire.

Montrons ensuite que l’application  $\Lambda$  est bien linéaire. Soient donc  $v_1, v_2 \in E$  et  $\lambda, \mu \in \mathbb{R}$ , alors, pour tout  $w \in E$ ,

$$\begin{aligned} \Lambda(\lambda v_1 + \mu v_2)(w) &= \Lambda_{\lambda v_1 + \mu v_2}(w) \\ &= \langle \lambda v_1 + \mu v_2, w \rangle \\ &= \lambda \langle v_1, w \rangle + \mu \langle v_2, w \rangle \\ &= \lambda \Lambda_{v_1}(w) + \mu \Lambda_{v_2}(w) \\ &= (\lambda \Lambda_{v_1} + \mu \Lambda_{v_2})(w) \end{aligned}$$

i.e.  $\Lambda(\lambda v_1 + \mu v_2) = \lambda \Lambda(v_1) + \mu \Lambda(v_2)$ .

Montrons enfin que  $\Lambda$  est injective : comme  $\dim(E) = \dim(E^*)$ , on obtiendra alors (par le théorème du rang) que  $\Lambda$  est bijective. Soit donc  $v \in E$  tel que  $\Lambda_v$  est l'application identiquement nulle. En particulier,  $\Lambda_v(v) = \langle v, v \rangle = 0$  donc  $v = 0_E$ . L'application  $\Lambda$  est donc bien injective, et  $\Lambda$  est donc un isomorphisme.  $\square$

*Remarque 2.4.2.* La surjectivité de  $\Lambda$  nous dit que pour toute forme linéaire  $\varphi \in E^*$ , il existe  $v \in E$  tel que, pour tout  $w \in E$ ,  $\varphi(w) = \langle v, w \rangle$ . De plus, comme  $\Lambda$  est bijective,  $v = \Lambda^{-1}(\varphi)$ .

Intéressant en soi, ce résultat nous permet également de mettre une évidence une correspondance entre les notions d'orthogonal et d'annulateur d'un sous-espace vectoriel d'un espace euclidien :

**Corollaire 2.4.3.** *Soient  $(E, \langle \cdot, \cdot \rangle)$  un espace euclidien et  $F$  un sous-espace vectoriel de  $E$ . On a  $\Lambda(F^\perp) = F^0$ .*

*Démonstration.* Montrons l'égalité équivalente  $\Lambda^{-1}(F^0) = F^\perp$ . On a

$$\begin{aligned} \Lambda^{-1}(F^0) &= \{v \in E \mid \Lambda(v) \in F^0\} \\ &= \{v \in E \mid \text{pour tout } w \in F, \langle v, w \rangle = 0\} \\ &= F^\perp \end{aligned}$$

$\square$

Soient  $(E, \langle \cdot, \cdot \rangle)$  un espace euclidien et  $F$  un sous-espace vectoriel de  $E$ . Le corollaire précédent affirme en particulier que  $\Lambda$  induit, par restriction, un isomorphisme linéaire  $F^\perp \rightarrow F^0$ , et donc que  $\dim(F^\perp) = \dim(F^0)$ . On obtient ainsi le résultat qui nous permet de terminer la preuve de la proposition 2.3.7 :

**Corollaire 2.4.4.** *On a  $\dim(E) = \dim(F) + \dim(F^\perp)$ .*

*Démonstration.* On a, en utilisant proposition 1.5.3 1.,

$$\dim(F^\perp) = \dim(F^0) = \dim(E) - \dim(F).$$

$\square$

## 2.5 Bases orthogonales et bases orthonormales

Soit  $(E, \langle \cdot, \cdot \rangle)$  un espace euclidien de dimension  $n \in \mathbb{N} \setminus \{0\}$  et soit  $\{v_1, \dots, v_n\}$  une base de  $E$ . Si  $v$  et  $w$  sont deux vecteurs de  $E$  de coordonnées respectives  $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$  et  $\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$  dans cette base, alors

$$\langle v, w \rangle = \left\langle \sum_{i=1}^n x_i v_i, \sum_{j=1}^n y_j v_j \right\rangle = \sum_{1 \leq i, j \leq n} x_i y_j \langle v_i, v_j \rangle.$$

On aimerait une base de  $E$  dans laquelle cette expression du produit scalaire  $\langle \cdot, \cdot \rangle$  sur  $E$  soit la plus simple possible. Cela nous mène à la notion de base orthogonale et de base orthonormale :

**Définition 2.5.1.** Soit  $\mathcal{B} = \{e_1, \dots, e_n\}$  une base de  $E$ .

1. On dit que  $\mathcal{B}$  est une base orthogonale de  $E$  si pour tous  $i, j \in \{1, \dots, n\}$  tels que  $i \neq j$ , on a  $\langle e_i, e_j \rangle = 0$ .
2. On dit que  $\mathcal{B}$  est une base orthonormale de  $E$  si  $\mathcal{B}$  est une base orthogonale de  $E$  et si tous les vecteurs de  $\mathcal{B}$  sont de norme euclidienne 1. Autrement dit,  $\mathcal{B}$  est une base orthonormale de  $E$  si et seulement si pour tous  $i, j \in \{1, \dots, n\}$ ,  $\langle e_i, e_j \rangle = \delta_{ij}$  (pour tout  $v \in E$ ,  $\|v\| = 1$  ssi  $\langle v, v \rangle = 1$ ).

*Exemple 2.5.2.* • La base canonique de  $\mathbb{R}^n$  est une base orthonormale pour le produit scalaire canonique sur  $\mathbb{R}^n$  (exemple 2.2.2 1.).

- Par définition, une base  $\mathcal{B}$  de  $E$  est une base orthonormale pour le produit scalaire associé  $\langle \cdot, \cdot \rangle_{\mathcal{B}}$  (exemple 2.2.2 2.).

*Remarque 2.5.3.* 1. Si  $\mathcal{B} = \{e_1, \dots, e_n\}$  est une base orthonormale pour  $\langle \cdot, \cdot \rangle$  et si  $v$  et  $w$  sont deux vecteurs de  $E$  de coordonnées respectives  $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$  et  $\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$  dans  $\mathcal{B}$ , alors

$$\langle v, w \rangle = \sum_{1 \leq i, j \leq n} x_i y_j \langle e_i, e_j \rangle = \sum_{i=1}^n x_i y_i.$$

En particulier, pour tout  $i \in \{1, \dots, n\}$ ,  $\langle v, e_i \rangle = x_i$  et donc

$$v = \sum_{i=1}^n \langle v, e_i \rangle e_i.$$

2. Si la famille  $\{\epsilon_1, \dots, \epsilon_n\}$  est une base orthogonale de  $E$ , alors la famille  $\{e_1, \dots, e_n\}$  avec, pour tout  $i \in \{1, \dots, n\}$ ,  $e_i := \frac{\epsilon_i}{\|\epsilon_i\|}$ , est une base orthonormale de  $E$ . En effet, pour tout  $i \in \{1, \dots, n\}$ ,  $\|e_i\| = \left\| \frac{\epsilon_i}{\|\epsilon_i\|} \right\| = \frac{1}{\|\epsilon_i\|} \|\epsilon_i\| = 1$ .
3. Plus généralement, on appellera famille orthonormale toute famille orthogonale  $\{v_1, \dots, v_p\}$  de  $E$  (remarque 2.3.2) telle que, pour tout  $i \in \{1, \dots, p\}$ ,  $\|v_i\| = 1$ .

Ainsi, avec les notations ci-dessus, si  $\{v_1, \dots, v_n\}$  est une base orthonormale de  $E$ , on a l'expression simple

$$\langle v, w \rangle = \sum_{i=1}^n x_i y_i.$$

du produit scalaire de  $v$  et  $w$ , qui rappelle l'expression du produit scalaire canonique sur  $\mathbb{R}^n$ .

Et un résultat fondamental de la théorie des espaces euclidiens est qu'il est toujours possible de construire, de façon algorithmique, une base orthonormale pour n'importe quel espace euclidien :



**Théorème 2.5.4** (Procédé d'orthonormalisation de Gram-Schmidt). *Soit  $\{v_1, \dots, v_p\}$  une famille libre de  $E$ . On peut construire, de façon algorithmique, une famille orthonormale  $\{e_1, \dots, e_p\}$  de  $E$  telle que, pour tout  $k \in \{1, \dots, p\}$ ,  $\text{Vect}\{e_1, \dots, e_k\} = \text{Vect}\{v_1, \dots, v_k\}$  (en particulier,  $\{e_1, \dots, e_p\}$  est donc une base orthonormale de  $\text{Vect}\{v_1, \dots, v_p\}$ ).*

*Démonstration.* On construit, de façon récursive, une base orthogonale  $\{\epsilon_1, \dots, \epsilon_p\}$  pour  $\text{Vect}\{v_1, \dots, v_p\}$ . On en déduit immédiatement une base orthonormale en “normalisant” les vecteurs obtenus (i.e. en les multipliant chacun par l'inverse de leur norme comme dans la remarque 2.5.3 2.).

Le procédé récursif est le suivant : pour  $1 \leq k \leq p-1$ , on suppose que l'on a déjà construit des vecteurs  $\epsilon_1, \dots, \epsilon_k \in E$  orthogonaux deux à deux tels que  $\text{Vect}\{\epsilon_1, \dots, \epsilon_k\} = \text{Vect}\{v_1, \dots, v_k\}$ . On recherche alors un vecteur  $\epsilon_{k+1}$  de  $E$  de la forme

$$\epsilon_{k+1} = v_{k+1} + \lambda_1 \epsilon_1 + \dots + \lambda_k \epsilon_k \in \text{Vect}\{\epsilon_1, \dots, \epsilon_k, v_{k+1}\} = \text{Vect}\{v_1, \dots, v_k, v_{k+1}\}$$

tel que, pour tout  $i \in \{1, \dots, k\}$ ,  $\langle \epsilon_{k+1}, \epsilon_i \rangle = 0$ .

Or, pour  $i \in \{1, \dots, k\}$ ,

$$\begin{aligned} \langle \epsilon_{k+1}, \epsilon_i \rangle = 0 &\Leftrightarrow \langle v_{k+1} + \lambda_1 \epsilon_1 + \dots + \lambda_k \epsilon_k, \epsilon_i \rangle = 0 \\ &\Leftrightarrow \langle v_{k+1}, \epsilon_i \rangle + \lambda_i \langle \epsilon_i, \epsilon_i \rangle = 0 \\ &\Leftrightarrow \lambda_i = -\frac{\langle v_{k+1}, \epsilon_i \rangle}{\|\epsilon_i\|^2} \end{aligned}$$

Ainsi, par construction, la famille  $\{\epsilon_1, \dots, \epsilon_{k+1}\}$  avec  $\epsilon_{k+1} = v_{k+1} - \sum_{i=1}^k \frac{\langle v_{k+1}, \epsilon_i \rangle}{\|\epsilon_i\|^2} \epsilon_i$  est orthogonale et engendre bien  $\text{Vect}\{v_1, \dots, v_k, v_{k+1}\}$  car  $\text{Vect}\{v_1, \dots, v_k, v_{k+1}\} = \text{Vect}\{\epsilon_1, \dots, \epsilon_k, v_{k+1}\} = \text{Vect}\{\epsilon_1, \dots, \epsilon_k, \epsilon_{k+1}\}$  (car  $\epsilon_{k+1} = v_{k+1} - \sum_{i=1}^k \frac{\langle v_{k+1}, \epsilon_i \rangle}{\|\epsilon_i\|^2} \epsilon_i \in \text{Vect}\{\epsilon_1, \dots, \epsilon_k, v_{k+1}\}$  et  $v_{k+1} = \epsilon_{k+1} + \sum_{i=1}^k \frac{\langle v_{k+1}, \epsilon_i \rangle}{\|\epsilon_i\|^2} \epsilon_i \in \text{Vect}\{\epsilon_1, \dots, \epsilon_k, \epsilon_{k+1}\}$ ).  $\square$

*Remarque 2.5.5.* Au terme de l'étape  $k+1$  du procédé ci-dessus, on peut remplacer  $\epsilon_{k+1}$  par n'importe quel vecteur non nul  $\epsilon'_{k+1}$  de la droite vectorielle engendrée par  $\epsilon_{k+1}$  : la famille  $\{\epsilon_1, \dots, \epsilon_k, \epsilon'_{k+1}\}$  ainsi obtenue reste orthogonale et engendre toujours  $\text{Vect}\{v_1, \dots, v_{k+1}\}$ . Cela peut être utile pour simplifier les calculs (notamment pour éviter de manipuler des fractions).

*Exemple 2.5.6.* On détermine une base orthonormée pour le sous-espace vectoriel  $F$  de  $\mathbb{R}^4$  engendré par les vecteurs  $v_1 = (1, 1, 0, 0)$ ,  $v_2 = (1, 0, -1, 1)$  et  $v_3 = (0, 1, 1, 1)$ . La famille  $\{v_1, v_2, v_3\}$  est libre (c'est donc une base de  $F = \text{Vect}\{v_1, v_2, v_3\}$ ) et on applique le procédé d'orthonormalisation de Gram-Schmidt pour déterminer une base orthonormale de  $F$ .

On pose

$$\begin{aligned} \epsilon_1 &:= v_1 = (1, 1, 0, 0) \\ \epsilon_2 &:= v_2 - \frac{\langle v_2, \epsilon_1 \rangle}{\|\epsilon_1\|^2} \epsilon_1 = v_2 - \frac{1}{2} \epsilon_1 = \left( \frac{1}{2}, -\frac{1}{2}, -1, 1 \right) = \frac{1}{2} (1, -1, -2, 2) \\ \epsilon'_2 &:= (1, -1, -2, 2) \text{ (pour simplifier les calculs)} \\ \epsilon_3 &:= v_3 - \frac{\langle v_3, \epsilon_1 \rangle}{\|\epsilon_1\|^2} \epsilon_1 - \frac{\langle v_3, \epsilon'_2 \rangle}{\|\epsilon'_2\|^2} \epsilon'_2 = v_3 - \frac{1}{2} \epsilon_1 + \frac{1}{10} \epsilon'_2 = \left( -\frac{2}{5}, \frac{2}{5}, \frac{4}{5}, \frac{6}{5} \right) = \frac{2}{5} (-1, 1, 2, 3) \\ \epsilon'_3 &:= (-1, 1, 2, 3) \end{aligned}$$

et la famille  $\{\epsilon_1, \epsilon'_2, \epsilon'_3\}$  est alors une base orthogonale de  $F$ . La famille

$$\left\{ \frac{1}{\sqrt{2}}(1, 1, 0, 0), \frac{1}{\sqrt{10}}(1, -1, -2, 2), \frac{1}{\sqrt{15}}(-1, 1, 2, 3) \right\}$$

ensuite obtenue par normalisation est une base orthonormale de  $F$ .

**Corollaire 2.5.7.** *Il existe une base orthonormale pour l'espace euclidien  $(E, \langle \cdot, \cdot \rangle)$*

*Démonstration.* Soit  $\{v_1, \dots, v_n\}$  une base de  $E$ . En particulier, la famille  $\{v_1, \dots, v_n\}$  est libre et, d'après le théorème 2.5.4, on peut construire une base orthonormale pour  $\text{Vect}\{v_1, \dots, v_n\} = E$ .  $\square$

*Remarque 2.5.8.* Soit  $\mathcal{B} = \{e_1, \dots, e_n\}$  une base orthonormale pour  $\langle \cdot, \cdot \rangle$ . Alors  $\langle \cdot, \cdot \rangle = \langle \cdot, \cdot \rangle_{\mathcal{B}}$  (exemple 2.2.2 2.). En effet, pour tous vecteurs  $v$  et  $w$  de  $E$ , de coordonnées respectives

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \text{ et } \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \text{ dans la base } \mathcal{B}, \text{ on a } \langle v, w \rangle = \sum_{i=1}^n x_i y_i = \langle v, w \rangle_{\mathcal{B}}.$$

Le choix d'une base orthonormale pour l'espace euclidien  $(E, \langle \cdot, \cdot \rangle)$  permet d'identifier  $E$  avec l'espace euclidien  $(\mathbb{R}^n, \langle \cdot, \cdot \rangle_{\text{can}})$  muni du produit scalaire canonique. Précisément :

**Corollaire 2.5.9.** *Il existe un isomorphisme (non canonique)  $\psi : E \rightarrow \mathbb{R}^n$  tel que, pour tous vecteurs  $v$  et  $w$  de  $E$ ,  $\langle \psi(v), \psi(w) \rangle_{\text{can}} = \langle v, w \rangle$ .*

*Démonstration.* Soit  $\mathcal{B} = \{e_1, \dots, e_n\}$  une base orthonormale pour  $\langle \cdot, \cdot \rangle$  et notons  $\mathcal{B}' = \{e'_1, \dots, e'_n\}$  la base canonique de  $\mathbb{R}^n$ . Notons ensuite  $\psi_{\mathcal{B}}$  l'application linéaire de  $E$  dans  $\mathbb{R}^n$  qui, pour tout  $i \in \{1, \dots, n\}$ , associe  $e'_i$  à  $e_i$ . Autrement dit,  $\psi_{\mathcal{B}}$  est l'application qui à tout vecteur  $v$  de  $E$  de

coordonnées  $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$  dans la base  $\mathcal{B}$  associe le vecteur  $(x_1, \dots, x_n)$  de  $\mathbb{R}^n$ . Il s'agit d'un isomor-

phisme linéaire et, si  $v$  et  $w$  sont deux vecteurs de  $E$  de coordonnées respectives  $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$  et  $\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$

dans  $\mathcal{B}$ , on a

$$\langle \psi_{\mathcal{B}}(v), \psi_{\mathcal{B}}(w) \rangle_{\text{can}} = \langle (x_1, \dots, x_n), (y_1, \dots, y_n) \rangle_{\text{can}} = \sum_{i=1}^n x_i y_i = \langle v, w \rangle$$

(car  $\mathcal{B}$  est une base orthonormale pour  $\langle \cdot, \cdot \rangle$ ).  $\square$

## 2.6 Représentation matricielle du produit scalaire

Afin d'aider aux calculs, on cherche à représenter le produit scalaire de façon matricielle.

Soit  $(E, \langle \cdot, \cdot \rangle)$  un espace euclidien. Le point de départ pour définir une telle représentation est le suivant. Supposons que  $E$  soit de dimension  $n \in \mathbb{N} \setminus \{0\}$  et soit  $\mathcal{B} = \{e_1, \dots, e_n\}$  une base (quelconque) de  $E$ . On a vu au début de la section précédente que, si  $v$  et  $w$  sont deux vecteurs

de  $E$  de coordonnées respectives  $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$  et  $\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$  dans la base  $\mathcal{B}$ , alors

$$\langle v, w \rangle = \sum_{1 \leq i, j \leq n} x_i y_j \langle e_i, e_j \rangle.$$

Ecrivons  $X := \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \text{Mat}_{\mathcal{B}}(v)$ ,  $Y := \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \text{Mat}_{\mathcal{B}}(w)$  et  $A$  la matrice  $(\langle e_i, e_j \rangle)_{1 \leq i, j \leq n}$ .

On a alors

$$\begin{aligned} \langle v, w \rangle &= \sum_{1 \leq i, j \leq n} x_i y_j \langle e_i, e_j \rangle \\ &= \sum_{1 \leq i, j \leq n} x_i \langle e_i, e_j \rangle y_j \\ &= \sum_{i=1}^n x_i \left( \sum_{j=1}^n \langle e_i, e_j \rangle y_j \right) \\ &= \sum_{i=1}^n x_i (AY)_i \text{ (où } (AY)_i \text{ désigne la } i^{\text{ème}} \text{ coordonnée du vecteur colonne } AY) \\ &= {}^t X AY \end{aligned}$$

Cette écriture motive la définition suivante :

**Définition 2.6.1.** On appelle matrice représentative du produit scalaire  $\langle \cdot, \cdot \rangle$  dans la base  $\mathcal{B}$  la matrice

$$\text{Mat}_{\mathcal{B}}^{\text{ps}}(\langle \cdot, \cdot \rangle) := (\langle e_i, e_j \rangle)_{1 \leq i, j \leq n} = \begin{pmatrix} \langle e_1, e_1 \rangle & \cdots & \langle e_1, e_n \rangle \\ \vdots & & \vdots \\ \langle e_n, e_1 \rangle & \cdots & \langle e_n, e_n \rangle \end{pmatrix}$$

*Exemple 2.6.2.* • Si  $\mathcal{B}_{\text{can}}$  désigne la base canonique de  $\mathbb{R}^n$ , on a  $\text{Mat}_{\mathcal{B}_{\text{can}}}^{\text{ps}}(\langle \cdot, \cdot \rangle_{\text{can}}) = I_n$ .

• On a  $\text{Mat}_{\mathcal{B}}^{\text{ps}}(\langle \cdot, \cdot \rangle_{\mathcal{B}}) = I_n$ .

- Considérons sur l'espace vectoriel  $\mathbb{R}_2[X]$  de dimension 3 le produit scalaire  $\langle \cdot, \cdot \rangle$  défini dans l'exemple 2.2.2 4.. On note  $\mathcal{B}$  la base  $\{1, X, X^2\}$  de  $\mathbb{R}_2[X]$  et on calcule alors

$$\begin{aligned} \langle 1, 1 \rangle &= \int_0^1 1 \, dt = [t]_0^1 = 1 \\ \langle 1, X \rangle = \langle X, 1 \rangle &= \int_0^1 t \, dt = \left[ \frac{t^2}{2} \right]_0^1 = \frac{1}{2} \\ \langle 1, X^2 \rangle = \langle X^2, 1 \rangle &= \int_0^1 t^2 \, dt = \left[ \frac{t^3}{3} \right]_0^1 = \frac{1}{3} \\ \langle X, X \rangle &= \int_0^1 t^2 \, dt = \left[ \frac{t^3}{3} \right]_0^1 = \frac{1}{3} \\ \langle X, X^2 \rangle = \langle X^2, X \rangle &= \int_0^1 t^3 \, dt = \left[ \frac{t^4}{4} \right]_0^1 = \frac{1}{4} \\ \langle X^2, X^2 \rangle &= \int_0^1 t^4 \, dt = \left[ \frac{t^5}{5} \right]_0^1 = \frac{1}{5} \end{aligned}$$

Ainsi,

$$\text{Mat}_{\mathcal{B}}^{\text{ps}}(\langle \cdot, \cdot \rangle) = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \end{pmatrix}$$

*Remarque 2.6.3.* • Attention aux confusions avec la matrice représentative d'une application linéaire !

- Avec les notations du début de cette section, on a

$$\langle v, w \rangle = {}^t \text{Mat}_{\mathcal{B}}(v) \text{Mat}_{\mathcal{B}}^{\text{ps}}(\langle \cdot, \cdot \rangle) \text{Mat}_{\mathcal{B}}(w),$$

qui est aussi égal à  ${}^t \text{Mat}_{\mathcal{B}}(w) \text{Mat}_{\mathcal{B}}^{\text{ps}}(\langle \cdot, \cdot \rangle) \text{Mat}_{\mathcal{B}}(v) = \langle w, v \rangle$  car  $\langle \cdot, \cdot \rangle$  est symétrique.

- Comme le produit scalaire est symétrique, on a, pour tous  $i, j \in \{1, \dots, n\}$ ,  $\langle e_i, e_j \rangle = \langle e_j, e_i \rangle$  et la matrice  $\text{Mat}_{\mathcal{B}}^{\text{ps}}(\langle \cdot, \cdot \rangle)$  est donc symétrique i.e.  ${}^t \text{Mat}_{\mathcal{B}}^{\text{ps}}(\langle \cdot, \cdot \rangle) = \text{Mat}_{\mathcal{B}}^{\text{ps}}(\langle \cdot, \cdot \rangle)$ .

Le fait que le produit scalaire  $\langle \cdot, \cdot \rangle$  soit "défini" s'exprime dans le fait que  $\text{Mat}_{\mathcal{B}}^{\text{ps}}(\langle \cdot, \cdot \rangle)$  est inversible. En effet, si on note  $A := \text{Mat}_{\mathcal{B}}^{\text{ps}}(\langle \cdot, \cdot \rangle)$  et si  $X$  est un vecteur colonne de taille  $n$  quelconque tel que  $AX$  est le vecteur colonne nul de taille  $n$ , alors  ${}^t XAX = 0$  et donc  $\langle v, v \rangle = 0$  où  $v$  désigne le vecteur de coordonnées  $X$  dans la base  $\mathcal{B}$ . Par suite,  $v = 0_E$  et  $X$  est donc le vecteur colonne nul. Comme le noyau de la matrice carrée  $A$  est réduit au vecteur colonne nul,  $A$  est inversible.

- La base  $\mathcal{B}$  est orthonormale par rapport au produit scalaire  $\langle \cdot, \cdot \rangle$  si et seulement si  $\text{Mat}_{\mathcal{B}}^{\text{ps}}(\langle \cdot, \cdot \rangle) = I_n$ .

Ainsi, le corollaire 2.5.7 affirme qu'il existe toujours une base dans laquelle la matrice représentative du produit scalaire  $\langle \cdot, \cdot \rangle$  est la matrice identité.

On peut à présent se demander comment sont reliées les matrices représentatives de  $\langle \cdot, \cdot \rangle$  dans deux bases (quelconques) différentes, autrement dit s'intéresser à la question du changement de base pour la matrice représentative d'un produit scalaire.

Soit donc  $\mathcal{B}'$  une autre base de  $E$  et considérons la matrice de passage  $P_{\mathcal{B} \rightarrow \mathcal{B}'}$  de la base  $\mathcal{B}$  à la base  $\mathcal{B}'$ . On a l'égalité suivante :

**Proposition 2.6.4.** *On a*

$$\text{Mat}_{\mathcal{B}'}^{\text{ps}}(\langle \cdot, \cdot \rangle) = {}^t P_{\mathcal{B} \rightarrow \mathcal{B}'} \text{Mat}_{\mathcal{B}}^{\text{ps}}(\langle \cdot, \cdot \rangle) P_{\mathcal{B} \rightarrow \mathcal{B}'}$$

*Démonstration.* Soient  $v, w \in E$ . Comme au début de la section, notons  $X := \text{Mat}_{\mathcal{B}}(v)$ ,  $Y := \text{Mat}_{\mathcal{B}}(w)$  et  $A := \text{Mat}_{\mathcal{B}}^{\text{ps}}(\langle \cdot, \cdot \rangle)$ . Notons ensuite  $X' := \text{Mat}_{\mathcal{B}'}(v)$ ,  $Y' := \text{Mat}_{\mathcal{B}'}(w)$  et  $A' := \text{Mat}_{\mathcal{B}'}^{\text{ps}}(\langle \cdot, \cdot \rangle)$ .

On a  $X' = P_{\mathcal{B}' \rightarrow \mathcal{B}} X = P_{\mathcal{B} \rightarrow \mathcal{B}'}^{-1} X \Leftrightarrow X = P_{\mathcal{B} \rightarrow \mathcal{B}'} X'$  et  $Y' = P_{\mathcal{B}' \rightarrow \mathcal{B}} Y = P_{\mathcal{B} \rightarrow \mathcal{B}'}^{-1} Y \Leftrightarrow Y = P_{\mathcal{B} \rightarrow \mathcal{B}'} Y'$ . Ainsi,

$$\begin{aligned} \langle v, w \rangle &= {}^t X A Y \\ &= {}^t (P_{\mathcal{B} \rightarrow \mathcal{B}'} X') A (P_{\mathcal{B} \rightarrow \mathcal{B}'} Y') \\ &= {}^t X' ({}^t P_{\mathcal{B} \rightarrow \mathcal{B}'} A P_{\mathcal{B} \rightarrow \mathcal{B}'}) Y'. \end{aligned}$$

Remarquons maintenant que, si  $M = (m_{ij})_{1 \leq i, j \leq n}$  est une matrice carrée de taille  $n$  quelconque et si, pour  $i \in \{1, \dots, n\}$ ,  $X_i$  désigne le vecteur colonne avec coordonnées 1 à la ligne  $i$  et 0 sur les autres lignes, on a, pour tous  $i, j \in \{1, \dots, n\}$ ,  ${}^t X_i M X_j = m_{ij}$ .

Soient alors  $i, j \in \{1, \dots, n\}$  et notons  $\mathcal{B}' = \{e'_1, \dots, e'_n\}$ . On a  $\text{Mat}_{\mathcal{B}'}(e'_i) = X_i$  et  $\text{Mat}_{\mathcal{B}'}(e'_j) = X_j$  et, par l'égalité ci-dessus,

$$\langle e'_i, e'_j \rangle = {}^t X_i ({}^t P_{\mathcal{B} \rightarrow \mathcal{B}'} A P_{\mathcal{B} \rightarrow \mathcal{B}'}) X_j.$$

Autrement dit, le coefficient à la ligne  $i$  et la colonne  $j$  de la matrice  $A'$  est égal au coefficient à la ligne  $i$  et la colonne  $j$  de la matrice  ${}^t P_{\mathcal{B} \rightarrow \mathcal{B}'} A P_{\mathcal{B} \rightarrow \mathcal{B}'}$ . Par conséquent,  $A' = {}^t P_{\mathcal{B} \rightarrow \mathcal{B}'} A P_{\mathcal{B} \rightarrow \mathcal{B}'}$ .  $\square$

*Remarque 2.6.5.* Attention à ne surtout pas confondre ce changement de base pour les produits scalaires avec le changement de base pour les applications linéaires.

## 2.7 Endomorphisme adjoint

Soit  $(E, \langle \cdot, \cdot \rangle)$  un espace euclidien et soit  $f$  un endomorphisme de  $E$ .

**Proposition et Définition 2.7.1.** *Il existe une et une seule application  $f^* : E \rightarrow E$  tel que, pour tous vecteurs  $v$  et  $w$  de  $E$ ,*

$$\langle f(v), w \rangle = \langle v, f^*(w) \rangle.$$

*De plus,  $f^*$  est une application linéaire (et donc un endomorphisme de  $E$ ). On appelle  $f^*$  l'endomorphisme adjoint de  $f$ .*

*Démonstration.* On montre l'existence de  $f^*$  en utilisant l'isomorphisme canonique entre  $E$  et  $E^*$  du théorème 2.4.1. Soit  $w \in E$ , on considère l'application

$$\varphi_w : \begin{array}{ccc} E & \rightarrow & \mathbb{R} \\ v & \mapsto & \langle f(v), w \rangle \end{array}$$

Cette application est linéaire car  $f$  l'est et  $\varphi_w$  est donc un élément du dual  $E^*$  de  $E$ . D'après le théorème 2.4.1, il existe donc un unique vecteur  $f^*(w)$  de  $E$  tel que  $\varphi_w = \Lambda(f^*(w))$  i.e. tel que, pour tout  $v \in E$ ,

$$\langle f(v), w \rangle = \varphi_w(v) = \Lambda(f^*(w))(v) = \langle f^*(w), v \rangle = \langle v, f^*(w) \rangle.$$

On note alors  $f^*$  l'application qui à tout vecteur  $w$  de  $E$  associe  $f^*(w)$ . L'unicité évoquée ci-dessus montre l'unicité de l'application  $f^* : E \rightarrow E$  telle que, pour tous  $v, w \in E$ ,  $\langle f(v), w \rangle = \langle v, f^*(w) \rangle$ .

Montrons à présent sa linéarité : soient  $w_1, w_2 \in E$  et  $\lambda, \mu \in \mathbb{R}$  alors, pour tout  $v \in E$ ,

$$\begin{aligned} \langle v, f^*(\lambda w_1 + \mu w_2) \rangle &= \langle f(v), \lambda w_1 + \mu w_2 \rangle \\ &= \lambda \langle f(v), w_1 \rangle + \mu \langle f(v), w_2 \rangle \\ &= \lambda \langle v, f^*(w_1) \rangle + \mu \langle v, f^*(w_2) \rangle \\ &= \langle v, \lambda f^*(w_1) + \mu f^*(w_2) \rangle \end{aligned}$$

Ainsi,  $\Lambda(f^*(\lambda w_1 + \mu w_2)) = \Lambda(\lambda f^*(w_1) + \mu f^*(w_2))$ . Par injectivité de l'application  $\Lambda$ , on obtient alors l'égalité  $f^*(\lambda w_1 + \mu w_2) = \lambda f^*(w_1) + \mu f^*(w_2)$  et  $f^*$  est donc bien linéaire.  $\square$

Avant d'étudier le pendant matriciel de cette opération d'"adjonction" des endomorphismes de  $E$ , on en montre quelques propriétés de base :

**Proposition 2.7.2.** 1.  $(\text{Id}_E)^* = \text{Id}_E$ .

2. Si  $g \in \mathcal{L}(E)$  et  $\lambda, \mu \in \mathbb{R}$ ,  $(\lambda f + \mu g)^* = \lambda f^* + \mu g^*$  et  $(g \circ f)^* = f^* \circ g^*$ .

3. Si  $f$  est une application bijective, son adjoint  $f^*$  l'est également et  $(f^*)^{-1} = (f^{-1})^*$ .

4.  $(f^*)^* = f$ .

*Démonstration.* 1. Pour tous  $v, w \in E$ , on a  $\langle \text{Id}_E(v), w \rangle = \langle v, w \rangle = \langle v, \text{Id}_E(w) \rangle$  donc  $(\text{Id}_E)^* = \text{Id}_E$  (par unicité de l'adjoint).

2. Pour tous  $v, w \in E$ , on a

$$\begin{aligned} \langle (\lambda f + \mu g)(v), w \rangle &= \lambda \langle f(v), w \rangle + \mu \langle g(v), w \rangle \\ &= \lambda \langle v, f^*(w) \rangle + \mu \langle v, g^*(w) \rangle \\ &= \langle v, (\lambda f^* + \mu g^*)(w) \rangle \end{aligned}$$

donc  $(\lambda f + \mu g)^* = \lambda f^* + \mu g^*$ .

Pour tous  $v, w \in E$ , on a

$$\begin{aligned}\langle g \circ f(v), w \rangle &= \langle g(f(v)), w \rangle \\ &= \langle f(v), g^*(w) \rangle \\ &= \langle v, f^*(g^*(w)) \rangle \\ &= \langle v, f^* \circ g^*(w) \rangle\end{aligned}$$

donc  $(g \circ f)^* = f^* \circ g^*$ .

3. On a  $f^* \circ (f^{-1})^* = (f^{-1} \circ f)^* = (\text{Id}_E)^* = \text{Id}_E$  et  $(f^{-1})^* \circ f^* = (f \circ f^{-1})^* = (\text{Id}_E)^* = \text{Id}_E$ .

4. Pour tous  $v, w \in E$ , on a

$$\begin{aligned}\langle f^*(v), w \rangle &= \langle w, f^*(v) \rangle \\ &= \langle f(w), v \rangle \\ &= \langle v, f(w) \rangle\end{aligned}$$

donc  $(f^*)^* = f$ .

□

*Remarque 2.7.3.* On dit que l'endomorphisme  $f$  est auto-adjoint si  $f^* = f$ .  $\text{Id}_E$  est un exemple d'endomorphisme auto-adjoint de  $E$ .

Donnons également deux égalités intéressantes reliant noyaux et images de  $f$  et de son adjoint via l'opération "orthogonal d'un sous-ensemble" :

**Proposition 2.7.4.** *On a  $\text{Ker } f^* = (\text{Im } f)^\perp$  et  $\text{Im } f^* = (\text{Ker } f)^\perp$*

*Démonstration.* Soit  $w \in E$ . Supposons que  $w \in \text{Ker } f^*$ . Alors, pour tout  $v \in E$ ,

$$\langle f(v), w \rangle = \langle v, f^*(w) \rangle = \langle v, 0 \rangle = 0$$

donc  $w \in (\text{Im } f)^\perp$ . Réciproquement, si  $w \in (\text{Im } f)^\perp$  alors, pour tout  $v \in E$ ,

$$\langle v, f^*(w) \rangle = \langle f(v), w \rangle = 0,$$

en particulier  $\langle f^*(w), f^*(w) \rangle = 0$  donc  $f^*(w) = 0_E$  i.e.  $w \in \text{Ker } f^*$ .

Pour montrer la seconde égalité, on utilise l'égalité précédente ainsi que la proposition 2.3.7 3) et la proposition 2.7.2 3) :

$$\begin{aligned}\text{Im } f^* &= \left( (\text{Im } f^*)^\perp \right)^\perp \\ &= (\text{Ker } (f^*)^*)^\perp \\ &= (\text{Ker } f)^\perp\end{aligned}$$

□

Intéressons-nous à présent à la représentation matricielle des endomorphismes adjoints, plus précisément à leur représentation dans une base orthonormale de  $(E, \langle \cdot, \cdot \rangle)$  :

**Proposition 2.7.5.** *Soit  $\mathcal{B} = \{e_1, \dots, e_n\}$  une base orthonormale de  $E$ . Alors*

$$\text{Mat}_{\mathcal{B}}(f^*) = {}^t\text{Mat}_{\mathcal{B}}(f).$$

*Démonstration.* Pour tous  $v, w \in E$ , on a, puisque la base  $\mathcal{B}$  est orthonormale et d'après la définition de l'adjoint,

$$\langle f(v), w \rangle = \langle v, f^*(w) \rangle \Leftrightarrow {}^t(\text{Mat}_{\mathcal{B}}(f)\text{Mat}_{\mathcal{B}}(v))\text{Mat}_{\mathcal{B}}(w) = {}^t\text{Mat}_{\mathcal{B}}(v)(\text{Mat}_{\mathcal{B}}(f^*)\text{Mat}_{\mathcal{B}}(w))$$

i.e.

$${}^t\text{Mat}_{\mathcal{B}}(v){}^t\text{Mat}_{\mathcal{B}}(f)\text{Mat}_{\mathcal{B}}(w) = {}^t\text{Mat}_{\mathcal{B}}(v)\text{Mat}_{\mathcal{B}}(f^*)\text{Mat}_{\mathcal{B}}(w)$$

En prenant pour  $v$  et  $w$  les vecteurs de la base  $\mathcal{B}$ , on obtient, pour tous  $i, j \in \{1, \dots, n\}$  l'égalité

$${}^tX_i{}^t\text{Mat}_{\mathcal{B}}(f)X_j = {}^tX_i\text{Mat}_{\mathcal{B}}(f^*)X_j$$

(voir la démonstration de la proposition 2.6.4) et on a donc bien  ${}^t\text{Mat}_{\mathcal{B}}(f) = \text{Mat}_{\mathcal{B}}(f^*)$ .  $\square$

Deux conséquences directes de la représentation matricielle dans une base orthonormale de l'“adjonction” par la transposition sont les suivantes :

**Corollaire 2.7.6.** *Le rang de l'adjoint  $f^*$  de  $f$  est égal au rang de  $f$  et le déterminant de  $f^*$  est égal au déterminant de  $f$ .*

*Démonstration.* Soit  $\mathcal{B}$  une base orthonormale de  $E$  alors

$$\text{rg}(f^*) = \text{rg}(\text{Mat}_{\mathcal{B}}(f^*)) = \text{rg}({}^t\text{Mat}_{\mathcal{B}}(f)) = \text{rg}(\text{Mat}_{\mathcal{B}}(f)) = \text{rg}(f)$$

et

$$\det(f^*) = \det(\text{Mat}_{\mathcal{B}}(f^*)) = \det({}^t\text{Mat}_{\mathcal{B}}(f)) = \det(\text{Mat}_{\mathcal{B}}(f)) = \det(f).$$

$\square$

La proposition 2.7.5 nous donne également une nouvelle interprétation vectorielle de la transposition matricielle : si  $A$  est une matrice carrée de  $M_n(\mathbb{R})$ , représentant un endomorphisme  $f$  de  $\mathbb{R}^n$  dans la base canonique, alors  ${}^tA$  est la matrice représentative de l'endomorphisme  $f^*$  adjoint de  $f$  par rapport au produit scalaire canonique (pour lequel la base canonique est une base orthonormale).

Cette remarque nous invite à rechercher un lien entre les interprétations “duale” (proposition 1.6.4) et “euclidienne” de la transposition matricielle. Un tel lien est induit par la relation suivante :

**Proposition 2.7.7.** *On a*

$$f^* = \Lambda^{-1} \circ {}^t f \circ \Lambda.$$



*Démonstration.* Pour tous vecteurs  $v$  et  $w$  de  $E$ , on a

$$\langle f(v), w \rangle = (\Lambda(w) \circ f)(v) = {}^t f(\Lambda(w))(v) = ({}^t f \circ \Lambda(w))(v)$$

et

$$\langle v, f^*(w) \rangle = (\Lambda(f^*(w)))(v) = (\Lambda \circ f^*(w))(v),$$

d'où l'égalité  ${}^t f \circ \Lambda = \Lambda \circ f^*$  et le résultat ( $\Lambda$  est un isomorphisme).  $\square$

**Corollaire 2.7.8.** *Soit  $\mathcal{B}$  une base orthonormale de  $E$ . Alors*

$$\text{Mat}_{\mathcal{B}}(f^*) = \text{Mat}_{\mathcal{B}^*}({}^t f).$$

*Démonstration.* D'après la propriété précédente, on a

$$\text{Mat}_{\mathcal{B}}(f^*) = \text{Mat}_{\mathcal{B}, \mathcal{B}^*}(\Lambda)^{-1} \text{Mat}_{\mathcal{B}^*}({}^t f) \text{Mat}_{\mathcal{B}, \mathcal{B}^*}(\Lambda).$$

Or, si  $i \in \{1, \dots, n\}$  et si  $v$  est un vecteur de  $E$  de coordonnées  $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$  dans  $\mathcal{B}$ , on a

$$\Lambda(e_i)(v) = \left\langle e_i, \sum_{j=1}^n x_j e_j \right\rangle = \sum_{j=1}^n x_j \langle e_i, e_j \rangle = x_i = e_i^*(v)$$

donc  $\Lambda(e_i) = e_i^*$  et la matrice représentative de  $\Lambda$  dans les bases  $\mathcal{B}$  et  $\mathcal{B}^*$  est la matrice identité  $I_n$ . Ainsi, on a bien

$$\text{Mat}_{\mathcal{B}}(f^*) = \text{Mat}_{\mathcal{B}^*}({}^t f).$$

$\square$

*Remarque 2.7.9.* La proposition 2.7.7 nous donne une raison "intrinsèque" de l'égalité  $\text{Mat}_{\mathcal{B}}(f^*) = \text{Mat}_{\mathcal{B}^*}({}^t f)$  que l'on pouvait déjà obtenir à partir de la proposition 2.7.5.

## 2.8 Endomorphismes orthogonaux et matrices orthogonales

Soit  $(E, \langle \cdot, \cdot \rangle)$  un espace euclidien de dimension  $n \in \mathbb{N} \setminus \{0\}$  et soit  $f$  un endomorphisme de  $E$ .

**Définition 2.8.1.** *On dit que  $f$  est un endomorphisme orthogonal de  $E$  si  $f^* \circ f = \text{Id}_E$ .*

*Exemple 2.8.2.* L'identité  $\text{Id}_E$  de  $E$  est un endomorphisme orthogonal.

*Remarque 2.8.3.* Si  $f$  est orthogonal alors  $f$  est bijectif et  $f^{-1} = f^*$ . En effet, l'égalité  $f^* \circ f = \text{Id}_E$  implique l'injectivité de  $f$  (si  $v \in E$  vérifie  $f(v) = 0_E$  alors  $v = f^*(f(v)) = 0_E$ ) et donc sa bijectivité car  $f$  est une application linéaire de  $E$  dans  $E$ . L'égalité  $f^* \circ f = \text{Id}_E$  est alors équivalente à  $f^* = f^{-1}$ . En particulier, on a également l'égalité  $f \circ f^* = \text{Id}_E$ .

Remarquons que, réciproquement, l'égalité  $f \circ f^* = (f^*)^* \circ f^* = \text{Id}_E$  implique  $f^* \circ (f^*)^* = f^* \circ f = \text{Id}_E$ .

On déduit également de ces considérations que  $f$  est orthogonal ssi son adjoint  $f^*$  est orthogonal.

L'orthogonalité d'un endomorphisme se caractérise géométriquement :

**Proposition 2.8.4.**  *$f$  est un endomorphisme orthogonal si et seulement si  $f$  conserve le produit scalaire  $\langle \cdot, \cdot \rangle$  si et seulement si  $f$  conserve la norme euclidienne  $\|\cdot\|$  associée à  $\langle \cdot, \cdot \rangle$  si et seulement si  $f$  conserve la distance euclidienne  $d$  associée à  $\langle \cdot, \cdot \rangle$ .*

*Démonstration.* On montre que les assertions

1.  $f$  est orthogonal,
2. pour tous  $v, w \in E$ ,  $\langle f(v), f(w) \rangle = \langle v, w \rangle$ ,
3. pour tout  $v \in E$ ,  $\|f(v)\| = \|v\|$ ,
4. pour tous  $v, w \in E$ ,  $d(f(v), f(w)) = d(v, w)$ ,

sont équivalentes.

Montrons tout d'abord  $1 \Rightarrow 2$  : supposons que  $f$  est orthogonal, alors, pour tous  $v, w \in E$ ,

$$\langle f(v), f(w) \rangle = \langle v, f^* \circ f(w) \rangle = \langle v, w \rangle.$$

Montrons ensuite  $2 \Rightarrow 3$  : si  $f$  préserve le produit scalaire  $\langle \cdot, \cdot \rangle$ , alors, pour tout  $v \in E$ ,

$$\|f(v)\| = \sqrt{\langle f(v), f(v) \rangle} = \sqrt{\langle v, v \rangle} = \|v\|.$$

Montrons également  $3 \Rightarrow 4$  : si  $f$  préserve la norme euclidienne associée à  $\langle \cdot, \cdot \rangle$ , alors, pour tous  $v, w \in E$ ,

$$d(f(v), f(w)) = \|f(w) - f(v)\| = \|f(w - v)\| = \|w - v\| = d(v, w).$$

Montrons enfin  $4 \Rightarrow 1$  : Commençons par remarquer que  $4 \Rightarrow 3 \Rightarrow 2$  car  $f$  est linéaire et, pour tout  $u \in E$ ,  $d(u, 0) = \|u\|$  et, pour tous  $v, w \in E$ ,  $\langle v, w \rangle = \frac{1}{2} (\|v + w\|^2 - \|v\|^2 - \|w\|^2)$  (remarque 2.2.8). Si  $f$  préserve la distance euclidienne,  $f$  préserve donc également le produit scalaire et, pour tous  $v, w \in E$ , on a alors

$$\langle f^* \circ f(v), w \rangle = \langle f(v), f(w) \rangle = \langle v, w \rangle$$

et donc  $\langle f^* \circ f(v) - v, w \rangle = 0$ . En particulier, pour tout  $v \in E$ ,  $\langle f^* \circ f(v) - v, f^* \circ f(v) - v \rangle = 0$  donc  $f^* \circ f(v) - v = 0_E$  i.e.  $f^* \circ f(v) = v$ . Ainsi  $f^* \circ f = \text{Id}_E$  et  $f$  est orthogonal.  $\square$

*Remarque 2.8.5.* Un endomorphisme orthogonal est également appelé isométrie.

Une autre caractérisation d'un endomorphisme orthogonal est qu'il transforme toute base orthonormale en une base orthonormale :

**Proposition 2.8.6.** *L'endomorphisme  $f$  est orthogonal si et seulement si  $f$  associe à toute base orthonormale de  $E$  une base orthonormale de  $E$ .*

*Démonstration.* Supposons que  $f$  est orthogonal et soit  $\{e_1, \dots, e_n\}$  une base orthonormale de  $E$ . Comme  $f$  est orthogonal, on a, par la proposition précédente,  $\langle f(e_i), f(e_j) \rangle = \langle e_i, e_j \rangle = \delta_{ij}$ , donc la famille de  $n$  vecteurs  $\{f(e_1), \dots, f(e_n)\}$  est orthonormale : il s'agit donc d'une base orthonormale de  $E$ .

Réciproquement, soit  $\mathcal{B} = \{e_1, \dots, e_n\}$  une base orthonormale de  $E$  et supposons que la famille  $\{f(e_1), \dots, f(e_n)\}$  est orthonormale. Soient alors  $v, w \in E$ , de coordonnées respectives

$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$  et  $\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$  dans la base  $\mathcal{B}$ . On a

$$\langle f(v), f(w) \rangle = \left\langle f \left( \sum_{i=1}^n x_i e_i \right), f \left( \sum_{j=1}^n y_j e_j \right) \right\rangle = \sum_{1 \leq i, j \leq n} x_i y_j \langle f(e_i), f(e_j) \rangle = \sum_{i=1}^n x_i y_i = \langle v, w \rangle.$$

Par la proposition précédente,  $f$  est donc orthogonal. □

*Remarque 2.8.7.* Remarquons que, d'après la démonstration ci-dessus, il suffit qu'il existe une base orthonormale  $\{e_1, \dots, e_n\}$  de  $E$  telle que la famille  $\{f(e_1), \dots, f(e_n)\}$  soit orthonormale pour que l'endomorphisme  $f$  soit orthogonal.

Passons maintenant à la caractérisation matricielle de l'orthogonalité. Soit  $\mathcal{B}$  une base orthonormale de  $E$ .

**Proposition 2.8.8.** *Soit  $A := \text{Mat}_{\mathcal{B}}(f)$  la matrice représentative de  $f$  dans la base  $\mathcal{B}$ . L'endomorphisme  $f$  de  $E$  est orthogonal si et seulement si  ${}^tAA = I_n$ .*

*Démonstration.* D'après la proposition 2.7.5, la matrice représentative de l'adjoint  $f^*$  de  $f$  dans  $\mathcal{B}$  est  ${}^tA$  et donc  $f^* \circ f = \text{Id}_E$  ssi  ${}^tAA = I_n$  □

Ce résultat motive la définition suivante :

**Définition 2.8.9.** *On dit qu'une matrice  $A$  de  $M_n(\mathbb{R})$  est orthogonale si  ${}^tAA = I_n$ .*

*Remarque 2.8.10.* L'endomorphisme  $f$  de  $E$  est donc orthogonal si et seulement si la matrice  $\text{Mat}_{\mathcal{B}}(f)$  est orthogonale. « Réciproquement », une matrice  $A$  de  $M_n(\mathbb{R})$  est orthogonale ssi l'endomorphisme de  $\mathbb{R}^n$  représenté par  $f$  dans la base canonique est orthogonal

*Exemple 2.8.11.*

1. La matrice

$$A := \frac{1}{3} \begin{pmatrix} 2 & -1 & 2 \\ 2 & 2 & -1 \\ -1 & 2 & 2 \end{pmatrix}$$

est orthogonale.

2. Une symétrie orthogonale est un endomorphisme orthogonal (remarque 2.3.14). Une projection orthogonale sur un sous-espace vectoriel strict n'est pas un endomorphisme orthogonal (remarque 2.3.9).

Un point de vue supplémentaire donné par l'égalité " ${}^tAA = I_n$ " est le suivant : une matrice  $A \in M_n(\mathbb{R})$  est orthogonale si et seulement si les vecteurs colonnes la composant forment une base orthonormale de  $\mathbb{R}^n$  muni du produit scalaire canonique. Dans ce cas, la matrice  $A$  est la matrice de passage de la base canonique de  $\mathbb{R}^n$  à la base orthonormale formée par ses vecteurs colonnes. On peut généraliser cela de la façon suivante :

**Proposition 2.8.12.** *Soit  $\mathcal{B}'$  une base de  $E$ . Alors  $\mathcal{B}'$  est orthonormale si et seulement si la matrice de passage de  $\mathcal{B}$  à  $\mathcal{B}'$  est orthogonale.*

*Démonstration.* La matrice de passage  $P := P_{\mathcal{B} \rightarrow \mathcal{B}'}$  de  $\mathcal{B} = \{e_1, \dots, e_n\}$  à  $\mathcal{B}' = \{e'_1, \dots, e'_n\}$  est formée, dans l'ordre, des vecteurs coordonnés des vecteurs  $e'_i$ ,  $i = 1, \dots, n$ , dans la base  $\mathcal{B}$ . On a alors, comme  $\mathcal{B}$  est orthonormale,

$${}^tPP = \begin{pmatrix} \langle e'_1, e'_1 \rangle & \cdots & \langle e'_1, e'_n \rangle \\ \vdots & & \vdots \\ \langle e'_n, e'_1 \rangle & \cdots & \langle e'_n, e'_n \rangle \end{pmatrix}$$

Ainsi, la matrice de passage  $P = P_{\mathcal{B} \rightarrow \mathcal{B}'}$  est orthogonale ssi la famille de vecteurs  $\{e'_1, \dots, e'_n\}$  est orthonormale.  $\square$

*Remarque 2.8.13.* En particulier, toute matrice de passage d'une base orthonormale à une autre est orthogonale.

Remarquons ensuite que l'égalité " ${}^tAA = I_n$ " permet le calcul du déterminant d'une matrice orthogonale et donc d'un endomorphisme orthogonal :

**Proposition et Définition 2.8.14.** *Soit  $A$  une matrice orthogonale de  $M_n(\mathbb{R})$ . Alors  $\det(A) = 1$  ou  $\det(A) = -1$ . Ainsi, si  $f$  est orthogonal,  $\det(f) = 1$  ou  $\det(f) = -1$ .*

*Si  $\det(A) = 1$ , resp.  $\det(f) = 1$ , on dit que  $A$ , resp.  $f$ , est une matrice orthogonale directe, resp. endomorphisme orthogonal direct. Si  $\det(A) = -1$ , resp.  $\det(f) = -1$ , on dit que  $A$ , resp.  $f$ , est une matrice orthogonale indirecte, resp. endomorphisme orthogonal indirect.*

*Démonstration.* On a  ${}^tAA = I_n$  donc  $1 = \det({}^tAA) = \det({}^tA) \det(A) = (\det(A))^2$  d'où  $\det(A) = 1$  ou  $\det(A) = -1$

Ainsi, si  $f$  est orthogonal, comme sa matrice représentative dans une base orthonormale est orthogonale (proposition 2.8.8), on a  $\det(f) = 1$  ou  $\det(f) = -1$ .  $\square$

*Exemple 2.8.15.* 1. La matrice orthogonale  $A$  de l'exemple 2.8.11 1. est directe.

2. Une symétrie orthogonale par rapport à un sous-espace vectoriel  $F$  de  $E$  est directe ou indirecte suivant la parité de la "codimension" de  $F$  : le déterminant d'une telle symétrie orthogonale est  $(-1)^{n-p}$  si  $p$  est la dimension de  $F$  (remarque 2.3.14).

Terminons cette section en remarquant que l'ensemble des endomorphismes orthogonaux de  $E$  muni de la composition forme un groupe. On a déjà vu plus haut que l'identité  $\text{Id}_E$  de  $E$  est orthogonale et que, si  $f$  est un endomorphisme orthogonal, il en est de même pour son adjoint. Enfin, la composition d'endomorphismes orthogonaux est également orthogonale :

**Proposition 2.8.16.** *Soit  $g$  un endomorphisme de  $E$  et supposons que  $f$  et  $g$  sont orthogonaux. Alors la composition  $f \circ g$  (ainsi que la composition  $g \circ f$ ) est orthogonale.*

*Démonstration.* On a  $(f \circ g)^* \circ (f \circ g) = g^* \circ f^* \circ f \circ g = g^* \circ \text{Id}_E \circ g = g^* \circ g = \text{Id}_E$ .  $\square$

**Corollaire et Définition 2.8.17.** *L'ensemble, noté  $\mathcal{O}(E, \langle \cdot, \cdot \rangle)$  ou simplement  $\mathcal{O}(E)$  lorsque le contexte est clair, des endomorphismes orthogonaux de l'espace euclidien  $(E, \langle \cdot, \cdot \rangle)$  est un sous-groupe du groupe  $(\mathcal{GL}(E), \circ)$  des automorphismes linéaires de  $E$  muni de la composition (appelé groupe linéaire de  $E$ ). On a appelé  $(\mathcal{O}(E), \circ)$  le groupe orthogonal de  $(E, \langle \cdot, \cdot \rangle)$ .*

De manière équivalente, l'ensemble  $\text{O}_n(\mathbb{R})$  des matrices orthogonales de  $\text{M}_n(\mathbb{R})$  est un sous-groupe du groupe  $(\text{GL}_n(\mathbb{R}), \cdot)$  des matrices réelles inversibles de taille  $n$  muni du produit matriciel.  $\text{O}_n(\mathbb{R})$  est appelé groupe orthogonal de  $\text{M}_n(\mathbb{R})$ .

*Remarque 2.8.18.* Le sous-ensemble de  $\mathcal{O}(E)$  des endomorphismes orthogonaux directs est un sous-groupe de  $(\mathcal{O}(E), \circ)$  appelé groupe spécial orthogonal de  $E$  et noté  $\mathcal{SO}(E)$ .

Le sous-ensemble de  $\text{O}_n(\mathbb{R})$  des matrices orthogonales directes est un sous-groupe de  $(\text{O}_n(\mathbb{R}), \cdot)$  appelé groupe spécial orthogonal de  $\text{M}_n(\mathbb{R})$  et noté  $\text{SO}_n(\mathbb{R})$ .

## 2.9 Décomposition QR d'une matrice inversible

Soit à nouveau  $(E, \langle \cdot, \cdot \rangle)$  un espace euclidien.

Pour terminer ce chapitre, revenons au procédé d'orthonormalisation de Gram-Schmidt (théorème 2.5.4). Ce procédé permet notamment de construire, à partir d'une base  $\{v_1, \dots, v_n\}$  quelconque de  $E$ , une base orthonormale  $\{e_1, \dots, e_n\}$  de  $E$ .

Nous allons étudier le pendant matriciel de ce procédé. Précisément, soit  $n \in \mathbb{N} \setminus \{0\}$  et considérons une matrice  $A$  de  $\text{M}_n(\mathbb{R})$  de rang maximal (i.e.  $\text{rg}(A) = n$ ) i.e.  $A \in \text{GL}_n(\mathbb{R})$ . Notons  $v_1, \dots, v_n$  les vecteurs colonnes qui, dans l'ordre, forment la matrice  $A$ . La famille  $\mathcal{B}' := \{v_1, \dots, v_n\}$ , considérée comme famille de vecteurs de  $\mathbb{R}^n$ , est une base de  $\mathbb{R}^n$  : si  $\mathcal{B}_0$  désigne la base canonique de  $\mathbb{R}^n$ ,  $A$  est alors la matrice de passage  $P_{\mathcal{B}_0 \rightarrow \mathcal{B}'}$ . Notons ensuite  $\mathcal{B} = \{e_1, \dots, e_n\}$  la base orthonormale de  $\mathbb{R}^n$  (par rapport au produit scalaire canonique) obtenue à partir du procédé d'orthonormalisation de Gram-Schmidt. On note  $P := P_{\mathcal{B} \rightarrow \mathcal{B}'}$  la matrice de passage de la base  $\{v_1, \dots, v_n\}$  à la base  $\{e_1, \dots, e_n\}$  de  $E$ .

Dans le procédé de Gram-Schmidt, pour  $k \in \{1, \dots, n\}$ , le vecteur  $e_k$  est défini comme une combinaison linéaire des vecteurs  $\{v_1, \dots, v_k\}$ . En particulier, la matrice de passage  $P$  est triangulaire supérieure. De par cette construction, on sait même que les coefficients diagonaux de  $P$  sont (strictement) positifs.

Enfin, si  $Q$  désigne la matrice de passage de  $\mathcal{B}_0$  à  $\mathcal{B}'$  (i.e.  $Q$  est la matrice formée, dans l'ordre, par les vecteurs colonnes  $e_1, \dots, e_n$ ), on a  $Q = AP$  ( $P_{\mathcal{B}_0 \rightarrow \mathcal{B}'} = P_{\mathcal{B}_0 \rightarrow \mathcal{B}} P_{\mathcal{B} \rightarrow \mathcal{B}'}$ ). On réécrit cette égalité sous la forme  $A = QR$  avec  $R := P^{-1} = P_{\mathcal{B}' \rightarrow \mathcal{B}}$ . Remarquons enfin que, comme les vecteurs colonnes de la matrice  $Q$  forment une base orthonormale de  $\mathbb{R}^n$  (par rapport au produit scalaire canonique),  $Q$  est orthogonale.

Au total, nous avons (partiellement) montré le résultat suivant :

**Théorème 2.9.1** (Décomposition  $QR$  d'une matrice inversible). *Soit  $n \in \mathbb{N} \setminus \{0\}$  et soit  $A \in \text{GL}_n(\mathbb{R})$ . Il existe une unique matrice orthogonale  $Q \in \text{O}_n(\mathbb{R})$  et une unique matrice  $R \in \text{GL}_n(\mathbb{R})$  triangulaire supérieure à coefficients diagonaux strictement positifs telles que  $A = QR$ . On appelle cette écriture la décomposition  $QR$  de  $A$ .*

*Démonstration.* Il nous reste à montrer l'unicité de la décomposition. Soient donc  $Q_1, Q_2 \in \text{O}_n(\mathbb{R})$  et  $R_1, R_2$  deux matrices triangulaires supérieures de  $\text{GL}_n(\mathbb{R})$  à coefficients diagonaux strictement positifs telles que  $A = Q_1 R_1 = Q_2 R_2$ .

On a alors  $R_1 R_2^{-1} = Q_1^{-1} Q_2$  et  $R_1 R_2^{-1} = Q_1^{-1} Q_2$  est alors une matrice orthogonale ( $(\text{O}_n(\mathbb{R}), \cdot)$  est un groupe) et triangulaire supérieure à coefficients strictement positifs (l'inverse d'une matrice triangulaire supérieure à coefficients strictement positifs est une matrice triangulaire supérieure à coefficients strictement positifs). Il s'agit donc de la matrice  $I_n$  d'après le lemme qui suit cette preuve.

Ainsi,  $R_1 R_2^{-1} = Q_1^{-1} Q_2 = I_n$  et donc  $Q_1 = Q_2$  et  $R_1 = R_2$ . □

**Lemme 2.9.2.** *Soit  $A \in \text{GL}_n(\mathbb{R})$  une matrice orthogonale et triangulaire supérieure à coefficients strictement positifs. Alors  $A = I_n$ .*

*Démonstration.* Notons  $v_1, \dots, v_n$  les vecteurs colonnes qui, dans l'ordre, forment la matrice  $A = (a_{ij})_{1 \leq i, j \leq n}$ . Le fait que  $A$  soit orthogonale signifie que, pour tout  $i, j \in \{1, \dots, n\}$ ,  $\langle v_i, v_j \rangle = \delta_{ij}$ . En particulier,  $\langle v_1, v_1 \rangle = 1$ . Mais, comme  $A$  est triangulaire supérieure,  $\langle v_1, v_1 \rangle = a_{11}^2$  et, comme  $a_{11}$  est strictement positif, on a nécessairement  $a_{11} = 1$ . Pour tout  $j \in \{2, \dots, n\}$ , on a alors  $\langle v_1, v_j \rangle = a_{1j} = 0$  (autrement dit, la première ligne n'a que des coefficients nuls sauf le coefficient  $a_{11}$  qui est égal à 1).

Soit  $k \in \{1, n-1\}$  et supposons que l'on a déjà montré que, pour tout  $i \in \{1, \dots, k\}$  et tout  $j \in \{1, \dots, n\}$ ,  $a_{ij} = \delta_{ij}$  (autrement dit que pour les  $k$  premières lignes, tous les coefficients d'une ligne  $i$  donnée sont nuls sauf le coefficient  $a_{ii}$  qui est égal à 1). On considère alors le vecteur colonne  $v_{k+1}$  dont, par hypothèse de récurrence, la seule coordonnée non nulle est  $a_{k+1, k+1} > 0$ . Comme  $\langle v_{k+1}, v_{k+1} \rangle = a_{k+1, k+1}^2 = 1$ , on a  $a_{k+1, k+1} = 1$  et, par suite, pour tout  $j \in \{k+2, \dots, n\}$ ,  $\langle v_{k+1}, v_j \rangle = a_{k+1, j} = 0$  : la ligne  $k+1$  n'a donc que des coefficients nuls sauf le coefficient  $a_{k+1, k+1}$  qui est égal à 1. □

*Exemple 2.9.3.* Considérons la matrice

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

de  $M_3(\mathbb{R})$ . On note  $v_1 := (1, 1, 0)$ ,  $v_2 := (1, 2, 0)$ ,  $v_3 := (0, 0, 1) \in \mathbb{R}^3$  et on applique le procédé

d'orthonormalisation de Gram-Schmidt à la base  $\{v_1, v_2, v_3\}$  de  $\mathbb{R}^3$ . On pose

$$\begin{aligned}\epsilon_1 &:= v_1 = (1, 1, 0) \\ \epsilon_2 &:= v_2 - \frac{\langle v_2, \epsilon_1 \rangle}{\|\epsilon_1\|^2} \epsilon_1 = v_2 - \frac{3}{2} \epsilon_1 = \left( -\frac{1}{2}, \frac{1}{2}, 0 \right) = \frac{1}{2}(-1, 1, 0) \\ \epsilon'_2 &:= (-1, 1, 0) \\ \epsilon_3 &:= v_3 - \frac{\langle v_3, \epsilon_1 \rangle}{\|\epsilon_1\|^2} \epsilon_1 - \frac{\langle v_3, \epsilon'_2 \rangle}{\|\epsilon'_2\|^2} \epsilon'_2 = v_3 = (0, 0, 1) \\ e_1 &:= \frac{1}{\|\epsilon_1\|} \epsilon_1 = \frac{1}{\sqrt{2}}(1, 1, 0) \\ e_2 &:= \frac{1}{\|\epsilon'_2\|} \epsilon'_2 = \frac{1}{\sqrt{2}}(-1, 1, 0) \\ e_3 &:= \frac{1}{\|\epsilon_3\|} \epsilon_3 = (0, 0, 1)\end{aligned}$$

Comme

$$\begin{aligned}v_1 &= \epsilon_1 = \sqrt{2}e_1 \\ v_2 &= \frac{3}{2}\epsilon_1 + \epsilon_2 = \frac{3}{2}\epsilon_1 + \frac{1}{2}\epsilon'_2 = \frac{3}{\sqrt{2}}e_1 + \frac{1}{\sqrt{2}}e_2 \\ v_3 &= \epsilon_3 = e_3\end{aligned}$$

la matrice de passage de la base  $\{e_1, e_2, e_3\}$  à la base  $\{v_1, v_2, v_3\}$  est

$$R := \begin{pmatrix} \sqrt{2} & \frac{3}{\sqrt{2}} & 0 \\ 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Si on note  $Q$  la matrice

$$\begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

formée par les coordonnées (dans la base canonique de  $\mathbb{R}^3$ ) des vecteurs  $e_1, e_2, e_3$ , l'égalité

$$A = QR = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \sqrt{2} & \frac{3}{\sqrt{2}} & 0 \\ 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

est la décomposition  $QR$  de  $A$ .





## Chapitre 3

# Rappels et compléments sur la réduction des endomorphismes

### 3.1 Introduction

On donne dans ce chapitre des rappels sur la théorie de réduction des endomorphismes, c'est-à-dire l'étude des bases dans lesquelles un endomorphisme donné possède la représentation matricielle la plus "simple" possible (la plus "réduite" possible).

On étudiera notamment les critères nécessaires et suffisants classiques de diagonalisabilité, directs (via la recherche des espaces propres) ou via les polynômes d'endomorphismes. On étudiera également la triangularisabilité et la réduction la plus aboutie des endomorphismes triangularisables, à savoir la réduction de Jordan pour laquelle nous donnerons une méthode systématique de réduction.

La première partie de ce chapitre étant constituée de rappels, les assertions seront la plupart du temps données sans preuve (on renvoie au cours de l'année passée pour les démonstrations). Nous les illustrerons cependant, ainsi que les méthodes, par des exemples. La preuve de l'existence de la réduction de Jordan pour les endomorphismes triangularisables sera elle donnée, notamment afin de dégager une méthode systématique de réduction.

Tout au long de ce chapitre,  $\mathbb{K}$  désigne un corps commutatif quelconque et  $E$  désigne un espace vectoriel sur  $\mathbb{K}$  de dimension finie.

### 3.2 Valeurs propres et espaces propres

Soit  $f$  un endomorphisme de  $E$ .

**Définition 3.2.1.** Soit  $\lambda \in \mathbb{K}$ . On dit que  $\lambda$  est une valeur propre de  $f$  s'il existe un vecteur non nul  $v$  de  $E$  tel que  $f(v) = \lambda v \Leftrightarrow (f - \lambda \text{Id}_E)(v) = 0_E$ , autrement dit si l'endomorphisme  $f - \lambda \text{Id}_E$  n'est pas injectif.

Si  $\lambda \in \mathbb{K}$  est une valeur propre de  $f$ , on note  $E_\lambda := \text{Ker}(f - \lambda \text{Id}_E)$  : il s'agit d'un sous-espace vectoriel de  $E$  que l'on appelle sous-espace propre de  $f$  associé à la valeur propre  $\lambda$  et tout vecteur non nul de  $E_\lambda$  est appelé vecteur propre de  $f$  associé à la valeur propre  $\lambda$ . On dira

qu'un vecteur  $v$  de  $E$  est un vecteur propre de  $f$  s'il existe une valeur propre  $\lambda \in \mathbb{K}$  de  $f$  telle que  $v$  est un vecteur propre de  $f$  associé à  $\lambda$ .

*Exemple 3.2.2.* 2 est une valeur propre de l'endomorphisme

$$f : \begin{array}{ccc} \mathbb{R}^2 & \rightarrow & \mathbb{R}^2 \\ (x, y) & \mapsto & (x + 2y, -x + 4y) \end{array}$$

car, par exemple,  $f(2, 1) = (4, 2) = 2(2, 1)$ , et on a

$$E_2 = \text{Ker} (f - 2\text{Id}_{\mathbb{R}^2}) = \{(x, y) \in \mathbb{R}^2 \mid x = 2y\}.$$

Soit  $n \in \mathbb{N} \setminus \{0\}$  et soit  $A$  une matrice carrée de  $M_n(\mathbb{K})$ . On peut, de façon analogue, définir une notion de valeur propre, d'espace propre et de vecteur propre pour  $A$  : un scalaire  $\lambda \in \mathbb{K}$  est une valeur propre de  $A$  s'il existe un vecteur colonne non nul  $X$  de  $M_{n,1}(\mathbb{K})$  tel que  $AX = \lambda X$ , autrement dit si le sous-espace vectoriel  $E_\lambda := \text{Ker} (A - \lambda I_n)$  de  $M_{n,1}(\mathbb{K})$  n'est pas réduit au vecteur colonne nul, et, dans ce cas,  $E_\lambda$  est appelé sous-espace propre de  $A$  associé à la valeur propre  $\lambda$  et tout vecteur colonne non nul de  $E_\lambda$  est appelé vecteur propre de  $A$  associé à la valeur propre  $\lambda$ .

*Remarque 3.2.3.* Supposons que  $\dim(E) = n$  et soit  $\mathcal{B}$  une base de  $E$ . Soient  $\lambda \in \mathbb{K}$  et  $v \in E$ . Si  $A = \text{Mat}_{\mathcal{B}}(f)$ , alors  $\lambda$  est une valeur propre de  $f$  ssi  $\lambda$  est une valeur propre de  $A$  et, dans ce cas,  $v$  est un vecteur propre de  $f$  associé à  $\lambda$  ssi  $\text{Mat}_{\mathcal{B}}(v)$  est un vecteur propre de  $A$  associé à  $\lambda$ .

**Définition 3.2.4.** L'ensemble des valeurs propres de  $f$ , resp.  $A$ , dans  $\mathbb{K}$  est appelé spectre de  $f$ , resp. spectre de  $A$ , et noté  $\text{Sp}(f)$ , resp.  $\text{Sp}(A)$ .

### 3.3 Polynôme caractéristique

Soient  $f \in \mathcal{L}(E)$  et  $\lambda \in \mathbb{K}$ .

**Proposition 3.3.1.**  $\lambda$  est une valeur propre de  $f$  ssi  $\det(f - \lambda \text{Id}_E) = 0$ .

On note  $\chi_f$  le polynôme

$$\det(f - X \text{Id}_E) \in \mathbb{K}[X],$$

appelé polynôme caractéristique de  $f$ . Ainsi,  $\lambda \in \text{Sp}(f)$  ssi  $\lambda$  est une racine (dans  $\mathbb{K}$ ) de  $\chi(f)$ .

*Exemple 3.3.2.* Reprenons l'endomorphisme  $f$  de l'exemple 3.2.2. La matrice représentative de  $f$  dans la base canonique de  $\mathbb{R}^2$  est

$$\begin{pmatrix} 1 & 2 \\ -1 & 4 \end{pmatrix}$$

et donc

$$\chi_f = \det(f - X \text{Id}_E) = \det\left(\begin{pmatrix} 1 & 2 \\ -1 & 4 \end{pmatrix} - X I_2\right) = \det\begin{pmatrix} 1-X & 2 \\ -1 & 4-X \end{pmatrix} = (3-X)(2-X)$$

d'où  $\text{Sp}(f) = \{2; 3\}$ .

Soit  $A \in M_n(\mathbb{K})$ , on définit de façon analogue  $\chi_A := \det(A - XI_n) \in \mathbb{K}[X]$  et le spectre de  $A$  est alors l'ensemble des racines de  $\chi_A$  dans  $\mathbb{K}$ .

*Remarque 3.3.3.* Attention au corps de base : si l'on considère par exemple la matrice

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

on a

$$\chi_A = \det \begin{pmatrix} -X & 1 \\ -1 & -X \end{pmatrix} = X^2 + 1.$$

Ainsi, si  $A$  est considérée comme une matrice de  $M_n(\mathbb{C})$ , on a  $\text{Sp}(A) = \text{Sp}_{\mathbb{C}}(A) = \{-i; i\}$ , et si  $A$  est considérée comme une matrice de  $M_n(\mathbb{R})$ , on a  $\text{Sp}(A) = \text{Sp}_{\mathbb{R}}(A) = \emptyset$ .

**Corollaire 3.3.4.** *Si  $n = \dim(E)$ , le polynôme  $\chi_f \in \mathbb{K}[X]$  est de degré  $n$  et  $f$  possède donc au plus  $n$  valeurs propres distinctes.*

De la même façon  $A$  possède au plus  $n$  valeurs propres distinctes.

*Remarque 3.3.5.* D'après le théorème de d'Alembert-Gauss, tout polynôme de  $\mathbb{C}[X]$  est scindé. Ainsi, tout endomorphisme sur  $\mathbb{C}$  admet au moins une valeur propre.

### 3.4 Diagonalisabilité et diagonalisation

Soit  $f \in \mathcal{L}(E)$ . On note  $n := \dim(E)$ .

La forme "la plus simple" pour une matrice carrée est la forme diagonale. Nous allons rappeler dans cette section des conditions suffisantes, voire nécessaires et suffisantes, sous lesquelles  $f$  possède une matrice représentative diagonale (diagonalisabilité) et, dans ce cas, chercher à déterminer une base de  $E$  dans laquelle la matrice représentative de  $f$  est diagonale (diagonalisation).

On commence par donner la définition précise de la diagonalisabilité de  $f$  :

**Définition 3.4.1.** *On dit que l'endomorphisme  $f$  est diagonalisable s'il existe une base  $\mathcal{B}$  de  $E$  et des scalaires  $\lambda_1, \dots, \lambda_n \in \mathbb{K}$  tels que*

$$\text{Mat}_{\mathcal{B}}(f) = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

Remarquons que  $f$  est donc diagonalisable ssi il existe une base de  $E$  formée de vecteurs propres de  $f$ . Mais on peut énoncer une caractérisation plus précise et pratique. Pour cela, commençons par énoncer le fait suivant :

**Proposition 3.4.2.** *Soient  $\lambda_1, \dots, \lambda_k$ ,  $k \in \mathbb{N} \setminus \{0\}$ , des valeurs propres deux à deux distinctes de  $f$ . Alors les sous-espaces propres  $E_{\lambda_1}, \dots, E_{\lambda_k}$  correspondants sont en somme directe.*

En conséquence, si  $\lambda_1, \dots, \lambda_p$ ,  $p \in \mathbb{N}$ , désignent les valeurs propres deux à deux distinctes de  $f$  :

**Théorème 3.4.3.**  *$f$  est diagonalisable ssi  $E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_n} = E$  ssi  $\sum_{i=1}^p \dim(E_{\lambda_i}) = \dim(E)$ .*

*Exemple 3.4.4.* Reprenons l'endomorphisme  $f$  des exemples 3.2.2 et 3.3.2. On avait déterminé que  $\text{Sp}(f) = \{2; 3\}$  et  $E_2 = \{(x, y) \in \mathbb{R}^2 \mid x = 2y\} = \text{Vect}\{(2, 1)\}$ . Enfin,

$$E_3 = \text{Ker}(f - 3\text{Id}_{\mathbb{R}^2}) = \{(x, y) \in \mathbb{R}^2 \mid x = y\} = \text{Vect}\{(1, 1)\}.$$

Ainsi,  $\dim(E_1) + \dim(E_2) = 2 = \dim(\mathbb{R}^2)$  et  $f$  est donc diagonalisable. De plus, la famille  $\mathcal{B} := \{(2, 1), (1, 1)\}$  est une base de  $E$  formée de vecteurs propres de  $f$  et on a

$$\text{Mat}_{\mathcal{B}}(f) = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$$

(on dit qu'on a diagonalisé  $f$ ).

*Remarque 3.4.5.* Diagonaliser un endomorphisme diagonalisable  $f$  de  $E$ , c'est déterminer une base  $\mathcal{B}$  de  $E$  dans laquelle la matrice représentative de  $f$  est diagonale et exprimer  $\text{Mat}_{\mathcal{B}}(f)$ .

On peut également exprimer la condition sur les dimensions du théorème 3.4.3 à l'aide du polynôme caractéristique et plus particulièrement à l'aide des multiplicités des valeurs propres en tant que racines du polynôme caractéristique :

**Définition 3.4.6.** *Soit  $\lambda \in \mathbb{K}$  une valeur propre de  $f$ . On note  $m_\lambda$  la multiplicité de  $\lambda$  en tant que racine du polynôme  $\chi_f \in \mathbb{K}[X]$ .*

On peut tout de suite remarquer que, si  $\lambda_1, \dots, \lambda_p$  désignent les valeurs propres de  $f$ , alors la somme  $m_{\lambda_1} + \cdots + m_{\lambda_p}$  est inférieure ou égale à  $\deg(\chi_f) = n = \dim(E)$ , et qu'il y a égalité ssi  $\chi_f$  est scindé sur  $\mathbb{K}$ . De plus :

**Proposition 3.4.7.** *Si  $\lambda \in \text{Sp}(f)$ , alors*

$$1 \leq \dim(E_\lambda) \leq m_\lambda.$$

Ainsi :

**Théorème 3.4.8.**  *$f$  est diagonalisable ssi  $\chi_f$  est scindé et, pour tout  $\lambda \in \text{Sp}(f)$ ,  $\dim(E_\lambda) = m_\lambda$ .*

*Exemple 3.4.9.* Si  $f$  admet  $n = \dim(E)$  valeurs propres deux à deux distinctes, alors  $f$  est diagonalisable.

La diagonalisation d'un endomorphisme correspond à un changement de base vers une base dans laquelle la matrice représentative de l'endomorphisme considéré est diagonale. L'analogie matriciel du changement de base est l'opération de "conjugaison" par une matrice inversible. Soit  $A$  une matrice de  $M_n(\mathbb{K})$ .

**Définition 3.4.10.** *On dit que  $A$  est diagonalisable s'il existe une matrice inversible  $P \in \text{GL}_n(\mathbb{K})$  et une matrice diagonale  $D$  de  $M_n(\mathbb{K})$  telles que*

$$P^{-1}AP = D.$$

*Remarque 3.4.11.* Si  $B \in M_n(\mathbb{K})$ , on dit que  $A$  est semblable à  $B$  s'il existe une matrice inversible  $P \in GL_n(\mathbb{K})$  telle que  $P^{-1}AP = B$  (la relation de similitude sur  $M_n(\mathbb{K})$  est une relation d'équivalence).

Ainsi,  $A$  est diagonalisable ssi  $A$  est semblable à une matrice diagonale.

Les résultats de diagonalisabilité d'un endomorphisme énoncés ci-dessus ont leurs analogues matriciels, à savoir :

**Théorème 3.4.12.** Soient  $\lambda_1, \dots, \lambda_p$ ,  $p \in \mathbb{N}$ , les valeurs propres deux à deux distinctes de  $A$  et, pour  $i \in \{1, \dots, p\}$ , notons  $m_{\lambda_i}$  la multiplicité de  $\lambda_i$  en tant que racine de  $\chi_A$ . Alors  $A$  est diagonalisable ssi  $\sum_{i=1}^p \dim(E_{\lambda_i}) = n$  ssi ( $\chi_A$  est scindé et, pour tout  $i \in \{1, \dots, p\}$ ,  $\dim(E_{\lambda_i}) = m_{\lambda_i}$ ).

*Remarque 3.4.13.* • Diagonaliser une matrice diagonalisable  $A$  de  $M_n(\mathbb{K})$ , c'est déterminer une matrice inversible  $P \in GL_n(\mathbb{K})$  telle que la matrice  $P^{-1}AP$  soit diagonale et exprimer  $P^{-1}AP$ .

- Diagonaliser une matrice diagonalisable permet entre autres de calculer ses puissances, comme présenté dans l'exemple ci-dessous.

*Exemple 3.4.14.* On considère la matrice

$$A := \begin{pmatrix} 1 & -1 \\ 2 & 4 \end{pmatrix}$$

de  $M_2(\mathbb{R})$ . Son polynôme caractéristique  $\chi_A = (X-2)(X-3)$  est scindé à racines simples donc  $A$  est diagonalisable.

Une base de  $E_2$  est  $\left\{ \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\}$ , une base de  $E_3$  est  $\left\{ \begin{pmatrix} 1 \\ -2 \end{pmatrix} \right\}$  et on pose

$$P := \begin{pmatrix} 1 & 1 \\ -1 & -2 \end{pmatrix}.$$

On a alors

$$\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} = P^{-1}AP$$

donc  $A = P \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} P^{-1}$  et, par associativité du produit matriciel, pour  $k \in \mathbb{N} \setminus \{0\}$ ,  $A^k = P \begin{pmatrix} 2^k & 0 \\ 0 & 3^k \end{pmatrix} P^{-1}$ .

Or  $P^{-1} = \begin{pmatrix} 2 & 1 \\ -1 & -1 \end{pmatrix}$  donc

$$A^k = \begin{pmatrix} 2^{k+1} - 3^k & 2^k + 3^k \\ -2^{k+1} + 2 \cdot 3^k & -2^k - 2 \cdot 3^k \end{pmatrix}.$$

### 3.5 Polynômes d'endomorphismes, polynômes annulateurs et diagonalisabilité

On va dans cette section présenter des conditions nécessaires et suffisantes de diagonalisabilité, exprimées à l'aide de la notion de polynôme d'endomorphisme :

**Définition 3.5.1.** Soient  $P = a_N X^N + a_{N-1} X^{N-1} + \dots + a_1 X + a_0 \in \mathbb{K}[X]$  et  $f \in \mathcal{L}(E)$ . On note

$$P(f) := a_N f^N + a_{N-1} f^{N-1} + \dots + a_1 f + a_0 \text{Id}_E \in \mathcal{L}(E)$$

(où, pour  $k \in \{1, \dots, N\}$ ,  $f^k$  désigne la composée  $k^{\text{ème}}$  de l'endomorphisme  $f$  avec lui-même). Un endomorphisme de  $E$  de cette forme est appelé polynôme d'endomorphisme.

On peut définir de façon analogue la notion de polynôme de matrice : si l'on reprend les notations de la définition ci-dessus et si  $A$  est une matrice de  $M_n(\mathbb{K})$ , on définit

$$P(A) := a_N A^N + a_{N-1} A^{N-1} + \dots + a_1 A + a_0 I_n \in M_n(\mathbb{K}),$$

où, pour  $k \in \{1, \dots, N\}$ ,  $A^k$  désigne la puissance  $k^{\text{ème}}$  de  $A$ .

Tous les énoncés et notions présentés ci-après sur les polynômes d'endomorphismes ont leur analogues immédiats pour les polynômes de matrices.

*Exemple 3.5.2.* Considérons la matrice

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

de  $M_2(\mathbb{R})$  et le polynôme  $P := X(X - 1)$  de  $\mathbb{R}[X]$ . On a

$$P(A) = A(A - I_2) = \begin{pmatrix} 1 & 0 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ -1 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Soient  $f \in \mathcal{L}(E)$  et  $P \in \mathbb{K}[X]$ .

**Définition 3.5.3.** On dit que  $P$  est un polynôme annulateur de  $f$  si  $P(f)$  est l'endomorphisme identiquement nul de  $E$ .

*Exemple 3.5.4.* Si l'on reprend les notations de l'exemple précédent 3.5.2,  $P$  est un polynôme annulateur de  $A$ .

Un premier pas entre les polynômes annulateurs et la réduction des endomorphismes est donné par le résultat suivant :

**Lemme 3.5.5.** Soit  $\lambda \in \text{Sp}(f)$ . Si  $P$  est un polynôme annulateur de  $f$ , alors  $\lambda$  est une racine de  $P$ .

*Exemple 3.5.6.* • Reprenons la matrice  $A$  de l'exemple 3.5.2. Comme  $P = X(X - 1)$  est un polynôme annulateur de  $A$ ,  $\text{Sp}(A) \subset \{0; 1\}$ .

- Si l'endomorphisme  $f$  vérifie  $f^3 = f$  i.e. le polynôme  $X^3 - X = X(X - 1)(X + 1)$  annule  $f$ , alors  $\text{Sp}(f) \subset \{0; -1; 1\}$ .

On en vient à un premier critère, nécessaire et suffisant, de diagonalisabilité d'un endomorphisme mettant en jeu la notion de polynôme annulateur :

**Théorème 3.5.7.** *L'endomorphisme  $f$  est diagonalisable si et seulement s'il existe un polynôme annulateur de  $f$  qui soit scindé à racines simples.*

*Exemple 3.5.8.* Si  $f$  vérifie alors  $f^3 = f$  alors  $f$  est diagonalisable car le polynôme  $X^3 - X = X(X - 1)(X + 1)$ , annulateur de  $f$ , est scindé à racines simples.

*Remarque 3.5.9.* Il est à noter que ce critère permet, si l'on trouve un tel polynôme annulateur de  $f$  scindé à racines simples, de montrer que  $f$  est diagonalisable sans avoir à calculer les dimensions des espaces propres de  $f$ .

### 3.6 Polynôme minimal

Soit  $f \in \mathcal{L}(E)$ . Remarquons que l'ensemble  $I_f$  des polynômes de  $\mathbb{K}[X]$  qui annulent  $f$  est un idéal de l'anneau  $\mathbb{K}[X]$  : le polynôme nul annule  $f$  et, si  $P$  et  $Q$  sont deux polynômes annulant  $f$  et si  $R$  est un polynôme de  $\mathbb{K}[X]$ ,  $(P - Q)(f) = P(f) - Q(f) = 0$  et  $(RP)(f) = R(f)P(f) = 0$ .

L'anneau  $\mathbb{K}[X]$  étant principal, l'idéal  $I_f$  peut être engendré par un seul élément de  $I_f$ , que l'on peut de plus supposer unitaire. En fait, il y a un seul générateur de ce type :

**Proposition et Définition 3.6.1.** *Il existe un unique polynôme unitaire  $\mu_f$  tel que  $I_f = (\mu_f)$ . On appelle  $\mu_f$  le polynôme minimal de  $f$ .*

*Remarque 3.6.2.* • Par définition, les polynômes annulateurs de  $f$  sont les multiples de  $\mu_f$ .  
En particulier, si l'on connaît  $\mu_f$ , on connaît alors tous les polynômes annulateurs de  $f$ .

- $\mu_f$  est le polynôme annulateur de  $f$  unitaire de plus petit degré.

Le polynôme minimal de  $f$  divise le polynôme caractéristique de  $f$  en vertu du théorème de Cayley-Hamilton :

**Théorème 3.6.3** (Théorème de Cayley-Hamilton). *Le polynôme caractéristique de  $f$  est un polynôme annulateur de  $f$ . Autrement dit  $\chi_f \in I_f$  i.e.  $\mu_f$  divise  $\chi_f$ .*

Ce théorème nous donne en particulier des informations supplémentaires sur  $\mu_f$  :

**Corollaire 3.6.4.** •  $1 \leq \deg(\mu_f) \leq n$ .

- Les racines de  $\mu_f$  dans  $\mathbb{K}$  (resp. dans la clôture algébrique de  $\mathbb{K}$ ) sont exactement les racines de  $\chi_f$  dans  $\mathbb{K}$  (resp. dans la clôture algébrique de  $\mathbb{K}$ ), i.e. les valeurs propres de  $f$  dans  $\mathbb{K}$  (resp. dans la clôture algébrique de  $\mathbb{K}$ ), avec multiplicités différentes a priori.

Le théorème de Cayley-Hamilton nous donne ainsi un moyen de déterminer le polynôme minimal de  $f$  à partir de la donnée du polynôme caractéristique :

*Exemple 3.6.5.* 1. On considère la matrice

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

de  $M_3(\mathbb{R})$ . Son polynôme caractéristique est  $\chi_A = -(X+1)(X+2)(X-3)$ . Ainsi, nécessairement,  $\mu_A = (X+1)(X+2)(X-3)$ .

2. On considère la matrice

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

de  $M_3(\mathbb{R})$ . Son polynôme caractéristique est  $\chi_A = -(X-1)(X+2)^2$ . Ainsi, nécessairement,  $\mu_A = (X-1)(X+2)$  ou  $\mu_A = (X-1)(X+2)^2$ . Comme  $\deg((X-1)(X+2)) < \deg((X-1)(X+2)^2)$ , on commence par tester si le polynôme  $(X-1)(X+2)$  annule  $A$ . On a

$$(A - I_3)(A + 2I_3) = \begin{pmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

et donc  $(X-1)(X+2)$  est le polynôme minimal de  $A$ .

3. On considère la matrice

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

de  $M_3(\mathbb{R})$ . Son polynôme caractéristique est  $\chi_A = -(X-1)(X-2)^2$ . Ainsi, nécessairement,  $\mu_A = (X-1)(X-2)$  ou  $\mu_A = (X-1)(X-2)^2$ . Comme  $\deg((X-1)(X-2)) < \deg((X-1)(X-2)^2)$ , on commence par tester si le polynôme  $(X-1)(X-2)$  annule  $A$ . Or on constate que la matrice  $(A - I_3)(A - 2I_3)$  n'est pas la matrice nulle de  $M_3(\mathbb{R})$  donc, nécessairement,  $\mu_A = (X-1)(X-2)^2$ .

On termine cette section par un critère de diagonalisabilité permettant de décider, à partir de la donnée du polynôme minimal de  $f$ , si  $f$  est diagonalisable ou non :

**Théorème 3.6.6.** *L'endomorphisme  $f$  est diagonalisable si et seulement si son polynôme minimal  $\mu_f$  est scindé à racines simples.*

*Exemple 3.6.7.* Les matrices des exemples 3.6.5 1. et 2. sont diagonalisables, la matrice de l'exemple 3.6.5 3. n'est pas diagonalisable.

*Remarque 3.6.8.* Le théorème 3.6.6 permet de déterminer si un endomorphisme est diagonalisable ou non sans passer par le calcul des dimensions de ses espaces propres.

### 3.7 Triangularisabilité et triangularisation

Soit  $f \in \mathcal{L}(E)$ .

Si  $f$  n'est pas diagonalisable, on peut chercher à déterminer si  $f$  peut être réduit sous forme triangulaire :

**Définition 3.7.1.** *On dit que l'endomorphisme  $f$  est triangularisable s'il existe une base  $\mathcal{B}$  de  $E$  telle que  $\text{Mat}_{\mathcal{B}}(f)$  est triangulaire (supérieure ou inférieure).*



De manière analogue, on dira qu'une matrice  $A$  de  $M_n(\mathbb{K})$  est triangularisable si  $A$  est semblable à une matrice triangulaire.

*Remarque 3.7.2.* • Si  $T$  est une matrice triangulaire représentant  $f$ , les coefficients de la diagonale de  $T$  sont les valeurs propres de  $f$ , apparaissant suivant leurs multiplicités dans  $\chi_f$ .

- Un endomorphisme diagonalisable est en particulier triangularisable.

On donne dès à présent le critère essentiel de triangularisabilité d'un endomorphisme :

**Théorème 3.7.3.** *L'endomorphisme  $f$  est triangularisable si et seulement si son polynôme caractéristique  $\chi_f$  est scindé sur  $\mathbb{K}$  (ssi son polynôme minimal  $\mu_f$  est scindé sur  $\mathbb{K}$ ).*

**Corollaire 3.7.4.** *Tout endomorphisme d'un  $\mathbb{C}$ -espace vectoriel est triangularisable.*

*Toute matrice de  $M_n(\mathbb{C})$  est semblable à une matrice triangulaire.*

*Exemple 3.7.5.* Reprenons la matrice

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

de l'exemple 3.6.5 3. On a vu plus haut que  $\chi_A = -(X-1)(X-2)^2$  est scindé sur  $\mathbb{R}$ . Donc, même si  $A$  n'est pas diagonalisable,  $A$  est triangularisable d'après le théorème ci-dessus. Cherchons une matrice triangulaire  $T$  de  $M_3(\mathbb{R})$  semblable à  $A$ .

Commençons par remarquer que chaque espace propre de  $A$  est de dimension 1. On a

$$A - I_3 = \begin{pmatrix} 2 & -1 & 1 \\ 2 & -1 & 1 \\ 1 & -1 & 1 \end{pmatrix} \text{ donc}$$

$$E_1 = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in M_3(\mathbb{R}) \mid \begin{cases} 2x - y + z = 0 \\ x - y + z = 0 \end{cases} \right\} = \text{Vect} \left\{ \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right\}$$

$$\text{et } A - 2I_3 = \begin{pmatrix} 1 & -1 & 1 \\ 2 & -2 & 1 \\ 1 & -1 & 0 \end{pmatrix} \text{ donc}$$

$$E_2 = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in M_3(\mathbb{R}) \mid \begin{cases} x - y + z = 0 \\ x - y = 0 \end{cases} \right\} = \text{Vect} \left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \right\}$$

Si on complétait la famille  $\left\{ \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \right\}$  de  $M_{3,1}(\mathbb{R})$  ainsi obtenue en une base, i.e. de façon à ce que la matrice

$$P := \begin{pmatrix} 0 & 1 & \star \\ 1 & 1 & \star \\ 1 & 0 & \star \end{pmatrix}$$

soit inversible, on aurait alors

$$P^{-1}AP = \begin{pmatrix} 1 & 0 & \star \\ 0 & 2 & \star \\ 0 & 0 & 2 \end{pmatrix}$$

(deux matrices semblables ont les mêmes polynômes caractéristiques).

Ainsi, si on note  $X_1 := \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$  et  $X_2 := \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$ , n'importe quel vecteur colonne  $X_3$  de  $M_{3,1}(\mathbb{R})$

n'appartenant à  $\text{Vect}\{X_1, X_2\}$  convient. On choisit, par exemple,  $X_3 := \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ , on pose alors

$$P := \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

est bien inversible et on a

$$P^{-1}AP = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$$

*Remarque 3.7.6.* Si on avait posé  $X_3 := \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$  et  $P := \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}$ , on aura eu  $P^{-1}AP =$

$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$ . En particulier, on peut donc triangulariser  $A$  de plusieurs façons.

*Remarque 3.7.7.* 1. Triangulariser un endomorphisme triangularisable  $f$  de  $E$ , c'est déterminer une base  $\mathcal{B}$  de  $E$  dans laquelle la matrice représentative de  $f$  est triangulaire et exprimer  $\text{Mat}_{\mathcal{B}}(f)$ .

2. Triangulariser une matrice triangularisable  $A$  de  $M_n(\mathbb{K})$ , c'est déterminer une matrice inversible  $P \in \text{GL}_n(\mathbb{K})$  telle que la matrice  $P^{-1}AP$  soit triangulaire et exprimer  $P^{-1}AP$ .

Si l'endomorphisme  $f$  est triangularisable, il existe des triangularisations plus "simples" que d'autres. Commençons par la triangularisation suivant les sous-espaces caractéristiques, qui est à la fois une triangularisation et une "diagonalisation par blocs" :

**Définition 3.7.8.** On suppose que  $f$  est triangularisable, i.e.  $\chi_f = (-1)^n \prod_{i=1}^p (X - \lambda_i)^{m_{\lambda_i}}$  avec  $\lambda_1, \dots, \lambda_p$  les valeurs propres deux à deux distinctes de  $f$ .

Pour  $i \in \{1, \dots, p\}$ , on appelle sous-espace caractéristique de  $f$  associé à la valeur propre  $\lambda_i$  le sous-espace vectoriel

$$N_{\lambda_i} := \text{Ker} (f - \lambda_i \text{Id}_E)^{m_{\lambda_i}}$$

de  $E$ .

*Remarque 3.7.9.* Avec les notations ci-dessus, pour tout  $i \in \{1, \dots, p\}$ ,  $E_{\lambda_i} \subset N_{\lambda_i}$  et  $f(N_{\lambda_i}) \subset N_{\lambda_i}$ .

*Exemple 3.7.10.* Reprenons la matrice

$$A = \begin{pmatrix} 3 & -1 & 1 \\ 2 & 0 & 1 \\ 1 & -1 & 2 \end{pmatrix} \in M_2(\mathbb{R})$$

de l'exemple 3.7.5. Son polynôme caractéristique est  $\chi_A = -(X-1)(X-2)^2$ . Le sous-espace caractéristique de  $A$  associé à la valeur propre 1 est  $N_1 = \text{Ker}(A - I_3) = E_1$  et le sous-espace caractéristique de  $A$  associé à la valeur propre 2 est  $N_2 = \text{Ker}(A - 2I_3)^2$ . Or

$$(A - 2I_3)^2 = \begin{pmatrix} 0 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & 1 & 0 \end{pmatrix}$$

$$\text{donc } N_2 = \text{Ker}(A - 2I_3)^2 = \text{Vect} \left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

Dans la suite de cette section, on supposera que  $f$  est triangularisable i.e. que son polynôme caractéristique est scindé et on gardera les notations de la définition précédente.

**Proposition 3.7.11.** *Les sous-espaces caractéristiques de  $f$  sont en somme directe et, si l'on reprend les notations de la définition ci-dessus,*

$$E = N_{\lambda_1} \oplus \cdots \oplus N_{\lambda_p}.$$

*Remarque 3.7.12.* Cette proposition est une conséquence du lemme des noyaux : si  $P_1, \dots, P_m \in \mathbb{K}[X]$  sont des polynômes premiers entre eux deux à deux et si  $P := P_1 \cdots P_m$ , alors

$$\text{Ker } P(f) = \text{Ker } P_1(f) \oplus \cdots \oplus \text{Ker } P_m(f).$$

On en vient à une triangularisation sous forme diagonale par blocs de l'endomorphisme triangularisable  $f$ , suivant ses sous-espaces caractéristiques :

**Théorème 3.7.13.** *Pour tout  $i \in \{1, \dots, p\}$ ,  $\dim(N_{\lambda_i}) = m_{\lambda_i}$  et il existe une base  $\mathcal{B}_i$  de  $N_{\lambda_i}$  telle que, si  $\mathcal{B} := \{\mathcal{B}_1, \dots, \mathcal{B}_p\}$ ,  $\text{Mat}_{\mathcal{B}}(f)$  est la matrice-blocs*

$$\begin{pmatrix} M_{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & M_{\lambda_p} \end{pmatrix}$$

où, pour tout  $i \in \{1, \dots, p\}$ ,  $M_{\lambda_i} \in M_{m_{\lambda_i}}(\mathbb{K})$  est la matrice  $\text{Mat}_{\mathcal{B}_i}(f|_{N_{\lambda_i}})$  et est de la forme

$$\begin{pmatrix} \lambda_i & & \star \\ & \ddots & \\ 0 & & \lambda_i \end{pmatrix}.$$

*Exemple 3.7.14.* Reprenons la matrice  $A$  de l'exemple 3.7.10. La famille  $\mathcal{B}_1 := \left\{ \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right\}$  est une

base de  $N_1 = E_1$ , la famille  $\mathcal{B}_2 := \left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$  est une base de  $N_2$  et, si  $P$  désigne la matrice inversible

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix},$$

on a

$$P^{-1}AP = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$$

*Remarque 3.7.15.* Si on avait pris  $\mathcal{B}_2 := \left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 3 \end{pmatrix} \right\}$  et  $P := \begin{pmatrix} 0 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 3 \end{pmatrix}$ , on aurait eu

$$P^{-1}AP = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 3 \\ 0 & 0 & 2 \end{pmatrix}$$

Mais on peut aller encore plus loin dans la simplification de la représentation triangulaire de l'endomorphisme triangularisable  $f$  : cette simplification "ultime" appelée réduction de Jordan est l'objet de la section suivante. En particulier, on décrira un algorithme systématique de triangularisation d'un endomorphisme triangularisable  $f$  sous "sa" forme dite de Jordan.

### 3.8 Réduction de Jordan

Soit  $f$  un endomorphisme triangularisable de  $E$  et soit  $\chi_f = (-1)^n \prod_{i=1}^p (X - \lambda_i)^{m_{\lambda_i}}$  son polynôme caractéristique scindé sur  $\mathbb{K}$ , avec  $\lambda_1, \dots, \lambda_p$  ses valeurs propres deux à deux distinctes.

Nous allons construire, par un processus algorithmique, une base de  $f$  dans laquelle la matrice représentative de  $f$  est une matrice diagonale par blocs dont les blocs diagonaux sont des blocs de Jordan :

**Définition 3.8.1.** Soit  $\lambda \in \mathbb{K}$ . Pour  $m \in \mathbb{N} \setminus \{0\}$ , on appelle  $\lambda$ -bloc de Jordan de taille  $m$  la matrice carrée

$$J_m(\lambda) := \begin{pmatrix} \lambda & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda \end{pmatrix}$$

de  $M_m(\mathbb{K})$  (par convention,  $J_1(\lambda) = (\lambda)$ ).

Pour  $m_1, \dots, m_k \in \mathbb{N} \setminus \{0\}$ , on note  $J_{m_1, \dots, m_k}(\lambda)$  la matrice diagonale par blocs

$$\begin{pmatrix} J_{m_1}(\lambda) & & 0 \\ & \ddots & \\ 0 & & J_{m_k}(\lambda) \end{pmatrix}$$

de  $M_{m_1 + \dots + m_k}(\mathbb{K})$ .

Dans cette section, nous allons précisément montrer le résultat de réduction suivant :

**Théorème 3.8.2.** *Il existe une base  $\mathcal{B}$  de  $E$  et, pour tout  $i \in \{1, \dots, p\}$ , des entiers  $m_1^i, \dots, m_{k_i}^i \in \mathbb{N} \setminus \{0\}$  telle que*

$$\text{Mat}_{\mathcal{B}}(f) = \begin{pmatrix} J_{m_1^1, \dots, m_{k_1}^1}(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & J_{m_1^p, \dots, m_{k_p}^p}(\lambda_p) \end{pmatrix}.$$

De plus, à permutation près, les entiers  $m_j^i$ ,  $1 \leq i \leq p$ ,  $1 \leq j \leq k_i$ , sont uniques et on appelle les blocs de Jordan  $J_{m_j^i}(\lambda_i)$ ,  $1 \leq i \leq p$ ,  $1 \leq j \leq k_i$ , les blocs de Jordan de l'endomorphisme triangularisable  $f$ . La matrice  $\text{Mat}_{\mathcal{B}}(f)$  est appelée la forme de Jordan de  $f$  (“la” forme de Jordan de  $f$  est unique à permutation près des blocs de Jordan).

*Exemple 3.8.3.* Par unicité des blocs de Jordan, la matrice  $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$  obtenue dans l'exemple 3.7.14 est donc la forme de Jordan de la matrice  $A = \begin{pmatrix} 3 & -1 & 1 \\ 2 & 0 & 1 \\ 1 & -1 & 2 \end{pmatrix}$ .

Un résultat de “classification” découlant du théorème précédent est le suivant :

**Corollaire 3.8.4.** *Deux matrices triangularisables de  $M_n(\mathbb{K})$  sont semblables si et seulement si elles ont les mêmes blocs de Jordan.*

Dans le reste de cette section, nous allons montrer l'existence de la réduction de Jordan de  $f$  par un procédé constructif algorithmique. Nous ne montrerons pas l'unicité des blocs de Jordan de  $f$ .

### 3.8.1 Etape 1

La première étape de cette réduction est une réduction suivant les sous-espaces caractéristiques de  $f$ . Pour  $i \in \{1, \dots, p\}$ , soit  $\mathcal{B}_i$  une base de  $N_{\lambda_i}$  et notons  $\mathcal{B}_0 := \{\mathcal{B}_1, \dots, \mathcal{B}_p\}$ . On a alors, comme  $E = N_{\lambda_1} \oplus \dots \oplus N_{\lambda_p}$  (proposition 3.7.11) et comme chaque sous-espace caractéristique de  $f$  est stable par  $f$  (remarque 3.7.9),

$$\text{Mat}_{\mathcal{B}_0}(f) = \begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_p \end{pmatrix}$$

où, pour tout  $i \in \{1, \dots, p\}$ ,  $A_i := \text{Mat}_{\mathcal{B}_i}(f|_{N_{\lambda_i}})$  (cette réduction n'est pas nécessairement une triangularisation comme dans le théorème 3.7.13 : il s'agit simplement d'une "diagonalisation par blocs").

Remarquons à présent que si, pour tout  $i \in \{1, \dots, p\}$ , on trouve une base  $\mathcal{B}'_i$  de  $N_{\lambda_i}$  telle que  $\text{Mat}_{\mathcal{B}'_i}(f|_{N_{\lambda_i}}) = J_{m_1^i, \dots, m_{k_i}^i}(\lambda_i)$  avec  $m_1^i, \dots, m_{k_i}^i \in \mathbb{N} \setminus \{0\}$ , alors, en notant  $\mathcal{B} := \{\mathcal{B}'_1, \dots, \mathcal{B}'_p\}$ , on a

$$\text{Mat}_{\mathcal{B}}(f) = \begin{pmatrix} J_{m_1^1, \dots, m_{k_1}^1}(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & J_{m_1^p, \dots, m_{k_p}^p}(\lambda_p) \end{pmatrix}.$$

On va donc rechercher une telle base  $\mathcal{B}'_i$  pour tout  $i \in \{1, \dots, p\}$ .

### 3.8.2 Etape 2

Soit donc  $\lambda$  une valeur propre de  $f$ . Nous allons montrer qu'il existe une base  $\mathcal{B}'$  de  $N_\lambda$  et des entiers  $m_1, \dots, m_k \in \mathbb{N} \setminus \{0\}$  tels que  $\text{Mat}_{\mathcal{B}'}(f|_{N_\lambda}) = J_{m_1, \dots, m_k}(\lambda)$ . Pour simplifier encore un peu plus les écritures, notons  $N := N_\lambda$ .

On écrit

$$f|_N = \lambda \text{Id}_N + f|_N - \lambda \text{Id}_N.$$

Nous allons en fait montrer qu'il existe une base  $\mathcal{B}'$  de  $N$  et des entiers  $m_1, \dots, m_k \in \mathbb{N} \setminus \{0\}$  tels que  $\text{Mat}_{\mathcal{B}'}(f|_N - \lambda \text{Id}_N) = J_{m_1, \dots, m_k}^0$  où

$$J_{m_1, \dots, m_k}^0 := \begin{pmatrix} J_{m_1}^0 & & 0 \\ & \ddots & \\ 0 & & J_{m_k}^0 \end{pmatrix}$$

et, si  $m \in \mathbb{N} \setminus \{0\}$ ,

$$J_m^0 := \begin{pmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & 0 \end{pmatrix} = J_m(0) \in M_m(\mathbb{K})$$

( $J_{m_1, \dots, m_k}^0 = J_{m_1, \dots, m_k}(0)$  et  $J_1^0 = (0)$ ).

On aura alors

$$\text{Mat}_{\mathcal{B}'}(f|_N) = \text{Mat}_{\mathcal{B}'}(\lambda \text{Id}_N) + \text{Mat}_{\mathcal{B}'}(f|_N - \lambda \text{Id}_N) = \lambda I_{m_\lambda} + J_{m_1, \dots, m_k}^0 = J_{m_1, \dots, m_k}(\lambda)$$

(rappelons que  $m_\lambda = \dim(N_\lambda)$ ).

Pour montrer l'existence d'une telle base  $\mathcal{B}'$ , on commence par remarquer que l'endomorphisme  $u := f|_N - \lambda \text{Id}_N$  de  $N$  est nilpotent, i.e. qu'il existe  $l \in \mathbb{N} \setminus \{0\}$  tel que  $u^l$  est l'endomorphisme nul de  $N$ . En effet, si  $v \in N = \overline{\text{Ker}(f - \lambda \text{Id}_E)}^{m_\lambda}$ , alors

$$u^{m_\lambda}(v) = (f|_N - \lambda \text{Id}_N)^{m_\lambda}(v) = (f - \lambda \text{Id}_E)^{m_\lambda}(v) = 0$$

(à noter que l'indice de nilpotence de  $u$ , i.e. le plus petit entier  $l \in \mathbb{N} \setminus \{0\}$  tel que  $u^l \equiv 0$ , est donc inférieur ou égal à  $m_\lambda$ ).

Nous allons montrer, de façon algorithmique, que tout endomorphisme nilpotent peut être réduit à une forme (de Jordan)  $J_{m_1, \dots, m_k}^0$  avec  $m_1, \dots, m_k \in \mathbb{N} \setminus \{0\}$ .

*Remarque 3.8.5.* Si  $u$  est un endomorphisme nilpotent quelconque de  $E$ , alors 0 est une valeur propre de  $u$  et c'est la seule. En effet, comme il existe  $l \in \mathbb{N} \setminus \{0\}$  tel que  $u^l \equiv 0$ , le polynôme  $X^l$  est un polynôme annulateur de  $u$  et donc le polynôme minimal de  $u$  est de la forme  $\mu_u = X^\nu$  avec  $1 \leq \nu \leq l$ . 0 est donc l'unique valeur propre de  $u$ . De plus, le degré  $\nu$  du polynôme minimal  $X^\nu$  de  $u$  est l'indice de nilpotence de  $u$ .

**Théorème 3.8.6** (Réduction des endomorphismes nilpotents à la forme de Jordan). *Soit  $u$  un endomorphisme nilpotent quelconque de  $E$ . Il existe une base  $\mathcal{B}$  de  $E$  et des entiers  $m_1, \dots, m_k \in \mathbb{N} \setminus \{0\}$  tels que*

$$\text{Mat}_{\mathcal{B}}(u) = J_{m_1, \dots, m_k}^0.$$

*Démonstration.* On prouve le résultat par récurrence sur la dimension. Précisément, on prouve que pour tout  $n \in \mathbb{N} \setminus \{0\}$ , pour tout espace vectoriel  $E$  sur  $\mathbb{K}$  de dimension  $n$ , pour tout endomorphisme nilpotent  $u$  de  $E$ , il existe une base  $\mathcal{B}$  de  $E$  et des entiers  $m_1, \dots, m_k \in \mathbb{N} \setminus \{0\}$  tels que  $\text{Mat}_{\mathcal{B}}(u) = J_{m_1, \dots, m_k}^0$ .

Pour  $n = 1$ , le résultat est vrai. En effet, soit  $E$  un espace vectoriel sur  $\mathbb{K}$  de dimension 1 et soit  $u$  un endomorphisme nilpotent de  $E$ . Soit  $v_0 \in E$  un vecteur engendrant  $E$ . Comme  $u$  est un endomorphisme de  $E$ , il existe  $\alpha \in \mathbb{K}$  tel que  $u(v_0) = \alpha v_0$ . Soit maintenant  $l \in \mathbb{N} \setminus \{0\}$  tel que  $u^l$  soit identiquement nul, alors  $0 = u^l(v_0) = \alpha^l v_0$  et donc  $\alpha = 0$  car  $v_0$  engendre  $E$  qui est de dimension 1. Ainsi, si on note  $\mathcal{B} := \{v_0\}$ , on a  $\text{Mat}_{\mathcal{B}}(u) = (0) = J_1^0$ .

Supposons maintenant le résultat vrai pour tout entier naturel non nul strictement inférieur à un entier  $n \in \mathbb{N} \setminus \{0\}$  fixé, et soient  $E$  un espace vectoriel sur  $\mathbb{K}$  de dimension  $n$  et  $u$  un endomorphisme nilpotent de  $E$ .

On note  $\nu$  l'indice de nilpotence de  $u$ . Si  $\nu = 1$ ,  $u$  est identiquement nul et, dans toute base  $\mathcal{B}$  de  $E$ ,  $\text{Mat}_{\mathcal{B}}(u) = J_{1, \dots, 1}^0$ . Si  $\nu > 1$ ,  $\text{Ker } u^{\nu-1} \neq E$  ( $\nu$  est le plus petit entier naturel non plus tel que  $u^\nu \equiv 0$  donc  $u^{\nu-1}$  n'est pas identiquement nul).

Soit alors  $v \in E \setminus \text{Ker } u^{\nu-1}$  : nous allons montrer que la famille  $\{u^{\nu-1}(v), u^{\nu-2}(v), \dots, u(v), v\}$  est libre. Soient  $\lambda_0, \dots, \lambda_{\nu-1} \in \mathbb{K}$  tels que

$$\lambda_0 v + \lambda_1 u(v) + \dots + \lambda_{\nu-1} u^{\nu-1}(v) = 0_E.$$

En appliquant  $u^{\nu-1}$  à cette égalité, on obtient (puisque  $u^\nu \equiv 0$ )

$$\lambda_0 u^{\nu-1}(v) = 0_E$$

et donc  $\lambda_0 = 0$  car  $u^{\nu-1}(v) \neq 0_E$  par hypothèse. On applique ensuite  $u^{\nu-2}$  à l'égalité

$$\lambda_1 u(v) + \dots + \lambda_{\nu-1} u^{\nu-1}(v) = 0_E$$

pour obtenir  $\lambda_1 u^{\nu-1}(v) = 0_E$  et donc  $\lambda_1 = 0$ . De proche en proche, on obtient ainsi que  $\lambda_0 = \lambda_1 = \dots = \lambda_{\nu-1} = 0$  et la famille  $\mathcal{B}' := \{u^{\nu-1}(v), \dots, u(v), v\}$  est donc libre. Il s'agit donc d'une base du sous-espace vectoriel  $F := \text{Vect} \{u^{\nu-1}(v), \dots, u(v), v\}$  de  $E$ .

Remarquons ensuite que  $F$  est stable par  $u$  (i.e.  $u(F) \subset F$ ) et

$$\text{Mat}_{\mathcal{B}'}(u|_F) = \begin{pmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & 0 \end{pmatrix} = J_{\nu}^0.$$

Si  $\nu = n$ , alors  $F = E$  et  $\text{Mat}_{\mathcal{B}'}(u|_F) = \text{Mat}_{\mathcal{B}'}(u) = J_{\nu}^0$ . Supposons à présent que  $\nu < n$ . Nous allons construire un supplémentaire  $G$  de  $F$  dans  $E$  qui soit également stable par  $u$ . Comme  $\dim(G) < \dim(E)$ , on pourra alors appliquer l'hypothèse de récurrence à  $G$  et  $u|_G$  et considérer une base  $\mathcal{B}''$  de  $G$  et des entiers  $\nu_1, \dots, \nu_l \in \mathbb{N} \setminus \{0\}$  tels que  $\text{Mat}_{\mathcal{B}''}(u|_G) = J_{\nu_1, \dots, \nu_l}^0$ . De sorte que, si  $\mathcal{B} := \{\mathcal{B}', \mathcal{B}''\}$ ,

$$\text{Mat}_{\mathcal{B}}(u) = \begin{pmatrix} \text{Mat}_{\mathcal{B}'}(u|_F) & 0 \\ 0 & \text{Mat}_{\mathcal{B}''}(u|_G) \end{pmatrix} = \begin{pmatrix} J_{\nu}^0 & 0 \\ 0 & J_{\nu_1, \dots, \nu_l}^0 \end{pmatrix} = J_{\nu, \nu_1, \dots, \nu_l}^0.$$

On construit un tel espace  $G$  en utilisant la dualité linéaire. Soit  $\varphi \in E^*$  tel que  $\varphi(u^{\nu-1}(v)) \neq 0$  (un tel  $\varphi$  existe car, sinon,  $u^{\nu-1}(v)$  serait nécessairement le vecteur nul par la proposition 1.3.6 2., ce qui n'est pas le cas par hypothèse sur  $v$ ) et montrons que la famille  $\{\varphi, {}^t u(\varphi), \dots, {}^t u^{\nu-1}(\varphi)\}$  de  $E^*$  est libre ( ${}^t u$  est la transposée de  $u$  : cf définition 1.6.1). Soient  $\mu_0, \dots, \mu_{\nu-1} \in \mathbb{K}$  tels que

$$\mu_0 \varphi + \mu_1 {}^t u(\varphi) + \dots + \mu_{\nu-1} {}^t u^{\nu-1}(\varphi) \equiv 0.$$

On applique cette égalité d'endomorphismes au vecteur  $u^{\nu-1}(v)$  : on obtient

$$\begin{aligned} 0 &= \mu_0 \varphi(u^{\nu-1}(v)) + \mu_1 {}^t u(\varphi)(u^{\nu-1}(v)) + \dots + \mu_{\nu-1} {}^t u^{\nu-1}(\varphi)(u^{\nu-1}(v)) \\ &= \mu_0 \varphi(u^{\nu-1}(v)) + \mu_1 \varphi \circ u(u^{\nu-1}(v)) + \dots + \mu_{\nu-1} \varphi \circ u^{\nu-1}(u^{\nu-1}(v)) \\ &= \mu_0 \varphi(u^{\nu-1}(v)) \end{aligned}$$

(car  $u^{\nu} \equiv 0$ ). Comme  $\varphi(u^{\nu-1}(v)) \neq 0$  par hypothèse, nécessairement  $\mu_0 = 0$ . En appliquant l'égalité

$$\mu_1 {}^t u(\varphi) + \dots + \mu_{\nu-1} {}^t u^{\nu-1}(\varphi) \equiv 0.$$

à  $u^{\nu-2}(v)$ , on déduit ensuite que  $\mu_1 = 0$ . De proche en proche, on obtient ainsi que  $\mu_0 = \mu_1 = \dots = \mu_{\nu-1} = 0$  et la famille  $\mathcal{C} := \{\varphi, {}^t u(\varphi), \dots, {}^t u^{\nu-1}(\varphi)\}$  de  $E^*$  est donc libre. Il s'agit d'une base du sous-espace vectoriel  $W := \text{Vect} \{\varphi, {}^t u(\varphi), \dots, {}^t u^{\nu-1}(\varphi)\}$  de  $E^*$ .

On considère ensuite l'annulateur  $W^0$  de  $W$  (définition 1.5.1). Montrons que  $F \cap W^0 = \{0_E\}$  : soit  $w = \lambda_0 v + \lambda_1 u(v) + \dots + \lambda_{\nu-1} u^{\nu-1}(v)$ ,  $\lambda_0, \dots, \lambda_{\nu-1} \in \mathbb{K}$ , un vecteur de  $F$  annulé par toutes les formes linéaires de  $W$ . En particulier,

$$\begin{aligned} 0 &= {}^t u^{\nu-1}(\varphi)(w) \\ &= \lambda_0 {}^t u^{\nu-1}(\varphi)(v) + \lambda_1 {}^t u^{\nu-1}(\varphi)(u(v)) + \dots + \lambda_{\nu-1} {}^t u^{\nu-1}(\varphi)(u^{\nu-1}(v)) \\ &= \lambda_0 \varphi(u^{\nu-1}(v)) \end{aligned}$$



( $u^\nu \equiv 0$ ) et donc, comme  $\varphi(u^{\nu-1}(v)) \neq 0$ ,  $\lambda_0 = 0$  et  $w = \lambda_1 u(v) + \dots + \lambda_{\nu-1} u^{\nu-1}(v)$ . Le vecteur  $w$  est également annulé par  ${}^t u^{\nu-2}(\varphi)$  et on en déduit de façon analogue que  $\lambda_1 = 0$ . De proche en proche, on obtient ainsi que  $\lambda_0 = \lambda_1 = \dots = \lambda_{\nu-1} = 0$  et donc  $w = 0_E$ . Les sous-espaces vectoriels  $F$  et  $W^0$  de  $E$  sont donc en somme directe.

De plus,  $\dim(W^0) = \dim(E) - \dim(W) = n - \nu$  (proposition 1.5.3) donc  $\dim(F) + \dim(W^0) = \nu + n - \nu = n = \dim(E)$  et  $F$  et  $W^0$  sont donc des sous-espaces vectoriels supplémentaires dans  $E$ .

Montrons enfin que  $W^0$  est stable par  $u$  : soit  $w \in W^0$  et soit  $\psi = \mu_0 \varphi + \mu_1 {}^t u(\varphi) + \dots + \mu_{\nu-1} {}^t u^{\nu-1}(\varphi)$ ,  $\mu_0, \dots, \mu_{\nu-1} \in \mathbb{K}$ , une forme linéaire de  $W$ . Alors

$$\begin{aligned} \psi(u(w)) &= \mu_0 \varphi(u(w)) + \mu_1 {}^t u(\varphi)(u(w)) + \dots + \mu_{\nu-1} {}^t u^{\nu-1}(\varphi)(u(w)) \\ &= \mu_0 {}^t u(\varphi)(w) + \mu_1 {}^t u^2(\varphi)(w) + \dots + \mu_{\nu-2} {}^t u^{\nu-1}(\varphi)(w) + 0 \\ &= 0 \end{aligned}$$

car  $w \in W^0$  et  $W = \text{Vect} \{ \varphi, {}^t u(\varphi), \dots, {}^t u^{\nu-1}(\varphi) \}$ .

On pose alors  $G := W^0$  :  $G$  est un supplémentaire de  $F$  dans  $E$  de dimension  $n - \nu < n$  et stable par  $u$ . La restriction de  $u$  à  $G$  reste nilpotente et, par hypothèse de récurrence, il existe donc une base  $\mathcal{B}''$  de  $G$  et des entiers  $\nu_1, \dots, \nu_l \in \mathbb{N} \setminus \{0\}$  tels que  $\text{Mat}_{\mathcal{B}''}(u|_G) = J_{\nu_1, \dots, \nu_l}^0$ . Ainsi, en posant  $\mathcal{B} := \{ \mathcal{B}', \mathcal{B}'' \}$ , on a

$$\text{Mat}_{\mathcal{B}}(u) = \begin{pmatrix} \text{Mat}_{\mathcal{B}'}(u|_F) & 0 \\ 0 & \text{Mat}_{\mathcal{B}''}(u|_G) \end{pmatrix} = \begin{pmatrix} J_\nu^0 & 0 \\ 0 & J_{\nu_1, \dots, \nu_l}^0 \end{pmatrix} = J_{\nu, \nu_1, \dots, \nu_l}^0.$$

□

### 3.8.3 Description matricielle de la méthode de réduction à la forme de Jordan

Résumons la méthode décrite ci-dessus en appliquant son pendant matriciel à une matrice triangularisable  $A \in M_n(\mathbb{K})$ . On suppose que l'on a déjà écrit le polynôme caractéristique  $\chi_A$  de  $A$  comme un produit

$$\chi_A = (-1)^n \prod_{i=1}^p (X - \lambda_i)^{m_{\lambda_i}}$$

dans  $\mathbb{K}[X]$ , où  $\lambda_1, \dots, \lambda_p$  sont les valeurs propres (deux à deux distinctes) de  $A$ .

**Etape 1 :** Pour tout  $i \in \{1, \dots, p\}$ , calculer la matrice  $(A - \lambda_i I_n)^{m_{\lambda_i}}$  et déterminer une base  $\mathcal{B}_i$  de  $N_{\lambda_i} = \text{Ker} (A - \lambda_i I_n)^{m_{\lambda_i}} \subset M_{n,1}(\mathbb{K})$ . Considérer la base  $\mathcal{B}_0 := \{ \mathcal{B}_1, \dots, \mathcal{B}_p \}$  de  $M_{n,1}(\mathbb{K})$  et la matrice  $P_0$  dont les colonnes sont, dans l'ordre, les vecteurs colonnes de la base  $\mathcal{B}_0$ . Calculer la matrice  $P_0^{-1} A P_0$  : elle est de la forme

$$\begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_p \end{pmatrix}$$

où, pour tout  $i \in \{1, \dots, p\}$ ,  $A_i \in M_{m_{\lambda_i}}(\mathbb{K})$  (et  $\chi_{A_i} = (-1)^{m_{\lambda_i}} (X - \lambda_i)^{m_{\lambda_i}}$ ).

**Etape 2 :** Pour chaque  $i \in \{1, \dots, p\}$ , calculer la matrice  $U_i := A_i - \lambda_i I_{m_{\lambda_i}}$  de  $M_{m_{\lambda_i}}(\mathbb{K})$  puis appliquer à la matrice nilpotente  $U_i$  la méthode décrite ci-après de réduction des matrices nilpotentes à la forme de Jordan : on obtient des entiers  $m_1^i, \dots, m_{k_i}^i \in \mathbb{N} \setminus \{0\}$  et une matrice inversible  $Q_i$  de taille  $m_{\lambda_i}$  tels que  $Q_i^{-1}U_i Q_i = J_{m_1^i, \dots, m_{k_i}^i}^0$

**Etape 3 :** On note

$$\tilde{P} := \begin{pmatrix} Q_1 & & 0 \\ & \ddots & \\ 0 & & Q_p \end{pmatrix} \in \text{GL}_n(\mathbb{K})$$

et alors

$$\begin{aligned} \tilde{P}^{-1} \begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_p \end{pmatrix} \tilde{P} &= \begin{pmatrix} \lambda_1 I_{m_{\lambda_1}} & & 0 \\ & \ddots & \\ 0 & & \lambda_p I_{m_{\lambda_p}} \end{pmatrix} + \begin{pmatrix} J_{m_1^1, \dots, m_{k_1}^1}^0 & & 0 \\ & \ddots & \\ 0 & & J_{m_1^p, \dots, m_{k_p}^p}^0 \end{pmatrix} \\ &= \begin{pmatrix} J_{m_1^1, \dots, m_{k_1}^1}(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & J_{m_1^p, \dots, m_{k_p}^p}(\lambda_p) \end{pmatrix}. \end{aligned}$$

En posant  $P := P_0 \tilde{P}$ , on obtient donc

$$P^{-1}AP = \begin{pmatrix} J_{m_1^1, \dots, m_{k_1}^1}(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & J_{m_1^p, \dots, m_{k_p}^p}(\lambda_p) \end{pmatrix}.$$

L'algorithme récursif permettant de réduire à la forme de Jordan une matrice nilpotente  $U$  quelconque de  $M_m(\mathbb{K})$ ,  $m \in \mathbb{N} \setminus \{0\}$  (la matrice dans la base canonique d'un endomorphisme nilpotent  $u$  de  $\mathbb{K}^m$ ), est le suivant :

**Etape a :** Déterminer l'indice de nilpotence de  $U$  : il s'agit de la plus petite puissance  $\nu \in \mathbb{N} \setminus \{0\}$  telle que  $U^\nu$  est la matrice nulle.

**Etape b :** Choisir un vecteur colonne  $Y \in M_{m,1}(\mathbb{K})$  (le vecteur colonne des coordonnées dans la base canonique d'un vecteur  $v$  de  $\mathbb{K}^m$ ) tel que le vecteur colonne  $U^{\nu-1}Y$  n'est pas nul, et constituer la famille libre  $\{U^{\nu-1}Y, \dots, UY, Y\}$  de  $M_{m,1}(\mathbb{K})$ .

**Etape c :** Si  $\nu = m$ , on note  $Q$  la matrice de  $\text{GL}_m(\mathbb{K})$  dont les colonnes sont, dans l'ordre, les coordonnées des vecteurs colonnes de la base  $\{U^{\nu-1}Y, \dots, UY, Y\}$  de  $M_{m,1}(\mathbb{K})$ , et on a alors

$$Q^{-1}UQ = J_m^0.$$

Si  $\nu \neq m$ , passer à l'étape d.

**Etape d :** Si  $\nu \neq m$ , choisir un vecteur colonne  $Z \in M_{m,1}(\mathbb{K})$  (le vecteur colonne des coordonnées dans la base duale de la base canonique d'une forme linéaire  $\varphi$  de  $(\mathbb{K}^m)^*$ ) tel que la quantité  ${}^t Z U^{\nu-1} Y$  (qui est la quantité  $\varphi(u^{\nu-1}(v))$ ) ne soit pas nulle, et constituer la famille libre  $\{Z, {}^t U Z, \dots, {}^t U^{\nu-1} Z\}$  de  $M_{m,1}(\mathbb{K})$  (correspondant à la famille  $\{\varphi, {}^t u(\varphi), \dots, {}^t u^{\nu-1}(\varphi)\}$  de  $(\mathbb{K}^m)^*$ ).

**Etape e :** Déterminer une famille libre  $\{X_1, \dots, X_{m-n}\}$  de  $M_{m,1}(\mathbb{K})$  telle que, pour tout  $i \in \{1, \dots, m-n\}$  et tout  $j \in \{0, \dots, \nu-1\}$ ,  ${}^t ({}^t U^j Z) X_i = {}^t Z U^j X_i = 0$  (la famille libre  $\{X_1, \dots, X_{m-n}\}$  correspond à une base de l'annulateur de  $\text{Vect}\{\varphi, {}^t u(\varphi), \dots, {}^t u^{\nu-1}(\varphi)\}$ ).

**Etape f :** On note  $Q_0$  la matrice de  $\text{GL}_m(\mathbb{K})$  dont les colonnes sont, dans l'ordre, les coordonnées des vecteurs colonnes de la base  $\{Y, UY, \dots, U^{\nu-1}Y, X_1, \dots, X_{m-n}\}$  de  $M_{m,1}(\mathbb{K})$ . Calculer la matrice  $Q_0^{-1} A Q_0$  : elle est de la forme

$$\begin{pmatrix} J_\nu^0 & 0 \\ 0 & \tilde{U} \end{pmatrix}$$

où  $\tilde{U}$  est une matrice nilpotente de  $M_{m-\nu,1}(\mathbb{K})$ .

**Etape g :** Appliquer la méthode à la matrice nilpotente  $\tilde{U}$  (à partir de l'étape a). En appliquant ce procédé récursif, on obtient une matrice inversible  $\tilde{Q} \in \text{GL}_{m-\nu}(\mathbb{K})$  et des entiers  $\nu_1, \dots, \nu_l \in \mathbb{N} \setminus \{0\}$  tels que  $\tilde{Q}^{-1} \tilde{U} \tilde{Q} = J_{\nu_1, \dots, \nu_l}^0$ .

**Etape h :** On note

$$Q := Q_0 \begin{pmatrix} I_\nu & 0 \\ 0 & \tilde{Q} \end{pmatrix} \in \text{GL}_m(\mathbb{K}),$$

et on a

$$Q^{-1} U Q = J_{\nu, \nu_1, \dots, \nu_l}^0.$$

*Exemple 3.8.7.* On considère la matrice

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

de  $M_4(\mathbb{R})$ .

Étape 0 : On calcule le polynôme caractéristique  $\chi_A$  de  $A$  :

$$\begin{aligned}
 \chi_A = \det(A - XI_4) &= \begin{vmatrix} 1-X & 0 & 0 & 0 \\ -1 & 4-X & 1 & -2 \\ 2 & 1 & 2-X & -1 \\ 1 & 2 & 1 & -X \end{vmatrix} \\
 &= (1-X) \begin{vmatrix} 4-X & 1 & -2 \\ 1 & 2-X & -1 \\ 2 & 1 & -X \end{vmatrix} \\
 &\stackrel{C_1 \leftarrow C_1 + C_3}{=} (1-X) \begin{vmatrix} 2-X & 1 & -2 \\ 0 & 2-X & -1 \\ 2-X & 1 & -X \end{vmatrix} \\
 &= (1-X)(2-X) \begin{vmatrix} 1 & 1 & -2 \\ 0 & 2-X & -1 \\ 1 & 1 & -X \end{vmatrix} \\
 &\stackrel{L_3 \leftarrow L_3 - L_1}{=} (1-X)(2-X) \begin{vmatrix} 1 & 1 & -2 \\ 0 & 2-X & -1 \\ 0 & 0 & 2-X \end{vmatrix} \\
 &= (1-X)(2-X)^3
 \end{aligned}$$

Étape 1 : La multiplicité de la valeur propre 1 dans  $\chi_A$  étant 1,  $N_1 = E_1$ , et

$$E_1 = \text{Ker}(A - I_4) = \text{Ker} \begin{pmatrix} 0 & 0 & 0 & 0 \\ -1 & 3 & 1 & -2 \\ 2 & 1 & 1 & -1 \\ 1 & 2 & 1 & -1 \end{pmatrix} = \text{Vect} \left\{ \begin{pmatrix} 1 \\ 1 \\ -4 \\ -1 \end{pmatrix} \right\}.$$

La multiplicité de la valeur propre 2 dans  $\chi_A$  est 3. Pour déterminer  $N_2 = \text{Ker}(A - 2I_4)^3$ , on commence par calculer la matrice  $(A - 2I_4)^3$ . On a

$$(A - 2I_4)^3 = \begin{pmatrix} -1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 4 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

et donc

$$N_2 = \text{Vect} \left\{ \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

On note  $P_0$  la matrice inversible

$$\begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -4 \\ 0 & 0 & 1 & -1 \end{pmatrix}$$

et on obtient

$$P_0^{-1}AP_0 = \begin{pmatrix} 4 & 1 & -2 & 0 \\ 1 & 2 & -1 & 0 \\ 2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}$$

Etape 2 : Le bloc  $A_2 = (1) \in M_1(\mathbb{R})$  est déjà un bloc de Jordan  $J_1(1)$ .

On note  $U_1$  la matrice

$$A_1 - 2I_3 = \begin{pmatrix} 4 & 1 & -2 \\ 1 & 2 & -1 \\ 2 & 1 & 0 \end{pmatrix} - 2I_3 = \begin{pmatrix} 2 & 1 & -2 \\ 1 & 0 & -1 \\ 2 & 1 & -2 \end{pmatrix} \in M_3(\mathbb{R})$$

et on applique la méthode de réduction à la forme de Jordan des matrices nilpotentes à  $U_1$  :

Etape a : Déterminons l'indice de nilpotence de  $U_1$  : on a

$$U_1^2 = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & -1 \end{pmatrix}$$

et  $U_1^3$  est la matrice nulle de  $M_3(\mathbb{R})$ . Ainsi, l'indice de nilpotence de  $U_1$  est 3.

Etape b : On choisit ensuite un vecteur colonne qui ne soit pas dans le noyau de  $U_1^2$  : on prend par exemple le vecteur colonne  $Y := \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$  et on calcule  $U_1 Y = \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}$  et  $U_1^2 Y = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$ .

Etape c : Comme l'indice de nilpotence de  $U_1$  est égal à la multiplicité de la valeur propre 2 dans  $\chi_A$ , la famille libre  $\{U_1^2 Y, U_1 Y, Y\}$  est une base de  $M_{3,1}(\mathbb{R})$  et on pose

$$Q_1 := \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 0 \\ 1 & 2 & 0 \end{pmatrix}.$$

On a alors

$$Q_1^{-1}U_1 Q_1 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

Étape 3 : On note

$$\tilde{P} := \begin{pmatrix} Q_1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

et

$$P := P_0 \tilde{P} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 2 & 1 & 1 \\ 0 & 1 & 0 & -4 \\ 1 & 2 & 0 & -1 \end{pmatrix},$$

et on a

$$P^{-1}AP = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} J_3(2) & 0 \\ 0 & J_1(1) \end{pmatrix}.$$

*Exemple 3.8.8.* On considère la matrice

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

de  $M_5(\mathbb{R})$ .

Etape 0 : Calculons le polynôme caractéristique  $\chi_A$  de  $A$  :

$$\begin{aligned}
\chi_A = \det(A - XI_5) &= \begin{vmatrix} 5-X & 0 & 4 & -2 & -3 \\ -2 & 3-X & -3 & 2 & 4 \\ 0 & 0 & 3-X & 0 & 0 \\ 0 & 0 & 0 & 3-X & 1 \\ 1 & 0 & 2 & -1 & 1-X \end{vmatrix} \\
&= (3-X) \begin{vmatrix} 5-X & 0 & -2 & -3 \\ -2 & 3-X & 2 & 4 \\ 0 & 0 & 3-X & 1 \\ 1 & 0 & -1 & 1-X \end{vmatrix} \\
&= (3-X)^2 \begin{vmatrix} 5-X & -2 & -3 \\ 0 & 3-X & 1 \\ 1 & -1 & 1-X \end{vmatrix} \\
&\stackrel{C_1 \leftarrow C_1 + C_2}{=} (3-X)^2 \begin{vmatrix} 3-X & -2 & -3 \\ 3-X & 3-X & 1 \\ 0 & -1 & 1-X \end{vmatrix} \\
&= (3-X)^3 \begin{vmatrix} 1 & -2 & -3 \\ 1 & 3-X & 1 \\ 0 & -1 & 1-X \end{vmatrix} \\
&\stackrel{L_2 \leftarrow L_2 - L_1}{=} (3-X)^3 \begin{vmatrix} 1 & -2 & -3 \\ 0 & 5-X & 4 \\ 0 & -1 & 1-X \end{vmatrix} \\
&= (3-X)^3 \begin{vmatrix} 5-X & 4 \\ -1 & 1-X \end{vmatrix} \\
&= (3-X)^3 [(5-X)(1-X) + 4] \\
&= (3-X)^3 (X^2 - 6X + 9) \\
&= (3-X)^5
\end{aligned}$$

3 est donc l'unique valeur propre de  $A$ .

Remarquons que

$$A - 3I_5 = \begin{pmatrix} 2 & 0 & 4 & -2 & -3 \\ -2 & 0 & -3 & 2 & 4 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 2 & -1 & -2 \end{pmatrix}$$

et que l'espace propre  $E_3 = \text{Ker}(A - 3I_5)$  est de dimension 2.

Etape 1 : 3 étant l'unique valeur propre de  $A$ , le sous-espace caractéristique  $N_3 = \text{Ker}(A - 3I_5)^5$  est  $M_{5,1}(\mathbb{R})$  tout entier (on "pose"  $P_0 = I_5$  et  $A_1 := A$ ).

Etape 2 : On note  $U := A - 3I_5$ .

Etape a : On calcule l'indice de nilpotence de  $U$ . On a

$$U^2 = \begin{pmatrix} 1 & 0 & 2 & -1 & -2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 2 & -1 & -2 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

et  $U^3$  est la matrice nulle de  $M_5(\mathbb{R})$  donc l'indice de nilpotence de  $U$  est 3.

Etape b : On choisit à présent un vecteur colonne qui ne soit pas dans le noyau de  $U^2$ , par exemple  $Y := \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$ , puis on calcule  $UY = \begin{pmatrix} 2 \\ -2 \\ 0 \\ 0 \\ 1 \end{pmatrix}$  et  $U^2Y = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}$ .

Etape c : L'indice de nilpotence de  $U$  est strictement inférieur à 5.

Etape d : On choisit un vecteur colonne  $Z \in M_{5,1}(\mathbb{R})$  tel que  ${}^tZU^2Y \neq 0$ , par exemple  $Z := \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$  (attention  $Z$  est "moralement différent de  $Y$ " :  $Z$  correspond à une forme linéaire),

et on calcule  $Z_1 := {}^tUZ = \begin{pmatrix} 2 \\ 0 \\ 4 \\ -2 \\ -3 \end{pmatrix}$  et  $Z_2 := {}^tU^2Z = \begin{pmatrix} 1 \\ 0 \\ 2 \\ -1 \\ -2 \end{pmatrix}$ .

Etape e : On détermine à présent une base du sous-espace vectoriel de  $M_{5,1}(\mathbb{R})$  des vecteurs colonnes  $X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix}$  qui vérifient  ${}^tZX = {}^tZ_1X = {}^tZ_2X = 0$ , i.e.

$$\begin{cases} x_1 & = 0 \\ 2x_1 + 4x_3 - 2x_4 - 3x_5 & = 0 \\ x_1 + 2x_3 - x_4 - 2x_5 & = 0 \end{cases}$$



par exemple la famille  $\left\{ \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 2 \\ 0 \end{pmatrix} \right\}$ .

Etape f : On note

$$Q_0 := \begin{pmatrix} 1 & 2 & 1 & 0 & 0 \\ 0 & -2 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 2 \\ 0 & 1 & 0 & 0 & 0 \end{pmatrix}$$

et on a

$$Q_0^{-1}UQ_0 = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} = J_{3,2}^0.$$

Etape 3 : On note  $P := Q_0$  et on a

$$P^{-1}AP = \begin{pmatrix} 3 & 1 & 0 & 0 & 0 \\ 0 & 3 & 1 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 3 \end{pmatrix} = J_{3,2}(3).$$

*Remarque 3.8.9.* Le fait que, dans l'exemple 3.8.8 ci-dessus, la matrice  $Q_0^{-1}UQ_0$  ait directement été de la forme voulue est un "heureux hasard". Si l'on avait choisi, pour base du sous-espace

vectorel  $\{X \in M_{5,1}(\mathbb{R}) \mid {}^tZX = {}^tZ_1X = {}^tZ_2X\}$  de  $M_{5,1}(\mathbb{R})$ , la famille  $\left\{ \begin{pmatrix} 0 \\ 1 \\ 1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 2 \\ 0 \end{pmatrix} \right\}$ , on aurait

posé

$$Q_0 := \begin{pmatrix} 1 & 2 & 1 & 0 & 0 \\ 0 & -2 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 2 & 2 \\ 0 & 1 & 0 & 0 & 0 \end{pmatrix}.$$

Pour cette matrice  $Q_0$ , on a

$$Q_0^{-1}UQ_0 = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & -1 & -1 \end{pmatrix} = \begin{pmatrix} J_3^0 & 0 \\ 0 & \tilde{U} \end{pmatrix}.$$

Dans ce cas, on passe alors à l'étape g : on applique le procédé récursif à la matrice nilpotente  $\tilde{U} = \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}$  de  $M_2(\mathbb{R})$  : l'ordre de nilpotence de  $\tilde{U}$  est 2, le vecteur  $\tilde{Y} := \begin{pmatrix} 1 \\ 0 \end{pmatrix}$  n'est pas dans son noyau et  $\tilde{U}\tilde{Y} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ . Si l'on note  $\tilde{Q}$  la matrice  $\begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix}$ , on a  $\tilde{Q}^{-1}\tilde{U}\tilde{Q} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = J_2^0$ .

On passe ensuite à l'étape h : on note  $Q := Q_0 \begin{pmatrix} I_3 & 0 \\ 0 & \tilde{Q} \end{pmatrix}$  et on a

$$\begin{aligned} Q^{-1}UQ &= \begin{pmatrix} J_3^0 & 0 \\ 0 & J_2^0 \end{pmatrix} = J_{3,2}^0 \\ &= \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \end{aligned}$$

A l'étape 3 de la méthode appliquée à  $A$ , on note alors  $P := Q$  et on a

$$P^{-1}AP = \begin{pmatrix} 3 & 1 & 0 & 0 & 0 \\ 0 & 3 & 1 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 3 \end{pmatrix}.$$

*Remarque 3.8.10.* La réduction de Jordan permet entre autres de calculer les puissances successives d'une matrice triangularisable. Précisément, soit  $A$  une matrice triangularisable de  $M_n(\mathbb{K})$ , soient  $\lambda_1, \dots, \lambda_p$  les valeurs propres deux à deux distinctes de  $A$  et soient  $P \in GL_n(\mathbb{K})$  et  $m_1^i, \dots, m_{k_i}^i \in \mathbb{N} \setminus \{0\}$ ,  $i \in \{1, \dots, p\}$ , tels que  $P^{-1}AP$  soit de forme de Jordan

$$\begin{pmatrix} J_{m_1^1, \dots, m_{k_1}^1}(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & J_{m_1^p, \dots, m_{k_p}^p}(\lambda_p) \end{pmatrix} = \begin{pmatrix} I_{m_{\lambda_1}} & & 0 \\ & \ddots & \\ 0 & & I_{m_{\lambda_p}} \end{pmatrix} + \begin{pmatrix} J_{m_1^1, \dots, m_{k_1}^1}^0 & & 0 \\ & \ddots & \\ 0 & & J_{m_1^p, \dots, m_{k_p}^p}^0 \end{pmatrix}.$$

Comme les deux matrices de cette dernière somme commutent, on peut calculer  $(P^{-1}AP)^k = P^{-1}A^kP$  – et donc  $A^k$  – pour tout  $k \in \mathbb{N}$  à l'aide du binôme de Newton. De plus, la nilpotence de la matrice de droite simplifie l'expression du développement.

*Exemple 3.8.11.* Reprenons la matrice  $A = \begin{pmatrix} 3 & -1 & 1 \\ 2 & 0 & 1 \\ 1 & -1 & 2 \end{pmatrix} \in M_3(\mathbb{R})$  de l'exemple 3.7.14. Pour

$P := \begin{pmatrix} 0 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \in GL_3(\mathbb{R})$ , on a

$$P^{-1}AP = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

Soit  $k \in \mathbb{N}$ . Comme les matrices  $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$  et  $\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$  commutent, on a

$$\begin{aligned}
 P^{-1}A^kP &= \sum_{i=0}^k \binom{i}{k} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}^{k-i} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}^i \\
 &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}^k + k \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}^{k-1} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \quad (\text{car la matrice } \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}^2 \text{ est nulle)} \\
 &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2^k & 0 \\ 0 & 0 & 2^k \end{pmatrix} + k \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 2^{k-1} \\ 0 & 0 & 0 \end{pmatrix} \\
 &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2^k & k2^{k-1} \\ 0 & 0 & 2^k \end{pmatrix}
 \end{aligned}$$

Enfin,  $P^{-1} = \begin{pmatrix} -1 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & -1 & 1 \end{pmatrix}$  donc

$$A^k = P \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2^k & k2^{k-1} \\ 0 & 0 & 2^k \end{pmatrix} P^{-1} = \begin{pmatrix} 2^k + k2^{k-1} & -k2^{k-1} & k2^{k-1} \\ 2^k + k2^{k-1} - 1 & -k2^{k-1} + 1 & k2^{k-1} \\ 2^k - 1 & -2^k + 1 & 2^k \end{pmatrix}.$$



# Chapitre 4

## Exponentielle de matrices

### 4.1 Introduction

On introduit dans ce chapitre une généralisation de la fonction exponentielle aux espaces de matrices. Cette “exponentielle de matrices” permet notamment de résoudre les systèmes différentiels linéaires du premier ordre à coefficients constants, et peut se calculer à l’aide de la réduction de Jordan.

Dans tout ce chapitre,  $\mathbb{K}$  désigne les corps  $\mathbb{R}$  ou  $\mathbb{C}$ , et  $m$  est un entier naturel non nul.

### 4.2 Norme de matrices

On commence ce chapitre en munissant le  $\mathbb{K}$ -espace vectoriel  $M_m(\mathbb{K})$  d’une norme. Tout d’abord, si  $p, q \in \mathbb{N} \setminus \{0\}$  et si  $M = (m_{ij})_{1 \leq i \leq p, 1 \leq j \leq q} \in M_{p,q}(\mathbb{C})$ , on note  $\bar{M}$  la matrice  $(\overline{m_{ij}})_{1 \leq i \leq p, 1 \leq j \leq q}$  de  $M_{p,q}(\mathbb{C})$ .

**Définition et Proposition 4.2.1.** *Pour tout  $A = (a_{ij})_{1 \leq i, j \leq m} \in M_m(\mathbb{K})$ , on définit*

$$\|A\| := \sqrt{\sum_{1 \leq i, j \leq m} |a_{ij}|^2} = \sqrt{\text{Tr}({}^t \bar{A} A)} \in [0, +\infty[.$$

L’application  $\|\cdot\| : \begin{array}{ccc} M_m(\mathbb{K}) & \rightarrow & [0, +\infty[ \\ A & \mapsto & \|A\| \end{array}$  est une norme et le couple  $(M_m(\mathbb{K}), \|\cdot\|)$  est donc un  $\mathbb{K}$ -espace vectoriel normé.

*Démonstration.* Si  $\mathbb{K} = \mathbb{R}$ , l’application  $\|\cdot\| : \begin{array}{ccc} M_m(\mathbb{R}) & \rightarrow & [0, +\infty[ \\ A & \mapsto & \|A\| = \sqrt{\text{Tr}({}^t A A)} \end{array}$  est la norme euclidienne associée au produit scalaire  $\langle \cdot, \cdot \rangle : \begin{array}{ccc} M_m(\mathbb{R}) \times M_m(\mathbb{R}) & \rightarrow & \mathbb{R} \\ (A, B) & \mapsto & \text{Tr}({}^t AB) \end{array}$  défini dans l’exemple 2.2.2 3.

Si  $\mathbb{K} = \mathbb{C}$ , l'application  $\|\cdot\| : M_m(\mathbb{C}) \rightarrow [0, +\infty[$  est la norme induite de façon analogue par l'application  $M_m(\mathbb{C}) \times M_m(\mathbb{C}) \rightarrow \mathbb{C}$  qui est un produit scalaire hermitien.

$$\begin{aligned} & \begin{array}{ccc} M_m(\mathbb{C}) & \rightarrow & [0, +\infty[ \\ A & \mapsto & \|A\| = \sqrt{\text{Tr}({}^t\bar{A}A)} \end{array} \\ & \begin{array}{ccc} M_m(\mathbb{C}) \times M_m(\mathbb{C}) & \rightarrow & \mathbb{C} \\ (A, B) & \mapsto & \text{Tr}({}^t\bar{A}B) \end{array} \end{aligned}$$

□

*Remarque 4.2.2.* • Si  $E$  est un  $\mathbb{C}$ -espace vectoriel, une application  $\langle \cdot, \cdot \rangle : E \times E \rightarrow \mathbb{C}$  est appelée produit scalaire hermitien si

1. pour tous  $v_1, v_2, w \in E$  et tous  $\lambda, \mu \in \mathbb{C}$ ,  $\langle \lambda v_1 + \mu v_2, w \rangle = \lambda \langle v_1, w \rangle + \mu \langle v_2, w \rangle$ ,
2. pour tous  $v, w \in E$ ,  $\langle w, v \rangle = \overline{\langle v, w \rangle}$  (en particulier, pour tout  $v \in E$ ,  $\langle v, v \rangle \in \mathbb{R}$ ),
3. pour tout  $v \in E$ ,  $\langle v, v \rangle \geq 0$ , et  $\langle v, v \rangle = 0$  si et seulement si  $v = 0_E$ ,

et, dans ce cas, elle induit une application  $\|\cdot\| : E \rightarrow [0; +\infty[$  ;  $v \mapsto \langle v, v \rangle$  qui vérifie l'inégalité de Cauchy-Schwarz et est une norme, appelée norme hermitienne.

- Soit  $n \in \mathbb{N} \setminus \{0\}$ , pour tous  $v = (x_1, \dots, x_n)$  et  $w = (y_1, \dots, y_n)$  dans  $\mathbb{C}^n$ , on définit

$$\langle v, w \rangle_{\text{can}} := x_1 \bar{y}_1 + \dots + x_n \bar{y}_n = \sum_{i=1}^n x_i \bar{y}_i.$$

L'application  $\langle \cdot, \cdot \rangle_{\text{can}} : \mathbb{C}^n \times \mathbb{C}^n \rightarrow \mathbb{C}$  est alors un produit scalaire hermitien sur  $\mathbb{C}^n$ , appelé produit scalaire hermitien canonique sur  $\mathbb{C}^n$ .

Cette norme  $\|\cdot\|$  sur  $M_m(\mathbb{K})$  possède une propriété qui n'est, en général, pas vérifiée par la norme d'un espace vectoriel normé quelconque :

**Lemme 4.2.3.** Soient  $A, B \in M_m(\mathbb{K})$ . On a  $\|AB\| \leq \|A\| \|B\|$ .

*Démonstration.* On note  $A = (a_{ij})_{1 \leq i, j \leq m}$  et  $B = (b_{ij})_{1 \leq i, j \leq m}$ . Pour  $i, j \in \{1, \dots, m\}$ , on note également  $v_i := (a_{i1}, \dots, a_{im})$ ,  $w_j := (b_{1j}, \dots, b_{mj}) \in \mathbb{K}^m$  (il s'agit respectivement de la  $i^{\text{ème}}$  ligne de  $A$  et de la transposée de la  $j^{\text{ème}}$  colonne de  $B$ ). On a alors

$$\begin{aligned} \|AB\|^2 &= \sum_{1 \leq i, j \leq m} \left| \sum_{k=1}^m a_{ik} b_{kj} \right|^2 \\ &= \sum_{1 \leq i, j \leq m} |\langle v_i, \bar{w}_j \rangle_{\text{can}}|^2 \\ &\leq \sum_{1 \leq i, j \leq m} \|v_i\|_{\text{can}}^2 \|\bar{w}_j\|_{\text{can}}^2 \\ &= \sum_{1 \leq i, j \leq m} \left( \sum_{k=1}^m |a_{ik}|^2 \right) \left( \sum_{l=1}^m |\bar{b}_{lj}|^2 \right) \\ &= \left( \sum_{1 \leq i, k \leq m} |a_{ik}|^2 \right) \left( \sum_{1 \leq j, l \leq m} |b_{lj}|^2 \right) \\ &= \|A\|^2 \|B\|^2 \end{aligned}$$

□

*Remarque 4.2.4.* En particulier, si  $A \in M_m(\mathbb{K})$  et  $n \in \mathbb{N}$ ,  $\|A^n\| \leq \|A\|^n$ .

En tant qu'espace vectoriel normé, on peut définir sur  $(M_m(\mathbb{K}), \|\cdot\|)$  une notion de limite et donc également une notion de continuité. A titre d'exemples, si  $A \in M_m(\mathbb{K})$ , les applications  $M_m(\mathbb{K}) \rightarrow M_m(\mathbb{K}) ; M \mapsto AM$  et  $M_m(\mathbb{K}) \rightarrow M_m(\mathbb{K}) ; M \mapsto MA$  sont continues : en effet, si  $M_0, M \in M_m(\mathbb{K})$ ,

$$\|AM - AM_0\| = \|A(M - M_0)\| \leq \|A\| \|M - M_0\|$$

et

$$\|MA - M_0A\| = \|(M - M_0)A\| \leq \|M - M_0\| \|A\|.$$

On peut également définir une notion de convergence pour les suites et les séries de matrices. Nous allons, dans la section suivante, associer à toute matrice de  $M_n(\mathbb{K})$  la somme d'une série de matrices absolument convergente, dont l'expression généralise le développement en série entière de la fonction exponentielle sur  $\mathbb{R}$  (et  $\mathbb{C}$ ).

### 4.3 Définition et propriétés de base de l'exponentielle de matrices

**Proposition et Définition 4.3.1.** Soit  $A \in M_m(\mathbb{K})$ . La série  $\sum_n \frac{A^n}{n!}$  est absolument convergente (en particulier convergente) et on note  $\exp(A) := \sum_{n=0}^{+\infty} \frac{A^n}{n!}$  sa somme.

*Démonstration.* Rappelons tout d'abord que la notation  $\sum_n \frac{A^n}{n!}$  désigne la suite des sommes partielles  $\left( \sum_{n=0}^k \frac{A^n}{n!} \right)_{k \in \mathbb{N}}$ . Une série est dite convergente si sa suite des sommes partielles est convergente et, dans ce cas, on appelle somme de la série la limite de la suite des sommes partielles.

Nous allons montrer que la série  $\sum_n \frac{A^n}{n!}$  est absolument convergente i.e. que la série numérique  $\sum_n \left\| \frac{A^n}{n!} \right\|$  est convergente. Rappelons qu'une série absolument convergente est en particulier convergente.

Pour tout  $n \in \mathbb{N}$ , on a  $0 \leq \left\| \frac{A^n}{n!} \right\| \leq \frac{\|A\|^n}{n!}$  (par le lemme 4.2.3), or la série (numérique)  $\sum_n \frac{\|A\|^n}{n!}$  est (absolument) convergente (sa somme est l'exponentielle de  $\|A\|$ ), donc la série  $\sum_n \left\| \frac{A^n}{n!} \right\|$  converge également. Ainsi, la série  $\sum_n \frac{A^n}{n!}$  est absolument convergente. □

**Définition 4.3.2.** On appelle l'application

$$\begin{array}{ccc} M_n(\mathbb{K}) & \rightarrow & M_n(\mathbb{K}) \\ \exp : & A & \mapsto \exp(A) = \sum_{n=0}^{+\infty} \frac{A^n}{n!} \end{array}$$

l'application exponentielle de  $M_n(\mathbb{K})$ .

*Remarque 4.3.3.* • Si  $A \in M_n(\mathbb{K})$ , on note également  $e^A := \exp(A)$ .

- Pour  $m = 1$ , on retrouve l'expression (du développement en série entière) de la fonction exponentielle  $\exp : \mathbb{K} \rightarrow \mathbb{K}$ .

*Exemple 4.3.4.* 1. On calcule l'exponentielle d'une matrice diagonale quelconque de  $M_m(\mathbb{K})$  :

soient  $a_1, \dots, a_m \in \mathbb{K}$ , alors, comme pour tout  $n \in \mathbb{N}$ ,

$$\begin{pmatrix} a_1 & & 0 \\ & \ddots & \\ 0 & & a_m \end{pmatrix}^n = \begin{pmatrix} a_1^n & & 0 \\ & \ddots & \\ 0 & & a_m^n \end{pmatrix},$$

on a

$$\begin{aligned} \exp \begin{pmatrix} a_1 & & 0 \\ & \ddots & \\ 0 & & a_m \end{pmatrix} &= \sum_{n=0}^{+\infty} \begin{pmatrix} \frac{a_1^n}{n!} & & 0 \\ & \ddots & \\ 0 & & \frac{a_m^n}{n!} \end{pmatrix} \\ &= \begin{pmatrix} \sum_{n=0}^{+\infty} \frac{a_1^n}{n!} & & 0 \\ & \ddots & \\ 0 & & \sum_{n=0}^{+\infty} \frac{a_m^n}{n!} \end{pmatrix} \\ &= \begin{pmatrix} e^{a_1} & & 0 \\ & \ddots & \\ 0 & & e^{a_m} \end{pmatrix} \end{aligned}$$

En particulier, si  $\lambda \in \mathbb{K}$ ,  $\exp(\lambda I_m) = e^\lambda I_m$ ,  $\exp(I_m) = e I_m$  et, si  $0_m$  désigne la matrice nulle de  $M_m(\mathbb{K})$ ,  $\exp(0_m) = I_m$ .

2. De manière analogue, si  $A$  est une matrice diagonale par blocs de la forme  $A = \begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_r \end{pmatrix}$

de  $M_m(\mathbb{K})$ , alors, pour tout  $n \in \mathbb{N}$ ,  $A^n = \begin{pmatrix} A_1^n & & 0 \\ & \ddots & \\ 0 & & A_r^n \end{pmatrix}$  et donc

$$e^A = \begin{pmatrix} e^{A_1} & & 0 \\ & \ddots & \\ 0 & & e^{A_r} \end{pmatrix}.$$



3. Soit  $J \in M_m(\mathbb{K})$  une matrice nilpotente d'indice de nilpotence  $\nu \in \mathbb{N} \setminus \{0\}$ . Alors, pour tout  $n \in \mathbb{N}$ , si  $n \geq \nu$ ,  $J^n = 0_m$  et on a donc

$$e^J = \sum_{n=0}^{\nu} \frac{J^n}{n!} = I_m + J + \frac{J^2}{2!} + \cdots + \frac{J^{\nu-1}}{(\nu-1)!}.$$

Par exemple, si  $J$  désigne la matrice nilpotente  $\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$  ( $= J_3^0$ ) de  $M_3(\mathbb{R})$ , on a  $J^2 =$

$$\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, J^3 = 0_3 \text{ et donc}$$

$$\begin{aligned} e^J &= I_3 + J + \frac{J^2}{2!} \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 1 & \frac{1}{2} \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}. \end{aligned}$$

Avant d'énoncer la proposition suivante, on rappelle un résultat relatif au produit de Cauchy de séries d'un espace vectoriel normé, appliqué ici à  $(M_m(\mathbb{K}), \|\cdot\|)$  : si  $\sum_n A_n$  et  $\sum_n B_n$  sont deux séries absolument convergentes de  $M_m(\mathbb{K})$ , alors la série  $\sum_n (\sum_{k=0}^n A_k B_{n-k})$  est absolument convergente et a pour somme  $(\sum_{n=0}^{+\infty} A_n) (\sum_{n=0}^{+\infty} B_n)$ .

**Proposition 4.3.5.** *Soient  $A, B \in M_m(\mathbb{K})$  et supposons que  $AB = BA$ . Alors*

$$e^{A+B} = e^A e^B.$$

*Démonstration.* Comme  $A$  et  $B$  commutent, on peut appliquer la formule du binôme de Newton : pour tout  $n \in \mathbb{N}$ , on a  $(A+B)^n = \sum_{k=0}^n \binom{n}{k} A^k B^{n-k}$  et donc

$$\begin{aligned} e^{A+B} &= \sum_{n=0}^{+\infty} \frac{(A+B)^n}{n!} = \sum_{n=0}^{+\infty} \sum_{k=0}^n \frac{1}{n!} \binom{n}{k} A^k B^{n-k} \\ &= \sum_{n=0}^{+\infty} \sum_{k=0}^n \frac{1}{k!(n-k)!} A^k B^{n-k} \\ &= \sum_{n=0}^{+\infty} \sum_{k=0}^n \frac{A^k}{k!} \frac{B^{n-k}}{(n-k)!} \\ &= \left( \sum_{n=0}^{+\infty} \frac{A^n}{n!} \right) \left( \sum_{n=0}^{+\infty} \frac{B^n}{n!} \right) \text{ (les deux séries sont absolument convergentes)} \\ &= e^A e^B \end{aligned}$$

□

*Remarque 4.3.6.* Si  $A, B \in M_m(\mathbb{K})$  commutent, on a  $e^A e^B = e^{A+B} = e^{B+A} = e^B e^A$ , en particulier les matrices  $e^A$  et  $e^B$  commutent également.

**Corollaire 4.3.7.** Soit  $A \in M_m(\mathbb{K})$ . La matrice  $e^A \in M_m(\mathbb{K})$  est inversible et  $(e^A)^{-1} = e^{-A}$ .

*Démonstration.* Les matrices  $A$  et  $-A$  commutent :  $A(-A) = A(-I_m)A = (-A)A$ . On a donc, par la proposition précédente,  $e^{-A} e^A = e^A e^{-A} = e^{A+(-A)} = e^{0_m} = I_m$ . □

Soit  $A \in M_m(\mathbb{K})$  et soit  $P \in GL_m(\mathbb{K})$  une matrice inversible. L'exponentielle de la conjuguée de  $A$  par  $P$  est la conjuguée de  $\exp(A)$  par  $P$  :

**Proposition 4.3.8.** On a

$$\exp(P^{-1}AP) = P^{-1}e^A P.$$

*Démonstration.*  $\exp(P^{-1}AP)$  est la somme de la série  $\sum_n \frac{(P^{-1}AP)^n}{n!}$ . Or, si  $k \in \mathbb{N}$ ,

$$\sum_{n=0}^k \frac{(P^{-1}AP)^n}{n!} = \sum_{n=0}^k P^{-1} \frac{A^n}{n!} P = P^{-1} \left( \sum_{n=0}^k \frac{A^n}{n!} \right) P,$$

donc

$$\begin{aligned} \exp(P^{-1}AP) &= \lim_{k \rightarrow +\infty} \sum_{n=0}^k \frac{(P^{-1}AP)^n}{n!} \\ &= \lim_{k \rightarrow +\infty} P^{-1} \left( \sum_{n=0}^k \frac{A^n}{n!} \right) P \\ &= P^{-1} \left( \lim_{k \rightarrow +\infty} \left( \sum_{n=0}^k \frac{A^n}{n!} \right) \right) P \quad (\text{l'application } \begin{array}{ccc} M_m(\mathbb{K}) & \rightarrow & M_m(\mathbb{K}) \\ M & \mapsto & P^{-1}MP \end{array} \text{ est continue}) \\ &= P^{-1}e^A P. \end{aligned}$$

□

Cette “compatibilité” de l'exponentielle de matrices avec le changement de base nous permet en particulier de montrer l'égalité suivante :

**Corollaire 4.3.9.** On a

$$\det(e^A) = e^{\text{Tr}(A)}.$$

*Démonstration.* Considérons  $A$  comme une matrice de  $M_m(\mathbb{C})$ . En tant que telle, elle est triangularisable dans  $M_m(\mathbb{C})$  (corollaire 3.7.4) : il existe une matrice  $Q \in M_m(\mathbb{C})$  et une matrice

triangulaire  $B = \begin{pmatrix} \lambda_1 & & \star \\ & \ddots & \\ 0 & & \lambda_m \end{pmatrix} \in M_m(\mathbb{C})$  telles que  $B = Q^{-1}AQ$  (les scalaires  $\lambda_1, \dots, \lambda_m \in \mathbb{C}$

sont les valeurs propres complexes, non nécessairement deux à deux distinctes, de  $A$ ).

Remarquons que, pour tout  $n \in \mathbb{N}$ ,  $B^n = \begin{pmatrix} \lambda_1^n & & \star \\ & \ddots & \\ 0 & & \lambda_m^n \end{pmatrix}$  (le produit de deux matrices triangulaires supérieures reste triangulaire supérieur) et donc

$$\begin{aligned} e^B &= \sum_{n=0}^{+\infty} \frac{B^n}{n!} \\ &= \sum_{n=0}^{+\infty} \begin{pmatrix} \frac{\lambda_1^n}{n!} & & \star \\ & \ddots & \\ 0 & & \frac{\lambda_m^n}{n!} \end{pmatrix} \\ &= \begin{pmatrix} \sum_{n=0}^{+\infty} \frac{\lambda_1^n}{n!} & & \star \\ & \ddots & \\ 0 & & \sum_{n=0}^{+\infty} \frac{\lambda_m^n}{n!} \end{pmatrix} \\ &= \begin{pmatrix} e^{\lambda_1} & & \star \\ & \ddots & \\ 0 & & e^{\lambda_m} \end{pmatrix} \end{aligned}$$

Ainsi, comme  $A = QBQ^{-1}$ , on a

$$\begin{aligned} e^A &= Qe^BQ^{-1} \quad (\text{par la proposition précédente}) \\ &= Q \begin{pmatrix} e^{\lambda_1} & & \star \\ & \ddots & \\ 0 & & e^{\lambda_m} \end{pmatrix} Q^{-1} \end{aligned}$$

et donc

$$\det(e^A) = \det \begin{pmatrix} e^{\lambda_1} & & \star \\ & \ddots & \\ 0 & & e^{\lambda_m} \end{pmatrix} = \prod_{j=1}^m e^{\lambda_j} = e^{\sum_{j=1}^m \lambda_j} = e^{\text{Tr}(B)} = e^{\text{Tr}(A)}$$

(la trace est invariante par changement de base).  $\square$

## 4.4 Calcul de l'exponentielle d'une matrice via la réduction de Jordan

Soit  $A$  une matrice de  $M_m(\mathbb{K})$ . Nous allons détailler une méthode pour calculer l'exponentielle de  $A$  à partir de sa réduction de Jordan en tant que matrice de  $M_m(\mathbb{C})$ .

Soient  $\lambda_1, \dots, \lambda_p \in \mathbb{C}$  les valeurs propres complexes deux à deux distinctes de  $A$ . D'après le théorème 3.8.2, il existe une matrice inversible  $P \in GL_m(\mathbb{C})$  et, pour tout  $i \in \{1, \dots, p\}$ , des entiers  $m_1^i, \dots, m_{k_i}^i \in \mathbb{N} \setminus \{0\}$  tels que

$$P^{-1}AP = \begin{pmatrix} J_{m_1, \dots, m_{k_1}}^1(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & J_{m_1^p, \dots, m_{k_p}^p}(\lambda_p) \end{pmatrix}.$$

Alors

$$\begin{aligned} \exp(A) &= \exp \left( P \begin{pmatrix} J_{m_1, \dots, m_{k_1}}^1(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & J_{m_1^p, \dots, m_{k_p}^p}(\lambda_p) \end{pmatrix} P^{-1} \right) \\ &= P \exp \begin{pmatrix} J_{m_1, \dots, m_{k_1}}^1(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & J_{m_1^p, \dots, m_{k_p}^p}(\lambda_p) \end{pmatrix} P^{-1} \quad (\text{par proposition 4.3.8}) \\ &= P \begin{pmatrix} \exp \left( J_{m_1, \dots, m_{k_1}}^1(\lambda_1) \right) & & 0 \\ & \ddots & \\ 0 & & \exp \left( J_{m_1^p, \dots, m_{k_p}^p}(\lambda_p) \right) \end{pmatrix} P^{-1} \quad (\text{exemple 4.3.4 3.}). \end{aligned}$$

Ainsi, pour pouvoir calculer  $\exp(A)$ , il nous suffit de savoir calculer  $\exp(J_{m_1, \dots, m_k}(\lambda))$  pour tous  $k \in \mathbb{N} \setminus \{0\}$ ,  $m_1, \dots, m_k \in \mathbb{N} \setminus \{0\}$  et  $\lambda \in \mathbb{C}$ . Soient donc  $k \in \mathbb{N} \setminus \{0\}$ ,  $m_1, \dots, m_k \in \mathbb{N} \setminus \{0\}$  et  $\lambda \in \mathbb{C}$ . Notons  $m_0 := m_1 + \dots + m_k$ . On a  $J_{m_1, \dots, m_k}(\lambda) = \lambda I_{m_0} + J_{m_1, \dots, m_k}^0$ . Or les matrices  $\lambda I_{m_0}$  et  $J_{m_1, \dots, m_k}^0$  commutent et la matrice  $J_{m_1, \dots, m_k}^0$  est nilpotente. Si on note  $\nu$  l'indice de nilpotence de cette dernière matrice que l'on note simplement  $J$ , on a alors

$$\begin{aligned} \exp(J_{m_1, \dots, m_k}(\lambda)) &= \exp(\lambda I_{m_0} + J) \\ &= \exp(\lambda I_{m_0}) \exp(J) \\ &= e^\lambda I_{m_0} \left( I_{m_0} + J + \frac{J^2}{2!} + \dots + \frac{J^{\nu-1}}{(\nu-1)!} \right) \\ &= e^\lambda \left( I_{m_0} + J + \frac{J^2}{2!} + \dots + \frac{J^{\nu-1}}{(\nu-1)!} \right) \end{aligned}$$

*Remarque 4.4.1.* Si  $A \in M_m(\mathbb{R})$ , alors  $e^A \in M_m(\mathbb{R})$  : même si l'on considère la réduction de Jordan "complexe" de  $A$  pour calculer  $e^A$ , la matrice que l'on obtient à la fin du calcul ne possède que des coefficients réels.

*Exemple 4.4.2.* 1. On considère la matrice  $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$  de  $M_2(\mathbb{R})$ . Son polynôme caractéristique est  $\chi_A = X^2 + 1$  qui est scindé à racines simples  $i$  et  $-i$  sur  $\mathbb{C}$  : si on note  $P := \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix}$ , on a

$$P^{-1}AP = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$$

(il s'agit de la réduction de Jordan de  $A$ ) et

$$\exp(A) = P \exp \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} P^{-1} = \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix} \begin{pmatrix} e^i & 0 \\ 0 & e^{-i} \end{pmatrix} \begin{pmatrix} \frac{1}{2} & \frac{-i}{2} \\ \frac{1}{2} & \frac{i}{2} \end{pmatrix} = \frac{1}{2} \begin{pmatrix} e^i + e^{-i} & -ie^i + ie^{-i} \\ ie^i - ie^{-i} & e^i + e^{-i} \end{pmatrix}.$$

Or,

$$e^i + e^{-i} = \sum_{k=0}^{+\infty} \frac{i^k}{k!} + \sum_{k=0}^{+\infty} \frac{(-i)^k}{k!} = \sum_{k=0}^{+\infty} \frac{i^k + (-i)^k}{k!} = \sum_{p=0}^{+\infty} \frac{2(-1)^p}{(2p)!} = 2 \sum_{p=0}^{+\infty} \frac{(-1)^p}{(2p)!} = 2 \cos(1)$$

et

$$ie^i - ie^{-i} = i \sum_{k=0}^{+\infty} \frac{i^k}{k!} - i \sum_{k=0}^{+\infty} \frac{(-i)^k}{k!} = \sum_{k=0}^{+\infty} \frac{i^{k+1} + (-i)^{k+1}}{k!} = \sum_{p=0}^{+\infty} \frac{2(-1)^{p+1}}{(2p+1)!} = -2 \sum_{p=0}^{+\infty} \frac{(-1)^p}{(2p+1)!} = -2 \sin(1).$$

Ainsi,

$$\exp(A) = \begin{pmatrix} \cos(1) & \sin(1) \\ -\sin(1) & \cos(1) \end{pmatrix}.$$

2. Reprenons la matrice  $A = \begin{pmatrix} 3 & -1 & 1 \\ 2 & 0 & 1 \\ 1 & -1 & 2 \end{pmatrix} \in M_3(\mathbb{R})$  de l'exemple 3.8.11. Pour  $P := \begin{pmatrix} 0 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \in$

$GL_3(\mathbb{R})$ , on a

$$P^{-1}AP = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix} = \begin{pmatrix} J_1(1) & 0 \\ 0 & J_2(2) \end{pmatrix}$$

donc

$$\exp(A) = P \begin{pmatrix} \exp(I_1) & 0 \\ 0 & \exp(J_2(2)) \end{pmatrix} P^{-1}.$$

Or  $J_2(2) = 2I_2 + J_2^0$  et, si on note  $J := J_2^0 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ ,  $J^2 = 0_2$  donc

$$\exp(J_2(2)) = e^2 (I_2 + J) = e^2 \left( \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \right) = e^2 \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Ainsi,

$$\exp(A) = P \begin{pmatrix} e & 0 & 0 \\ 0 & e^2 & e^2 \\ 0 & 0 & e^2 \end{pmatrix} P^{-1}.$$

Comme  $P^{-1} = \begin{pmatrix} -1 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & -1 & 1 \end{pmatrix}$ , on a enfin

$$\exp(A) = \begin{pmatrix} 2e^2 & -e^2 & e^2 \\ -e + 2e^2 & e - e^2 & e^2 \\ -e + e^2 & e - e^2 & e^2 \end{pmatrix}.$$



**Proposition 4.5.1.** *La fonction  $\varphi$  est dérivable sur  $\mathbb{R}$  et, pour tout  $t \in \mathbb{R}$ ,  $\varphi'(t) = Ae^{tA}$ .*

*Démonstration.* Soit  $t \in \mathbb{R}$ . On montre que l'expression

$$\frac{e^{(t+h)A} - e^{tA}}{h}$$

possède une limite finie quand  $h$  tend vers 0 et que cette limite est égale à  $Ae^{tA}$ .

Soit donc  $h \in \mathbb{R}^*$ . Remarquons tout d'abord que, puisque les matrices  $tA$  et  $hA$  commutent,

$$\frac{e^{(t+h)A} - e^{tA}}{h} = \frac{e^{hA} e^{tA} - e^{tA}}{h} = \frac{e^{hA} - I_m}{h} e^{tA},$$

et nous allons en fait montrer que la quantité  $\frac{e^{hA} - I_m}{h}$  tend vers  $A$  quand  $h$  tend vers 0, i.e. que la quantité  $\left\| \frac{e^{hA} - I_m - hA}{h} \right\|$  tend vers 0 quand  $h$  tend vers 0.

Comme  $e^{hA} = \sum_{n=0}^{+\infty} \frac{(hA)^n}{n!} = I_m + hA + \sum_{n=2}^{+\infty} \frac{(hA)^n}{n!}$ , on a

$$\left\| e^{hA} - I_m - hA \right\| = \left\| \sum_{n=2}^{+\infty} \frac{(hA)^n}{n!} \right\| \leq \sum_{n=2}^{+\infty} \frac{\|hA\|^n}{n!} = e^{\|hA\|} - 1 - \|hA\| = e^{|h|\|A\|} - 1 - |h|\|A\|,$$

donc

$$\left\| \frac{e^{hA} - I_m - hA}{h} \right\| \leq \frac{e^{|h|\|A\|} - 1}{|h|} - \|A\|.$$

Or, quand  $h$  tend vers 0,  $|h|$  tend vers 0 et la quantité  $\frac{e^{|h|\|A\|} - 1}{|h|}$  tend alors vers la dérivée de la fonction  $\mathbb{R} \rightarrow \mathbb{R}$ ;  $t \mapsto e^{t\|A\|}$  en 0, c'est-à-dire  $\|A\|$ . Ainsi, la quantité  $\left\| \frac{e^{hA} - I_m - hA}{h} \right\|$  tend donc bien vers 0 quand  $h$  tend vers 0.

Au total, la quantité  $\frac{e^{(t+h)A} - e^{tA}}{h} = \frac{e^{hA} - I_m}{h} e^{tA}$  tend donc bien vers  $Ae^{tA}$  quand  $h$  tend vers 0 : l'application  $\varphi$  est donc bien dérivable en  $t$  et  $\varphi'(t) = Ae^{tA}$ . □

On s'intéresse à présent à la résolution du système différentiel linéaire (S)  $X' = AX$  de fonction vectorielle dérivable inconnue  $X : \mathbb{R} \rightarrow \mathbb{K}^m$ . Les solutions de (S) sont données par l'exponentielle de matrices :

**Proposition 4.5.2.** *Les solutions de (S) sont les fonctions vectorielles de la forme*

$$X : \begin{array}{l} \mathbb{R} \rightarrow \mathbb{K}^m \\ t \rightarrow e^{tA} X_0 \end{array}$$

avec  $X_0 \in \mathbb{K}^m$ .

*Démonstration.* On montre tout d'abord que, si  $X_0 \in \mathbb{K}^m$ , la fonction dérivable  $X : \mathbb{R} \rightarrow \mathbb{K}^m$  ;  $t \mapsto e^{tA}X_0$  est une solution de (S) : pour tout  $t \in \mathbb{R}$ , on a, d'après la proposition précédente,

$$X'(t) = Ae^{tA}X_0 = AX(t).$$

Réciproquement, soit  $X : \mathbb{R} \rightarrow \mathbb{K}^m$  une solution de (S) et considérons la fonction dérivable  $Y : \mathbb{R} \rightarrow \mathbb{K}^m$  ;  $t \mapsto e^{-tA}X(t)$ . Pour tout  $t \in \mathbb{R}$ , on a

$$Y'(t) = -Ae^{-tA}X(t) + e^{-tA}X'(t) = e^{-tA}(X'(t) - AX(t)) = 0$$

donc il existe un vecteur  $X_0 \in \mathbb{R}^m$  tel que, pour tout  $t \in \mathbb{R}$ ,  $Y(t) = X_0 \Leftrightarrow X(t) = e^{tA}X_0$ .  $\square$

*Remarque 4.5.3.* • Si  $X_0 \in \mathbb{K}^m$ , la fonction  $X : \mathbb{R} \rightarrow \mathbb{K}^m$  ;  $t \mapsto e^{tA}X_0$  est l'unique solution de (S) prenant pour valeur  $X_0$  en 0 (on a  $X(0) = e^{0A}X_0 = I_m X_0 = X_0$ ). L'égalité  $X(0) = X_0$  est appelée condition initiale en 0.

- Pour  $m = 1$ , on retrouve la résolution des équations différentielles linéaires du premier ordre  $x'(t) = ax(t)$  avec  $a \in \mathbb{K}$  et  $x : \mathbb{R} \rightarrow \mathbb{R}$ .

*Exemple 4.5.4.* On considère la matrice

$$A := \begin{pmatrix} 3 & -1 & 1 \\ 2 & 0 & 1 \\ 1 & -1 & 2 \end{pmatrix} \in M_3(\mathbb{R})$$

de l'exemple 4.4.2 2. et le système différentiel linéaire (S)  $X' = AX$ , avec  $X : \mathbb{R} \rightarrow \mathbb{R}^3$ , dont on souhaite déterminer l'ensemble des solutions.

D'après la proposition précédente, les solutions de (S) sont les fonctions de la forme  $X : \mathbb{R} \rightarrow \mathbb{K}^m$  ;  $t \mapsto e^{tA}X_0$  avec  $X_0 \in \mathbb{R}^3$ . Soit  $t \in \mathbb{R}$  et calculons donc  $e^{tA}$ .

En reprenant les notations de l'exemple 4.4.2 2., on a

$$A = P \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix} P^{-1} = P \begin{pmatrix} 1 & 0 \\ 0 & J_2(2) \end{pmatrix} P^{-1}$$

donc

$$tA = P \begin{pmatrix} t & 0 & 0 \\ 0 & 2t & t \\ 0 & 0 & 2t \end{pmatrix} P^{-1} = P \begin{pmatrix} t & 0 \\ 0 & tJ_2(2) \end{pmatrix} P^{-1}$$

et

$$\exp(tA) = P \begin{pmatrix} \exp(t) & 0 \\ 0 & \exp(tJ_2(2)) \end{pmatrix} P^{-1}.$$

Or  $tJ_2(2) = 2tI_2 + tJ_2^0$  et, notant  $J := J_2^0$ ,  $J^2 = 0_2$  donc  $(tJ)^2 = t^2J^2 = 0_2$  et

$$\exp(tJ_2(2)) = \exp(2tI_2) \exp(tJ) = e^{2t} (I_2 + tJ) = e^{2t} \left( \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & t \\ 0 & 0 \end{pmatrix} \right) = e^{2t} \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}.$$



Ainsi,

$$\begin{aligned} \exp(tA) &= P \begin{pmatrix} e^t & 0 & 0 \\ 0 & e^{2t} & te^{2t} \\ 0 & 0 & e^{2t} \end{pmatrix} P^{-1} \\ &= \begin{pmatrix} e^{2t} + te^{2t} & -te^{2t} & te^{2t} \\ -e^t + e^{2t} + te^{2t} & e^t - te^{2t} & te^{2t} \\ -e^t + e^{2t} & e^t - e^{2t} & e^{2t} \end{pmatrix} \end{aligned}$$

Les solutions du système différentiel (S) sont donc les fonctions de la forme

$$\begin{aligned} \mathbb{R} &\rightarrow \mathbb{R}^3 \\ t &\mapsto \begin{pmatrix} e^{2t} + te^{2t} & -te^{2t} & te^{2t} \\ -e^t + e^{2t} + te^{2t} & e^t - te^{2t} & te^{2t} \\ -e^t + e^{2t} & e^t - e^{2t} & e^{2t} \end{pmatrix} X_0 \end{aligned}$$

avec  $X_0 \in \mathbb{R}^3$ .



# Chapitre 5

## Orthogonalité et réduction

### 5.1 Introduction

Dans ce chapitre, on aborde la question de la réductibilité de certaines classes d'endomorphismes des espaces euclidiens. En particulier, on montre que tout endomorphisme auto-adjoint est diagonalisable dans une base orthonormale, et que tout endomorphisme orthogonal est diagonalisable par blocs, suivant des blocs de rotations vectorielles de dimension 2 ou 1.

On abordera également la notion de positivité d'une matrice symétrique, montrant notamment qu'une matrice symétrique positive possède une "racine carrée". Cela nous permettra d'établir l'existence d'une "décomposition polaire" pour toute matrice à coefficients réels inversible.

### 5.2 Diagonalisabilité et diagonalisation des endomorphismes auto-adjoints

Soit  $(E, \langle \cdot, \cdot \rangle)$  un espace euclidien de dimension  $n \in \mathbb{N} \setminus \{0\}$  et soit  $f$  un endomorphisme de  $E$ . On rappelle tout d'abord la définition d'un endomorphisme auto-adjoint de  $E$ , donnée pour la première fois à la remarque 2.7.3 :

**Définition 5.2.1.** *On dit que l'endomorphisme  $f$  de  $E$  est auto-adjoint ou symétrique si  $f^* = f$ .*

*Remarque 5.2.2.* Suivant la définition de l'adjoint d'un endomorphisme d'un espace euclidien (cf proposition et définition 2.7.1), l'endomorphisme  $f$  de  $E$  est symétrique si et seulement si, pour tous vecteurs  $v$  et  $w$  de  $E$ ,

$$\langle f(v), w \rangle = \langle v, f(w) \rangle.$$

L'endomorphisme  $f$  est symétrique si et seulement si sa matrice dans une base orthonormale de  $(E, \langle \cdot, \cdot \rangle)$  est symétrique :

**Proposition 5.2.3.** *Soit  $\mathcal{B}$  une base orthonormale de  $E$  et notons  $A := \text{Mat}_{\mathcal{B}}(f)$  la matrice représentative de  $f$  dans  $\mathcal{B}$ . Alors  $f$  est symétrique si et seulement si la matrice  $A$  est symétrique i.e.  ${}^tA = A$ .*

*Démonstration.* D'après la proposition 2.7.5, la matrice représentative de  $f^*$  dans  $\mathcal{B}$  est la transposée de  $A$ . Or  $f^* = f$  si et seulement si les matrices représentatives de  $f$  et  $f^*$  dans  $\mathcal{B}$  sont égales i.e. si et seulement si  ${}^tA = A$ .  $\square$

*Exemple 5.2.4.* L'endomorphisme  $h : \mathbb{R}^2 \rightarrow \mathbb{R}^2 ; (x, y) \mapsto (x + 2y, 2x + y)$  de  $\mathbb{R}^2$  est auto-adjoint par rapport au produit scalaire canonique de  $\mathbb{R}^2$ .

Une propriété remarquable des endomorphismes auto-adjoints est qu'ils sont diagonalisables, et ce, dans une base orthonormale. Avant d'énoncer ce résultat, donnons la définition suivante : deux sous-espaces vectoriels  $F_1$  et  $F_2$  de  $E$  sont dits orthogonaux si pour tous vecteurs  $v_1$  de  $F_1$  et  $v_2$  de  $F_2$ ,  $\langle v_1, v_2 \rangle = 0$ .

**Théorème 5.2.5.** *On suppose que l'endomorphisme  $f$  est auto-adjoint. Alors*

- $f$  est diagonalisable,
- les sous-espaces propres de  $f$  sont orthogonaux deux à deux.

*En particulier, si l'on considère, pour chacun de ces espaces propres, une base orthonormale, la réunion de ces bases est une base orthonormale de  $E$  :  $f$  est donc diagonalisable dans une base orthonormale de  $E$ .*

*Démonstration.* On commence par montrer que le polynôme caractéristique de  $f$  est scindé sur  $\mathbb{R}$ . Pour cela, soit  $\mathcal{B}_0$  une base orthonormale de  $E$  et considérons la matrice  $A := \text{Mat}_{\mathcal{B}_0}(f)$  de  $f$  dans  $\mathcal{B}_0$ .  $A$  est une matrice de  $M_n(\mathbb{R})$  et peut être également considérée comme une matrice de  $M_n(\mathbb{C})$ . Nous allons montrer que les racines complexes du polynôme caractéristique de  $A$  (autrement dit les valeurs propres complexes de  $A$ ) sont toutes réelles :  $\chi_f = \chi_A$  sera donc également scindé sur  $\mathbb{R}$ .

Soit donc une valeur propre complexe  $\lambda \in \mathbb{C}$  de  $A$  et montrons que  $\lambda \in \mathbb{R}$ . Soit  $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in$

$M_{n,1}(\mathbb{C})$  un vecteur propre de  $A$  associé à  $\lambda$  : on a donc  $AX = \lambda X$ . Nous allons appliquer la conjugaison complexe à cette dernière égalité : si  $M = (m_{ij})_{1 \leq i \leq p, 1 \leq j \leq q}$  est une matrice de  $M_{p,q}(\mathbb{C})$ , la matrice conjuguée  $\overline{M}$  de  $M$  est la matrice  $(\overline{m_{ij}})_{1 \leq i, j \leq n}$  de  $M_n(\mathbb{C})$ . On obtient alors

$$\overline{AX} = \overline{\lambda X} \Leftrightarrow \overline{AX} = \overline{\lambda} \overline{X} \Leftrightarrow A\overline{X} = \overline{\lambda} \overline{X}$$

(les coefficients de  $A$  sont réels).

On considère d'autre part l'égalité  ${}^t(AX)\overline{X} = {}^tXA\overline{X}$ , satisfaite car  ${}^tA = A$  (car  $f$  est symétrique : proposition 5.2.3). On y remplace  $AX$  par  $\lambda X$  ( $X$  est un vecteur propre associé à  $\lambda$ ) pour obtenir  ${}^t(\lambda X)\overline{X} = {}^tX\overline{\lambda}\overline{X}$  i.e.  $\lambda \sum_{i=1}^n |x_i|^2 = \overline{\lambda} \sum_{i=1}^n |x_i|^2$  et donc  $\lambda = \overline{\lambda}$  car  $X$  n'est pas le vecteur colonne nul ( $X$  est un vecteur propre). Ainsi,  $\lambda \in \mathbb{R}$ .

Le polynôme caractéristique de l'endomorphisme symétrique  $f$  est donc scindé sur  $\mathbb{R}$ . En particulier, le spectre de  $f$  est non-vide.

On montre à présent que  $f$  est diagonalisable. On le montre par récurrence sur la dimension  $n$  de  $E$ . Précisément, on montre par récurrence l'assertion suivante : pour tout  $n \in \mathbb{N} \setminus \{0\}$ , pour tout espace euclidien  $(E, \langle \cdot, \cdot \rangle)$  de dimension  $n$ , pour tout endomorphisme symétrique  $f$  de  $E$ ,  $f$  est diagonalisable.

Toute matrice carrée de taille 1 étant diagonale, le résultat est vrai pour  $n = 1$ .

Supposons à présent la propriété vraie au rang  $n - 1$  pour  $n \in \mathbb{N} \setminus \{0, 1\}$  fixé, et montrons-la pour l'endomorphisme symétrique  $f$  de l'espace euclidien  $E$  de dimension  $n$ . Soit  $\lambda \in \mathbb{R}$  une valeur propre de  $f$  (un tel  $\lambda$  existe car  $\text{Sp}_{\mathbb{R}}(f) \neq \emptyset$ ). Soit ensuite  $v$  un vecteur propre de  $f$  associé à  $\lambda$  et notons  $F := \{v\}^{\perp} = (\text{Vect}\{v\})^{\perp}$ . Montrons que  $F$  est stable par  $f$  : soit  $w \in F$  alors

$$\begin{aligned} \langle f(w), v \rangle &= \langle w, f(v) \rangle \quad (f \text{ est symétrique}) \\ &= \langle w, \lambda v \rangle \quad (v \in E_{\lambda}) \\ &= \lambda \langle w, v \rangle \\ &= 0 \quad (\text{car } w \in F = \{v\}^{\perp}). \end{aligned}$$

Ainsi  $F$  est stable par  $F$  et on peut donc restreindre  $f$  en un endomorphisme  $f|_F$  de  $F$ , qui est également symétrique (par rapport à la restriction du produit scalaire  $\langle \cdot, \cdot \rangle$  sur  $F$ ). Comme  $\dim(F) = \dim((\text{Vect}\{v\})^{\perp}) = \dim(E) - \dim(\text{Vect}\{v\}) = n - 1$ , on peut ensuite appliquer l'hypothèse de récurrence à l'endomorphisme symétrique  $f|_F$  de  $F$  :  $f|_F$  est diagonalisable i.e. il existe une base de  $F$  formée de vecteurs propres  $e_2, \dots, e_n$  pour  $f|_F$ . Les sous-espaces vectoriels  $\text{Vect}\{x\}$  et  $F$  étant en somme directe, la famille  $\{v, e_2, \dots, e_n\}$  est alors une base de  $E$ , formée de vecteurs propres pour  $f$  donc  $f$  est diagonalisable.

On montre enfin que les sous-espaces propres de  $f$  sont deux à deux orthogonaux. Soient donc  $\lambda_1$  et  $\lambda_2$  deux valeurs propres distinctes de  $f$  et soient  $v_1 \in E_{\lambda_1}, v_2 \in E_{\lambda_2}$ . Montrons que les vecteurs  $v_1$  et  $v_2$  sont orthogonaux. On a d'une part

$$\langle f(v_1), v_2 \rangle = \langle \lambda_1 v_1, v_2 \rangle = \lambda_1 \langle v_1, v_2 \rangle$$

et, d'autre part, comme  $f$  est symétrique,

$$\langle f(v_1), v_2 \rangle = \langle v_1, f(v_2) \rangle = \langle v_1, \lambda_2 v_2 \rangle = \lambda_2 \langle v_1, v_2 \rangle.$$

Ainsi,  $\lambda_1 \langle v_1, v_2 \rangle = \lambda_2 \langle v_1, v_2 \rangle$  i.e.  $(\lambda_1 - \lambda_2) \langle v_1, v_2 \rangle = 0$  et donc  $\langle v_1, v_2 \rangle = 0$  car  $\lambda_1 \neq \lambda_2$ . Les vecteurs  $v_1$  et  $v_2$  sont donc orthogonaux.  $\square$

*Exemple 5.2.6.* On considère l'endomorphisme

$$f : \begin{array}{ccc} \mathbb{R}^3 & \rightarrow & \mathbb{R}^3 \\ (x, y, z) & \mapsto & (5x - y + 2z, -x + 5y + 2z, 2x + 2y + 2z) \end{array}$$

de  $\mathbb{R}^3$ . La matrice représentative de  $f$  dans la base canonique de  $\mathbb{R}^3$  (qui est une base orthonormale pour le produit scalaire canonique de  $\mathbb{R}^3$ ) est

$$A := \begin{pmatrix} 5 & -1 & 2 \\ -1 & 5 & 2 \\ 2 & 2 & 2 \end{pmatrix}.$$

La matrice  $A$  étant symétrique, l'endomorphisme  $f$  est auto-adjoint.

Le polynôme caractéristique de  $f$  est  $\chi_f = (6 - X)^2(-X)$  et les valeurs propres de  $f$  sont donc 6 et 0.

On a

$$E_6 = \text{Ker} (f - 6\text{Id}_{\mathbb{R}^3}) = \{(x, y, z) \in \mathbb{R}^3 \mid -x - y + 2z = 0\}$$

et

$$E_0 = \text{Ker} f.$$

La famille  $\{(2, 0, 1), (1, 1, 0)\}$  est une base de  $E_6$ . En appliquant le procédé d'orthonormalisation de Gram-Schmidt à cette famille libre de  $\mathbb{R}^3$ , on obtient la base orthonormale  $\left\{ \frac{1}{\sqrt{5}}(2, 0, 1), \frac{1}{\sqrt{30}}(1, -5, -2) \right\}$  de  $E_6$ .

D'autre part, le vecteur  $\frac{1}{\sqrt{6}}(-1, -1, 2)$  de norme 1 engendre  $E_0$ .

Si l'on note  $\mathcal{B} := \left\{ \frac{1}{\sqrt{5}}(2, 0, 1), \frac{1}{\sqrt{30}}(1, -5, -2), \frac{1}{\sqrt{6}}(-1, -1, 2) \right\}$ , la famille  $\mathcal{B}$  est alors une base orthonormale de  $\mathbb{R}^3$  et la matrice représentative de  $f$  dans  $\mathcal{B}$  est

$$\begin{pmatrix} 6 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Le pendant matriciel du théorème 5.2.5 consiste en l'énoncé suivant :

**Corollaire 5.2.7.** *Soit  $A$  une matrice de  $M_n(\mathbb{R})$ . On suppose que  $A$  est symétrique i.e.  ${}^tA = A$ . Alors il existe une matrice orthogonale  $O \in O_n(\mathbb{R})$  et une matrice diagonale  $D \in M_n(\mathbb{R})$  telles que*

$$D = O^{-1}AO = {}^tOAO$$

*Démonstration.*  $A$  est la matrice représentative d'un endomorphisme  $h$  de  $\mathbb{R}^n$  dans la base canonique  $\mathcal{B}_0$  de  $\mathbb{R}^n$ , qui est une base orthonormale pour le produit scalaire canonique de  $\mathbb{R}^n$ . Comme  $A$  est symétrique,  $h$  est auto-adjoint et, par le théorème précédent, il existe une base orthonormale  $\mathcal{B}$  et une matrice diagonale  $D \in M_n(\mathbb{R})$  telle que  $\text{Mat}_{\mathcal{B}}(h) = D$ .

Les bases  $\mathcal{B}_0$  et  $\mathcal{B}$  étant des bases orthonormales de  $\mathbb{R}^n$ , la matrice de passage  $O := P_{\mathcal{B}_0 \rightarrow \mathcal{B}}$  est une matrice orthogonale (proposition 2.8.12) et on a alors

$${}^tOAO = O^{-1}AO = P_{\mathcal{B} \rightarrow \mathcal{B}_0} \text{Mat}_{\mathcal{B}_0}(h) P_{\mathcal{B}_0 \rightarrow \mathcal{B}} = \text{Mat}_{\mathcal{B}}(h) = D.$$

□

*Exemple 5.2.8.* Si l'on reprend la matrice  $A := \begin{pmatrix} 5 & -1 & 2 \\ -1 & 5 & 2 \\ 2 & 2 & 2 \end{pmatrix}$  définie dans l'exemple 5.2.6

précédent et si l'on note

$$O := \begin{pmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{30}} & \frac{1}{\sqrt{6}} \\ 0 & \frac{-5}{\sqrt{30}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{5}} & \frac{-2}{\sqrt{30}} & \frac{-2}{\sqrt{6}} \end{pmatrix},$$

la matrice  $O$  est orthogonale et

$$O^{-1}AO = {}^tOAO = \begin{pmatrix} 6 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

### 5.3 Matrices symétriques positives

Soit  $n \in \mathbb{N} \setminus \{0\}$ . Dans cette section, nous allons considérer des matrices symétriques réelles particulières. On commence par noter  $S_n(\mathbb{R})$  l'ensemble des matrices symétriques de  $M_n(\mathbb{R})$ .

**Définition 5.3.1.** Soit  $A \in S_n(\mathbb{R})$ . On dit que  $A$  est

- positive si, pour tout  $X \in M_{n,1}(\mathbb{R})$ ,  ${}^tXAX \geq 0$ ,
- définie positive si, pour tout  $X \in M_{n,1}(\mathbb{R}) \setminus \left\{ \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \right\}$ ,  ${}^tXAX > 0$ .

*Exemple 5.3.2.* 1. La matrice symétrique  $A := \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} \in M_3(\mathbb{R})$  est définie positive :

pour tout  $X = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in M_{3,1}(\mathbb{R})$ ,

$${}^tXAX = x^2 + 2y^2 + 3z^2 + 2xy = (x + y)^2 + y^2 + 3z^2 \geq 0,$$

et cette quantité est égale à 0 ssi  $x = y = z = 0$  ssi  $X = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ .

2. La matrice symétrique  $B := \begin{pmatrix} 2 & -3 & 0 \\ -3 & 11 & 0 \\ 0 & 0 & 0 \end{pmatrix} \in M_3(\mathbb{R})$  est positive car, pour tous  $X =$

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \in M_{3,1}(\mathbb{R}),$$

$${}^tXBX = 2x^2 + 11y^2 - 6xy = (x - 3y)^2 + x^2 + 2y^2 \geq 0,$$

mais elle n'est pas définie positive car  $(0 \ 0 \ 1) B \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = 0$ .

*Remarque 5.3.3.* Une matrice symétrique  $A$  de  $S_n(\mathbb{R})$  est définie positive si et seulement si elle représente un produit scalaire. En effet, notons tout d'abord  $\mathcal{B} = \{e_1, \dots, e_n\}$  la base canonique de  $\mathbb{R}^n$  et définissons l'application

$$\langle \cdot, \cdot \rangle : \left( v = \sum_{i=1}^n x_i e_i, w = \sum_{j=1}^n y_j e_j \right) \mapsto {}^t(x_1 \ \dots \ x_n) A \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

Cette application est bilinéaire, symétrique (car  $A$  est symétrique), et elle est définie positive ssi  $A$  est définie positive. L'application  $\langle \cdot, \cdot \rangle$  est donc un produit scalaire sur  $\mathbb{R}^n$  ssi  $A$  est définie positive et, dans ce cas,  $A = \text{Mat}_{\mathcal{B}}^{\text{ps}}(\langle \cdot, \cdot \rangle)$  car, pour tout  $i, j \in \{1, \dots, n\}$ ,  ${}^t X_i A X_j = \langle e_i, e_j \rangle$  (cf preuve de la proposition 2.6.4).

Réciproquement, si  $A$  représente un produit scalaire  $\langle \cdot, \cdot \rangle$  d'un espace euclidien  $E$  de dimension  $n$  dans une base  $\mathcal{B}$ , alors  $A$  est une matrice symétrique définie positive : pour tous  $X, Y \in M_{n,1}(\mathbb{R})$ ,  ${}^t X A Y = \langle v, w \rangle$  où  $v$  et  $w$  sont respectivement les vecteurs de  $E$  de coordonnées les coordonnées des vecteurs colonnes  $X$  et  $Y$  dans la base  $\mathcal{B}$ , et le produit scalaire  $\langle \cdot, \cdot \rangle$  est défini positif.

Soit  $A \in S_n(\mathbb{R})$ . Nous allons à présent exhiber une caractérisation du caractère positif, resp. défini positif, de  $A$  en termes de ses valeurs propres. Rappelons que, comme  $A$  est symétrique, son polynôme caractéristique est scindé sur  $\mathbb{R}$  (cf preuve du théorème 5.2.5).

**Proposition 5.3.4.** *La matrice symétrique  $A$  est positive, resp. définie positive, si et seulement si toutes ses valeurs propres sont positives ou nulles, resp. strictement positives.*

*Démonstration.* Comme  $A$  est symétrique, d'après le corollaire 5.2.7, il existe une matrice orthogonale  $O \in O_n(\mathbb{R})$  et une matrice diagonale  $D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \in M_n(\mathbb{R})$  telles que  ${}^t O A O = D$ .

En particulier, les scalaires  $\lambda_1, \dots, \lambda_n$  sont les valeurs propres (non nécessairement deux à deux distinctes) de  $A$ .

Alors

$$\begin{aligned} \text{pour tout } X \in M_{n,1}(\mathbb{R}), {}^t X A X \geq 0 & \quad \text{ssi} \quad \text{pour tout } Y \in M_{n,1}(\mathbb{R}), {}^t (OY) A (OY) \geq 0 \quad (O \in GL_n(\mathbb{R})) \\ & \quad \text{ssi} \quad \text{pour tout } Y \in M_{n,1}(\mathbb{R}), {}^t Y ({}^t O A O) Y \geq 0 \\ & \quad \text{ssi} \quad \text{pour tout } Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in M_{n,1}(\mathbb{R}), {}^t Y D Y \geq 0 \\ & \quad \text{ssi} \quad \text{pour tous } y_1, \dots, y_n \in \mathbb{R}, \sum_{i=1}^n \lambda_i y_i^2 \geq 0. \end{aligned}$$

Or, pour tous  $y_1, \dots, y_n \in \mathbb{R}$ ,  $\sum_{i=1}^n \lambda_i y_i^2 \geq 0$  ssi pour tout  $i \in \{1, \dots, n\}$ ,  $\lambda_i \geq 0$  (car, si  $j \in$



$$\{1, \dots, n\}, \lambda_j = \sum_{i=1}^n \lambda_i \delta_{ij}^2).$$

Ainsi, la matrice symétrique  $A$  est positive si et seulement si toutes ses valeurs propres sont positives ou nulles.

De façon analogue,

$$\begin{aligned} A \text{ est définie positive} & \quad \text{ssi} \quad \text{pour tout } X \in M_{n,1}(\mathbb{R}) \setminus \left\{ \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \right\}, {}^t X A X > 0 \\ & \quad \text{ssi} \quad \text{pour tout } (y_1, \dots, y_n) \in \mathbb{R}^n \setminus \{(0, \dots, 0)\}, \sum_{i=1}^n \lambda_i y_i^2 > 0 \\ & \quad \text{ssi} \quad \text{pour tout } i \in \{1, \dots, n\}, \lambda_i > 0 \end{aligned}$$

(si pour tout  $i \in \{1, \dots, n\}, \lambda_i > 0$ , alors, pour tout  $(y_1, \dots, y_n) \in \mathbb{R}^n, \sum_{i=1}^n \lambda_i y_i^2 = 0$  ssi pour tout  $i \in \{1, \dots, n\}, y_i = 0$ ).

□

*Exemple 5.3.5.* Reprenons les exemples de l'exemple 5.3.2.

1. On a  $\chi_A = (X^2 - 3X + 1)(3 - X)$  donc  $\text{Sp}(A) = \left\{ \frac{3-\sqrt{5}}{2}, \frac{3+\sqrt{5}}{2}, 3 \right\}$ , et on peut donc directement en déduire que la matrice symétrique  $A$  est définie positive.
2. On a  $\chi_B = (X^2 - 13X + 13)(-X)$  donc  $\text{Sp}(A) = \left\{ \frac{13-\sqrt{117}}{2}, \frac{13+\sqrt{117}}{2}, 0 \right\}$ , et on peut donc directement en déduire que la matrice symétrique  $B$  est positive, non définie positive.

*Remarque 5.3.6.* Une matrice symétrique définie positive est en particulier inversible.

Nous allons montrer plus bas qu'une matrice symétrique positive possède une "racine carrée". Pour pouvoir démontrer ce résultat, nous aurons besoin de la propriété de "diagonalisation simultanée" suivante :

**Proposition 5.3.7.** *Soient  $\mathbb{K}$  un corps commutatif et  $E$  un espace vectoriel sur  $\mathbb{K}$  de dimension finie  $n$ . Soient  $f, g \in \mathcal{L}(E)$ . Si  $f$  et  $g$  sont diagonalisables et si  $f \circ g = g \circ f$ , alors il existe une base de  $E$  dans laquelle les matrices représentatives de  $f$  et  $g$  sont toutes deux diagonales (on dit que  $f$  et  $g$  sont co-diagonalisables).*

*Démonstration.* Supposons donc que les endomorphismes  $f$  et  $g$  de  $E$  sont diagonalisables et commutent. Soient  $\lambda_1, \dots, \lambda_p \in \mathbb{K}$  les valeurs propres deux à deux distinctes de  $f$ .

Soit  $i \in \{1, \dots, p\}$  et montrons que le sous-espace propre  $E_{\lambda_i}$  de  $f$  associé à la valeur propre  $\lambda_i$  est stable par  $g$ . Soit donc  $v \in E_{\lambda_i}$  et montrons que  $g(v) \in E_{\lambda_i}$  : on a

$$\begin{aligned}
 (f - \lambda_i \text{Id}_E)(g(v)) &= (f \circ g)(v) - \lambda_i g(v) \\
 &= (g \circ f)(v) - \lambda_i g(v) \quad (\text{car } f \text{ et } g \text{ commutent}) \\
 &= g(f(v) - \lambda_i v) \quad (g \text{ est linéaire}) \\
 &= g(0_E) \quad (\text{car } v \in E_{\lambda_i}) \\
 &= 0_E
 \end{aligned}$$

donc  $g(v)$  appartient bien à  $E_{\lambda_i}$ . Comme  $E_{\lambda_i}$  est un sev de  $E$  stable par l'endomorphisme diagonalisable  $g$ , la restriction  $g|_{E_{\lambda_i}}$  de  $g$  à  $E_{\lambda_i}$  est un endomorphisme diagonalisable de  $E_{\lambda_i}$  : il existe donc une base  $\mathcal{B}_i$  de  $E_{\lambda_i}$  dans laquelle la matrice représentative de  $g|_{E_{\lambda_i}}$  est diagonale. Mais comme  $E_{\lambda_i}$  est le sous-espace propre de  $f$  associé à la valeur propre  $\lambda_i$ ,  $E_{\lambda_i}$  est stable par  $f$  et la matrice de la restriction  $f|_{E_{\lambda_i}}$  est également diagonale dans cette base  $\mathcal{B}_i$ .

Si l'on pose enfin  $\mathcal{B} := \{\mathcal{B}_1, \dots, \mathcal{B}_p\}$ , les matrices représentatives de  $f$  et  $g$  dans la base  $\mathcal{B}$  de  $E$  sont toutes deux diagonales.  $\square$

*Remarque 5.3.8.* • Dans la preuve ci-dessus, on a utilisé la propriété remarquable suivante : si  $\mathbb{K}$  est un corps commutatif,  $E$  est un  $\mathbb{K}$ -espace vectoriel,  $F$  est un sev de  $E$  et  $f$  est un endomorphisme diagonalisable de  $E$  tel que  $f(F) \subset F$ , alors la restriction de  $f$  à  $F$  est un endomorphisme diagonalisable de l'espace vectoriel  $F$ . Pour prouver cette assertion, on peut utiliser le théorème 3.6.6 : le polynôme minimal  $\mu_{f|_F}$  de la restriction  $f|_F$  divise le polynôme minimal  $\mu_f$  de  $f$  (car  $\mu_f$  est un polynôme annulateur de l'endomorphisme  $f|_F$  de  $F$ ), or,  $f$  étant diagonalisable,  $\mu_f$  est un polynôme scindé à racines simples sur  $\mathbb{K}$ .  $\mu_{f|_F}$  est donc également scindé à racines simples et la restriction  $f|_F$  est un endomorphisme diagonalisable de  $F$ .

- Le pendant matriciel de la proposition 5.3.7 est l'énoncé suivant. Soient  $A, B \in M_n(\mathbb{R})$ . Si  $A$  et  $B$  sont diagonalisables et si  $AB = BA$ , alors il existe une matrice inversible  $P \in \text{GL}_n(\mathbb{R})$  et des matrices diagonales  $D_1, D_2 \in M_n(\mathbb{R})$  telles que  $P^{-1}AP = D_1$  et  $P^{-1}BP = D_2$  (on dit que les matrices  $A$  et  $B$  sont co-diagonalisables).
- Avec les notations de la proposition 5.3.7, si  $E$  est un  $\mathbb{R}$ -espace vectoriel euclidien et  $f$  et  $g$  sont deux endomorphismes auto-adjoints (donc en particulier diagonalisables) qui commutent, alors il existe une base orthonormale dans laquelle les matrices représentatives de  $f$  et  $g$  sont toutes deux diagonales. En effet, conservant les notations de la preuve ci-dessus, il suffit dans ce cas de considérer, pour chaque  $i \in \{1, \dots, p\}$ , une base  $\mathcal{B}_i$  de  $E_{\lambda_i}$  qui soit orthonormale (la restriction de  $g$  à  $E_{\lambda_i}$  est également un endomorphisme auto-adjoint de  $E_{\lambda_i}$ ) : les sous-espaces propres de  $f$  étant orthogonaux, la base  $\mathcal{B}$  est alors une base orthonormale de  $E$ .
- Le pendant matriciel de la remarque précédente est l'énoncé suivant. Soient  $A, B \in S_n(\mathbb{R})$ . Si  $AB = BA$ , alors il existe une matrice orthogonale  $O \in O_n(\mathbb{R})$  et des matrices diagonales  $D_1, D_2 \in M_n(\mathbb{R})$  telles que  ${}^tOAO = D_1$  et  ${}^tOBO = D_2$ .

**Théorème et Définition 5.3.9.** *Soit  $A \in S_n(\mathbb{R})$  une matrice positive. Il existe une unique matrice  $R \in S_n(\mathbb{R})$  positive telle que  $A = R^2 = RR$ . De plus, si  $A$  est définie positive,  $R$  l'est également. On appelle  $R$  la racine carrée de  $A$ .*

*Démonstration.* Montrons tout d'abord l'existence de cette décomposition. Comme  $A$  est une matrice symétrique positive, il existe une matrice orthogonale  $O \in O_n(\mathbb{R})$  et une matrice dia-

gonale  $D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \in M_n(\mathbb{R})$  telles que  ${}^tOAO = D$  (car  $A$  est symétrique : corollaire 5.2.7), avec, pour tout  $i \in \{1, \dots, n\}$ ,  $\lambda_i \geq 0$  (car  $A$  est positive : proposition précédente 5.3.4).

Pour  $i \in \{1, \dots, n\}$ , notons  $\mu_i := \sqrt{\lambda_i} \geq 0$  et posons

$$R := O \begin{pmatrix} \mu_1 & & 0 \\ & \ddots & \\ 0 & & \mu_n \end{pmatrix} {}^tO.$$

$R \in M_n(\mathbb{R})$  est une matrice symétrique ( ${}^tR = R$ ) de valeurs propres  $\mu_1, \dots, \mu_n$  positives ou nulles donc, par la proposition 5.3.4,  $R$  est positive.

Enfin,

$$\begin{aligned} R^2 = RR &= O \begin{pmatrix} \mu_1 & & 0 \\ & \ddots & \\ 0 & & \mu_n \end{pmatrix} ({}^tOO) \begin{pmatrix} \mu_1 & & 0 \\ & \ddots & \\ 0 & & \mu_n \end{pmatrix} {}^tO \\ &= O \begin{pmatrix} \mu_1 & & 0 \\ & \ddots & \\ 0 & & \mu_n \end{pmatrix} \begin{pmatrix} \mu_1 & & 0 \\ & \ddots & \\ 0 & & \mu_n \end{pmatrix} {}^tO \quad (O \text{ est orthogonale}) \\ &= O \begin{pmatrix} \mu_1^2 & & 0 \\ & \ddots & \\ 0 & & \mu_n^2 \end{pmatrix} {}^tO \\ &= O \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} {}^tO \\ &= OD{}^tO \\ &= A. \end{aligned}$$

Remarquons que si  $A$  est définie positive alors, pour tout  $i \in \{1, \dots, p\}$ ,  $\lambda_i > 0$  donc  $\mu_i > 0$ , et  $R$  est donc également définie positive.

Montrons ensuite l'unicité de la matrice  $R$  en tant que matrice symétrique positive dont le carré est  $A$ . Soit  $\tilde{R} \in S_n(\mathbb{R})$  une matrice positive telle que  $\tilde{R}^2 = A$ . En particulier,  $\tilde{R}$  commute avec  $A$  donc avec tout polynôme de matrice en  $A$ . Or, si l'on considère un polynôme  $L \in \mathbb{R}[X]$

tel que, pour tout  $i \in \{1, \dots, n\}$ ,  $L(\lambda_i) = \mu_i$  (par exemple le polynôme donné par l'interpolation de Lagrange), on a

$$\begin{aligned}
 L(A) &= L(OD^tO) \\
 &= OL(D)^tO \\
 &= O \begin{pmatrix} L(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & L(\lambda_n) \end{pmatrix}^t O \\
 &= O \begin{pmatrix} \mu_1 & & 0 \\ & \ddots & \\ 0 & & \mu_n \end{pmatrix}^t O \\
 &= R
 \end{aligned}$$

et donc  $\tilde{R}$  commute avec  $R = L(A)$ . Comme, de plus, les matrices  $R$  et  $\tilde{R}$  sont diagonalisables (car symétriques), elles sont co-diagonalisables d'après la proposition 5.3.7 : il existe une matrice inversible  $P \in \text{GL}_n(\mathbb{R})$  et des matrices diagonales  $D_1, D_2 \in M_n(\mathbb{R})$  de coefficients diagonaux positifs ou nuls ( $R$  et  $\tilde{R}$  sont positives) telles que  $P^{-1}RP = D_1$  et  $P^{-1}\tilde{R}P = D_2$ . Alors

$$D_1^2 = P^{-1}R^2P = P^{-1}AP = P^{-1}\tilde{R}^2P = D_2^2.$$

Les matrices  $D_1$  et  $D_2$  étant toutes deux diagonales à coefficients positifs ou nuls, on obtient finalement l'égalité  $D_1 = D_2$  et donc  $R = PD_1P^{-1} = PD_2P^{-1} = \tilde{R}$ .  $\square$

La preuve de l'existence de la racine carrée d'une matrice symétrique positive nous donne une méthode pour la calculer :

*Exemple 5.3.10.* On considère la matrice symétrique

$$A := \begin{pmatrix} 11 & -5 & 5 \\ -5 & 3 & -3 \\ 5 & -3 & 3 \end{pmatrix} \in \text{S}_3(\mathbb{R}).$$

On a

$$\begin{aligned}
\chi_A = \det(A - XI_3) &= \begin{vmatrix} 11-X & -5 & 5 \\ -5 & 3-X & -3 \\ 5 & -3 & 3-X \end{vmatrix} \\
&\stackrel{C_2 \leftarrow C_2 + C_3}{=} \begin{vmatrix} 11-X & 0 & 5 \\ -5 & -X & -3 \\ 5 & -X & 3-X \end{vmatrix} \\
&= (-X) \begin{vmatrix} 11-X & 0 & 5 \\ -5 & 1 & -3 \\ 5 & 1 & 3-X \end{vmatrix} \\
&\stackrel{L_3 \leftarrow L_3 - L_2}{=} (-X) \begin{vmatrix} 11-X & 0 & 5 \\ -5 & 1 & -3 \\ 10 & 0 & 6-X \end{vmatrix} \\
&= (-X) \begin{vmatrix} 11-X & 5 \\ 10 & 6-X \end{vmatrix} \\
&\stackrel{C_1 \leftarrow C_1 + C_2}{=} (-X) \begin{vmatrix} 16-X & 5 \\ 16-X & 6-X \end{vmatrix} \\
&= (-X)(16-X) \begin{vmatrix} 1 & 5 \\ 1 & 6-X \end{vmatrix} \\
&= (-X)(16-X)(1-X).
\end{aligned}$$

En particulier, comme les valeurs propres de  $A$  sont positives ou nulles,  $A$  est positive (non définie positive car 0 est une valeur propre de  $A$ ).

Par ailleurs,  $\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \in E_0 (= \text{Ker } A)$ ,  $\begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix} \in E_{16}$  et  $\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} \in E_1$  donc  $\left\{ \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} \right\}$  est une base orthonormale de  $E_0$ ,  $\left\{ \begin{pmatrix} \frac{2}{\sqrt{6}} \\ -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \end{pmatrix} \right\}$  est une base orthonormale de  $E_{16}$  et  $\left\{ \begin{pmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{3}} \end{pmatrix} \right\}$  est une base orthonormale de  $E_1$ . La matrice

$$P := \begin{pmatrix} 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{3}} \end{pmatrix}$$

est alors une matrice orthogonale et on a

$$P^{-1}AP = {}^tPAP = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 16 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

La racine carrée de  $A$  est donc

$$R := P \begin{pmatrix} 0 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{pmatrix} {}^t P = \begin{pmatrix} 3 & -1 & 1 \\ -1 & 1 & -1 \\ 1 & -1 & 1 \end{pmatrix}.$$

## 5.4 Décomposition polaire

Soit  $n \in \mathbb{N} \setminus \{0\}$ . L'existence d'une racine carrée pour toute matrice symétrique positive va nous permettre de montrer le théorème de décomposition suivant :

**Théorème 5.4.1** (Décomposition polaire). *Soit  $A \in \text{GL}_n(\mathbb{R})$  une matrice inversible. Il existe une matrice orthogonale  $O \in \text{O}_n(\mathbb{R})$  et une matrice symétrique définie positive  $S \in \text{S}_n(\mathbb{R})$  telles que  $A = OS$ . De plus, le couple  $(O, S)$  est unique, et l'égalité  $A = OS$  est appelée la décomposition polaire de  $A$ .*

Pour démontrer ce résultat, nous aurons besoin du lemme suivant :

**Lemme 5.4.2.** *Soit  $M \in \text{M}_n(\mathbb{R})$ . La matrice  ${}^tMM$  de  $\text{M}_n(\mathbb{R})$  est symétrique positive. Si  $M$  est de plus inversible, la matrice symétrique  ${}^tMM$  est alors définie positive.*

*Démonstration.* La matrice  ${}^tMM$  est symétrique car

$${}^t({}^tMM) = {}^tM {}^t({}^tM) = {}^tMM.$$

Montrons à présent que toutes les valeurs propres de la matrice symétrique  ${}^tMM \in \text{S}_n(\mathbb{R})$  sont positives. Pour cela, notons  $f$  l'endomorphisme de  $\mathbb{R}^n$  représenté par  $M$  dans la base canonique. Alors la matrice symétrique  ${}^tMM$  est la matrice représentative de l'endomorphisme auto-adjoint  $f^* \circ f$  de  $(\mathbb{R}^n, \langle \cdot, \cdot \rangle_{\text{can}})$  dans la base canonique.

Soit donc  $\lambda \in \mathbb{R}$  une valeur propre de  ${}^tMM$  et soit  $v \in \mathbb{R}^n$  un vecteur propre de  $f$  associé à  $\lambda$ . D'une part, on a

$$\langle f^* \circ f(v), v \rangle_{\text{can}} = \langle f(v), f(v) \rangle_{\text{can}} = \|f(v)\|^2$$

et, d'autre part,

$$\langle f^* \circ f(v), v \rangle_{\text{can}} = \langle \lambda v, v \rangle_{\text{can}} = \lambda \|v\|^2.$$

Ainsi,  $\lambda = \frac{\|f(v)\|^2}{\|v\|^2} \geq 0$ .

Si l'on suppose de plus que  $M$  est inversible, alors  $f$  est un isomorphisme et, avec les notations ci-dessus,  $f(v) \neq 0_{\mathbb{R}^n}$ . On a donc  $\|f(v)\|^2 \neq 0$  et  $\lambda \neq 0$ . La matrice symétrique  ${}^tMM$  est donc dans ce cas définie positive.  $\square$

Nous montrons maintenant le théorème 5.4.1 :

*Démonstration du théorème 5.4.1.* On considère la matrice  ${}^tAA \in \text{M}_n(\mathbb{R})$ , qui est symétrique définie positive par le lemme précédent. On note ensuite  $S$  la racine carrée de  ${}^tAA$  :  $S$  est également symétrique définie positive.

On pose ensuite  $O := AS^{-1}$ . On a

$$\begin{aligned}
 {}^tOO &= {}^t(AS^{-1})AS^{-1} \\
 &= {}^tS^{-1}{}^tAAS^{-1} \\
 &= {}^tS^{-1}S^2S^{-1} \text{ (} S \text{ est la racine carrée de } {}^tAA \text{)} \\
 &= {}^tS^{-1}S \\
 &= S^{-1}S \text{ (} S \text{ est symétrique donc } {}^tS = S \text{)} \\
 &= I_n
 \end{aligned}$$

donc  $O \in O_n(\mathbb{R})$ , et on a  $A = OS$ .

Montrons l'unicité de ce couple  $(O, S)$ . Soient donc  $\tilde{O} \in O_n(\mathbb{R})$  une matrice orthogonale et  $\tilde{S} \in S_n(\mathbb{R})$  une matrice définie positive telle que  $A = \tilde{O}\tilde{S}$ . Alors

$$\begin{aligned}
 {}^tAA &= {}^t(\tilde{O}\tilde{S})\tilde{O}\tilde{S} \\
 &= {}^t\tilde{S}{}^t\tilde{O}\tilde{O}\tilde{S} \\
 &= {}^t\tilde{S}\tilde{S} \text{ (} \tilde{O} \text{ est orthogonale donc } {}^t\tilde{O}\tilde{O} = I_n \text{)} \\
 &= \tilde{S}^2 \text{ (} \tilde{S} \text{ est symétrique donc } {}^t\tilde{S} = \tilde{S} \text{)}
 \end{aligned}$$

donc  $\tilde{S}$  est la racine carrée de  ${}^tAA$  donc  $\tilde{S} = S$ . Enfin,  $\tilde{O} = AS^{-1} = O$ . □

La preuve de l'existence de la décomposition polaire pour une matrice inversible nous donne une méthode pour déterminer cette décomposition :

*Exemple 5.4.3.* On considère la matrice inversible

$$A := \begin{pmatrix} 1 & 2 & 1 \\ -2 & -1 & -1 \\ -1 & -1 & -2 \end{pmatrix} \in \text{GL}_3(\mathbb{R}).$$

Pour déterminer sa décomposition polaire, on commence par calculer la racine carrée de

${}^tAA$  : on a  ${}^tAA = \begin{pmatrix} 6 & 5 & 5 \\ 5 & 6 & 5 \\ 5 & 5 & 6 \end{pmatrix}$  et

$$\begin{aligned} \chi_A = \det(A - XI_3) &= \begin{vmatrix} 6-X & 5 & 5 \\ 5 & 6-X & 5 \\ 5 & 5 & 6-X \end{vmatrix} \\ &\stackrel{C_1 \leftarrow C_1 + C_2 + C_3}{=} \begin{vmatrix} 16-X & 5 & 5 \\ 16-X & 6-X & 5 \\ 16-X & 5 & 6-X \end{vmatrix} \\ &= (16-X) \begin{vmatrix} 1 & 5 & 5 \\ 1 & 6-X & 5 \\ 1 & 5 & 6-X \end{vmatrix} \\ &\stackrel{L_2 \leftarrow L_2 - L_1, L_3 \leftarrow L_3 - L_1}{=} (16-X) \begin{vmatrix} 1 & 5 & 5 \\ 0 & 1-X & 0 \\ 0 & 0 & 1-X \end{vmatrix} \\ &= (16-X)(1-X)^2 \end{aligned}$$

Une base orthonormale de  $E_{16}$  est  $\left\{ \begin{pmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{pmatrix} \right\}$ . Une base de  $E_1$  est  $\left\{ \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \right\}$ . En appliquant le procédé d'orthonormalisation à cette dernière famille libre de  $M_{3,1}(\mathbb{R})$ , on obtient la

base orthonormale  $\left\{ \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \\ -\frac{2}{\sqrt{6}} \end{pmatrix} \right\}$  de  $E_1$ .

En posant  $P := \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & -\frac{2}{\sqrt{6}} \end{pmatrix} \in O_3(\mathbb{R})$ , on a alors

$${}^tAA = P \begin{pmatrix} 16 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} {}^tP$$

et la racine carrée de  ${}^tAA$  est donc

$$S := P \begin{pmatrix} 4 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} {}^tP = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}.$$

Enfin,  $S^{-1} = P \begin{pmatrix} \frac{1}{4} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} {}^tP = \frac{1}{4} \begin{pmatrix} 3 & -1 & -1 \\ -1 & 3 & -1 \\ -1 & -1 & 3 \end{pmatrix}$  et on calcule  $O := AS^{-1} = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} \in$

$O_3(\mathbb{R})$ .





- Les isométries directes de  $\mathbb{R}^3$  sont les rotations autour d'un axe, chacune de matrice  $\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{pmatrix}$ ,  $\theta \in \mathbb{R}$ , dans une base orthonormale correspondante (quand  $\theta = \pi$ , on parle de retournement).

Les isométries indirectes de  $\mathbb{R}^3$  sont les compositions d'une symétrie orthogonale par rapport à un plan (une telle transformation est également appelée réflexion) et d'une rotation autour de l'axe orthogonal à ce plan (par rapport au produit scalaire canonique de  $\mathbb{R}^3$ ),

chacune de matrice représentative  $\begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{pmatrix}$ ,  $\theta \in \mathbb{R}$ , dans une base orthonormale correspondante  $\begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ .

Pour montrer le théorème 5.5.1, nous utiliserons le lemme suivant :

**Lemme 5.5.3.** *Soit  $F$  un sev de  $E$ . Si  $F$  est stable par l'endomorphisme orthogonal  $f$ , alors l'orthogonal  $F^\perp$  de  $F$  est également stable par  $f$ .*

*Démonstration.* Soit  $v \in F^\perp$ . On montre que  $f(v) \in F^\perp$ . Soit donc  $w \in F$ . Comme, pour tous  $w_1, w_2 \in F$ , on a  $\langle f(w_1), f(w_2) \rangle = \langle w_1, w_2 \rangle$  (car  $f$  est orthogonal), la restriction  $f|_F$  de  $f$  à  $F$  ( $F$  est stable par  $f$ ) est un endomorphisme orthogonal de  $F$ , en particulier, il est bijectif : il existe donc  $\tilde{w} \in F$  tel que  $w = f(\tilde{w})$ .

On a alors

$$\begin{aligned} \langle f(v), w \rangle &= \langle f(v), f(\tilde{w}) \rangle \\ &= \langle v, \tilde{w} \rangle \quad (\text{car } f \text{ est orthogonal}) \\ &= 0 \quad (\text{car } v \in F^\perp \text{ et } \tilde{w} \in F). \end{aligned}$$

□

*Démonstration du théorème 5.5.1.* On procède par récurrence sur la dimension  $n$  de  $E$  : on montre par récurrence que pour tout  $n \in \mathbb{N} \setminus \{0\}$ , tout espace euclidien  $(E, \langle \cdot, \cdot \rangle)$  de dimension  $n$  et tout endomorphisme orthogonal  $f$  de  $E$ , il existe une base orthonormale de  $E$  dans laquelle la matrice représentative de  $f$  est de la forme voulue.

Pour  $n = 1$ , soit  $E = \text{Vect}\{v\}$  un  $\mathbb{R}$ -espace vectoriel de dimension 1 muni d'un produit scalaire  $\langle \cdot, \cdot \rangle$  et soit  $f$  un endomorphisme orthogonal de  $E$ . Il existe  $\alpha \in \mathbb{R}$  tel que  $f(v) = \alpha v$  et alors  $\|f(v)\| = \|\alpha v\| = |\alpha| \|v\|$ . De plus, comme  $f$  est orthogonal, on a  $\|f(v)\| = \|v\|$  donc  $|\alpha| \|v\| = \|v\|$  et donc  $|\alpha| = 1$  ( $\|v\| \neq 0$  car  $v \neq 0_E$  car  $v$  engendre  $E$ ). Ainsi  $\alpha \in \{+1; -1\}$  et la matrice de  $f$  dans la base  $\{v\}$ , ainsi que dans la base orthonormale  $\left\{ \frac{v}{\|v\|} \right\}$ , est  $(\alpha)$ .

Supposons maintenant la propriété vérifiée pour tout entier naturel non nul strictement plus petit que  $n$  avec  $n \in \mathbb{N} \setminus \{0\}$  fixé et considérons l'endomorphisme orthogonal  $f$  de l'espace euclidien  $E$  de dimension  $n$ .



D'autre part, le sev  $F := \text{Vect}\{v, f(v)\}$  de  $E$  de dimension 2 est stable par  $f$  : on a  $h(v) = \lambda v$  i.e.  $(f + f^*)(v) = \lambda v$ , donc

$$\lambda f(v) = f(\lambda v) = f \circ (f + f^*)(v) = f^2(v) + f \circ f^*(v) = f^2(v) + v$$

( $f$  est orthogonal donc  $f \circ f^* = \text{Id}_E$ ) d'où  $f^2(v) = \lambda f(v) - v \in F$ .

Considérons alors une base orthonormale  $\mathcal{B}'$  de  $F$ . La matrice représentative de l'endomorphisme orthogonal  $f|_F$  de  $F$  dans la base orthonormale  $\mathcal{B}'$  est une matrice orthogonale de  $O_2(\mathbb{R})$  (proposition 2.8.8).  $\text{Mat}_{\mathcal{B}'}(f)$  est donc de la forme  $\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$  ou  $\begin{pmatrix} \cos \theta & \sin \theta \\ \sin \theta & -\cos \theta \end{pmatrix}$  avec  $\theta \in \mathbb{R}$  (cf feuille de TD 2, exercice 7). Mais, comme  $f$  n'admet pas de valeur propre réelle,  $\text{Mat}_{\mathcal{B}'}(f)$  ne peut pas être symétrique donc ne peut pas être de la seconde forme. Ainsi, il existe  $\theta \in \mathbb{R}$  tel que

$$\text{Mat}_{\mathcal{B}'}(f) = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} = R(\theta),$$

avec  $\theta \notin \{k\pi \mid k \in \mathbb{Z}\}$  car  $\text{Mat}_{\mathcal{B}'}(f)$  n'est pas symétrique ( $R(0) = I_2$ ,  $R(\pi) = -I_2$ ).

Enfin,  $F^\perp$  est également stable par  $f$  par le lemme 5.5.3 donc, par hypothèse de récurrence (puisque  $\dim(F^\perp) < n$ ), il existe une base orthonormale  $\mathcal{B}_0$  de  $F^\perp$  telle que

$$\text{Mat}_{\mathcal{B}'}(f|_{F^\perp}) = \begin{pmatrix} R(\theta_1) & & 0 \\ & \ddots & \\ 0 & & R(\theta_s) \end{pmatrix}$$

où, pour tout  $j \in \{1, \dots, s\}$ ,  $R(\theta_j) := \begin{pmatrix} \cos \theta_j & -\sin \theta_j \\ \sin \theta_j & \cos \theta_j \end{pmatrix}$  avec  $\theta_j \in \mathbb{R} \setminus \{k\pi \mid k \in \mathbb{Z}\}$  ( $f$  ne possède pas de valeur propre réelle). En notant  $\mathcal{B} := \{\mathcal{B}', \mathcal{B}_0\}$ , la famille  $\mathcal{B}$  est alors une base orthonormale de  $E$  et on a

$$\text{Mat}_{\mathcal{B}}(f) = \begin{pmatrix} R(\theta) & & & 0 \\ & R(\theta_1) & & \\ & & \ddots & \\ 0 & & & R(\theta_s) \end{pmatrix}$$

□

*Remarque 5.5.4.* • Au cours de la preuve, on a montré en particulier que les seules valeurs propres réelles possibles pour un endomorphisme orthogonal sont 1 et  $-1$ .

- Si, avec les notations ci-dessus, l'endomorphisme orthogonal  $f$  possède 1 et  $-1$  comme valeurs propres, alors les sous-espaces propres associés  $E_1$  et  $E_{-1}$  sont orthogonaux. En effet, si  $v \in E_1 \setminus \{0_E\}$  et  $w \in E_{-1} \setminus \{0_E\}$ , on a d'une part

$$\langle f(v), f(w) \rangle = \langle v, -w \rangle = -\langle v, w \rangle,$$

et d'autre part

$$\langle f(v), f(w) \rangle = \langle v, w \rangle$$

(car  $f$  est orthogonal) et donc  $\langle v, w \rangle = -\langle v, w \rangle$  donc  $\langle v, w \rangle = 0$ .

Par ailleurs, le sous-espace vectoriel  $E_1 \oplus E_{-1}$  de  $E$  est stable par  $f$  donc  $(E_1 \oplus E_{-1})^\perp$  également.

Au total, pour calculer une réduction de  $f$  comme dans le théorème 5.5.1, on peut donc commencer par déterminer une base orthonormale  $\mathcal{B}'$  de  $E_1$ , une base orthonormale  $\mathcal{B}''$  de  $E_{-1}$  puis une base du sous-espace stable  $\{\mathcal{B}', \mathcal{B}''\}^\perp$  (donné par des équations linéaires). Dans la représentation matricielle correspondante de  $f$ , le bloc correspondant à la restriction de  $f$  à ce dernier sous-espace stable est orthogonal sans valeur propre réelle et on peut alors lui appliquer la méthode (algorithmique) du théorème pour ce cas.

*Exemple 5.5.5.* Considérons la matrice orthogonale

$$A := \frac{1}{3} \begin{pmatrix} 2 & -1 & 2 \\ 2 & 2 & -1 \\ -1 & 2 & 2 \end{pmatrix}$$

de l'exemple 2.8.11 1. Calculons son polynôme caractéristique : on a

$$\begin{aligned} \chi_A = \det(A - XI_3) &= \begin{vmatrix} \frac{2}{3} - X & -\frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{2}{3} - X & -\frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} & \frac{2}{3} - X \end{vmatrix} \\ &= \begin{vmatrix} 1 - X & -\frac{1}{3} & \frac{2}{3} \\ 1 - X & \frac{2}{3} - X & -\frac{1}{3} \\ 1 - X & \frac{2}{3} & \frac{2}{3} - X \end{vmatrix} \\ &= (1 - X) \begin{vmatrix} 1 & -\frac{1}{3} & \frac{2}{3} \\ 1 & \frac{2}{3} - X & -\frac{1}{3} \\ 1 & \frac{2}{3} & \frac{2}{3} - X \end{vmatrix} \\ &= (1 - X) \begin{vmatrix} 1 & -\frac{1}{3} & \frac{2}{3} \\ 0 & 1 - X & -1 \\ 0 & 1 & -X \end{vmatrix} \\ &= (1 - X)(X^2 - X + 1) \end{aligned}$$

On a  $E_1 = \text{Ker} \left( \frac{1}{3} \begin{pmatrix} -1 & -1 & 2 \\ 2 & -1 & -1 \\ -1 & 2 & -1 \end{pmatrix} \right)$  et le vecteur colonne  $Y_1 := \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$  de norme 1 engendre  $E_1$ .

De plus,  $\left( \text{Vect} \left\{ \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\} \right)^\perp = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in M_{3,1}(\mathbb{R}) \mid x + y + z = 0 \right\}$ , dont une base ortho-normale est formée des vecteurs  $Y_2 := \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$  et  $Y_3 := \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}$ .

On a

$$AY_2 = A \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ -\frac{1}{\sqrt{2}} \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = \frac{1}{2\sqrt{2}} (\sqrt{2} Y_2 + \sqrt{6} Y_3) = \frac{1}{2} Y_2 + \frac{\sqrt{3}}{2} Y_3$$

et

$$AY_3 = A \begin{pmatrix} \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \\ -\frac{1}{\sqrt{6}} \end{pmatrix} = \begin{pmatrix} -\frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{6}} \end{pmatrix} = \frac{1}{\sqrt{6}} \begin{pmatrix} -1 \\ 2 \\ 1 \\ -1 \end{pmatrix} = \frac{1}{2\sqrt{6}} (-3\sqrt{2} Y_2 + \sqrt{6} Y_3) = -\frac{\sqrt{3}}{2} Y_2 + \frac{1}{2} Y_3.$$

Ainsi, si on note

$$P := \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{3}} & 0 & -\frac{1}{\sqrt{2}} \end{pmatrix} \in O_3(\mathbb{R}),$$

on a

$$P^{-1}AP = {}^tPAP = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ 0 & \frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\left(\frac{\pi}{3}\right) & -\sin\left(\frac{\pi}{3}\right) \\ 0 & \sin\left(\frac{\pi}{3}\right) & \cos\left(\frac{\pi}{3}\right) \end{pmatrix}$$

La matrice  $A$  est donc la matrice représentative dans la base canonique de  $\mathbb{R}^3$  de la rotation d'axe la droite vectorielle engendrée par le vecteur  $(1, 1, 1)$  et d'angle  $\frac{\pi}{3}$ .

## Chapitre 6

# Normes matricielles subordonnées, rayon spectral, conditionnement

### 6.1 Introduction

Dans ce chapitre, on étudie des normes particulières sur les espaces de matrices carrées à coefficients réels ou complexes : les normes dites subordonnées. Ces normes possèdent des propriétés adaptées à l'étude des matrices dans différents aspects et utilisations.

On fait également le lien avec la notion de rayon spectral : il s'agit du plus grand module des valeurs propres complexes d'une matrice carrée complexe. On verra notamment que la donnée du rayon spectral d'une matrice permet de déterminer si la suite de ses puissances successives converge vers la matrice nulle ou non.

Enfin, on aborde la question de la sensibilité de la solution d'un système linéaire inversible aux erreurs d'approximations sur les données, des perturbations que l'on maîtrise à l'aide d'une quantité nommée conditionnement.

Dans tout ce chapitre,  $\mathbb{K}$  désigne les corps  $\mathbb{R}$  ou  $\mathbb{C}$ , et  $n$  est un entier naturel non nul.

### 6.2 Normes matricielles subordonnées

Soit  $A = (a_{ij})_{1 \leq i, j \leq n}$  une matrice de  $M_n(\mathbb{K})$ . Dans la définition et proposition 4.2.1, on avait défini la norme

$$\sqrt{\sum_{1 \leq i, j \leq n} |a_{ij}|^2} = \sqrt{\text{Tr}({}^t \bar{A} A)} \in [0, +\infty[.$$

de  $A$ . Dans la suite, on la note  $\|A\|_2$ .

La norme  $\|\cdot\|_2$  sur  $M_n(\mathbb{K})$ , appelée norme de Frobenius sur  $M_n(\mathbb{K})$ , est induite, si  $\mathbb{K} = \mathbb{R}$ , par le produit scalaire  $\langle \cdot, \cdot \rangle$  :

$$\begin{array}{ccc} M_n(\mathbb{R}) \times M_n(\mathbb{R}) & \rightarrow & \mathbb{R} \\ (A, B) & \mapsto & \text{Tr}({}^t AB) \end{array} \quad \text{sur } M_n(\mathbb{R}) \text{ et, si } \mathbb{K} = \mathbb{C},$$

par le produit scalaire hermitien

$$\begin{array}{ccc} M_n(\mathbb{C}) \times M_n(\mathbb{C}) & \rightarrow & \mathbb{C} \\ (A, B) & \mapsto & \text{Tr}({}^t \bar{A} B) \end{array} \quad (\text{cf preuve de la définition})$$

et proposition 4.2.1), et nous avons montré qu'il s'agissait, dans les deux cas, d'une norme matricielle :

**Définition 6.2.1.** Soit  $\|\cdot\| : M_n(\mathbb{K}) \rightarrow [0, +\infty[$  une norme sur  $M_n(\mathbb{K})$ . On dit que  $\|\cdot\|$  est une norme matricielle si pour tous  $A, B \in M_n(\mathbb{K})$ ,  $\|AB\| \leq \|A\|\|B\|$ .

Exemple 6.2.2. • La norme

$$\|\cdot\|_1 : \begin{array}{ccc} M_n(\mathbb{K}) & \rightarrow & [0, +\infty[ \\ A & \mapsto & \|A\|_1 \end{array},$$

avec, pour  $A = (a_{ij})_{1 \leq i, j \leq n} \in M_n(\mathbb{K})$ ,  $\|A\|_1 := \sum_{1 \leq i, j \leq n} |a_{ij}|$ , est également une norme matricielle. En effet, si  $A = (a_{ij})_{1 \leq i, j \leq n}$ ,  $B = (b_{ij})_{1 \leq i, j \leq n} \in M_n(\mathbb{K})$ , on a

$$\begin{aligned} \|AB\|_1 &= \sum_{1 \leq i, j \leq n} \left| \sum_{k=1}^n a_{ik} b_{kj} \right| \\ &\leq \sum_{1 \leq i, j \leq n} \sum_{k=1}^n |a_{ik} b_{kj}| \\ &= \sum_{1 \leq i, j \leq n} \sum_{k=1}^n |a_{ik}| |b_{kj}| \\ &\leq \sum_{1 \leq i, j \leq n} \sum_{k=1}^n |a_{ik}| \left( \sum_{l=1}^n |b_{lj}| \right) \\ &= \left( \sum_{1 \leq i, k \leq n} |a_{ik}| \right) \left( \sum_{1 \leq j, l \leq n} |b_{lj}| \right) \\ &= \|A\|_1 \|B\|_1 \end{aligned}$$

• La norme

$$\|\cdot\|_\infty : \begin{array}{ccc} M_n(\mathbb{K}) & \rightarrow & [0, +\infty[ \\ A & \mapsto & \|A\|_\infty \end{array},$$

avec, pour  $A = (a_{ij})_{1 \leq i, j \leq n} \in M_n(\mathbb{K})$ ,  $\|A\|_\infty := \max_{1 \leq i, j \leq n} |a_{ij}|$ , n'est pas une norme matricielle. En effet, si l'on considère par exemple les matrices  $\begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}$  et  $\begin{pmatrix} 1 & 0 \\ -1 & 0 \end{pmatrix}$  de  $M_2(\mathbb{R}) \subset M_2(\mathbb{C})$ , on a

$$\left\| \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -1 & 0 \end{pmatrix} \right\|_\infty = \left\| \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} \right\|_\infty = 2$$

alors que  $\left\| \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix} \right\|_\infty = \left\| \begin{pmatrix} 1 & 0 \\ -1 & 0 \end{pmatrix} \right\|_\infty = 1$ .



Remarquons que  $\|I_n\|_2 = \sqrt{n}$  et  $\|I_n\|_1 = n$ . Pour différents usages, on aimerait construire des normes matricielles pour lesquelles la norme de la matrice identité est 1. Les normes dites “subordonnées” satisfont cette condition. Soit  $\|\cdot\| : \mathbb{K}^n \rightarrow [0, +\infty[$  une norme sur  $\mathbb{K}^n$ .

**Proposition 6.2.3.** *Soit  $A \in M_n(\mathbb{K})$ . L'application*

$$\psi : \begin{array}{ccc} \mathbb{K}^n \setminus \{(0, \dots, 0)\} & \rightarrow & [0, +\infty[ \\ v & \mapsto & \frac{\|Av\|}{\|v\|} \end{array}$$

admet un maximum et on note

$$\|A\| := \max_{v \in \mathbb{K}^n \setminus \{(0, \dots, 0)\}} \frac{\|Av\|}{\|v\|}.$$

*Démonstration.* Commençons par remarquer que, si  $v \in \mathbb{K}^n \setminus \{(0, \dots, 0)\}$ ,

$$\psi(v) = \frac{\|Av\|}{\|v\|} = \left\| \frac{1}{\|v\|} Av \right\| = \left\| A \left( \frac{v}{\|v\|} \right) \right\|$$

et que  $\frac{v}{\|v\|}$  appartient à la sphère unité  $\mathbb{S}_{\|\cdot\|}^{n-1} := \{w \in \mathbb{K}^n \mid \|w\| = 1\}$ .

Considérons alors l'application

$$f : \begin{array}{ccc} \mathbb{S}_{\|\cdot\|}^{n-1} & \rightarrow & [0, +\infty[ \\ w & \mapsto & \|Aw\| \end{array}$$

$f$  est une application continue (comme composée des applications continues  $\|\cdot\| : \mathbb{K}^n \rightarrow [0, +\infty[$  et  $\mathbb{K}^n \rightarrow \mathbb{K}^n ; w \mapsto Aw$ ) sur le compact  $\mathbb{S}_{\|\cdot\|}^{n-1}$  de  $\mathbb{K}^n$  (la sphère  $\mathbb{S}_{\|\cdot\|}^{n-1}$  est fermée bornée dans le  $\mathbb{K}$ -espace vectoriel normé de dimension finie  $(\mathbb{K}^n, \|\cdot\|)$ ) :  $f$  est donc bornée et atteint ses bornes.

Ensuite, pour tout  $v \in \mathbb{K}^n \setminus \{(0, \dots, 0)\}$ ,

$$\psi(v) = \left\| A \left( \frac{v}{\|v\|} \right) \right\| = f \left( \frac{v}{\|v\|} \right) \leq \max_{w \in \mathbb{S}_{\|\cdot\|}^{n-1}} f(w)$$

donc l'application  $\psi$  est bornée et  $\sup_{v \in \mathbb{K}^n \setminus \{(0, \dots, 0)\}} \psi(v) \leq \max_{w \in \mathbb{S}_{\|\cdot\|}^{n-1}} f(w)$ . Enfin, si  $w \in \mathbb{S}_{\|\cdot\|}^{n-1}$ ,  $f(w) = \psi(w)$  donc

$$\max_{w \in \mathbb{S}_{\|\cdot\|}^{n-1}} f(w) = \sup_{v \in \mathbb{K}^n \setminus \{(0, \dots, 0)\}} \psi(v) = \max_{v \in \mathbb{K}^n \setminus \{(0, \dots, 0)\}} \psi(v).$$

□

*Remarque 6.2.4.* Dans la démonstration précédente, on a montré au passage que

$$\|A\| = \max_{v \in \mathbb{K}^n \setminus \{(0, \dots, 0)\}} \frac{\|Av\|}{\|v\|} = \max_{v \in \mathbb{K}^n, \|v\|=1} \|Av\|.$$

**Proposition et Définition 6.2.5.** *L'application*

$$\| \cdot \| : M_n(\mathbb{K}) \rightarrow [0, +\infty[ \\ A \mapsto \|A\|$$

est une norme matricielle sur  $M_n(\mathbb{K})$ , appelée norme subordonnée à la norme  $\| \cdot \|$ . De plus,  $\|I_n\| = 1$ .

*Démonstration.* Commençons par remarquer qu'en effet

$$\|I_n\| = \max_{v \in \mathbb{K}^n \setminus \{(0, \dots, 0)\}} \frac{\|I_n v\|}{\|v\|} = \max_{v \in \mathbb{K}^n \setminus \{(0, \dots, 0)\}} \frac{\|v\|}{\|v\|} = 1.$$

Montrons ensuite que  $\| \cdot \| : M_n(\mathbb{K}) \rightarrow [0, +\infty[$  est une norme sur  $M_n(\mathbb{K})$  :

- soient  $A \in M_n(\mathbb{K})$  et  $\lambda \in \mathbb{K}$ , on a

$$\|\lambda A\| = \max_{v \in \mathbb{K}^n \setminus \{(0, \dots, 0)\}} \frac{\|(\lambda A)v\|}{\|v\|} = \max_{v \in \mathbb{K}^n \setminus \{(0, \dots, 0)\}} |\lambda| \frac{\|Av\|}{\|v\|} = |\lambda| \|A\|.$$

- soit  $A \in M_n(\mathbb{K})$  telle que  $\|A\| = \max_{v \in \mathbb{K}^n \setminus \{(0, \dots, 0)\}} \frac{\|Av\|}{\|v\|} = 0$ , alors, pour tout  $v \in \mathbb{K}^n \setminus \{(0, \dots, 0)\}$ ,  $\frac{\|Av\|}{\|v\|} = 0$  donc  $\|Av\| = 0$  donc  $Av = (0, \dots, 0)$  (car  $\| \cdot \|$  est une norme sur  $\mathbb{K}^n$ ). En particulier, si  $\{e_1, \dots, e_n\}$  désigne la base canonique de  $\mathbb{K}^n$ , pour tout  $i \in \{1, \dots, n\}$ ,  $Ae_i = 0$  i.e. la  $i^{\text{ème}}$  colonne de  $A$  est nulle, et donc  $A = 0_n$ .
- soient  $A, B \in M_n(\mathbb{K})$ , alors, pour tout  $v \in \mathbb{S}_{\| \cdot \|}^{n-1}$ ,

$$\begin{aligned} \|(A+B)v\| &= \|Av + Bv\| \\ &\leq \|Av\| + \|Bv\| \text{ (car } \| \cdot \| \text{ est une norme sur } \mathbb{K}^n) \\ &\leq \|A\| + \|B\| \end{aligned}$$

et donc

$$\|A+B\| = \max_{v \in \mathbb{S}_{\| \cdot \|}^{n-1}} \|(A+B)v\| \leq \|A\| + \|B\|.$$

Montrons enfin que la norme  $\| \cdot \|$  sur  $M_n(\mathbb{K})$  est une norme matricielle. Soient donc  $A, B \in M_n(\mathbb{K})$  et soit  $v \in \mathbb{K}^n \setminus \{(0, \dots, 0)\}$ . Si  $Bv = (0, \dots, 0)$ , on a  $\frac{\|(AB)v\|}{\|v\|} = \frac{\|A(Bv)\|}{\|v\|} = 0 \leq \|A\| \|B\|$ . Si  $Bv \neq (0, \dots, 0)$ , on a

$$\frac{\|(AB)v\|}{\|v\|} = \frac{\|(AB)v\|}{\|Bv\|} \times \frac{\|Bv\|}{\|v\|} = \frac{\|A(Bv)\|}{\|Bv\|} \times \frac{\|Bv\|}{\|v\|} \leq \|A\| \|B\|.$$

Ainsi,  $\|AB\| = \max_{v \in \mathbb{K}^n \setminus \{(0, \dots, 0)\}} \frac{\|(AB)v\|}{\|v\|} \leq \|A\| \|B\|$ . □

Soit  $A = (a_{ij})_{1 \leq i, j \leq n} \in M_n(\mathbb{K})$ . On va donner les expressions de  $\|A\|_1$  et  $\|A\|_\infty$  en fonction des coefficients de  $A$ . On rappelle que, si  $v = (v_1, \dots, v_n) \in \mathbb{K}^n$ ,  $\|v\|_1 = \sum_{i=1}^n |v_i|$  et  $\|v\|_\infty = \max_{1 \leq i \leq n} |v_i|$ .

**Théorème 6.2.6.** *On a*

- $\|A\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^n |a_{ij}|$ ,
- $\|A\|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}|$ .

*Remarque 6.2.7.* • Pour  $i$  et  $j$  dans  $\{1, \dots, n\}$ , si on note  $L_i$  la  $i^{\text{ème}}$  ligne de  $A$  et  $C_j$  la  $j^{\text{ème}}$  colonne de  $A$  et si on les considère comme des vecteurs de  $\mathbb{K}^n$ , on a

$$\|A\|_1 = \max_{1 \leq j \leq n} \|C_j\|_1 \quad \text{et} \quad \|A\|_\infty = \max_{1 \leq i \leq n} \|L_i\|_1.$$

- Attention : en général  $\|A\|_1 \neq \|A\|_\infty$  (resp.  $\|A\|_\infty \neq \|A\|_1$ ).

*Démonstration du théorème 6.2.6.* Montrons tout d'abord l'égalité  $\|A\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^n |a_{ij}|$ .

Soit  $v = (x_1, \dots, x_n) \in \mathbb{K}^n \setminus \{(0, \dots, 0)\}$ , on a  $Av = \left( \sum_{k=1}^n a_{1k}x_k, \dots, \sum_{k=1}^n a_{nk}x_k \right)$  donc

$$\begin{aligned} \|Av\|_1 &= \sum_{i=1}^n \left| \sum_{k=1}^n a_{ik}x_k \right| \\ &\leq \sum_{i=1}^n \sum_{k=1}^n |a_{ik}x_k| \\ &= \sum_{i=1}^n \sum_{k=1}^n |a_{ik}| |x_k| \\ &= \sum_{k=1}^n \left( \sum_{i=1}^n |a_{ik}| \right) |x_k| \\ &\leq \sum_{k=1}^n \left( \max_{1 \leq j \leq n} \sum_{i=1}^n |a_{ij}| \right) |x_k| \\ &= \left( \max_{1 \leq j \leq n} \sum_{i=1}^n |a_{ij}| \right) \sum_{k=1}^n |x_k| \\ &= \left( \max_{1 \leq j \leq n} \sum_{i=1}^n |a_{ij}| \right) \|v\|_1 \end{aligned}$$

et ainsi  $\frac{\|Av\|_1}{\|v\|_1} \leq \max_{1 \leq j \leq n} \sum_{i=1}^n |a_{ij}|$  et  $\|A\|_1 \leq \max_{1 \leq j \leq n} \sum_{i=1}^n |a_{ij}|$ .

Mais, si l'on note  $j_0$  l'indice de  $\{1, \dots, n\}$  tel que  $\sum_{i=1}^n |a_{ij_0}| = \max_{1 \leq j \leq n} \sum_{i=1}^n |a_{ij}|$  et si  $\{e_1, \dots, e_n\}$  désigne la base canonique de  $\mathbb{K}^n$ , on a

$$\frac{\|Ae_{j_0}\|_1}{\|e_{j_0}\|_1} = \|Ae_{j_0}\|_1 = \sum_{i=1}^n |a_{ij_0}| = \max_{1 \leq j \leq n} \sum_{i=1}^n |a_{ij}|$$

et donc

$$\|A\|_1 \leq \max_{1 \leq j \leq n} \sum_{i=1}^n |a_{ij}| = \frac{\|Ae_{j_0}\|_1}{\|e_{j_0}\|_1} \leq \|A\|_1,$$

d'où l'égalité  $\|A\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^n |a_{ij}|$ .

Montrons à présent l'égalité  $\|A\|_\infty = \max_{1 \leq i \leq n} \sum_{k=1}^n |a_{ik}|$ .

Avec les notations ci-dessus, on a

$$\begin{aligned} \|Av\|_\infty &= \max_{1 \leq i \leq n} \left| \sum_{k=1}^n a_{ik} x_k \right| \\ &\leq \max_{1 \leq i \leq n} \sum_{k=1}^n |a_{ik} x_k| \\ &= \max_{1 \leq i \leq n} \sum_{k=1}^n |a_{ik}| |x_k| \\ &\leq \max_{1 \leq i \leq n} \sum_{k=1}^n |a_{ik}| \left( \max_{1 \leq j \leq n} |x_j| \right) \\ &= \max_{1 \leq i \leq n} \sum_{k=1}^n |a_{ik}| \|v\|_\infty \\ &= \left( \max_{1 \leq i \leq n} \sum_{k=1}^n |a_{ik}| \right) \|v\|_\infty \end{aligned}$$

et ainsi  $\frac{\|Av\|_\infty}{\|v\|_\infty} \leq \max_{1 \leq i \leq n} \sum_{k=1}^n |a_{ik}|$  et  $\|A\|_\infty \leq \max_{1 \leq i \leq n} \sum_{k=1}^n |a_{ik}|$ .

Notons ensuite  $i_0$  l'indice de  $\{1, \dots, n\}$  tel que  $\sum_{k=1}^n |a_{i_0 k}| = \max_{1 \leq i \leq n} \sum_{k=1}^n |a_{ik}|$  et notons  $v_0$  le vecteur de  $\mathbb{K}^n$  dont, pour tout  $j \in \{1, \dots, n\}$ , la  $j^{\text{ème}}$  coordonnée notée  $y_j$  est

$$\begin{cases} e^{-i \operatorname{Arg}(a_{i_0 j})} & \text{si } a_{i_0 j} \neq 0 \text{ (si } a_{i_0 j} \in \mathbb{R}, e^{-i \operatorname{Arg}(a_{i_0 j})} \in \{-1; 1\}), \\ 0 & \text{si } a_{i_0 j} = 0, \end{cases}$$

$$\text{alors } \sum_{k=1}^n a_{i_0 k} y_k = \sum_{k=1}^n |a_{i_0 k}|.$$

Si  $v_0$  est le vecteur nul de  $\mathbb{K}^n$ , cela signifie que tous les coefficients de la matrice  $A$  sont nuls et, dans ce cas, on a bien l'égalité  $\|A\|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}|$ . Si  $v_0$  n'est pas le vecteur nul, alors, comme les coefficients non nuls de  $v_0$  sont de module 1,  $v_0$  appartient à la sphère unité  $\mathbb{S}_{\|\cdot\|_\infty}^{n-1} = \{w \in \mathbb{K}^n \mid \|w\|_\infty = 1\}$  et on a

$$\|A\|_\infty \leq \max_{1 \leq i \leq n} \sum_{k=1}^n |a_{ik}| = \sum_{k=1}^n |a_{i_0 k}| = \sum_{k=1}^n a_{i_0 k} y_k \leq \left| \sum_{k=1}^n a_{i_0 k} y_k \right| \leq \max_{1 \leq i \leq n} \left| \sum_{k=1}^n a_{ik} y_k \right| = \|Av_0\|_\infty \leq \|A\|_\infty.$$

$$\text{D'où l'égalité } \|A\|_\infty = \max_{1 \leq i \leq n} \sum_{k=1}^n |a_{ik}|. \quad \square$$

*Exemple 6.2.8.* Pour  $A := \begin{pmatrix} 1 & 0 & -6 \\ -2 & -4 & 3 \\ -1 & 5 & 2 \end{pmatrix} \in M_3(\mathbb{R})$ , on a  $\|A\|_1 = \max\{1+2+1, 0+4+5, 6+3+2\} = 11$  et  $\|A\|_\infty = \max\{1+0+6, 2+4+3, 1+5+2\} = 9$ .

*Remarque 6.2.9.* Si  $A$  est une matrice symétrique  $\|A\|_1 = \|A\|_\infty$ .

### 6.3 Rayon spectral

Soit  $A \in M_n(\mathbb{K}) \subset M_n(\mathbb{C})$ . On considère dans cette partie les valeurs propres complexes de  $A$  : on note  $\text{Sp}_{\mathbb{C}}(A)$  le spectre de  $A$  en tant que matrice de  $M_n(\mathbb{C})$ .

**Définition 6.3.1.** On appelle rayon spectral de  $A$  la quantité  $\rho(A) := \max_{\lambda \in \text{Sp}_{\mathbb{C}}(A)} |\lambda| \in [0, +\infty[$ .

*Exemple 6.3.2.* • Le rayon spectral de la matrice  $\begin{pmatrix} 1 & 0 \\ 0 & -2 \end{pmatrix} \in M_2(\mathbb{R})$  est 2.

- Le rayon spectral de la matrice  $\begin{pmatrix} -3i & 0 \\ 0 & i \end{pmatrix} \in M_2(\mathbb{C})$  est 3.
- Si  $A := \begin{pmatrix} 1 & 2 \\ -1 & 4 \end{pmatrix} \in M_2(\mathbb{R})$ ,  $\text{Sp}_{\mathbb{C}}(A) = \text{Sp}_{\mathbb{R}}(A) = \{2; 3\}$  (exemple 3.3.2) donc  $\rho(A) = 3$ .
- Si  $A := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in M_2(\mathbb{R})$ ,  $\text{Sp}_{\mathbb{C}}(A) = \{-i; i\}$  (remarque 3.3.3) donc  $\rho(A) = 1$ .
- Si  $A := \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -2 \\ 0 & 1 & 0 \end{pmatrix} \in M_3(\mathbb{R})$ ,  $\chi_A = (1-X)(X^2+2) = (1-X)(\sqrt{2}i-X)(-\sqrt{2}i-X)$  donc  $\text{Sp}_{\mathbb{C}}(A) = \{1, \sqrt{2}i, -\sqrt{2}i\}$  donc  $\rho(A) = \sqrt{2}$ .

*Remarque 6.3.3.* Si  $A \in M_n(\mathbb{R})$  et  $\chi_A$  est scindé sur  $\mathbb{R}$ ,  $\text{Sp}_{\mathbb{C}}(A) = \text{Sp}_{\mathbb{R}}(A)$  et donc  $\rho(A) = \max_{\lambda \in \text{Sp}_{\mathbb{R}}(A)} |\lambda|$ .

Dans cette section, nous allons établir deux résultats importants en relation avec le rayon spectral. Le premier consiste en une expression de la norme matricielle subordonnée à la norme euclidienne de  $\mathbb{R}^n$  mettant en jeu le rayon spectral. Le second est un lien entre le rayon spectral d'une matrice complexe et la convergence de la suite de ses puissances successives.

Commençons par énoncer le premier résultat. Ci-dessous, la notation  $\|\cdot\|_2$  désigne la norme matricielle sur  $M_n(\mathbb{R})$  subordonnée à la norme euclidienne  $\|\cdot\|_2$  sur  $\mathbb{R}^n$  i.e. l'application qui à tout vecteur  $v$  de  $\mathbb{R}^n$  associe  $\|v\|_2 := \sqrt{\langle v, v \rangle_{\text{can}}} = \sqrt{{}^t v v}$  (ici et ci-dessous, on identifie un vecteur de  $\mathbb{R}^n$  avec le vecteur colonne de ses coordonnées dans la base canonique de  $\mathbb{R}^n$ ).

**Théorème 6.3.4.** *Supposons que  $A \in M_n(\mathbb{R})$ . Alors  $\|A\|_2 = \sqrt{\rho({}^t A A)}$ .*

La preuve de ce théorème reposera sur le lemme technique suivant :

**Lemme 6.3.5.** *Soit  $S \in S_n(\mathbb{R})$  une matrice symétrique positive. On considère l'application*

$$R_S : \begin{array}{ccc} \mathbb{R}^n \setminus \{(0, \dots, 0)\} & \rightarrow & \mathbb{R} \\ v & \mapsto & \frac{{}^t v S v}{{}^t v v} \end{array}$$

(appelée quotient de Rayleigh). L'application  $R_S$  est bornée et

$$\sup_{v \in \mathbb{R}^n \setminus \{(0, \dots, 0)\}} R_S(v) = \max_{v \in \mathbb{R}^n \setminus \{(0, \dots, 0)\}} R_S(v) = \rho(S).$$

*Démonstration.* Comme  $S$  est une matrice symétrique positive, d'après le corollaire 5.2.7 et la proposition 5.3.4, il existe une matrice orthogonale  $O \in O_n(\mathbb{R})$  et une matrice diagonale

$$D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \in M_n(\mathbb{R}) \text{ avec } \lambda_1, \dots, \lambda_n \in [0, +\infty[ \text{ telles que } {}^t O S O = D. \text{ On peut}$$

supposer, quitte à permuter les colonnes de la matrice  $O$  (ce qui modifie pas son caractère orthogonal), que  $0 \leq \lambda_1 \leq \dots \leq \lambda_n$ . En particulier,  $\rho(S) = \lambda_n$ .

Soit maintenant  $v \in \mathbb{R}^n \setminus \{(0, \dots, 0)\}$  et notons  $w = (w_1, \dots, w_n) := {}^t O v \in \mathbb{R}^n \setminus \{(0, \dots, 0)\}$  (les matrices  $O$  et  ${}^t O$  sont inversibles). On a  ${}^t w w = {}^t v O {}^t O v = {}^t v v$  et

$$R_S(v) = \frac{{}^t v S v}{{}^t v v} = \frac{{}^t v O D {}^t O v}{{}^t v v} = \frac{{}^t w D w}{{}^t w w} = \frac{\sum_{i=1}^n \lambda_i w_i^2}{\sum_{i=1}^n w_i^2} \leq \frac{\sum_{i=1}^n \lambda_n w_i^2}{\sum_{i=1}^n w_i^2} = \lambda_n.$$

La fonction  $R_S$  est donc bornée. De plus, pour  $v_0 := O e_n$ , avec  $e_n$  le  $n^{\text{ème}}$  vecteur  $(0, \dots, 0, 1)$  de la base canonique de  $\mathbb{R}^n$ , on a  ${}^t O v_0 = e_n$  et

$$R_S(v_0) = \frac{{}^t v_0 S v_0}{{}^t v_0 v_0} = \frac{{}^t e_n D e_n}{{}^t e_n e_n} = {}^t e_n D e_n = \lambda_n$$

donc la borne  $\lambda_n$  est atteinte et

$$\sup_{v \in \mathbb{R}^n \setminus \{(0, \dots, 0)\}} R_S(v) = \max_{v \in \mathbb{R}^n \setminus \{(0, \dots, 0)\}} R_S(v) = \lambda_n = \rho(S).$$

□

*Démonstration du théorème 6.3.4.* Soit  $v \in \mathbb{R}^n \setminus \{(0, \dots, 0)\}$ , on a

$$\left( \frac{\|Av\|_2}{\|v\|_2} \right)^2 = \frac{{}^t(Av)(Av)}{{}^tvv} = \frac{{}^tv {}^tAA v}{{}^tvv} = R_{{}^tAA}(v).$$

Ainsi, comme la matrice  ${}^tAA$  est symétrique positive (lemme 5.4.2),

$$\|A\|_2^2 = \left( \max_{v \in \mathbb{R}^n \setminus \{(0, \dots, 0)\}} \frac{\|Av\|_2}{\|v\|_2} \right)^2 = \max_{v \in \mathbb{R}^n \setminus \{(0, \dots, 0)\}} \left( \frac{\|Av\|_2}{\|v\|_2} \right)^2 = \max_{v \in \mathbb{R}^n \setminus \{(0, \dots, 0)\}} R_{{}^tAA}(v) = \rho({}^tAA)$$

et donc  $\|A\|_2 = \sqrt{\rho({}^tAA)}$ .

□

**Corollaire 6.3.6.** Si  $A \in S_n(\mathbb{R})$ , alors  $\|A\|_2 = \rho(A)$ .

*Démonstration.*  $A$  étant une matrice symétrique, il existe  $O \in O_n(\mathbb{R})$  et une matrice diagonale  $D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \in M_n(\mathbb{R})$  telles que  ${}^tOSO = D$ . En particulier,  $\text{Sp}_{\mathbb{C}}(A) = \text{Sp}_{\mathbb{R}}(A) = \{\lambda_1, \dots, \lambda_n\}$ . Ainsi,

$${}^tAA = A^2 = (OD {}^tO)^2 = OD^2 {}^tO = O \begin{pmatrix} \lambda_1^2 & & 0 \\ & \ddots & \\ 0 & & \lambda_n^2 \end{pmatrix} {}^tO$$

et  $\text{Sp}_{\mathbb{R}}(A^2) = \{\lambda_1^2, \dots, \lambda_n^2\}$ . Finalement,

$$\begin{aligned} \|A\|_2^2 = \rho({}^tAA) &= \max_{\mu \in \text{Sp}_{\mathbb{R}}(A^2)} |\mu| \\ &= \max_{\lambda \in \text{Sp}_{\mathbb{R}}(A)} |\lambda^2| \\ &= \max_{\lambda \in \text{Sp}_{\mathbb{R}}(A)} |\lambda|^2 \\ &= \left( \max_{\lambda \in \text{Sp}_{\mathbb{R}}(A)} |\lambda| \right)^2 \\ &= \rho(A)^2 \end{aligned}$$

d'où le résultat ( $\|A\|_2$  et  $\rho(A)$  sont deux quantités positives).

□

*Exemple 6.3.7.* • On considère la matrice

$$A := \begin{pmatrix} 1 & 2 & 1 \\ -2 & -1 & -1 \\ -1 & -1 & -2 \end{pmatrix} \in M_3(\mathbb{R})$$

de l'exemple 5.4.3. On a  $\text{Sp}_{\mathbb{C}}({}^tAA) = \text{Sp}_{\mathbb{R}}({}^tAA) = \{1; 16\}$  donc  $\|A\|_2 = \sqrt{\rho({}^tAA)} = \sqrt{16} = 4$ .

• On considère la matrice

$$S := \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \in M_3(\mathbb{R}).$$

Il s'agit d'une matrice symétrique donc  $\|S\|_2 = \rho(S)$ . Or  $\text{Sp}_{\mathbb{C}}(S) = \text{Sp}_{\mathbb{R}}(S) = \{0; 3\}$  donc  $\|S\|_2 = 3$ .

Nous allons à présent expliciter un lien entre le rayon spectral  $\rho(A)$  de la matrice  $A$  de  $M_n(\mathbb{K})$  et la convergence éventuelle de la suite  $(A^k)_{k \in \mathbb{N}}$ . Précisément, nous allons montrer le résultat suivant :

**Théorème 6.3.8.** *Les propriétés suivantes sont équivalentes :*

1. la suite  $(A^k)_{k \in \mathbb{N}}$  converge et  $\lim_{k \rightarrow +\infty} A^k = 0_n$ ,
2.  $\rho(A) < 1$ .

*Remarque 6.3.9.* Considérons, pour tout  $k \in \mathbb{N}$ , une matrice  $A_k = (a_{ij}^{(k)})_{1 \leq i, j \leq n}$  de  $M_n(\mathbb{K})$ . Alors la suite  $(A_k)_{k \in \mathbb{N}}$  converge ssi pour tous  $i, j \in \{1, \dots, n\}$ , la suite  $(a_{ij}^{(k)})_{k \in \mathbb{N}}$  converge, et  $(A_k)_{k \in \mathbb{N}}$  converge vers une matrice  $B = (b_{ij})_{1 \leq i, j \leq n}$  ssi pour tous  $i, j \in \{1, \dots, n\}$ , la suite  $(a_{ij}^{(k)})_{k \in \mathbb{N}}$  converge vers  $b_{ij}$ .

Cette convergence matricielle est au sens de n'importe quelle norme sur  $M_n(\mathbb{K})$  (en effet  $M_n(\mathbb{K})$  est un  $\mathbb{K}$ -espace vectoriel de dimension finie) et on peut par exemple montrer ces équivalences à l'aide de la norme  $\|\cdot\|_{\infty}$  sur  $M_n(\mathbb{K})$ .

Pour prouver le théorème 6.3.8, nous utiliserons une démonstration circulaire en faisant intervenir deux assertions supplémentaires : nous montrerons le théorème

**Théorème 6.3.10.** *Les propriétés suivantes sont équivalentes :*

1. la suite  $(A^k)_{k \in \mathbb{N}}$  converge et  $\lim_{k \rightarrow +\infty} A^k = 0_n$ ,
2. pour tout  $v \in \mathbb{C}^n$ , la suite  $(A^k v)_{k \in \mathbb{N}}$  de  $\mathbb{C}^n$  converge et  $\lim_{k \rightarrow +\infty} A^k v = (0, \dots, 0)$ ,
3.  $\rho(A) < 1$ ,
4. il existe une norme matricielle subordonnée  $\|\cdot\|$  sur  $M_n(\mathbb{C})$  telle que  $\|A\| < 1$ .



Dans la preuve de ce théorème, nous utiliserons le résultat qui suit. Nous ne démontrerons pas celui-ci : pour une preuve, on pourra par exemple consulter le livre de Philippe G. Ciarlet intitulé *Introduction à l'analyse numérique matricielle et à l'optimisation* (Théorème 1.4-3.).

**Théorème 6.3.11.** *Soit  $\|\cdot\|$  une norme matricielle sur  $M_n(\mathbb{C})$ . Alors  $\rho(A) \leq \|A\|$ . De plus, pour tout  $\epsilon > 0$  il existe une norme matricielle subordonnée  $\|\!\| \cdot \|\!$  sur  $M_n(\mathbb{C})$  telle que  $\|\!\| A \|\! \leq \rho(A) + \epsilon$ .*

*Remarque 6.3.12.* On peut néanmoins donner une justification rapide du premier fait énoncé. Soit  $\lambda \in \text{Sp}_{\mathbb{C}}(A)$  tel que  $\rho(A) = |\lambda|$  et soit  $v \in E_{\lambda} \setminus \{(0, \dots, 0)\}$ . Soit maintenant  $w \in \mathbb{C}^n$  tel que la matrice  $v^t w$  de  $M_n(\mathbb{C})$  ne soit pas nulle (par exemple le  $k^{\text{ème}}$  vecteur de la base canonique de  $\mathbb{C}^n$  si la  $k^{\text{ème}}$  coordonnée de  $v$  est non nulle). Alors

$$\rho(A) \|v^t w\| = |\lambda| \|v^t w\| = \|\lambda(v^t w)\| = \|(\lambda v)^t w\| = \|(Av)^t w\| = \|A(v^t w)\| \leq \|A\| \|v^t w\|$$

( $\|\cdot\|$  est une norme matricielle par hypothèse) et donc, comme  $\|v^t w\| \neq 0$  car  $v^t w \neq 0_n$  et  $\|\cdot\|$  est une norme,  $\rho(A) \leq \|A\|$ .

*Démonstration du théorème 6.3.10.*

- 1.  $\Rightarrow$  2. : On suppose que la suite  $(A^k)_{k \in \mathbb{N}}$  converge, et ce vers la matrice nulle  $0_n$ . Soient  $\|\cdot\|$  une norme sur  $\mathbb{C}^n$  et  $\|\!\| \cdot \|\!$  la norme subordonnée à la norme  $\|\cdot\|$ . Soit  $v \in \mathbb{C}^n$ . Si  $v$  est le vecteur nul, alors, pour tout  $k \in \mathbb{N}$ ,  $A^k v$  est le vecteur nul et donc la suite  $(A^k v)_{k \in \mathbb{N}}$  converge vers le vecteur nul. Si  $v$  n'est pas le vecteur nul, alors, pour tout  $k \in \mathbb{N}$ , on a  $\frac{\|A^k v\|}{\|v\|} \leq \|\!\| A^k \|\!$  donc

$$0 \leq \|A^k v\| \leq \|\!\| A^k \|\! \|v\|.$$

Or, par hypothèse, la suite réelle  $(\|\!\| A^k \|\!)_{k \in \mathbb{N}}$  converge vers 0 (la suite  $(A^k)_{k \in \mathbb{N}}$  converge vers  $0_n$  au sens de n'importe quelle norme sur le  $\mathbb{C}$ -espace vectoriel de dimension finie  $M_n(\mathbb{C})$ ), il en est donc de même pour la suite  $(\|\!\| A^k \|\! \|v\|)_{k \in \mathbb{N}}$  et, par théorème d'encadrement, pour la suite  $(\|A^k v\|)_{k \in \mathbb{N}}$ . La suite  $(A^k v)_{k \in \mathbb{N}}$  de  $\mathbb{C}^n$  converge donc vers le vecteur nul.

- 2.  $\Rightarrow$  3. : Montrons cette implication par contraposée : supposons que  $\rho(A) \geq 1$  et montrons qu'il existe un vecteur  $v \in \mathbb{C}^n$  tel que la suite  $(A^k v)_{k \in \mathbb{N}}$  de  $\mathbb{C}^n$  ne converge pas vers le vecteur nul. Comme  $\rho(A) \geq 1$ , il existe une valeur propre  $\lambda$  de  $\text{Sp}_{\mathbb{C}}(A)$  telle que  $|\lambda| \geq 1$ . Soit maintenant  $v \in E_{\lambda} \setminus \{(0, \dots, 0)\}$ . Pour tout  $k \in \mathbb{N}$ , on a alors  $A^k v = \lambda^k v$ , mais, comme  $|\lambda| \geq 1$ , la suite  $(\lambda^k)_{k \in \mathbb{N}}$  de  $\mathbb{C}$  ne converge pas vers 0 et la suite  $(A^k v)_{k \in \mathbb{N}} = (\lambda^k v)_{k \in \mathbb{N}}$  ne converge donc pas vers le vecteur nul.
- 3.  $\Rightarrow$  4. : On suppose  $\rho(A) < 1$ . Soit alors  $\epsilon > 0$  tel que  $\rho(A) + \epsilon < 1$  (par exemple,  $\epsilon = \frac{1-\rho(A)}{2}$  convient). D'après le théorème 6.3.11, il existe une norme subordonnée  $\|\!\| \cdot \|\!$  sur  $M_n(\mathbb{C})$  telle que  $\|\!\| A \|\! \leq \rho(A) + \epsilon < 1$ .
- 4.  $\Rightarrow$  1. : On suppose qu'il existe une norme matricielle subordonnée  $\|\!\| \cdot \|\!$  sur  $M_n(\mathbb{C})$  telle que  $\|\!\| A \|\! < 1$ . Alors, pour tout  $k \in \mathbb{N}$ ,  $0 \leq \|\!\| A^k \|\! \leq \|\!\| A \|\!^k$  (car  $\|\!\| \cdot \|\!$  est une norme matricielle) et la suite réelle  $(\|\!\| A^k \|\!)_{k \in \mathbb{N}}$  converge vers 0 car  $\|\!\| A \|\! < 1$ . Par théorème d'encadrement, la suite réelle  $(\|A^k v\|)_{k \in \mathbb{N}}$  converge vers 0 i.e. la suite  $(A^k v)_{k \in \mathbb{N}}$  converge vers la matrice nulle.

□

*Exemple 6.3.13.* • La suite des puissances successives de la matrice  $\begin{pmatrix} 1 & -3 \\ \frac{1}{8} & -\frac{1}{4} \end{pmatrix} \in M_2(\mathbb{R})$ , de spectre  $\{\frac{1}{2}, \frac{1}{4}\}$  et de rayon spectral  $\frac{1}{2}$ , converge vers la matrice nulle de taille 2.

- Pour chacune des matrices de l'exemple 6.3.2, la suite de ses puissances successives ne converge pas vers la matrice nulle.
- La suite  $(I_n^k)_{k \in \mathbb{N}}$  converge, vers la matrice identité.

## 6.4 Conditionnement

Pour amener et motiver la notion de “conditionnement” d’une matrice carrée inversible, on étudie en préambule l'exemple suivant issu du livre cité précédemment de Philippe G. Ciarlet (section 2.2).

On considère la matrice

$$A := \begin{pmatrix} 10 & 7 & 8 & 7 \\ 7 & 5 & 6 & 5 \\ 8 & 6 & 10 & 9 \\ 7 & 5 & 9 & 10 \end{pmatrix}$$

de  $M_4(\mathbb{R})$ . Le déterminant de  $A$  est 1 et  $A$  est donc inversible. En particulier, pour tout vecteur colonne  $B \in M_{4,1}(\mathbb{R})$ , le système

$$AX = B, \quad X \in M_{4,1}(\mathbb{R})$$

possède une unique solution  $X = A^{-1}B$ . Considérons par exemple le vecteur colonne  $B := \begin{pmatrix} 32 \\ 23 \\ 33 \\ 31 \end{pmatrix}$

et le système linéaire  $AX = B$  de solution  $X = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$ .

Nous allons “perturber” le système  $AX = B$  : on considère le vecteur  $B' := B + \begin{pmatrix} 0, 1 \\ -0, 1 \\ 0, 1 \\ -0, 1 \end{pmatrix}$

et la solution du système  $AX' = B'$  est alors le vecteur  $X' = \begin{pmatrix} 9, 2 \\ -12, 6 \\ 4, 5 \\ -1, 1 \end{pmatrix}$ . On constate ainsi que,

même si l’“erreur relative”  $\frac{\|B-B'\|_\infty}{\|B\|_\infty}$  de  $B'$  par rapport à  $B$  n’est “que” de  $\frac{0,1}{33} \simeq 0,003$ , l’erreur

relative de  $X'$  par rapport à  $X$  est elle de  $\frac{\|X-X'\|_\infty}{\|X\|_\infty} = \frac{13,6}{1} = 13,6$  : le rapport d'“amplification” de l'erreur est de  $\frac{13,6}{\frac{0,1}{33}} = 4488!$

Si l'on remplace à présent la matrice  $A$  par la matrice

$$A'' := A + \begin{pmatrix} 0 & 0 & 0,1 & 0,2 \\ 0,08 & 0,04 & 0 & 0 \\ 0 & -0,02 & -0,11 & 0 \\ -0,01 & -0,01 & 0 & -0,02 \end{pmatrix},$$

le système  $A''X'' = B$  a pour solution  $X'' = \begin{pmatrix} -81 \\ 137 \\ -34 \\ 22 \end{pmatrix}$ . L'erreur relative de  $A''$  par rapport à  $A$

est  $\frac{\|A-A''\|_\infty}{\|A\|_\infty} = \frac{0,2}{10} = 0,02$  et l'erreur relative de  $X''$  par rapport à  $X$  est  $\frac{\|X-X''\|_\infty}{\|X\|_\infty} = 136$ , d'où un rapport d'amplification de  $\frac{136}{0,02} = 6800!$

Perturber même légèrement les données du système  $AX = B$  peut donc entraîner des perturbations très importantes sur sa solution, alors même que la matrice  $A$  peut paraître “sympathique” (ici, la matrice est symétrique, son déterminant est 1, son inverse est  $A^{-1} =$

$$\begin{pmatrix} 25 & -41 & 10 & -6 \\ -41 & 68 & -17 & 10 \\ 10 & -17 & 5 & -3 \\ -6 & 10 & -3 & 2 \end{pmatrix}).$$

Nous allons définir une notion qui va permettre d'étudier et de maîtriser ce phénomène. Soit  $A$  une matrice inversible de  $M_n(\mathbb{K})$ .

**Définition 6.4.1.** Soit  $\|\cdot\|$  une norme sur  $\mathbb{K}^n$  et  $\|\|\cdot\|\|$  la norme subordonnée sur  $M_n(\mathbb{K})$  associée. Le conditionnement de  $A$  par rapport à la norme  $\|\cdot\|$  est la quantité

$$\|\|A\|\| \|A^{-1}\|,$$

que l'on note  $\text{cond}(A)$ .

*Remarque 6.4.2.* Le conditionnement de la matrice  $A$  dépend de la norme choisie sur  $\mathbb{K}^n$ . Usuellement, on note  $\text{cond}_1$ ,  $\text{cond}_2$  et  $\text{cond}_\infty$  les conditionnements respectifs par rapport aux normes  $\|\cdot\|_1$ ,  $\|\cdot\|_2$  et  $\|\cdot\|_\infty$  sur  $\mathbb{K}^n$ .

*Exemple 6.4.3.* • Pour la matrice  $A$  considérée ci-dessus, on a  $\text{cond}_1(A) = \text{cond}_\infty(A) = \|\|A\|\|_\infty \|A^{-1}\|_\infty = 33 \times 136 = 4488$  (ici, on a utilisé le théorème 6.2.6 ainsi que la remarque 6.2.9).

• Si  $A := \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \in M_3(\mathbb{R})$ ,  $A^{-1} = \begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}$  et donc

$$\text{cond}_\infty(A) = \|\|A\|\|_\infty \|A^{-1}\|_\infty = 2 \times 3 = 6$$

et

$$\text{cond}_1(A) = \|\|A\|\|_1 \|A^{-1}\|_1 = 2 \times 3 = 6.$$

- On considère la matrice symétrique  $A := \begin{pmatrix} 6 & 5 & 5 \\ 5 & 6 & 5 \\ 5 & 5 & 6 \end{pmatrix} \in \mathbb{S}_3(\mathbb{R})$ . On a  $\text{Sp}_{\mathbb{C}}(A) = \text{Sp}_{\mathbb{R}}(A) = \{1, 16\}$  (voir exemple 5.4.3) et  $\text{Sp}(A^{-1}) = \{\frac{1}{16}, 1\}$  donc

$$\text{cond}_2(A) = \|A\|_2 \|A^{-1}\|_2 = 16 \times 1 = 16$$

(on a utilisé ici le corollaire 6.3.6).

Soit  $\|\cdot\|$  une norme sur  $\mathbb{K}^n$ . Nous allons tout d'abord énoncer quelques propriétés de base du conditionnement  $\text{cond}$  associé :

**Proposition 6.4.4.** *On a*

1.  $\text{cond}(A) = \text{cond}(A^{-1})$ ,
2. pour tout  $\lambda \in \mathbb{K} \setminus \{0\}$ ,  $\text{cond}(\lambda A) = \text{cond}(A)$ ,
3.  $\text{cond}(A) \geq 1$ .

*Démonstration.* 1. On a

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

2. Soit  $\lambda \in \mathbb{K} \setminus \{0\}$ , on a

$$\text{cond}(\lambda A) = \|\lambda A\| \|(\lambda A)^{-1}\| = \|\lambda A\| \left\| \frac{1}{\lambda} A^{-1} \right\| = |\lambda| \left| \frac{1}{\lambda} \right| \|A\| \|A^{-1}\| = \|A\| \|A^{-1}\| = \text{cond}(A).$$

3. On a

$$1 = \|I_n\| = \|AA^{-1}\| \leq \|A\| \|A^{-1}\| = \text{cond}(A)$$

( $\|\cdot\|$  est une norme matricielle).

□

Etant donné un système  $AX = B$ , on utilise le conditionnement de  $A$  pour estimer et maîtriser l'erreur induite sur la solution du système par une perturbation ou une erreur d'approximation sur les données  $A$  ou  $B$  :

**Théorème 6.4.5.** *Soient  $B, B' \in \mathbb{M}_{n,1}(\mathbb{K})$  avec  $B$  différent du vecteur colonne nul. On note  $X$  la solution du système linéaire  $AX = B$  et  $X'$  la solution du système linéaire  $AX' = B'$ . On a*

$$\frac{\|X - X'\|}{\|X\|} \leq \text{cond}(A) \frac{\|B - B'\|}{\|B\|}.$$

*Démonstration.* On a d'une part  $A(X - X') = B - B'$  i.e.  $X - X' = A^{-1}(B - B')$  et donc  $\|X - X'\| \leq \|A^{-1}\| \|B - B'\|$ .

D'autre part,  $B = AX$  donc  $\|B\| \leq \|A\| \|X\|$  donc  $\frac{1}{\|X\|} \leq \frac{\|A\|}{\|B\|}$ .

Ainsi,

$$\frac{\|X - X'\|}{\|X\|} = \|X - X'\| \times \frac{1}{\|X\|} \leq \|A^{-1}\| \|B - B'\| \times \frac{\|A\|}{\|B\|} = \text{cond}(A) \frac{\|B - B'\|}{\|B\|}.$$

□

En conséquence, l'erreur relative  $\frac{\|X-X'\|}{\|X\|}$  sur la solution est d'autant plus petite que la conditionnement de la matrice  $A$  (ainsi que l'erreur relative sur la donnée du second membre  $\frac{\|B-B'\|}{\|B\|}$ ) est petit.

Le résultat relatif à la perturbation du premier membre autrement dit de la matrice  $A$  est le suivant :

**Théorème 6.4.6.** *Soit  $B \in M_{n,1}(\mathbb{K})$  différent du vecteur colonne nul. Soit  $A' \in GL_n(\mathbb{K})$  et notons  $X$  la solution du système linéaire  $AX = B$  et  $X'$  la solution du système linéaire  $A'X' = B$ . On a*

$$\frac{\|X - X'\|}{\|X'\|} \leq \text{cond}(A) \frac{\|A - A'\|}{\|A\|} \quad \text{et} \quad \frac{\|X - X'\|}{\|X\|} \leq \text{cond}(A) \frac{\|A - A'\|}{\|A\|} \frac{\|A'^{-1}\|}{\|A^{-1}\|}$$

*Démonstration.* On a  $AX = B = A'X'$  donc  $0_n = AX - A'X' = AX - AX' + AX' - A'X' = A(X - X') + (A - A')X'$  d'où  $X - X' = -A^{-1}(A - A')X'$  et donc

$$\begin{aligned} \|X - X'\| &= \|A^{-1}(A - A')X'\| \leq \|A^{-1}(A - A')\| \|X'\| \\ &\leq \|A^{-1}\| \|A - A'\| \|X'\| \\ &= \|A^{-1}\| \|A\| \frac{\|A - A'\|}{\|A\|} \|X'\| \\ &= \text{cond}(A) \frac{\|A - A'\|}{\|A\|} \|X'\| \end{aligned}$$

Enfinement,  $\frac{\|X-X'\|}{\|X'\|} \leq \text{cond}(A) \frac{\|A-A'\|}{\|A\|}$ .

Pour établir l'autre inégalité, on considère l'égalité  $0_n = AX - A'X' = AX - A'X + A'X - A'X' = (A - A')X + A'(X - X')$  d'où  $X - X' = -A'^{-1}(A - A')X$  et donc

$$\begin{aligned} \|X - X'\| &= \|A'^{-1}(A - A')X\| \leq \|A'^{-1}(A - A')\| \|X\| \\ &\leq \|A'^{-1}\| \|A - A'\| \|X\| \\ &= \|A'^{-1}\| \|A - A'\| \|X\| \frac{\text{cond}(A)}{\|A\| \|A^{-1}\|} \\ &= \text{cond}(A) \frac{\|A - A'\|}{\|A\|} \frac{\|A'^{-1}\|}{\|A^{-1}\|} \|X\| \end{aligned}$$

Enfinement,  $\frac{\|X-X'\|}{\|X\|} \leq \text{cond}(A) \frac{\|A-A'\|}{\|A\|} \frac{\|A'^{-1}\|}{\|A^{-1}\|}$ . □

*Remarque 6.4.7.* Par continuité de l'application qui à une matrice inversible de  $GL_n(\mathbb{K})$  associe son inverse, la quantité  $\frac{\|A'^{-1}\|}{\|A^{-1}\|}$  tend vers 1 quand  $A'$  tend vers  $A$  (i.e. quand  $\|A - A'\|$  tend vers 0).

Ainsi, le conditionnement de la matrice inversible  $A$  permet de majorer l'erreur (relative) sur la solution du système  $AX = B$  quand il y a perturbations sur les données du premier ou

du second membre du système. Si l'on maîtrise les erreurs d'approximations sur les données, on peut alors maîtriser les erreurs sur les solutions obtenues.

Un système  $AX = B$  dont la matrice  $A$  possède un conditionnement petit (proche de 1) sera d'autant plus robuste face aux perturbations (i.e. sa solution sera peu sensible aux erreurs sur les données) et on dira qu'un tel système est bien conditionné. Dans le cas contraire (si le conditionnement de  $A$  est grand), on dira que le système est mal conditionné.

## Chapitre 7

# Matrices stochastiques et théorèmes de Perron-Frobenius

### 7.1 Introduction

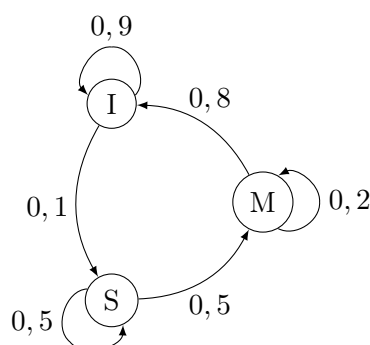
On commence par un exemple de situation qui motive la progression du chapitre et les résultats qui y sont énoncés. Ces derniers sont appliqués dans d'autres situations plus générales et complexes, comme l'algorithme de classification des pages web utilisé (en tout cas à ses débuts) par un célèbre moteur de recherche dont l'appellation commence par la septième lettre de l'alphabet latin.

Remarque : Cet exemple est repris du cours des années précédentes et toute ressemblance avec une situation existante lors de l'année en cours (2020) est absolument fortuite.

La situation que nous considérerons dans cet exemple introductif est celui de l'évolution d'une maladie donnée au sein d'une population donnée. Pour la maladie considérée, chaque individu de la population revêt l'un des trois états suivants :

- M : Malade
- I : Immunisé (sous-entendu pas malade)
- S : Sain (i.e. pas malade) mais non immunisé

D'une semaine sur l'autre, un individu peut "passer" d'un état à un autre : on parle de transition d'un état à un autre. On modélise l'évolution de la maladie semaine après semaine par la probabilité pour un individu de passer d'un état donné à un autre. Les différentes transitions ainsi que leurs probabilités d'avènement sont représentées par le graphe suivant :



Par exemple,

- la probabilité de passer, d'une semaine sur l'autre, de l'état I à l'état S (i.e. de perdre son immunité) est de 0,1,
- la probabilité de passer, d'une semaine sur l'autre, de l'état S à l'état M (i.e. de tomber malade) est de 0,5,
- la probabilité de passer, d'une semaine sur l'autre, de l'état M à l'état M (i.e. de rester malade) est de 0,2.

On peut rassembler les probabilités d'avènement des différentes transitions dans une matrice

$$\begin{array}{ccc}
 I & S & M \\
 \downarrow & \downarrow & \downarrow \\
 \begin{pmatrix} 0,9 & 0 & 0,8 \\ 0,1 & 0,5 & 0 \\ 0 & 0,5 & 0,2 \end{pmatrix} & \leftarrow I \\
 & \leftarrow S \\
 & \leftarrow M
 \end{array}$$

appelée matrice de transition.

Chaque colonne de la matrice de transition donne les probabilités de passer d'un état donné à l'un des trois états. Par exemple, la troisième colonne consiste en les probabilités de passer de l'état M à l'état I (0,8), de passer de l'état M à l'état S (0) et de passer de l'état M à l'état M (0,2), i.e. les probabilités de passer à l'un des états (I, S ou M) "sachant que" l'on est à l'état M (pour employer le langage des probabilités conditionnelles). Remarquons que les coefficients de la matrice de transition sont positifs ou nuls et que, sur chaque colonne de la matrice de transition, la somme des coefficients est égale à 1.

Notons  $A$  cette matrice. Elle va nous permettre de déterminer l'évolution semaine après semaine de l'état de la population que l'on considère, en les termes suivants. Notons  $x$  la proportion des individus qui sont immunisés (état I),  $y$  la proportion des individus sains (état S) et  $z$  la proportion des individus malades (état M). On note ensuite  $V := \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in M_{3,1}(\mathbb{R})$ .

Remarquons que les coefficients de  $V$  sont positifs ou nuls et que leur somme est égale à 1.



Le vecteur  $V$  représente l'“état” de la population à une semaine donnée (désignée semaine 0). L'état de la population à la semaine suivante est le vecteur

$$V_1 := AV = \begin{pmatrix} 0,9x + 0,8z \\ 0,1x + 0,5y \\ 0,5y + 0,2z \end{pmatrix}.$$

Par exemple, 90% des personnes immunisées sont restées immunisées et 80% des personnes malades sont devenues immunisées.

Si  $k \in \mathbb{N}$ , notons  $V_k$  le vecteur de l'“état” de la population après  $k$  semaines. On a la relation de récurrence  $V_{k+1} = AV_k$ ,  $k \in \mathbb{N}$ , et donc, pour tout  $k \in \mathbb{N}$ ,

$$V_k = A^k V.$$

L'état après  $k$  semaines dépend donc de la puissance  $k^{\text{ème}}$  de la matrice de transition  $A$  et de l'état initial  $V$  de la population.

Considérons la matrice

$$A^2 = \begin{array}{ccc} & \begin{array}{ccc} I & S & M \\ \downarrow & \downarrow & \downarrow \end{array} & \\ \begin{pmatrix} 0,81 & 0,4 & 0,88 \\ 0,14 & 0,25 & 0,08 \\ 0,05 & 0,35 & 0,04 \end{pmatrix} & \begin{array}{l} \leftarrow I \\ \leftarrow S \\ \leftarrow M \end{array} & \end{array}$$

Par exemple,

- le coefficient situé 0,14 sur la ligne 2 et la colonne 1 de  $A^2$  est la probabilité de passer, au bout de deux semaines, à l'état S sachant que l'on était à l'état I à la semaine 0, i.e. la probabilité de perdre son immunité au bout de deux semaines,
- le coefficient 0,05 situé sur la ligne 3 et la colonne 1 de  $A^2$  est la probabilité de passer, au bout de deux semaines, à l'état M sachant que l'on était à l'état I à la semaine 0, i.e. la probabilité de tomber malade au bout de deux semaines alors que l'on était immunisé lors de la semaine 0.

L'intérêt de la matrice de transition  $A$  et de l'étude de ses puissances successives est ainsi de pouvoir modéliser, comprendre, anticiper l'évolution de l'état de la population semaine après semaine, et à long terme.

Ici, une étude numérique nous apprend que la suite  $(A^k)_{k \in \mathbb{N}}$  semble avoir une limite de la forme

$$\begin{pmatrix} 0,7547\dots & 0,7547\dots & 0,7547\dots \\ 0,1509\dots & 0,1509\dots & 0,1509\dots \\ 0,0943\dots & 0,0943\dots & 0,0943\dots \end{pmatrix}$$

où les trois colonnes de la matrice sont identiques. Ainsi, au bout d'un temps “suffisamment

long”, i.e pour  $k$  assez grand, l’état de la population sera

$$\begin{aligned} \begin{pmatrix} 0,7547\dots & 0,7547\dots & 0,7547\dots \\ 0,1509\dots & 0,1509\dots & 0,1509\dots \\ 0,0943\dots & 0,0943\dots & 0,0943\dots \end{pmatrix} V &= \begin{pmatrix} 0,7547\dots & 0,7547\dots & 0,7547\dots \\ 0,1509\dots & 0,1509\dots & 0,1509\dots \\ 0,0943\dots & 0,0943\dots & 0,0943\dots \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \\ &= \begin{pmatrix} (0,7547\dots)x + (0,7547\dots)y + (0,7547\dots)z \\ (0,1509\dots)x + (0,1509\dots)y + (0,1509\dots)z \\ (0,0943\dots)x + (0,0943\dots)y + (0,0943\dots)z \end{pmatrix} \\ &= \begin{pmatrix} 0,7547\dots \\ 0,1509\dots \\ 0,0943\dots \end{pmatrix} \end{aligned}$$

(car  $x + y + z = 1$ ) et ce, quel que soit l’état initial de la population.

Ces constats soulèvent les questions suivantes :

- pourquoi la suite  $(A^k)_{k \in \mathbb{N}}$  admet-elle une limite et pourquoi toutes les colonnes de cette limite sont-elles identiques ?
- peut-on déterminer l’“état limite” de la population a priori, i.e. sans passer par le calcul des puissances successives de  $A$  ?

Nous allons répondre au cours de ce chapitre.

Tout au long de celui-ci,  $n$  désigne un entier naturel non nul.

## 7.2 Matrices stochastiques et vecteurs stochastiques

La matrice  $A$  de l’introduction vérifie une propriété particulière : ses coefficients sont tous positifs ou nuls et la somme des coefficients de chaque colonne est égale à 1. Il s’agit de la transposée d’une matrice dite stochastique :

**Définition 7.2.1.** Une matrice  $A$  de  $M_n(\mathbb{R})$  est dite stochastique si tous ses coefficients sont positifs ou nuls et si, sur chaque ligne de  $A$ , la somme des coefficients est égale à 1.

*Exemple 7.2.2.* La transposée de la matrice  $A$  considérée dans l’introduction est stochastique (la matrice  $A$  ne l’est pas).

*Remarque 7.2.3.* Même si cela peut paraître moins “naturel”, en théorie des probabilités, la matrice de transition est “traditionnellement” définie comme la transposée de la matrice que nous avons considérée. Le vecteur de l’état initial prend quant à lui la forme d’un vecteur ligne et on obtient l’état suivant en effectuant le produit à gauche de ce vecteur par la matrice de transition.

On commence par établir quelques propriétés des matrices stochastiques que nous utiliserons par la suite :

**Proposition 7.2.4.** *Soient  $A$  et  $B$  deux matrices stochastiques de  $M_n(\mathbb{R})$ . Alors le produit  $AB$  est également une matrice stochastique.*

*Démonstration.* Notons  $A = (a_{ij})_{1 \leq i, j \leq n}$  et  $B = (b_{ij})_{1 \leq i, j \leq n}$ . Comme les matrices  $A$  et  $B$  sont stochastiques, on a, pour tout  $i \in \{1, \dots, n\}$ ,  $\sum_{j=1}^n a_{ij} = 1$  et  $\sum_{j=1}^n b_{ij} = 1$ .

Soit maintenant  $i \in \{1, \dots, n\}$ . Si  $j \in \{1, \dots, n\}$ , la coefficient à la ligne  $i$  et la colonne  $j$  de  $AB$  est la somme  $\sum_{k=1}^n a_{ik}b_{kj}$ , qui est positive ou nulle, et la somme des coefficients de la ligne  $i$  de  $AB$  est donc

$$\begin{aligned} \sum_{j=1}^n \left( \sum_{k=1}^n a_{ik}b_{kj} \right) &= \sum_{k=1}^n a_{ik} \left( \sum_{j=1}^n b_{kj} \right) \\ &= \sum_{k=1}^n a_{ik} \quad (\text{car } B \text{ est stochastique donc, pour tout } k \in \{1, \dots, n\}, \sum_{j=1}^n b_{kj} = 1) \\ &= 1 \quad (\text{car } A \text{ est stochastique donc } \sum_{k=1}^n a_{ik} = 1). \end{aligned}$$

□

**Proposition 7.2.5.** *La limite d'une suite convergente de matrices stochastiques de  $M_n(\mathbb{R})$  est une matrice stochastique.*

*Démonstration.* L'ensemble

$$\left\{ (a_{ij})_{1 \leq i, j \leq n} \in M_n(\mathbb{R}) \mid \forall i, j \in \{1, \dots, n\}, a_{ij} \geq 0, \text{ et, } \forall i \in \{1, \dots, n\}, \sum_{j=1}^n a_{ij} = 1 \right\}$$

des matrices stochastiques de  $M_n(\mathbb{R})$  est un fermé de  $M_n(\mathbb{R})$ , d'où le résultat. □

**Corollaire 7.2.6.** *Soit  $A$  une matrice stochastique de  $M_n(\mathbb{R})$ . Si la suite  $(A^k)_{k \in \mathbb{N}}$  des puissances successives de  $A$  converge, alors sa limite est une matrice stochastique.*

*Démonstration.* En utilisant la proposition 7.2.4, on montre par récurrence que, pour tout  $k \in \mathbb{N}$ ,  $A^k$  est une matrice stochastique. La suite  $(A^k)_{k \in \mathbb{N}}$  est donc une suite de matrices stochastiques de  $M_n(\mathbb{R})$  convergente : la proposition 7.2.5 précédente permet de conclure. □

On définit également la notion de vecteur stochastique :

**Définition 7.2.7.** *Un vecteur de  $\mathbb{R}^n$  est dit stochastique si toutes ses coordonnées sont positives ou nulles et si leur somme est égale à 1.*

*Exemple 7.2.8.* Le vecteur  $V$  considéré dans l'introduction est stochastique.

*Remarque 7.2.9.* Deux vecteurs stochastiques de  $\mathbb{R}^n$  proportionnels sont égaux. En effet, soient  $v = (v_1, \dots, v_n)$  et  $w = (w_1, \dots, w_n)$  deux vecteurs stochastiques de  $\mathbb{R}^n$  et supposons qu'ils sont proportionnels : il existe  $\lambda \in \mathbb{R} \setminus \{0\}$  tel que  $w = \lambda v$  ( $\lambda$  ne peut être nul car  $w$  est stochastique donc en particulier différent du vecteur nul). On a alors

$$1 = \sum_{j=1}^n w_j = \sum_{j=1}^n \lambda v_j = \lambda \sum_{j=1}^n v_j = \lambda$$

( $v$  et  $w$  sont stochastiques) et donc  $w = v$ .

### 7.3 Matrices positives, strictement positives, primitives, irréductibles

Afin de comprendre et expliquer les phénomènes constatés dans l'introduction, nous appliquerons le théorème dit de Perron à des matrices stochastiques particulières. Afin de pouvoir énoncer ce théorème ainsi que le théorème dit de Frobenius, nous introduisons les notions suivantes :

**Définition 7.3.1.** Soit  $A = (a_{ij})_{1 \leq i, j \leq n} \in M_n(\mathbb{R})$ . On dit que  $A$  est :

- positive si pour tous  $i, j \in \{1, \dots, n\}$ ,  $a_{ij} \geq 0$ ,
- strictement positive si pour tous  $i, j \in \{1, \dots, n\}$ ,  $a_{ij} > 0$ ,
- primitive s'il existe  $k \in \mathbb{N}$  tel que la matrice  $A^k$  est strictement positive,
- irréductible si pour tous  $i, j \in \{1, \dots, n\}$ , il existe  $k \in \mathbb{N}$  ( $k$  dépend de  $i$  et  $j$ ) tel que  $(A^k)_{ij} > 0$  ( $(A^k)_{ij}$  désigne le coefficient à la ligne  $i$  et la colonne  $j$  de la matrice  $A^k$ ).

*Exemple 7.3.2.* • Toute matrice stochastique est positive.

- La matrice  $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$  de  $M_2(\mathbb{R})$  est positive mais non stochastique.
- Toute matrice strictement positive est positive.
- La matrice nulle  $0_n$  de  $M_n(\mathbb{R})$  est positive mais non strictement positive.
- Toute matrice strictement positive est primitive.
- La matrice  $\begin{pmatrix} 0 & -1 \\ -1 & -1 \end{pmatrix}$  est primitive (car son carré est la matrice  $\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$ ) mais non positive (et donc non strictement positive).
- Toute matrice primitive est irréductible.

- La matrice  $A := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$  de  $M_2(\mathbb{R})$  est irréductible, car  $(A)_{12} > 0$ ,  $(A)_{21} > 0$  et, comme  $A^2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ ,  $(A^2)_{11} > 0$ ,  $(A^2)_{22} > 0$ , mais  $A$  n'est pas primitive car, pour tout  $p \in \mathbb{N}$ , les matrices  $A^{2p} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$  et  $A^{2p+1} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$  ne sont pas strictement positives.
- La matrice  $A$  de l'introduction est une matrice primitive (son carré  $A^2$  est une matrice strictement positive) non strictement positive.

*Remarque 7.3.3.* Attention : il ne faut surtout pas confondre les notions de “positivité” de matrices introduites ci-dessus avec les notions de matrice symétrique positive ou définie positive.

A titre d'exemple, la matrice  $S := \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$  est une matrice symétrique (définie) positive mais n'est pas une matrice positive.

## 7.4 Les théorèmes de Perron-Frobenius

Le théorème de Perron porte sur le rayon spectral des matrices positives et primitives. Nous montrerons une partie de ses conclusions sous des hypothèses plus fortes. Dans la section suivante, nous appliquerons le théorème de Perron aux matrices stochastiques primitives.

A la fin de cette section, nous énoncerons également, sans le démontrer, le théorème de Frobenius, qui porte sur les matrices positives et irréductibles.

**Théorème 7.4.1** (Théorème de Perron). *Soit  $A$  une matrice positive et primitive de  $M_n(\mathbb{R})$ . Alors*

1. *le rayon spectral  $\rho(A)$  de  $A$  est une valeur propre (réelle) de  $A$ , de multiplicité 1 (ainsi  $\dim(E_{\rho(A)}) = 1$ ),*
2. *il existe un vecteur propre  $v$  de  $\mathbb{R}^n$  de coordonnées strictement positives pour la valeur propre réelle  $\rho(A)$  (on a  $E_{\rho(A)} = \text{Vect}\{v\}$ ),*
3. *pour tout  $\lambda \in \text{Sp}_{\mathbb{C}}(A) \setminus \{\rho(A)\}$ ,  $|\lambda| < \rho(A)$  (on dit que la valeur propre  $\rho(A)$  de  $A$  est dominante).*

Nous allons établir certaines de ces conclusions sous des hypothèses plus fortes. Précisément, nous montrerons le résultat suivant et nous admettrons le cas général énoncé ci-dessus (pour une preuve du théorème de Perron 7.4.1, on renvoie au document de Bachir Bekka intitulé “Le théorème de Perron-Frobenius, les chaînes de Markov et un célèbre moteur de recherche”, disponible sur sa page web).

**Théorème 7.4.2.** *Soit  $A$  une matrice strictement positive telle que  $\rho(A) = 1$ . Alors*

1.  *$\rho(A) = 1$  est une valeur propre (réelle) de  $A$ ,*
2. *il existe un vecteur propre  $v \in \mathbb{R}^n$  pour la valeur propre  $\rho(A) = 1$  de  $A$ , de coordonnées strictement positives.*

Dans la preuve ci-dessous, nous utiliserons les notations suivantes :

- si  $v = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$  et  $w = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$  sont deux vecteurs de  $\mathbb{R}^n$ , on note  $v \geq 0$  (resp.  $v > 0$ ) si, pour tout  $i \in \{1, \dots, n\}$ ,  $x_i \geq 0$  (resp.  $x_i > 0$ ), et  $v \geq w$  (resp.  $v > w$ ) si  $v - w \geq 0$  (resp.  $v - w > 0$ ) i.e. si, pour tout  $i \in \{1, \dots, n\}$ ,  $x_i \geq y_i$  (resp.  $x_i > y_i$ ),
- si  $u = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix}$  est un vecteur de  $\mathbb{C}^n$ , on note  $|u|$  le vecteur  $\begin{pmatrix} |z_1| \\ \vdots \\ |z_n| \end{pmatrix}$  de  $\mathbb{R}^n$  : on a  $|u| \geq 0$ .

*Démonstration du théorème 7.4.2.* Soit  $\lambda \in \text{Sp}_{\mathbb{C}}(A)$  telle que  $|\lambda| = 1$  (existe car  $\rho(A) = 1$ ) et soit  $u \in \mathbb{C}^n$  un vecteur propre (complexe) de  $A$  pour la valeur propre  $\lambda$ .

On montre tout d'abord que  $|u| \leq A|u|$ . On a, d'une part,  $Au = \lambda u$  donc, si  $z_1, \dots, z_n$  sont les coordonnées de  $u$  dans  $\mathbb{C}^n$ ,

$$|Au| = |\lambda u| = \begin{pmatrix} |\lambda z_1| \\ \vdots \\ |\lambda z_n| \end{pmatrix} = \begin{pmatrix} |\lambda| |z_1| \\ \vdots \\ |\lambda| |z_n| \end{pmatrix} = |\lambda| \begin{pmatrix} |z_1| \\ \vdots \\ |z_n| \end{pmatrix} = |u|$$

(car  $|\lambda| = 1$ ). D'autre part, si on note  $A = (a_{ij})_{1 \leq i, j \leq n}$ , pour tout  $i \in \{1, \dots, n\}$ , la  $i^{\text{ème}}$  coordonnée  $(Au)_i$  du vecteur  $Au$  est  $\sum_{j=1}^n a_{ij} z_j$  et on a

$$|(Au)_i| = |(Au)_i| = \left| \sum_{j=1}^n a_{ij} z_j \right| \leq \sum_{j=1}^n |a_{ij}| |z_j| = \sum_{j=1}^n a_{ij} |z_j| = (A|u|)_i$$

(les coefficients de  $A$  sont réels et positifs). Ainsi,

$$|u| = |Au| \leq A|u|.$$

Nous allons maintenant montrer que, nécessairement,  $|u| = A|u|$ . Pour cela, on procède par l'absurde : on suppose qu'il existe  $i \in \{1, \dots, n\}$  tel que  $(A|u| - |u|)_i > 0$ . Comme  $A$  est une matrice strictement positive, on a alors  $A(A|u| - |u|) > 0$  (appliquer une matrice de coefficients tous strictement positifs à un vecteur de coordonnées positives ou nulles avec au moins une coordonnée strictement positive donne un vecteur de coordonnées strictement positives). Il existe alors  $\epsilon > 0$  tel que  $A(A|u| - |u|) > \epsilon A|u|$  i.e.  $\left(\frac{1}{1+\epsilon} A\right) A|u| > A|u|$  (on peut par exemple choisir  $\epsilon$  de telle sorte que la plus grande coordonnée du vecteur  $\epsilon A|u|$  soit strictement plus petite que la plus petite coordonnée du vecteur  $A(A|u| - |u|)$ ).

On a ensuite, puisque la matrice  $\left(\frac{1}{1+\epsilon}A\right)$  est également à coefficients strictement positifs (et préserve donc la relation  $\left(\frac{1}{1+\epsilon}A\right)A|u| > A|u|$ ),

$$\left(\frac{1}{1+\epsilon}A\right)^2 A|u| > \left(\frac{1}{1+\epsilon}A\right) A|u| > A|u|$$

puis, par récurrence, pour tout  $k \in \mathbb{N}$ ,

$$\left(\frac{1}{1+\epsilon}A\right)^k A|u| > A|u|.$$

Or  $\rho\left(\frac{1}{1+\epsilon}A\right) = \frac{1}{1+\epsilon}\rho(A) = \frac{1}{1+\epsilon} < 1$ . D'après le théorème 6.3.8 du chapitre précédent, la suite  $\left(\left(\frac{1}{1+\epsilon}A\right)^k\right)_{k \in \mathbb{N}}$  converge donc vers 0. Le vecteur  $\left(\frac{1}{1+\epsilon}A\right)^k A|u|$  converge donc vers le vecteur nul et, en faisant tendre  $k$  vers  $+\infty$  dans l'inégalité de vecteurs  $\left(\frac{1}{1+\epsilon}A\right)^k A|u| > A|u|$ , on obtient que les coordonnées du vecteur  $A|u|$  sont négatives ou nulles, ce qui est impossible car

- $|u| \geq 0$ ,
- $u$  n'est pas le vecteur nul (car vecteur propre),
- la matrice  $A$  est strictement positive (i.e. tous ses coefficients sont strictement positifs).

Ainsi, nécessairement,  $|u| = A|u|$  i.e.  $A|u| = |u|$  et, comme  $|u|$  n'est pas le vecteur nul,  $1 (= \rho(A))$  est donc une valeur propre de  $A$ . De plus,  $|u| = A|u| > 0$  car  $|u| \geq 0$ ,  $u$  n'est pas le vecteur nul et  $A$  est strictement positive. Le vecteur  $|u|$  de  $\mathbb{R}^n$  est donc un vecteur propre pour la valeur propre  $1 = \rho(A)$  de coordonnées strictement positives.  $\square$

**Corollaire 7.4.3.** *Soit  $A$  une matrice strictement positive. Alors*

1.  $\rho(A)$  est une valeur propre (réelle) de  $A$ ,
2. il existe un vecteur propre  $v \in \mathbb{R}^n$  pour la valeur propre  $\rho(A)$  de  $A$ , de coordonnées strictement positives.

*Démonstration.* Comme  $A$  est une matrice strictement positive,  $\rho(A) > 0$ . En effet, si  $\rho(A) = 0$ , alors  $\text{Sp}_{\mathbb{C}}(A) = \{0\}$  et donc  $\chi_A = (-X)^n$ , en particulier  $A$  est nilpotente, ce qui est en contradiction avec le fait que  $A$  est strictement positive (car comme  $A$  est strictement positive, pour tout  $k \in \mathbb{N} \setminus \{0\}$ ,  $A^k$  est strictement positive).

On applique alors le théorème 7.4.2 à la matrice  $\frac{1}{\rho(A)}A$  dont le rayon spectral est  $\rho\left(\frac{1}{\rho(A)}A\right) = \frac{1}{\rho(A)}\rho(A) = 1$  : 1 est une valeur propre de  $\frac{1}{\rho(A)}A$  et il existe un vecteur propre  $v \in \mathbb{R}^n$  pour la valeur propre 1 de  $\frac{1}{\rho(A)}A$ , de coordonnées strictement positives. Ainsi  $\frac{1}{\rho(A)}Av = v \Leftrightarrow Av = \rho(A)v$ , et donc  $\rho(A)$  est une valeur propre de  $A$ , et  $v$  est un vecteur propre pour la valeur propre  $\rho(A)$  de  $A$ , de coordonnées strictement positives.  $\square$

*Remarque 7.4.4.* Les conclusions du théorème de Perron ne sont plus vraies si l'on retire l'hypothèse de positivité de la matrice  $A$ . Si l'on considère par exemple la matrice primitive  $M := \begin{pmatrix} 0 & -1 \\ -1 & -1 \end{pmatrix}$  de l'exemple 7.3.2, son polynôme caractéristique est  $\chi_M = X^2 + X - 1 = \left(X - \frac{-1-\sqrt{5}}{2}\right) \left(X - \frac{-1+\sqrt{5}}{2}\right)$  donc  $\rho(M) = \frac{1+\sqrt{5}}{2}$  n'est pas une valeur propre de  $M$ .

On énonce ensuite le théorème de Frobenius, théorème que nous admettrons également (pour une preuve du théorème de Frobenius, on renvoie aux références citées dans le document de Bachir Bekka déjà mentionné plus haut). Ce théorème généralise les deux premières conclusions du théorème de Perron aux matrices positives et irréductibles.

**Théorème 7.4.5** (Théorème de Frobenius). *Soit  $A$  une matrice positive et irréductible. Alors*

1. *le rayon spectral  $\rho(A)$  de  $A$  est une valeur propre de  $A$ , de multiplicité 1,*
2. *il existe un vecteur propre  $v$  de  $\mathbb{R}^n$  de coordonnées strictement positives pour la valeur propre réelle  $\rho(A)$ .*

*Remarque 7.4.6.* Si  $A$  est une matrice positive et irréductible, la valeur propre  $\rho(A)$  de  $A$  n'est pas nécessairement dominante. Si on considère par exemple la matrice positive et irréductible  $A := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$  de  $M_2(\mathbb{R})$ ,  $\text{Sp}(A) = \{-1; 1\}$  et la valeur propre  $\rho(A) = 1$  de  $A$  n'est pas dominante.

## 7.5 Le cas des matrices primitives stochastiques

Nous allons à présent appliquer le théorème de Perron au cas particulier des matrices primitives et stochastiques afin de répondre aux questions posées dans l'introduction.

Soit tout d'abord  $A$  une matrice stochastique quelconque de  $M_n(\mathbb{R})$ .

**Proposition 7.5.1.** *Le rayon spectral  $\rho(A)$  de la matrice stochastique  $A$  est 1.*

*Démonstration.* On a  $A \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$  (car sur chaque ligne de  $A$ , la somme des coefficients est égale à 1) donc 1 est une valeur propre de  $A$ . En particulier,  $\rho(A) \geq 1$ .

D'autre part,  $\rho(A) \leq \|A\|_\infty$ , par la première assertion du théorème 6.3.11, et  $\|A\|_\infty$ , qui est la plus grande somme des valeurs absolues des coefficients sur les lignes de  $A$  par le théorème 6.2.6, est égal à 1 (car les coefficients de  $A$  sont positifs ou nuls et, sur chaque ligne, la somme des coefficients est égale à 1).

Ainsi, on a

$$1 \leq \rho(A) \leq 1$$

i.e.  $\rho(A) = 1$ . □

Nous allons à présent nous intéresser aux matrices stochastiques primitives, pour lesquelles on peut appliquer le théorème de Perron :



**Théorème 7.5.2.** *Supposons que la matrice stochastique  $A$  est de plus primitive. Alors la suite  $(A^k)_{k \in \mathbb{N}}$  des puissances successives de  $A$  converge, vers une matrice stochastique de la forme*

$$\begin{pmatrix} x_1 & \cdots & x_n \\ \vdots & & \vdots \\ x_1 & \cdots & x_n \end{pmatrix}$$

où  $x_1, \dots, x_n \in \mathbb{R}$  forment un vecteur stochastique  $(x_1, \dots, x_n) \in \mathbb{R}^n$ .

*Remarque 7.5.3.* • Avec les notations ci-dessus, le vecteur  $(x_1, \dots, x_n)$ , présenté sous la forme d'un vecteur ligne  $(x_1 \ \cdots \ x_n) \in M_{1,n}(\mathbb{R})$ , est appelé état limite associé à  $A$ .

- Ce théorème répond à la première des deux questions de l'introduction.

*Démonstration du théorème 7.5.2.* On commence par triangulariser la matrice  $A$ , considérée en tant que matrice de  $M_n(\mathbb{C})$ , sous forme de Jordan : il existe  $P \in GL_n(\mathbb{C})$ ,  $s \in \mathbb{N}$  et pour tout  $i \in \{1, \dots, s\}$ ,  $m_i \in \mathbb{N} \setminus \{0\}$  et  $\lambda_i \in \mathbb{C}$  tels que

$$P^{-1}AP = \begin{pmatrix} 1 & & & 0 \\ & J_{m_1}(\lambda_1) & & \\ & & \ddots & \\ 0 & & & J_{m_s}(\lambda_s) \end{pmatrix}$$

De plus, pour tout  $i \in \{1, \dots, s\}$ ,  $|\lambda_i| < 1$  : en effet, comme  $A$  est stochastique,  $\rho(A) = 1$  (par proposition 7.5.1) et, comme  $A$  est primitive,  $\rho(A)$  est une valeur propre simple de  $A$  et cette valeur propre est dominante (par le théorème de Perron : théorème 7.4.1).

Afin de simplifier les écritures, notons, pour  $i \in \{1, \dots, s\}$ ,  $J_i := J_{m_i}(\lambda_i)$ . Alors, pour tout  $k \in \mathbb{N}$ , on a

$$A^k = P \begin{pmatrix} 1 & & & 0 \\ & J_1^k & & \\ & & \ddots & \\ 0 & & & J_s^k \end{pmatrix} P^{-1}.$$

Pour tout  $i \in \{1, \dots, s\}$ , comme  $\rho(J_i) = |\lambda_i| < 1$ , d'après le théorème 6.3.8, la suite  $(J_i^k)_{k \geq 0}$  de  $M_{m_i}(\mathbb{R})$  converge vers la matrice nulle  $0_{m_i}$ . Ainsi,

$$\begin{pmatrix} 1 & & & 0 \\ & J_1^k & & \\ & & \ddots & \\ 0 & & & J_s^k \end{pmatrix} \xrightarrow{k \rightarrow +\infty} \begin{pmatrix} 1 & & & 0 \\ & 0_{m_1} & & \\ & & \ddots & \\ 0 & & & 0_{m_s} \end{pmatrix}$$

et, par continuité de l'application  $M_n(\mathbb{K}) \rightarrow M_n(\mathbb{K})$ ,  
 $M \mapsto P^{-1}MP$ ,

$$A^k \xrightarrow[k \rightarrow +\infty]{} P \begin{pmatrix} 1 & & & 0 \\ & 0_{m_1} & & \\ & & \ddots & \\ 0 & & & 0_{m_s} \end{pmatrix} P^{-1}.$$

Déterminons enfin la forme de cette matrice limite. Comme la matrice  $A$  est stochastique, le vecteur  $\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$  est un vecteur propre de  $A$  pour la valeur propre  $1 = \rho(A)$ . Comme  $A$  est de plus primitive, le sous-espace propre associé à la valeur propre  $\rho(A)$  est de dimension 1 (par le théorème de Perron). Le vecteur  $\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$  constitue donc une base de  $E_1$  et on peut supposer sans

perdre de généralité que la première colonne de la matrice de passage  $P$  est le vecteur  $\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$ . On

a alors

$$P \begin{pmatrix} 1 & & & 0 \\ & 0_{m_1} & & \\ & & \ddots & \\ 0 & & & 0_{m_s} \end{pmatrix} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 1 & 0 & \cdots & 0 \end{pmatrix}$$

et, si  $(x_1 \ \cdots \ x_n)$  est la première ligne de  $P^{-1}$ ,

$$P \begin{pmatrix} 1 & & & 0 \\ & 0_{m_1} & & \\ & & \ddots & \\ 0 & & & 0_{m_s} \end{pmatrix} P^{-1} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 1 & 0 & \cdots & 0 \end{pmatrix} P^{-1} = \begin{pmatrix} x_1 & \cdots & x_n \\ \vdots & & \vdots \\ x_1 & \cdots & x_n \end{pmatrix}.$$

Notons  $L$  cette matrice limite  $\begin{pmatrix} x_1 & \cdots & x_n \\ \vdots & & \vdots \\ x_1 & \cdots & x_n \end{pmatrix}$ . Comme

- $A^k \xrightarrow[k \rightarrow +\infty]{} L$ ,
- pour tout  $k \in \mathbb{N}$ ,  $A^k \in M_n(\mathbb{R})$ ,
- $M_n(\mathbb{R})$  est un fermé de  $M_n(\mathbb{C})$ ,

la matrice limite  $L$  est dans  $M_n(\mathbb{R})$ . De plus, comme la matrice  $A$  est stochastique,  $L$  est également une matrice stochastique par le corollaire 7.2.6 : le vecteur  $(x_1, \dots, x_n)$  est donc un vecteur stochastique de  $\mathbb{R}^n$ .  $\square$

Avec les hypothèses et notations du théorème ci-dessus, on souhaiterait pouvoir déterminer l'état limite  $l := (x_1 \ \cdots \ x_n)$  de la matrice stochastique primitive  $A$  sans passer par la réduction de Jordan de  $A$ . La solution de ce problème est donnée par la proposition suivante. Il s'agit de la réponse à la deuxième question de l'introduction.

**Proposition 7.5.4.** *Supposons que la matrice stochastique  $A$  est primitive et notons  $l \in M_{1,n}(\mathbb{R})$  son état limite. La matrice  ${}^tA$  transposée de  $A$  possède un unique vecteur propre associé à la valeur propre 1 qui soit stochastique, et ce vecteur est le vecteur colonne  ${}^tl$ .*

*Démonstration.* Comme la matrice  $A$  est primitive, sa transposée  ${}^tA$  l'est également (pour tout  $k \in \mathbb{N}$ ,  $({}^tA)^k = {}^t(A^k)$ ). De plus, on a  $\chi_{{}^tA} = \chi_A$  donc  $1 = \rho(A) = \rho({}^tA)$  est une valeur propre simple (par le théorème de Perron 7.4.1) de  ${}^tA$ . L'espace propre  $E_1$  de  ${}^tA$  est donc de dimension 1. De plus, comme  ${}^tA$  est primitive, d'après le théorème de Perron 7.4.1, il existe un vecteur propre  $v$  pour la valeur propre  $\rho({}^tA) = 1$  de  ${}^tA$  qui soient de coordonnées strictement positives. Si on note  $v_1, \dots, v_n$  les coordonnées du vecteur  $v$ , le vecteur  $w := \frac{1}{\sum_{i=1}^n v_i} v$  est alors un vecteur

stochastique de  $\mathbb{R}^n$ , qui est également un vecteur propre pour la valeur propre 1 de  ${}^tA$ . Comme  $E_1$  est de dimension 1, tout vecteur propre stochastique de  $E_1$  sera proportionnel à  $w$  et donc égal à  $w$  (remarque 7.2.9).

Montrons à présent que  $w = {}^tl$ . On a  ${}^tAw = w$  et donc, pour tout  $k \in \mathbb{N}$ ,  $({}^tA)^k w = w \Leftrightarrow {}^t(A^k) w = w$ . Comme la transposition est une application continue sur  $M_n(\mathbb{R})$ , si  $L$  désigne la matrice limite de la suite des puissances successives de  $A$ , on a alors, par passage à la limite,  ${}^tLw = w$ , autrement dit  $w = {}^tLw$  et  $w$  est donc dans l'image de la matrice  ${}^tL$ . Or  ${}^tL = ({}^tl \mid \cdots \mid {}^tl)$  et donc  $w \in \text{Im } {}^tL = \text{Vect } \{{}^tl\}$ . Il existe donc un scalaire  $\alpha \in \mathbb{R}$  tel que  $w = \alpha {}^tl$ . Mais le vecteur  ${}^tl = (l_1, \dots, l_n)$  est, comme le vecteur  $w$ , un vecteur stochastique donc finalement  $w = {}^tl$  par la remarque 7.2.9.

□

Ainsi, si la matrice stochastique  $A$  est primitive, son état limite est l'unique vecteur ligne stochastique  $l$  tel que  ${}^tA^k l = {}^tl$ . Nous allons appliquer cette propriété à notre exemple introductif :

*Exemple 7.5.5* (Retour à l'exemple introductif). La matrice stochastique

$$A := \begin{pmatrix} 0,9 & 0,1 & 0 \\ 0 & 0,5 & 0,5 \\ 0,8 & 0 & 0,2 \end{pmatrix}$$

(il s'agit de la transposée de la matrice que nous avons considérée dans l'exemple) est primitive car la matrice  $A^2$  est strictement positive, et l'état limite de  $A$  est le vecteur ligne stochastique

$l = (l_1 \ l_2 \ l_3)$  tel que

$$\begin{aligned} {}^t A^t l = {}^t l &\Leftrightarrow \begin{pmatrix} 0,9 & 0 & 0,8 \\ 0,1 & 0,5 & 0 \\ 0 & 0,5 & 0,2 \end{pmatrix} \begin{pmatrix} l_1 \\ l_2 \\ l_3 \end{pmatrix} = \begin{pmatrix} l_1 \\ l_2 \\ l_3 \end{pmatrix} \\ &\Leftrightarrow \begin{cases} 0,9l_1 & + 0,8l_3 & = l_1, \\ 0,1l_1 + 0,5l_2 & & = l_2 \\ & 0,5l_2 + 0,2l_3 & = l_3 \end{cases} \\ &\Leftrightarrow \begin{cases} -0,1l_1 & + 0,8l_3 & = 0 \\ 0,1l_1 - 0,5l_2 & & = 0 \\ & 0,5l_2 - 0,8l_3 & = 0 \end{cases} \\ &\Leftrightarrow \begin{cases} l_1 & = 8l_3 \\ l_2 & = \frac{8}{5}l_3 \end{cases}, \quad l_3 \in \mathbb{R} \end{aligned}$$

De plus,

$$l_1 + l_2 + l_3 = 1 \Leftrightarrow 8l_3 + \frac{8}{5}l_3 + l_3 = 1 \Leftrightarrow \frac{53}{5}l_3 = 1 \Leftrightarrow l_3 = \frac{5}{53},$$

donc l'état limite de  $A$  est le vecteur ligne  $l = \left(\frac{40}{53} \ \frac{8}{53} \ \frac{5}{53}\right) = (0,7547\dots \ 0,1509\dots \ 0,0943\dots)$ .

Le théorème de Perron a également été utilisé pour le classement des pages web :

*Exemple 7.5.6.* Notons  $W := \{x_i, i \in I\}$  l'ensemble des pages web présentes sur le "World Wide Web", où  $I := \{1, \dots, N\}$  avec  $N$  entier supérieur à  $13 \times 10^{13}$ .

On forme un graphe avec ces pages webs : pour  $i, j \in I$ , on écrit  $x_i \rightarrow x_j$  si la page  $x_i$  contient un lien vers la page  $x_j$  (on dit alors que  $x_i$  pointe vers  $x_j$ ).

Si, pour  $i \in I$ , la page  $x_i$  contient au moins un lien, on fait l'hypothèse que chacun des liens présents sur la page  $x_i$  pointe vers une page différente et que, depuis la page  $x_i$ , la probabilité de cliquer sur l'un de ces liens est toujours la même : si on note  $d_i$  le nombre de liens présents sur une telle page  $x_i$ , cette probabilité est de  $\frac{1}{d_i}$ .

Pour tous  $i, j \in I$ , on pose alors

$$a_{ij} := \begin{cases} \frac{1}{d_i} & \text{si } x_i \text{ pointe vers } x_j, \\ 0 & \text{si } x_i \text{ ne pointe pas vers } x_j, \end{cases}$$

et on forme la matrice de transition  $A := (a_{ij})_{1 \leq i, j \leq N}$  (et, par exemple, si  $i, j \in \{1, \dots, n\}$ , le

coefficient  $(A^2)_{ij} = \sum_{k=1}^N a_{ik}a_{kj}$  de  $A^2$  est la probabilité, partant de la page  $x_i$ , d'aboutir à la page  $x_j$  en deux clics).

Soit  $i \in I$ . Remarquons que s'il y a au moins un lien sur la page  $x_i$ , alors la somme  $\sum_{j=1}^N a_{ij}$

des coefficients de la ligne  $i$  de  $A$  est  $\sum_{j \mid x_i \rightarrow x_j} \frac{1}{d_i} = d_i \times \frac{1}{d_i} = 1$  et que, s'il n'y a aucun lien sur la page  $x_i$ , tous les coefficients de la ligne  $i$  de  $A$  sont nuls. Afin de "rendre" cette matrice stochastique, on remplace tous les coefficients des lignes nulles de  $A$ , lignes qui correspondent à des pages sans lien, par  $\frac{1}{N}$ . On note  $\tilde{A}$  la matrice ainsi obtenue, qui est alors une matrice stochastique.

Cependant, la matrice  $\tilde{A}$  n'est pas nécessairement primitive. Pour remédier à cela, on considère la matrice

$$G_\alpha := \alpha \tilde{A} + (1 - \alpha)E$$

où  $\alpha \in ]0; 1[$  et  $E$  est la matrice de  $M_N(\mathbb{R})$  dont les coefficients sont tous égaux à  $\frac{1}{N}$ . Alors la matrice  $G_\alpha$  est stochastique et strictement positive (en particulier primitive) : on peut donc lui appliquer le théorème 7.5.2 et la proposition 7.5.4. En classant ensuite les coordonnées du vecteur d'état limite de cette matrice de la plus grande à la plus petite valeur, on obtient un classement des pages web, suivant "leur probabilité d'être visitée à la limite".

Dans la pratique, il faut choisir un nombre  $\alpha$  qui soit "proche" de 1 pour que la matrice  $G_\alpha$  soit "proche" de la matrice  $\tilde{A}$ , mais "pas trop proche" pour que le calcul de l'état limite ne soit "pas trop difficile". Dans les derniers documents publics détaillant cette méthode de classement des pages web,  $\alpha$  avait été choisi égal à 0,85.



## Chapitre 8

# Résolution de systèmes linéaires, décompositions $LU$ et décomposition de Cholesky

### 8.1 Introduction

Soit  $\mathbb{K} = \mathbb{R}$  ou  $\mathbb{C}$  et soit  $n$  un entier naturel non nul.

Soient une matrice  $A \in M_n(\mathbb{K})$  et un vecteur colonne  $B \in M_{n,1}(\mathbb{K})$ , et considérons le système linéaire

$$(S) \quad AX = B$$

de vecteur inconnu  $X \in M_{n,1}(\mathbb{R})$ . L'objectif de ce chapitre est de présenter des méthodes de résolution d'un tel système qui soient "peu" coûteuses en calculs pour  $n$  "grand".

Supposons tout d'abord que  $A$  est une matrice inversible. Dans ce cas, le système (S) possède une unique solution  $X = A^{-1}B$ . En particulier, le calcul de l'inverse  $A^{-1}$  de  $A$  permet de résoudre le système (S). Une méthode de calcul de cet inverse consiste à déterminer les vecteurs colonnes de la matrice  $A^{-1} = (Y_1 | \cdots | Y_n)$  à l'aide de la résolution des  $n$  systèmes linéaires

$$\begin{cases} AY_1 &= X_1 \\ &\vdots \\ AY_n &= X_n \end{cases}$$

de vecteurs inconnus  $Y_1, \dots, Y_n \in M_{n,1}(\mathbb{R})$ , où, pour  $i \in \{1, \dots, n\}$ ,  $X_i$  désigne le vecteur colonne de  $M_{n,1}(\mathbb{R})$  dont toutes les coordonnées sont nulles sauf la  $i^{\text{ème}}$  coordonnée qui est 1 : on a

$$AA^{-1} = I_n \text{ ssi } A(Y_1 | \cdots | Y_n) = (X_1 | \cdots | X_n) \text{ ssi } \forall i \in \{1, \dots, n\}, AY_i = X_i.$$

Dans la visée de la résolution du seul système (S), cette méthode est bien trop coûteuse en calculs. Il faut donc recourir à d'autres méthodes plus "efficaces".

Par exemple, lorsque  $A$ , en plus d'être inversible, est une matrice triangulaire supérieure, il existe une méthode permettant de résoudre le système (S) avec un minimum de calculs : la méthode dite de remontée. Cette méthode consiste à partir de la dernière équation du système (S) et puis "remonter" les équations une à une pour déterminer successivement les coordonnées du vecteur solution. Précisément, on procède de la manière suivante. Notons

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ & \ddots & \vdots \\ 0 & & a_{nn} \end{pmatrix}, B = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \text{ et } X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \text{ alors}$$

$$\begin{aligned} \text{(S) } AX = B &\Leftrightarrow \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ & \ddots & \vdots \\ 0 & & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \\ &\Leftrightarrow \begin{cases} a_{11}x_1 + \dots + a_{1n}x_n = b_1 \\ \vdots \\ a_{nn}x_n = b_n \end{cases} \\ &\Leftrightarrow \begin{cases} a_{11}x_1 + \dots + a_{1n}x_n = b_1 \\ \vdots \\ x_n = \frac{b_n}{a_{nn}} \end{cases} \\ &\Leftrightarrow \begin{cases} a_{11}x_1 + \dots + a_{1n}x_n = b_1 \\ \vdots \\ a_{n-1\ n-1}x_{n-1} + a_{n-1\ n}x_n = b_{n-1} \\ x_n = \frac{b_n}{a_{nn}} \end{cases} \\ &\Leftrightarrow \begin{cases} a_{11}x_1 + \dots + a_{1n}x_n = b_1 \\ \vdots \\ x_{n-1} = \frac{1}{a_{n-1\ n-1}} (b_{n-1} - (a_{n-1\ n}x_n)) \\ x_n = \frac{1}{a_{nn}} b_n \end{cases} \\ &\Leftrightarrow \begin{cases} x_1 = \frac{1}{a_{11}} (b_1 - (a_{12}x_2 + \dots + a_{1n}x_n)) \\ x_2 = \frac{1}{a_{22}} (b_2 - (a_{23}x_3 + \dots + a_{2n}x_n)) \\ \vdots \\ x_{n-1} = \frac{1}{a_{n-1\ n-1}} (b_{n-1} - a_{n-1\ n}x_n) \\ x_n = \frac{b_n}{a_{nn}} \end{cases} \end{aligned}$$

(il est ici à noter que, pour tout  $i \in \{1, \dots, n\}$ ,  $a_{ii} \neq 0$ , car  $A$  est inversible). On dit que l'on a obtenu la solution  $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$  du système (S) par "remontées successives" : on obtient une



coordonnée  $x_i$ ,  $i \in \{1, \dots, n\}$ , à partir des coordonnées  $x_j$ ,  $j > i$  déterminées “plus bas”. Les calculs mis en œuvre dans cette méthode sont en particulier simples et “peu” nombreux.

*Exemple 8.1.1.* On considère le système

$$(S) \begin{pmatrix} 1 & -2 & 5 \\ 0 & -4 & 3 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ 3 \end{pmatrix} \Leftrightarrow \begin{cases} x - 2y + 5z = 2 \\ -4y + 3z = 0 \\ -z = 3 \end{cases}$$

de vecteur inconnu  $\begin{pmatrix} x \\ y \\ z \end{pmatrix} \in M_{3,1}(\mathbb{R})$ . Alors

$$(S) \begin{cases} x - 2y + 5z = 2 \\ -4y + 3z = 0 \\ -z = 3 \end{cases} \Leftrightarrow \begin{cases} x - 2y + 5z = 2 \\ -4y + 3z = 0 \\ z = -3 \end{cases} \\ \Leftrightarrow \begin{cases} x - 2y + 5z = 2 \\ y = \frac{-3 \times (-3)}{-4} = -\frac{9}{4} \\ z = -3 \end{cases} \\ \Leftrightarrow \begin{cases} x = 2 + 2 \times \left(-\frac{9}{4}\right) - 5 \times (-3) = \frac{25}{2} \\ y = -\frac{9}{4} \\ z = -3 \end{cases}$$

*Remarque 8.1.2.* On peut adapter la méthode de remontée décrite ci-dessus dans le cas où  $A$  est une matrice triangulaire supérieure non inversible (i.e. au moins un coefficient diagonal de  $A$  est nul). Considérons par exemple les deux systèmes suivants.

Soit  $\begin{pmatrix} x \\ y \\ z \end{pmatrix} \in M_{3,1}(\mathbb{R})$ . Alors le système

$$\begin{pmatrix} 3 & 7 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & -5 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 7 \\ -2 \end{pmatrix} \Leftrightarrow \begin{cases} 3x + 7y = 1 \\ 2z = 7 \\ -5z = -2 \end{cases} \\ \Leftrightarrow \begin{cases} 3x + 7y = 1 \\ 2z = 7 \\ z = \frac{2}{5} \end{cases} \\ \Leftrightarrow \begin{cases} 3x + 7y = 1 \\ 0 = 7 - \frac{4}{5} \\ z = \frac{2}{5} \end{cases}$$

n'a pas de solution, et le système

$$\begin{aligned} \begin{pmatrix} 1 & -2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} &= \begin{pmatrix} 5 \\ 1 \\ 0 \end{pmatrix} \Leftrightarrow \begin{cases} x - 2y + 3z = 5 \\ 4y + 5z = 1 \\ 0 = 0 \end{cases} \\ &\Leftrightarrow \begin{cases} x - 2y + 3z = 5 \\ y = \frac{1-5z}{4} \\ 0 = 0 \end{cases} \\ &\Leftrightarrow \begin{cases} x = 5 + 2\left(\frac{1-5z}{4}\right) - 3z = \frac{11-11z}{2} \\ y = \frac{1-5z}{4} \end{cases} \end{aligned}$$

a pour ensemble de solutions

$$\left\{ \begin{pmatrix} \frac{11-11z}{2} \\ \frac{1-5z}{4} \\ z \end{pmatrix} = \begin{pmatrix} \frac{11}{2} \\ \frac{1}{4} \\ 0 \end{pmatrix} + z \begin{pmatrix} \frac{-11}{2} \\ \frac{-5}{4} \\ 1 \end{pmatrix} \mid z \in \mathbb{R} \right\}.$$

Si la matrice  $A$  est triangulaire inférieure, il existe une méthode dite de descente, analogue de la méthode de remontée pour les systèmes triangulaires supérieurs. On illustre la méthode

de descente avec le système triangulaire inférieur suivant : si  $\begin{pmatrix} x \\ y \\ z \end{pmatrix} \in M_{3,1}(\mathbb{R})$ , alors

$$\begin{aligned} \begin{pmatrix} 2 & 0 & 0 \\ -1 & 7 & 0 \\ 1 & 3 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} &= \begin{pmatrix} 3 \\ 2 \\ -1 \end{pmatrix} \Leftrightarrow \begin{cases} 2x & = 3 \\ -x + 7y & = 2 \\ x + 3y + 4z & = -1 \end{cases} \\ &\Leftrightarrow \begin{cases} x & = \frac{3}{2} \\ -x + 7y & = 2 \\ x + 3y + 4z & = -1 \end{cases} \\ &\Leftrightarrow \begin{cases} x & = \frac{3}{2} \\ y & = \frac{1}{7} \left( 2 + \frac{3}{2} \right) = \frac{1}{2} \\ x + 3y + 4z & = -1 \end{cases} \\ &\Leftrightarrow \begin{cases} x & = \frac{3}{2} \\ y & = \frac{1}{2} \\ z & = \frac{1}{4} \left( -1 - \frac{3}{2} - 3 \times \frac{1}{2} \right) = -1 \end{cases} \end{aligned}$$

Les méthodes de résolution des systèmes linéaires que nous allons présenter dans ce chapitre vont consister en des “factorisations matricielles” permettant de se ramener à des systèmes triangulaires, systèmes triangulaires que l'on résout ensuite à l'aide des méthodes de remontée et/ou de descente décrites plus haut.

Nous allons étudier une méthode qui permet de ramener la résolution du système (S) à la résolution d'un système triangulaire supérieur.

## 8.2 Méthode du pivot de Gauss pour la résolution de systèmes linéaires

Considérons le système (S)  $AX = B$  comme dans l'introduction, avec  $A \in M_n(\mathbb{R})$  quelconque. Une première méthode de résolution de ce système consiste à lui appliquer l'algorithme du pivot de Gauss : en utilisant des "pivots", on effectue des opérations sur les lignes de  $A$  et sur les coordonnées du vecteur colonne  $B$  (les mêmes), de façon à se ramener à un système triangulaire supérieur, pour lequel on peut alors employer la méthode de remontée.

On introduit cette méthode avec l'exemple suivant. On suppose que  $A$  est la matrice  $\begin{pmatrix} 5 & 2 & 1 \\ 5 & -6 & 2 \\ -4 & 2 & 1 \end{pmatrix}$  de  $M_3(\mathbb{R})$  et que  $B$  est le vecteur colonne  $\begin{pmatrix} 12 \\ -1 \\ 3 \end{pmatrix}$  de  $M_{3,1}(\mathbb{R})$ . Alors, si  $X = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in M_{3,1}(\mathbb{R})$ ,

$$\begin{aligned}
 \text{(S) } AX = B & \Leftrightarrow \begin{cases} 5x + 2y + z = 12 \\ 5x - 6y + 2z = -1 \\ -4x + 2y + z = 3 \end{cases} \Leftrightarrow \begin{pmatrix} 5 & 2 & 1 \\ 5 & -6 & 2 \\ -4 & 2 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 12 \\ -1 \\ 3 \end{pmatrix} \\
 & \Leftrightarrow_{L_2 \leftarrow L_2 - L_1, L_3 \leftarrow L_3 + \frac{4}{5}L_1} \begin{cases} 5x + 2y + z = 12 \\ -8y + z = -13 \\ \frac{18}{5}y + \frac{9}{5}z = \frac{63}{5} \end{cases} \Leftrightarrow \begin{pmatrix} 5 & 2 & 1 \\ 0 & -8 & 1 \\ 0 & \frac{18}{5} & \frac{9}{5} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 12 \\ -13 \\ \frac{63}{5} \end{pmatrix} \\
 & \Leftrightarrow_{L_3 \leftarrow L_3 + \frac{18}{5} \cdot \frac{1}{8} L_2} \begin{cases} 5x + 2y + z = 12 \\ -8y + z = -13 \\ \frac{9}{4}z = \frac{27}{4} \end{cases} \Leftrightarrow \begin{pmatrix} 5 & 2 & 1 \\ 0 & -8 & 1 \\ 0 & 0 & \frac{9}{4} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 12 \\ -13 \\ \frac{27}{4} \end{pmatrix}
 \end{aligned}$$

Ce dernier système étant triangulaire supérieure, on peut le résoudre par remontée et on a finalement

$$\text{(S)} \Leftrightarrow \begin{cases} x = \frac{1}{5}(12 - (2y + z)) = \frac{1}{5}(12 - (2 \times 2 + 3)) = 1 \\ y = -\frac{1}{8}(-13 - z) = -\frac{1}{8}(-13 - 3) = 2 \\ z = \frac{4}{9} \times \frac{27}{4} = 3 \end{cases}$$

et le système (S) possède donc une unique solution  $\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ .

Les opérations sur les lignes du système effectuées ci-dessus à chaque étape de l'algorithme du pivot de Gauss reviennent à multiplier à gauche la matrice  $A$  et le vecteur  $B$  par certaines matrices particulières, appelées matrices d'élimination :

**Définition 8.2.1.** On appelle matrice d'élimination toute matrice de la forme

$$\begin{array}{c} k \\ \downarrow \\ \left( \begin{array}{cccc} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & \alpha_{k+1} & \ddots \\ & & \vdots & & \ddots \\ & & \alpha_n & & & 1 \end{array} \right) \leftarrow k \end{array}$$

(où tous les coefficients non indiqués sont nuls) avec  $k \in \{1, \dots, n\}$  et  $\alpha_{k+1}, \dots, \alpha_n \in \mathbb{K}$ . La matrice ci-dessus est notée  $E_k(\alpha_{k+1}, \dots, \alpha_n)$ .

Reprenons notre matrice  $A$  quelconque de  $M_n(\mathbb{K})$  et notons  $L_1, \dots, L_n$  ses lignes (dans l'ordre). Alors :

**Lemme 8.2.2.** Soient  $k \in \{1, \dots, n\}$  et  $\alpha_{k+1}, \dots, \alpha_n \in \mathbb{K}$ . La matrice  $E_k(\alpha_{k+1}, \dots, \alpha_n)A$  est la matrice obtenue à partir de la matrice  $A$  en ajoutant, pour tout  $k \in \{l+1, \dots, n\}$ ,  $\alpha_l L_k$  à la ligne  $L_l$ .

*Démonstration.* Soit  $l \in \{k+1, \dots, n\}$  alors, pour tout  $j \in \{1, \dots, n\}$ , le coefficient situé à la ligne  $l$  et la colonne  $j$  de la matrice  $E_k(\alpha_{k+1}, \dots, \alpha_n)A$  est  $\alpha_l a_{kj} + a_{lj}$ .  $\square$

Dans l'exemple ci-dessus, la première étape de l'algorithme consistait à multiplier à gauche  $A$  et  $B$  par la matrice d'élimination  $E_1(-1, \frac{4}{5}) = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ \frac{4}{5} & 0 & 1 \end{pmatrix}$ , la deuxième à multiplier à gauche la matrice  $E_1(-1, \frac{4}{5})A$  et le vecteur  $E_1(-1, \frac{4}{5})B$  par la matrice d'élimination  $E_2(\frac{9}{20}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \frac{9}{20} & 1 \end{pmatrix}$ . Autrement dit, pour passer du système initial (S) au système triangulaire de la fin de l'algorithme, nous avons multiplié à gauche la matrice  $A$  et le vecteur  $B$  par la matrice

$$M := E_2\left(\frac{9}{20}\right) E_1\left(-1, \frac{4}{5}\right) = E_2\left(\frac{9}{20}\right) \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ \frac{4}{5} & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ \frac{7}{20} & \frac{9}{20} & 1 \end{pmatrix}$$

**Remarque 8.2.3.** • Pour  $k \in \{1, \dots, n\}$  et  $\alpha_{k+1}, \dots, \alpha_n \in \mathbb{K}$ , la matrice d'élimination  $E_k(\alpha_{k+1}, \dots, \alpha_n)$  est inversible d'inverse  $E_k(-\alpha_{k+1}, \dots, -\alpha_n)$ .

- Pour  $k \in \{1, \dots, n\}$  et  $\alpha_{k+1}, \dots, \alpha_n \in \mathbb{K}$ ,  $\det(E_k(\alpha_{k+1}, \dots, \alpha_n)A) = 1$ .
- La matrice identité  $I_n$  est une matrice d'élimination :  $I_n = E_1(0, \dots, 0)$ .

Lorsque l'on applique l'algorithme du pivot de Gauss pour résoudre un système linéaire, on peut également être amené à effectuer un échange de lignes pour "déplacer" un pivot à la "bonne place". Par exemple, dans le système

$$\begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \\ 5 \end{pmatrix}$$

le coefficient situé à la ligne 1 et la colonne 1 de la matrice est nulle et on échange alors, par exemple, les deux premières lignes de la matrice, afin de se ramener au système équivalent

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \\ 5 \end{pmatrix}$$

où le coefficient non nul situé à la ligne 1 et la colonne 1 peut être utilisé comme premier pivot.

Les échanges de deux lignes ainsi appliqués au cours de l'algorithme du pivot de Gauss correspondent à des multiplications à gauche par des matrices dites de transposition :

**Définition 8.2.4.** On appelle *matrice de transposition* toute matrice obtenue à partir de la matrice identité  $I_n$  en échangeant deux lignes. Pour  $i, j \in \{1, \dots, n\}$ , la matrice de transposition obtenue en échangeant les lignes  $i$  et  $j$  de  $I_n$  est notée  $T_{i,j}$ .

**Lemme 8.2.5.** Soient  $i, j \in \{1, \dots, n\}$ . La matrice  $T_{i,j}A$  est la matrice obtenue à partir de la matrice  $A$  en échangeant les lignes  $i$  et  $j$  de  $A$ .

*Démonstration.* Soit  $k, l \in \{1, \dots, n\}$ . Si  $k \notin \{i, j\}$ , le coefficient situé à la ligne  $k$  et la colonne  $l$  de la matrice  $T_{i,j}A$  est  $\sum_{m=1}^n \delta_{k,m} a_{ml} = a_{kl}$ . Le coefficient situé à la ligne  $i$  et la colonne  $l$  de la matrice  $T_{i,j}A$ , quant à lui, est  $a_{jl}$ . Enfin, le coefficient situé à la ligne  $j$  et la colonne  $l$  de la matrice  $T_{i,j}A$  est lui  $a_{il}$ .  $\square$

Dans l'exemple considéré plus haut, on a multiplié à gauche la matrice et le vecteur considérés par la matrice de transposition  $T_{1,2} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ .

*Remarque 8.2.6.* Soient  $i, j \in \{1, \dots, n\}$ . On a

- $T_{i,j} = T_{j,i}$ ,
- la matrice  $T_{i,j}$  est inversible et l'inverse de  $T_{i,j}$  est  $T_{i,j}$  elle-même,
- $\det(T_{i,j}) = -1$ .

Nous allons à présent montrer que la méthode du pivot de Gauss pour la résolution de systèmes linéaires fonctionne toujours, autrement dit qu'il est toujours possible, à partir d'un système (S)  $AX = B$  quelconque, de se ramener à un système triangulaire supérieur à l'aide d'opérations élémentaires sur les lignes, i.e. à l'aide de produits à gauche par des matrices d'éliminations et de transpositions :

**Théorème 8.2.7** (Méthode du pivot de Gauss). *Il existe une matrice  $M \in \text{GL}_n(\mathbb{K})$ , produit de matrices d'éliminations et de transpositions, telle que  $MA$  soit une matrice triangulaire supérieure.*

*Remarque 8.2.8.* Si  $M$  est une telle matrice alors, en particulier, le système (S)  $AX = B$  est équivalent au système  $MAX = MB$ , qui est triangulaire supérieur.

*Démonstration du théorème 8.2.7.* On montre le résultat par récurrence sur  $n$ . Précisément, on montre que pour tout  $n \in \mathbb{N} \setminus \{0\}$ , pour tout  $A \in M_n(\mathbb{K})$ , il existe une matrice  $M \in \text{GL}_n(\mathbb{K})$ , produit de matrices d'éliminations et de transpositions, telle que  $MA$  est une matrice triangulaire supérieure.

Le résultat est vrai pour  $n = 1$  car toute matrice carrée de taille 1 est en particulier triangulaire supérieure.

Supposons à présent la propriété vérifiée au rang  $n - 1$  pour  $n \in \mathbb{N} \setminus \{0, 1\}$  fixé et reprenons notre matrice quelconque  $A$  de  $M_n(\mathbb{K})$ .

Si la première colonne de  $A$  est nulle,  $A$  est de la forme

$$\begin{pmatrix} 0 & \star & \cdots & \star \\ 0 & & & \\ \vdots & & B & \\ 0 & & & \end{pmatrix}$$

où  $B \in M_{n-1}(\mathbb{K})$  : d'après l'hypothèse de récurrence, il existe alors une matrice  $N \in \text{GL}_{n-1}(\mathbb{K})$ , produit de matrices  $N_1, \dots, N_m$  où  $m \in \mathbb{N}$  et, pour tout  $s \in \{1, \dots, m\}$ ,  $N_s$  est une matrice d'élimination ou une matrice de transposition de  $M_{n-1}(\mathbb{K})$ , telle que  $NB$  soit une matrice triangulaire supérieure de  $M_{n-1}(\mathbb{K})$ . Si l'on note alors

$$M := \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & N & \\ 0 & & & \end{pmatrix} \in \text{GL}_n(\mathbb{K}).$$

et, pour tout  $s \in \{1, \dots, m\}$ ,

$$M_s := \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & N_s & \\ 0 & & & \end{pmatrix} \in \text{GL}_n(\mathbb{K})$$

on a  $M = \prod_{s=1}^m M_s$ . De plus, pour tout  $s \in \{1, \dots, m\}$ , si  $N_s$  est une matrice d'élimination, resp. de transition, de  $M_{n-1}(\mathbb{K})$ , alors  $M_s$  est une matrice d'élimination, resp. de transposition, de

$M_n(\mathbb{K})$ . Enfin, la matrice

$$MA = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & N & & \\ 0 & & & \end{pmatrix} \begin{pmatrix} 0 & \star & \cdots & \star \\ 0 & & & \\ \vdots & B & & \\ 0 & & & \end{pmatrix} = \begin{pmatrix} 0 & \star & \cdots & \star \\ 0 & & & \\ \vdots & NB & & \\ 0 & & & \end{pmatrix}$$

est triangulaire supérieure.

Supposons maintenant que la première colonne de  $A$  soit non nulle, et notons  $i_0$  le plus petit indice  $i \in \{1, \dots, n\}$  tel que  $a_{i1} \neq 0$ . Si  $i_0 \neq 1$ , on multiplie tout d'abord à gauche la matrice  $A$  par la matrice de transposition  $T_{i_0,1}$  (afin d'échanger les lignes  $i_0$  et 1 de  $A$ ) et on considère alors la matrice  $A' := T_{i_0,1}A$ . Si  $i_0 = 1$ , on pose  $A' := A$ . Ainsi, si on note  $A' = (a'_{ij})_{1 \leq i, j \leq n}$ , on a dans tous les cas  $a'_{11} \neq 0$  et on peut alors multiplier, à gauche, la matrice  $A'$  par la matrice d'élimination  $E := E_1 \left( -\frac{a'_{21}}{a'_{11}}, \dots, -\frac{a'_{n1}}{a'_{11}} \right)$  afin d'éliminer les autres coefficients de la première colonne de  $A'$  : on a

$$EA' = \begin{pmatrix} a'_{11} & \star & \cdots & \star \\ 0 & & & \\ \vdots & B & & \\ 0 & & & \end{pmatrix}$$

où  $B \in M_{n-1}(\mathbb{K})$ . On applique ensuite l'hypothèse de récurrence à  $B$  comme dans le cas précédent : reprenant les mêmes notations, le produit

$$MEA' = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & N & & \\ 0 & & & \end{pmatrix} \begin{pmatrix} a'_{11} & \star & \cdots & \star \\ 0 & & & \\ \vdots & B & & \\ 0 & & & \end{pmatrix} = \begin{pmatrix} a'_{11} & \star & \cdots & \star \\ 0 & & & \\ \vdots & NB & & \\ 0 & & & \end{pmatrix}$$

est une matrice triangulaire supérieure, et la matrice  $ME T_{i_0,1}$ , resp.  $ME$ , est bien une matrice inversible produit de matrices d'éliminations et de transpositions.  $\square$

### 8.3 La décomposition LU

La décomposition dite *LU* consiste en la "factorisation" de matrices vérifiant une certaine condition de "régularité" en le produit d'une matrice triangulaire inférieure ( $L$  pour "Lower") par une matrice triangulaire supérieure ( $U$  pour "Upper"). Cela permet de ramener la résolution de systèmes linéaires mettant en jeu ces matrices particulières à la résolution de deux systèmes triangulaires.

Précisément, la décomposition *LU* existe pour les matrices dont toutes les sous-matrices principales sont inversibles :

**Définition 8.3.1.** Soit  $A \in M_n(\mathbb{K})$  et soit  $i \in \{1, \dots, n\}$ . La sous-matrice principale de taille  $i$  de  $A$  est la sous-matrice de  $A$  obtenue en en supprimant les  $n - i$  dernières lignes et  $n - i$  dernières colonnes. On appelle également mineur principal d'ordre  $i$  de  $A$  le déterminant de la sous-matrice principale de taille  $i$  de  $A$ .

*Exemple 8.3.2.* Les sous-matrices principales de la matrice  $\begin{pmatrix} 5 & 2 & 1 \\ 5 & -6 & 2 \\ -4 & 2 & 1 \end{pmatrix}$  sont  $(5)$ ,  $\begin{pmatrix} 5 & 2 \\ 5 & -6 \end{pmatrix}$  et  $\begin{pmatrix} 5 & 2 & 1 \\ 5 & -6 & 2 \\ -4 & 2 & 1 \end{pmatrix}$ , et les mineurs principaux de  $A$  sont donc 5,  $-40$  et  $-90$ .

Soit  $A \in M_n(\mathbb{K})$ .

**Théorème 8.3.3** (Décomposition  $LU$ ). *On suppose que tous les mineurs principaux de  $A$  sont non nuls (i.e. toutes les sous-matrices principales de  $A$  sont inversibles). Alors il existe des matrices  $L$  et  $U$  de  $GL_n(\mathbb{K})$  uniques telles que*

- $L$  est une matrice triangulaire inférieure dont tous les coefficients diagonaux sont égaux à 1,
- $U$  est une matrice triangulaire supérieure,
- $A = LU$ .

*Remarque 8.3.4.* Si tous les mineurs principaux de la matrice  $A$  sont non nuls, alors  $A$  est en particulier inversible (car la sous-matrice principale d'ordre  $n$  de  $A$  est  $A$  elle-même). La réciproque est fautive : par exemple, le mineur principal d'ordre 1 de la matrice inversible  $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \in M_2(\mathbb{R})$  est égal à 0.

La démonstration de l'existence de la décomposition  $LU$  va consister à appliquer l'algorithme du pivot de Gauss. Dans la preuve du théorème 8.3.3, nous aurons également besoin du lemme suivant :

**Lemme 8.3.5.** *Supposons que tous les mineurs principaux de  $A$  sont non nuls, et soit  $E \in M_n(\mathbb{K})$  une matrice d'élimination. Alors tous les mineurs principaux de la matrice produit  $EA$  sont non nuls.*

*Démonstration.* Soit  $i \in \{1, \dots, n\}$ . Notons  $A_i$  la sous-matrice principale de taille  $i$  de  $A$ . Alors

$$A = \begin{pmatrix} A_i & B \\ C & D \end{pmatrix}$$

avec  $B \in M_{i, n-i}(\mathbb{K})$ ,  $C \in M_{n-i, i}(\mathbb{K})$  et  $D \in M_{n-i}(\mathbb{K})$ . Quant à la matrice d'élimination  $E$ , elle est de la forme

$$\begin{pmatrix} E' & 0_{i, n-i} \\ C' & D' \end{pmatrix}$$

où  $E' \in M_i(\mathbb{K})$  et  $D' \in M_{n-i}(\mathbb{K})$  sont également des matrices d'éliminations, et  $C' \in M_{n-i, i}(\mathbb{K})$ . On a alors

$$EA = \begin{pmatrix} E' & 0_{i, n-i} \\ C' & D' \end{pmatrix} \begin{pmatrix} A_i & B \\ C & D \end{pmatrix} = \begin{pmatrix} E'A_i + 0_{i, n-i}C & E'B + 0_{i, n-i}D \\ C'A_i + D'C & C'B + D'D \end{pmatrix} = \begin{pmatrix} E'A_i & E'B \\ C'A_i + D'C & C'B + D'D \end{pmatrix}$$



et la matrice principale de taille  $i$  de  $EA$  est donc la matrice  $E'A_i$ . Or

$$\det(E'A_i) = \det(E')\det(A_i) = \det(A_i) \neq 0.$$

□

*Démonstration du théorème 8.3.3.* On montre tout d'abord l'existence de la décomposition  $LU$  de  $A$ , par récurrence sur  $n$  : on montre que pour tout  $n \in \mathbb{N} \setminus \{0\}$ , toute matrice  $A \in M_n(\mathbb{K})$  dont les mineurs principaux sont tous non nuls admet une décomposition  $A = LU$  telle que  $L \in GL_n(\mathbb{K})$  est une matrice triangulaire inférieure dont tous les coefficients diagonaux sont égaux à 1,  $U \in GL_n(\mathbb{K})$  est une matrice triangulaire supérieure et  $A = LU$ .

Pour  $n = 1$ , si  $(a) \in M_1(\mathbb{K})$  est inversible (i.e.  $a \neq 0$ ), alors  $(a) = (1) (a)$  est une décomposition  $LU$  pour  $(a)$ .

Maintenant, supposons la propriété vérifiée au rang  $n-1$  pour  $n \in \mathbb{N} \setminus \{0, 1\}$  fixé, et reprenons notre matrice  $A \in M_n(\mathbb{K})$  dont tous les mineurs principaux sont supposés non nuls.

Notons  $A = (a_{ij})_{1 \leq i, j \leq n}$ . On applique la première étape de l'algorithme du pivot de Gauss à  $A$  en choisissant le coefficient  $a_{11}$  comme pivot :  $a_{11}$  est le mineur principal d'ordre 1 de  $A$  et est donc non nul. Si l'on note  $E_1 := E_1 \left( -\frac{a_{21}}{a_{11}}, \dots, -\frac{a_{n1}}{a_{11}} \right)$ , on a alors

$$E_1 A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & & & \\ \vdots & & A' & \\ 0 & & & \end{pmatrix}$$

où  $A' \in M_{n-1}(\mathbb{K})$ . Soit  $i \in \{1, \dots, n-1\}$  et notons  $A'_i$  la matrice principale d'ordre  $i$  de  $A'$  et  $(E_1 A)_{i+1}$  la matrice principale d'ordre  $i+1$  de  $E_1 A$ . On a

$$(E_1 A)_{i+1} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1i+1} \\ 0 & & & \\ \vdots & & A'_i & \\ 0 & & & \end{pmatrix}$$

et, d'après le lemme 8.3.5,  $\det((E_1 A)_{i+1}) \neq 0$ . Or  $\det((E_1 A)_{i+1}) = a_{11} \det(A'_i)$  donc  $\det(A'_i) \neq 0$ .

On a ainsi montré que tous les mineurs principaux de la matrice  $A'$  de  $M_{n-1}(\mathbb{K})$  étaient non nuls. On peut appliquer l'hypothèse de récurrence à  $A'$  : il existe une matrice triangulaire inférieure  $L' \in GL_{n-1}(\mathbb{K})$  de coefficients diagonaux tous égaux à 1 et une matrice triangulaire supérieure  $U' \in GL_{n-1}(\mathbb{K})$  telles que  $A' = L'U'$ . On a alors

$$E_1 A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & & & \\ \vdots & & L'U' & \\ 0 & & & \end{pmatrix} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & L' & \\ 0 & & & \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & & & \\ \vdots & & U' & \\ 0 & & & \end{pmatrix}.$$

et on pose

$$L := (E_1)^{-1} \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & L' & \\ 0 & & & \end{pmatrix} = E_1 \left( \frac{a_{21}}{a_{11}}, \dots, \frac{a_{n1}}{a_{11}} \right) \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & L' & \\ 0 & & & \end{pmatrix} \text{ et } U := \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & & & \\ \vdots & & U' & \\ 0 & & & \end{pmatrix}$$

La matrice  $L$  est une matrice triangulaire inférieure de  $\text{GL}_n(\mathbb{K})$  dont tous les coefficients diagonaux sont égaux à 1 (car  $L' \in \text{M}_{n-1}(\mathbb{K})$  et  $(E_1)^{-1} \in \text{M}_n(\mathbb{K})$  sont des matrices triangulaires inférieures de coefficients diagonaux tous égaux à 1) et  $U$  est une matrice triangulaire supérieure inversible de  $\text{M}_n(\mathbb{K})$  (car  $U'$  est une matrice triangulaire supérieure inversible de  $\text{M}_{n-1}(\mathbb{K})$  et  $a_{11} \neq 0$ ).

On montre enfin l'unicité de la décomposition  $LU$  de  $A$  : soit  $\tilde{L} \in \text{GL}_n(\mathbb{K})$  une matrice triangulaire inférieure dont tous les coefficients diagonaux sont égaux à 1 et soit  $\tilde{U} \in \text{GL}_n(\mathbb{K})$  une matrice triangulaire supérieure telles que  $A = \tilde{L}\tilde{U}$ . On montre que  $\tilde{L} = L$  et  $\tilde{U} = U$ .

On a  $LU = \tilde{L}\tilde{U}$  et donc  $\tilde{L}^{-1}L = \tilde{U}U^{-1}$ . Or le produit  $\tilde{L}^{-1}L$  est une matrice triangulaire inférieure dont tous les coefficients diagonaux sont égaux à 1 (car  $L, \tilde{L}$  et donc  $\tilde{L}^{-1}$  sont toutes de telles matrices) et le produit  $\tilde{U}U^{-1}$  est une matrice triangulaire supérieure (car  $\tilde{U}, U$  et  $U^{-1}$  sont toutes de telles matrices). Ainsi, nécessairement,  $\tilde{L}^{-1}L = \tilde{U}U^{-1} = I_n$ , et donc  $L = \tilde{L}$  et  $U = \tilde{U}$ .  $\square$

*Exemple 8.3.6.* On calcule la décomposition  $LU$  de la matrice  $A := \begin{pmatrix} 5 & 2 & 1 \\ 5 & -6 & 2 \\ -4 & 2 & 1 \end{pmatrix} \in \text{M}_3(\mathbb{R})$

dont tous les mineurs principaux sont non nuls (exemple 8.3.2).

Nous savons qu'il existe  $L := \begin{pmatrix} 1 & 0 & 0 \\ a & 1 & 0 \\ b & c & 1 \end{pmatrix} \in \text{GL}_3(\mathbb{R})$  et  $U := \begin{pmatrix} d & e & f \\ 0 & g & h \\ 0 & 0 & k \end{pmatrix} \in \text{GL}_3(\mathbb{R})$  telles que

$$A = \begin{pmatrix} 5 & 2 & 1 \\ 5 & -6 & 2 \\ -4 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ a & 1 & 0 \\ b & c & 1 \end{pmatrix} \begin{pmatrix} d & e & f \\ 0 & g & h \\ 0 & 0 & k \end{pmatrix}.$$

On a alors

1.  $d = 5, e = 2, f = 1$ , ainsi

$$\begin{pmatrix} 5 & 2 & 1 \\ 5 & -6 & 2 \\ -4 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ a & 1 & 0 \\ b & c & 1 \end{pmatrix} \begin{pmatrix} 5 & 2 & 1 \\ 0 & g & h \\ 0 & 0 & k \end{pmatrix}$$

2.  $5 = a \times 5$  donc  $a = 1$ , et  $-4 = b \times 5$  donc  $b = -\frac{4}{5}$ , ainsi

$$\begin{pmatrix} 5 & 2 & 1 \\ 5 & -6 & 2 \\ -4 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ -\frac{4}{5} & c & 1 \end{pmatrix} \begin{pmatrix} 5 & 2 & 1 \\ 0 & g & h \\ 0 & 0 & k \end{pmatrix}$$

3.  $-6 = 1 \times 2 + 1 \times g$  donc  $g = -8$ , et  $2 = 1 \times 1 + 1 \times h$  donc  $h = 1$ , ainsi

$$\begin{pmatrix} 5 & 2 & 1 \\ 5 & -6 & 2 \\ -4 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ -\frac{4}{5} & c & 1 \end{pmatrix} \begin{pmatrix} 5 & 2 & 1 \\ 0 & -8 & 1 \\ 0 & 0 & k \end{pmatrix}$$

4.  $2 = (-\frac{4}{5}) \times 2 + c \times (-8)$  donc  $c = -\frac{9}{20}$ , ainsi

$$\begin{pmatrix} 5 & 2 & 1 \\ 5 & -6 & 2 \\ -4 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ -\frac{4}{5} & -\frac{9}{20} & 1 \end{pmatrix} \begin{pmatrix} 5 & 2 & 1 \\ 0 & -8 & 1 \\ 0 & 0 & k \end{pmatrix}$$

5.  $1 = (-\frac{4}{5}) \times 1 + (-\frac{9}{20}) \times 1 + 1 \times k$  donc  $k = \frac{9}{4}$ , ainsi

$$A = \begin{pmatrix} 5 & 2 & 1 \\ 5 & -6 & 2 \\ -4 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ -\frac{4}{5} & -\frac{9}{20} & 1 \end{pmatrix} \begin{pmatrix} 5 & 2 & 1 \\ 0 & -8 & 1 \\ 0 & 0 & \frac{9}{4} \end{pmatrix}$$

et cette dernière expression est la décomposition  $LU$  de  $A$ .

Supposons que tous les mineurs principaux de la matrice  $A$  soient non nuls. Comme illustré par l'exemple ci-dessus, le calcul de la décomposition  $LU$  de  $A$  est peu coûteux en calculs. De plus, si  $B$  est un vecteur colonne de  $M_{n,1}(\mathbb{R})$ , cette factorisation nous permet de résoudre le système (S)  $AX = B$ , de vecteur inconnu  $X \in M_{n,1}(\mathbb{R})$ , de manière particulièrement efficace. En effet,

$$AX = B \text{ ssi } L(UX) = B.$$

Ainsi,  $X \in M_{n,1}(\mathbb{R})$  est l'unique solution du système (S) (unique car  $A$  est inversible) si et seulement si le vecteur  $UX$  est l'unique solution  $Y \in M_{n,1}(\mathbb{R})$  du système  $LY = B$  ( $L$  est inversible). Résoudre le système (S) revient donc à résoudre successivement le système  $LY = B$  puis le système  $UX = Y$  ( $U$  est également inversible), qui sont tous deux des systèmes triangulaires que l'on peut donc résoudre à l'aide des méthodes de remontée et de descente.

*Exemple 8.3.7.* On reprend la matrice  $A := \begin{pmatrix} 5 & 2 & 1 \\ 5 & -6 & 2 \\ -4 & 2 & 1 \end{pmatrix} \in M_3(\mathbb{R})$  de l'exemple 8.3.6 précédent

et on résout le système

$$AX = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

de vecteur inconnu  $X = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in M_{3,1}(\mathbb{R})$ .

La décomposition  $LU$  de  $A$  est

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ -\frac{4}{5} & -\frac{9}{20} & 1 \end{pmatrix} \begin{pmatrix} 5 & 2 & 1 \\ 0 & -8 & 1 \\ 0 & 0 & \frac{9}{4} \end{pmatrix} :$$

notons  $L := \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ -\frac{4}{5} & -\frac{9}{20} & 1 \end{pmatrix}$  et  $U := \begin{pmatrix} 5 & 2 & 1 \\ 0 & -8 & 1 \\ 0 & 0 & \frac{9}{4} \end{pmatrix}$ . Pour résoudre le système  $AX = B$ , on commence par résoudre le système  $LY = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$  de vecteur inconnu  $Y = \begin{pmatrix} a \\ b \\ c \end{pmatrix} \in M_{3,1}(\mathbb{R})$  : on a

$$\begin{aligned} LY = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} &\Leftrightarrow \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ -\frac{4}{5} & -\frac{9}{20} & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \\ &\Leftrightarrow \begin{cases} a & = 1 \\ a + b & = 2 \\ -\frac{4}{5}a - \frac{9}{20}b + c & = 3 \end{cases} \\ &\Leftrightarrow \begin{cases} a & = 1 \\ b & = 2 - 1 = 1 \\ -\frac{4}{5}a - \frac{9}{20}b + c & = 3 \end{cases} \\ &\Leftrightarrow \begin{cases} a & = 1 \\ b & = 1 \\ c & = 3 + \frac{4}{5} \times 1 + \frac{9}{20} \times 1 = \frac{17}{4} \end{cases} \end{aligned}$$

Puis on résout le système  $UX = \begin{pmatrix} 1 \\ 1 \\ \frac{17}{4} \end{pmatrix}$  : on a

$$\begin{aligned} UX = \begin{pmatrix} 1 \\ 1 \\ \frac{17}{4} \end{pmatrix} &\Leftrightarrow \begin{pmatrix} 5 & 2 & 1 \\ 0 & -8 & 1 \\ 0 & 0 & \frac{9}{4} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ \frac{17}{4} \end{pmatrix} \\ &\Leftrightarrow \begin{cases} 5x + 2y + z & = 1 \\ -8y + z & = 1 \\ \frac{9}{4}z & = \frac{17}{4} \end{cases} \\ &\Leftrightarrow \begin{cases} 5x + 2y + z & = 1 \\ -8y + z & = 1 \\ z & = \frac{17}{9} \end{cases} \\ &\Leftrightarrow \begin{cases} 5x + 2y + z & = 1 \\ y & = -\frac{1}{8} \left(1 - \frac{17}{9}\right) = \frac{1}{9} \\ z & = \frac{17}{9} \end{cases} \\ &\Leftrightarrow \begin{cases} x & = \frac{1}{5} \left(1 - 2 \times \frac{1}{9} - \frac{17}{9}\right) = -\frac{10}{45} = -\frac{2}{9} \\ y & = \frac{1}{9} \\ z & = \frac{17}{9} \end{cases} \end{aligned}$$

et le vecteur  $X = \frac{1}{9} \begin{pmatrix} -2 \\ 1 \\ 17 \end{pmatrix}$  est l'unique solution du système  $AX = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ .

## 8.4 La décomposition PLU

Une généralisation de la décomposition  $LU$  existe pour toute matrice de  $M_n(\mathbb{K})$ . Cette décomposition fait apparaître, en plus d'une matrice triangulaire inférieure de coefficients diagonaux tous égaux à 1 et d'une matrice triangulaire supérieure, une matrice dite de permutation, due aux éventuels échanges de lignes dans l'application de l'algorithme du pivot de Gauss.

**Définition 8.4.1.** *Une matrice de permutation de  $M_n(\mathbb{K})$  est une matrice dans laquelle chaque ligne et chaque colonne ne contient qu'un seul coefficient non nul, égal à 1.*

*Remarque 8.4.2.* • Une matrice de permutation est obtenue par permutation (au sens du groupe symétrique) des lignes de la matrice identité  $I_n$  i.e. en appliquant une permutation du groupe symétrique  $\mathfrak{S}_n$  à l'ensemble des lignes de la matrice  $I_n$ . Il est à noter que, une permutation de  $\mathfrak{S}_n$  étant une composition de transpositions et une matrice de transposition (définition 8.2.4) étant obtenue en appliquant une transposition (au sens du groupe symétrique) à l'ensemble des lignes de la matrice  $I_n$ , une matrice de permutation est un produit de matrices de transpositions.

- Si  $P \in M_n(\mathbb{K})$  est une matrice de permutation obtenue en appliquant une permutation  $\sigma \in \mathfrak{S}_n$  aux lignes de la matrice identité  $I_n$ ,  $\det(P) = \epsilon(\sigma)$  où  $\epsilon(\sigma)$  désigne la signature de la permutation  $\sigma$ . En particulier,  $P$  est inversible.
- $P \in M_n(\mathbb{K})$  est une matrice de permutation obtenue en appliquant une permutation  $\sigma \in \mathfrak{S}_n$  aux lignes de  $I_n$ , et si  $M \in M_n(\mathbb{K})$ , la matrice produit  $PM$  est la matrice obtenue à partir de  $M$  en appliquant la même permutation  $\sigma$  aux lignes de  $M$ .

Considérons donc maintenant une matrice  $A$  quelconque de  $M_n(\mathbb{K})$ . On a le résultat de décomposition/factorisation suivant :

**Théorème 8.4.3** (Décomposition PLU). *Il existe des matrices  $P$ ,  $L$  et  $U$  de  $M_n(\mathbb{K})$  telles que*

- $P$  est une matrice de permutation,
- $L$  est une matrice triangulaire inférieure dont tous les coefficients diagonaux sont égaux à 1,
- $U$  est une matrice triangulaire supérieure,
- $A = PLU$ .

Dans la preuve de ce théorème, on utilisera, comme dans la preuve du théorème 8.3.3 de décomposition  $LU$ , l'algorithme du pivot de Gauss mais en faisant, ici, également intervenir des échanges de lignes. On emploiera également le lemme suivant :

**Lemme 8.4.4.** Soient  $i, j, k \in \{1, \dots, n\}$  tels que  $k < i < j$ . Soient  $\alpha_{k+1}, \dots, \alpha_n \in \mathbb{K}$  et considérons les matrices de transposition  $T_{i,j}$  et d'élimination  $E_k(\alpha_{k+1}, \dots, \alpha_n)$  de  $M_n(\mathbb{K})$ . Alors

$$E_k(\alpha_{k+1}, \dots, \alpha_n) T_{i,j} = T_{i,j} E_k(\alpha_{k+1}, \dots, \alpha_j, \dots, \alpha_i, \dots, \alpha_n)$$

*Démonstration.* Commençons par remarquer que multiplier à droite une matrice  $M \in M_n(\mathbb{K})$  par une matrice de transposition  $T_{r,s}$ ,  $r, s \in \{1, \dots, n\}$ ,  $r \neq s$ , échange les colonnes  $r$  et  $s$  de la matrice  $M$ .

Considérons ensuite la matrice  $E_k(\alpha_{k+1}, \dots, \alpha_n)$ . Il s'agit de la matrice

$$\begin{array}{c} k \\ \downarrow \\ \left( \begin{array}{cccc} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & \alpha_{k+1} & \\ & & \vdots & \ddots \\ & & \alpha_n & & 1 \end{array} \right) \leftarrow k \end{array}$$

La matrice  $E_k(\alpha_{k+1}, \dots, \alpha_n) T_{i,j}$ , obtenue en échangeant les colonnes  $i$  et  $j$  de  $E_k(\alpha_{k+1}, \dots, \alpha_n)$ , est la matrice obtenue en échangeant les lignes  $i$  et  $j$  de la matrice  $E_k(\alpha_{k+1}, \dots, \alpha_j, \dots, \alpha_i, \dots, \alpha_n)$ , i.e. la matrice  $T_{i,j} E_k(\alpha_{k+1}, \dots, \alpha_j, \dots, \alpha_i, \dots, \alpha_n)$  ( $k < i < j$ ).  $\square$

*Démonstration du théorème 8.4.3.* Nous allons montrer le résultat suivant, par récurrence sur  $n \in \mathbb{N}$  : pour tout  $n \in \mathbb{N} \setminus \{0\}$ , pour toute matrice  $A$  dans  $M_n(\mathbb{K})$ , il existe une matrice triangulaire supérieure  $U \in M_n(\mathbb{K})$ , il existe  $r, s \in \mathbb{N}$  et des matrices de transposition  $T_1, \dots, T_r \in M_n(\mathbb{K})$  ainsi que des matrices d'élimination  $E_1, \dots, E_s \in M_n(\mathbb{K})$  telles que

$$A = \left( \prod_{i=1}^r T_i \right) \left( \prod_{j=1}^s E_j \right) U :$$

un tel produit  $\left( \prod_{i=1}^r T_i \right)$  forme une matrice de permutation et le produit  $\left( \prod_{j=1}^s E_j \right)$  forme une matrice triangulaire inférieure de coefficients diagonaux tous égaux à 1.

Le résultat est vrai pour  $n = 1$  pour la même raison que celle évoquée dans la preuve du théorème 8.3.3. Supposons donc maintenant le résultat vrai au rang  $n - 1$  pour  $n \in \mathbb{N} \setminus \{0, 1\}$  fixé, et considérons notre matrice quelconque  $A$  de  $M_n(\mathbb{K})$ .

Si la première colonne de  $A$  est nulle,  $A$  est de la forme

$$\begin{pmatrix} 0 & a_{12} & \cdots & a_{1n} \\ 0 & & & \\ \vdots & & B & \\ 0 & & & \end{pmatrix}$$

où  $B \in M_{n-1}(\mathbb{K})$  : d'après l'hypothèse de récurrence, il existe alors une matrice triangulaire supérieure  $U' \in M_{n-1}(\mathbb{K})$ , il existe des entiers naturels  $r$  et  $s$ , il existe des matrices de transposition  $T'_1, \dots, T'_r \in M_{n-1}(\mathbb{K})$  et des matrices d'élimination  $E'_1, \dots, E'_s \in M_{n-1}(\mathbb{K})$  telles que

$$B = \left( \prod_{i=1}^r T'_i \right) \left( \prod_{j=1}^s E'_j \right) U'.$$

On pose alors

$$U := \begin{pmatrix} 0 & a_{12} & \cdots & a_{1n} \\ 0 & & & \\ \vdots & & U' & \\ 0 & & & \end{pmatrix} \in M_n(\mathbb{K})$$

et, pour tout  $i \in \{1, \dots, r\}$  et tout  $j \in \{1, \dots, s\}$ ,

$$T_i := \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & T'_i & \\ 0 & & & \end{pmatrix} \text{ et } E_j := \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & E'_j & \\ 0 & & & \end{pmatrix}.$$

Pour  $i \in \{1, \dots, r\}$ , la matrice  $T_i$  est une matrice de transposition de  $M_n(\mathbb{K})$  et, pour  $j \in \{1, \dots, s\}$ , la matrice  $E_j$  est une matrice d'élimination de  $M_n(\mathbb{K})$ . Enfin,

$$A = \left( \prod_{i=1}^r T_i \right) \left( \prod_{j=1}^s E_j \right) U.$$

Si la première colonne de  $A$  est non nulle, notons  $i_0$  le plus petit indice  $i \in \{1, \dots, n\}$  tel que  $a_{i1} \neq 0$  : si  $i_0 \neq 1$ , on commence par multiplier à gauche la matrice  $A$  par la matrice de transposition  $T := T_{i_0,1}$  et on considère la matrice  $A' := TA$ , et, si  $i_0 = 1$ , on pose  $A' := A$ . Ainsi, si on note  $A' = (a'_{ij})_{1 \leq i, j \leq n}$ ,  $a'_{11} \neq 0$ , on peut ensuite multiplier à gauche la matrice  $A'$  par la matrice d'élimination  $E := E_1 \left( -\frac{a'_{21}}{a'_{11}}, \dots, -\frac{a'_{n1}}{a'_{11}} \right)$  afin d'éliminer les autres coefficients de la première colonne de  $A'$  : on a

$$EA' = \begin{pmatrix} a'_{11} & a'_{12} & \cdots & a'_{1n} \\ 0 & & & \\ \vdots & & B & \\ 0 & & & \end{pmatrix}$$

où  $B \in M_{n-1}(\mathbb{K})$ . En procédant de la même manière que dans le cas précédent (i.e. en appliquant l'hypothèse de récurrence à  $B$ ) et en conservant les mêmes notations, on obtient alors l'égalité

$$EA' = \left( \prod_{i=1}^r T_i \right) \left( \prod_{j=1}^s E_j \right) U$$

i.e.

$$A' = E^{-1} \left( \prod_{i=1}^r T_i \right) \left( \prod_{j=1}^s E_j \right) U.$$

Maintenant,  $E^{-1} = \left( E_1 \left( -\frac{a'_{21}}{a'_{11}}, \dots, -\frac{a'_{n1}}{a'_{11}} \right) \right)^{-1} = E_1 \left( \frac{a'_{21}}{a'_{11}}, \dots, \frac{a'_{n1}}{a'_{11}} \right)$ . Comme les matrices de transposition  $T_i$ ,  $i \in \{1, \dots, r\}$ , échangent des lignes d'indices strictement plus grands que 1, il existe, d'après le lemme 8.4.4, une matrice d'élimination  $\tilde{E} \in M_n(\mathbb{K})$  telle que  $E^{-1} \left( \prod_{i=1}^r T_i \right) = \left( \prod_{i=1}^r T_i \right) \tilde{E}$ , et alors

$$A' = \left( \prod_{i=1}^r T_i \right) \tilde{E} \left( \prod_{j=1}^s E_j \right) U.$$

Enfin, dans le cas où  $i_0 \neq 1$ ,  $T^{-1} = T$  et donc

$$A = T \left( \prod_{i=1}^r T_i \right) \tilde{E} \left( \prod_{j=1}^s E_j \right) U.$$

□

*Remarque 8.4.5.* Il n'y a pas unicité de la décomposition  $PLU$ . Par exemple :

$$\begin{pmatrix} 1 & 1 \\ 2 & 3 \end{pmatrix} = I_2 \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \frac{1}{2} & 1 \end{pmatrix} \begin{pmatrix} 2 & 3 \\ 0 & -\frac{1}{2} \end{pmatrix}.$$

*Exemple 8.4.6.* Considérons la matrice  $A := \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix} \in M_3(\mathbb{R})$ . On applique l'algorithme du pivot de Gauss pour déterminer une décomposition  $PLU$  de  $A$ .

On commence par échanger les deux premières lignes :

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

Puis on élimine le coefficient non nul de la première colonne de cette dernière matrice :

$$E_1(0, -1) T_{2,1} A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

Enfin, on utilise le coefficient situé sur la ligne 2 et la colonne 2 de cette dernière matrice comme pivot et on a :

$$E_2(-1) E_1(0, -1) T_{2,1} A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & -1 \end{pmatrix}.$$



On pose alors  $U := \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & -1 \end{pmatrix}$  (la matrice  $U \in M_3(\mathbb{R})$  est triangulaire supérieure) et on a :

$$\begin{aligned} A &= (T_{2,1})^{-1} (E_1(0, -1))^{-1} (E_2(-1))^{-1} U \\ &= T_{2,1} E_1(0, 1) E_2(1) U. \end{aligned}$$

Si on pose  $P := T_{2,1} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$  et  $L := E_1(0, 1) E_2(1) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}$ ,  $P$  est une matrice de permutation de  $M_3(\mathbb{R})$ ,  $L$  est une matrice triangulaire inférieure de  $M_3(\mathbb{R})$  de coefficients diagonaux tous égaux à 1, et on a :

$$A = PLU.$$

Une décomposition  $PLU$  d'une matrice  $A$  de  $M_n(\mathbb{K})$  permet notamment de résoudre efficacement tout système  $AX = B$  de vecteur inconnu  $X \in M_{n,1}(\mathbb{K})$ , où  $B$  est un vecteur colonne de  $M_{n,1}(\mathbb{K})$ . La résolution d'un tel système revient à la résolution successive des trois systèmes

1.  $PZ = B$ , de vecteur inconnu  $Z \in M_{n,1}(\mathbb{K})$ , système possédant une unique solution  $Z$  rapide à calculer car  $P$  est une matrice de permutation (les coordonnées de  $Z = P^{-1}B$  sont obtenues par permutation des coordonnées de  $B$ ),
2.  $LY = Z$ , de vecteur inconnu  $Y \in M_{n,1}(\mathbb{K})$ , système possédant une unique solution  $Y$  ( $L$  est inversible) et résoluble par la méthode de descente ( $L$  est triangulaire inférieure),
3.  $UX = Y$ , de vecteur inconnu  $X \in M_{n,1}(\mathbb{K})$ , qui est un système triangulaire supérieur et donc résoluble par la méthode de remontée.

En effet, si  $X \in M_{n,1}(\mathbb{K})$ ,

$$AX = B \text{ ssi } PLUX = B \text{ ssi } LUX = P^{-1}B = Z \text{ ssi } UX = L^{-1}Z = Y.$$

*Exemple 8.4.7.* Reprenons la matrice  $A$  de l'exemple précédent 8.4.6. Nous allons utiliser la décomposition  $PLU$  calculée alors pour déterminer la solution du système  $AX = \begin{pmatrix} 0 \\ -1 \\ 5 \end{pmatrix}$  de

vecteur inconnu  $X = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in M_{3,1}(\mathbb{R})$ .

On commence par résoudre le système  $PZ = \begin{pmatrix} 0 \\ -1 \\ 5 \end{pmatrix}$  de vecteur inconnu  $Z = \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} \in$

$M_{3,1}(\mathbb{R})$  : on a

$$\begin{aligned}
 PZ = \begin{pmatrix} 0 \\ -1 \\ 5 \end{pmatrix} &\Leftrightarrow \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \\ 5 \end{pmatrix} \\
 &\Leftrightarrow \begin{cases} \beta = 0 \\ \alpha = -1 \\ \gamma = 5 \end{cases} \\
 &\Leftrightarrow \begin{cases} \alpha = -1 \\ \beta = 0 \\ \gamma = 5 \end{cases}
 \end{aligned}$$

Puis on résout le système  $LY = \begin{pmatrix} -1 \\ 0 \\ 5 \end{pmatrix}$  de vecteur inconnu  $Y = \begin{pmatrix} a \\ b \\ c \end{pmatrix} \in M_{3,1}(\mathbb{R})$  : on a

$$\begin{aligned}
 LY = \begin{pmatrix} -1 \\ 0 \\ 5 \end{pmatrix} &\Leftrightarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \\ 5 \end{pmatrix} \\
 &\Leftrightarrow \begin{cases} a = -1 \\ b = 0 \\ a + b + c = 5 \end{cases} \\
 &\Leftrightarrow \begin{cases} a = -1 \\ b = 0 \\ c = 5 - (-1) - 0 = 6 \end{cases}
 \end{aligned}$$

Enfin, on résout le système  $UX = \begin{pmatrix} -1 \\ 0 \\ 6 \end{pmatrix}$  : on a

$$\begin{aligned}
 UX = \begin{pmatrix} -1 \\ 0 \\ 6 \end{pmatrix} &\Leftrightarrow \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \\ 6 \end{pmatrix} \\
 &\Leftrightarrow \begin{cases} x & + z & = -1 \\ & y + z & = 0 \\ & & -z & = 6 \end{cases} \\
 &\Leftrightarrow \begin{cases} x & + z & = -1 \\ & y + z & = 0 \\ & & z & = -6 \end{cases} \\
 &\Leftrightarrow \begin{cases} x & + z & = -1 \\ & y & = -(-6) = 6 \\ & & z & = -6 \end{cases} \\
 &\Leftrightarrow \begin{cases} x & = -1 - (-6) = 5 \\ y & = 6 \\ z & = -6 \end{cases}
 \end{aligned}$$

et le vecteur  $X = \begin{pmatrix} 5 \\ 6 \\ -6 \end{pmatrix}$  est l'unique solution du système  $AX = \begin{pmatrix} 0 \\ -1 \\ 5 \end{pmatrix}$ .

## 8.5 La décomposition de Cholesky

La décomposition de Cholesky est une factorisation des matrices symétriques définies positives (définition 5.3.1). Elle est construite à partir de la décomposition  $LU$  de ces matrices : les matrices symétriques définies positives vérifient en effet l'hypothèse de "régularité" du théorème 8.3.3.

**Proposition 8.5.1.** *Soit  $S \in S_n(\mathbb{R})$  une matrice symétrique définie positive. Alors tous les mineurs principaux de  $S$  sont strictement positifs.*

*Démonstration.* Soit  $i \in \{1, \dots, n\}$  et notons  $S_i$  la sous-matrice principale de taille  $i$  de  $S$ .

Remarquons tout d'abord que  $S_i \in S_i(\mathbb{R})$ . Soit maintenant  $\tilde{X} = \begin{pmatrix} x_1 \\ \vdots \\ x_i \end{pmatrix} \in M_{i,1}(\mathbb{R}) \setminus \left\{ \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \right\}$

et notons  $X := \begin{pmatrix} x_1 \\ \vdots \\ x_i \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in M_{n,1}(\mathbb{R}) \setminus \left\{ \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \right\}$ . On a alors  ${}^t \tilde{X} A_i \tilde{X} = {}^t X A X > 0$  et la matrice

symétrique  $S_i$  est donc définie positive. En particulier, la matrice  $S_i \in \mathcal{S}_i(\mathbb{R})$  est diagonalisable (théorème 5.2.5) et ses valeurs propres sont strictement positives (proposition 5.3.4) : le déterminant de  $S_i$  est alors égal au produit de ses valeurs propres (avec multiplicités) et est donc strictement positif.  $\square$

**Corollaire 8.5.2.** *Soit  $S \in \mathcal{S}_n(\mathbb{R})$  une matrice symétrique définie positive. Alors  $S$  admet une décomposition  $LU$ . De plus, les coefficients diagonaux de  $U$  sont strictement positifs.*

*Démonstration.* D'après la proposition précédente, tous les mineurs principaux de  $S$  sont strictement positifs, en particulier non nuls : on peut donc appliquer le théorème 8.3.3 à la matrice  $S$  qui possède alors une décomposition  $S = LU$  avec  $L = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ \star & & 1 \end{pmatrix} \in M_n(\mathbb{R})$  et

$$U = \begin{pmatrix} u_{11} & & \star \\ & \ddots & \\ 0 & & u_{nn} \end{pmatrix} \in M_n(\mathbb{R}).$$

Soit  $i \in \{1, \dots, n\}$  et notons  $S_i$ ,  $L_i$  et  $U_i$  les sous-matrices principales de taille  $i$  respectives de  $S$ ,  $L$  et  $U$  : on a  $L_i = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ \star & & 1 \end{pmatrix}$ ,  $U_i = \begin{pmatrix} u_{11} & & \star \\ & \ddots & \\ 0 & & u_{ii} \end{pmatrix} \in M_i(\mathbb{R})$  et

$$S = \begin{pmatrix} S_i & A \\ B & C \end{pmatrix}, L = \begin{pmatrix} L_i & 0_{i,n-i} \\ D & E \end{pmatrix}, U = \begin{pmatrix} U_i & F \\ 0_{n-i,i} & G \end{pmatrix}$$

avec  $A, F \in M_{i,n-i}(\mathbb{R})$ ,  $B, D \in M_{n-i,i}(\mathbb{R})$  et  $C, E, G \in M_{n-i}(\mathbb{R})$ . Alors

$$S = LU = \begin{pmatrix} L_i U_i & L_i F \\ D U_i & D F + E G \end{pmatrix}$$

et, en particulier,  $S_i = L_i U_i$  et donc  $\det(S_i) = \det(L_i) \det(U_i) = \prod_{j=1}^i u_{jj}$ . Or  $\det(S_i) > 0$  (par

la preuve de la proposition précédente) donc  $\prod_{j=1}^i u_{jj} > 0$ .

On a ainsi montré que, pour tout  $i \in \{1, \dots, n\}$ ,  $\prod_{j=1}^i u_{jj} > 0$ . En particulier  $u_{11} > 0$  et, si

$i \in \{2, \dots, n\}$ ,  $\prod_{j=1}^i u_{jj} > 0$  et  $\prod_{j=1}^{i-1} u_{jj} > 0$  donc, nécessairement,  $u_{ii} = \frac{\prod_{j=1}^i u_{jj}}{\prod_{j=1}^{i-1} u_{jj}} > 0$ .  $\square$

Soit  $S \in \mathcal{S}_n(\mathbb{R})$  une matrice symétrique définie positive. Considérons donc la décomposition  $LU$  de  $S$  suivant les notations de la preuve précédente. Nous allons utiliser cette factorisation pour écrire  $S$  comme le produit d'une matrice triangulaire inférieure de coefficients diagonaux strictement positifs et de sa transposée :

**Théorème 8.5.3** (Décomposition de Cholesky). *Il existe une unique matrice  $T \in \mathcal{M}_n(\mathbb{R})$  triangulaire inférieure à coefficients diagonaux strictement positifs (en particulier  $T$  est inversible) telle que*

$$S = T^t T.$$

*Démonstration.* On considère la décomposition  $S = LU$  de  $S$ . D'après le corollaire 8.5.2 ci-dessus, les coefficients diagonaux  $u_{11}, \dots, u_{nn}$  de  $U$  sont tous strictement positifs et on pose alors

$$D := \begin{pmatrix} \sqrt{u_{11}} & & 0 \\ & \ddots & \\ 0 & & \sqrt{u_{nn}} \end{pmatrix} \in \text{GL}_n(\mathbb{R}).$$

Remarquons que l'on a

$$S = LU = LDD^{-1}U.$$

On pose ensuite  $T := LD = \begin{pmatrix} \sqrt{u_{11}} & & 0 \\ & \ddots & \\ \star & & \sqrt{u_{nn}} \end{pmatrix}$  et  $\tilde{T} = D^{-1}U = \begin{pmatrix} \sqrt{u_{11}} & & \star \\ & \ddots & \\ 0 & & \sqrt{u_{nn}} \end{pmatrix}$

$(D^{-1} = \begin{pmatrix} \frac{1}{\sqrt{u_{11}}} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{\sqrt{u_{nn}}} \end{pmatrix})$ , et on a donc  $S = T\tilde{T}$ . Notons que  $T$  est une matrice triangulaire inférieure et que ses coefficients diagonaux sont tous strictement positifs.

Nous allons maintenant montrer que  $\tilde{T} = {}^t T$ . Comme  $S$  est une matrice symétrique, on a

$$T\tilde{T} = S = {}^t S = {}^t \tilde{T} {}^t T$$

et donc, comme la matrice  $\tilde{T}$  est inversible,

$${}^t \tilde{T}^{-1} T = {}^t T \tilde{T}^{-1}.$$

A présent, comme  $\tilde{T}^{-1} = \begin{pmatrix} \frac{1}{\sqrt{u_{11}}} & & \star \\ & \ddots & \\ 0 & & \frac{1}{\sqrt{u_{nn}}} \end{pmatrix}$ , on a  ${}^t \tilde{T}^{-1} T = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ \star & & 1 \end{pmatrix}$  et  ${}^t T \tilde{T}^{-1} =$

$$\begin{pmatrix} 1 & & \star \\ & \ddots & \\ 0 & & 1 \end{pmatrix}, \text{ d'où}$$

$${}^t\tilde{T}^{-1}T = {}^tT\tilde{T}^{-1} = I_n$$

et donc  $\tilde{T} = {}^tT$ .

Montrons enfin que la décomposition  $S = T{}^tT$ , avec  $T \in M_n(\mathbb{R})$  triangulaire inférieure à coefficients diagonaux strictement positifs, est unique. Soit donc  $T' = \begin{pmatrix} t_{11} & & 0 \\ & \ddots & \\ \star & & t_{nn} \end{pmatrix} \in$

$M_n(\mathbb{R})$  une matrice triangulaire inférieure à coefficients diagonaux strictement positifs tels que  $S = T'{}^tT'$  et montrons que  $T' = T$ . Commençons par noter  $D'$  la matrice diagonale inversible  $\begin{pmatrix} t_{11} & & 0 \\ & \ddots & \\ 0 & & t_{nn} \end{pmatrix} \in M_n(\mathbb{R})$ . On a

$$S = T'{}^tT' = T'D'^{-1}D'{}^tT' :$$

comme  $T'D'^{-1} \in M_n(\mathbb{R})$  est une matrice triangulaire inférieure de coefficients diagonaux tous égaux à 1 ( $D'^{-1} = \begin{pmatrix} \frac{1}{t_{11}} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{t_{nn}} \end{pmatrix}$ ) et  $D'{}^tT' \in M_n(\mathbb{R})$  est une matrice triangulaire supérieure, l'égalité  $S = (T'D'^{-1})(D'{}^tT')$  est la décomposition  $LU$  de  $S$ .

On obtient ainsi les égalités  $L = T'D'^{-1}$  et  $U = D'{}^tT'$  i.e.  $TD^{-1} = T'D'^{-1}$  ( $T = LD$ ) et  $D{}^tT = D'{}^tT'$  ( ${}^tT = D^{-1}U$ ). Or  $D{}^tT = \begin{pmatrix} u_{11} & & \star \\ & \ddots & \\ 0 & & u_{nn} \end{pmatrix}$  et  $D'{}^tT' = \begin{pmatrix} t_{11}^2 & & \star \\ & \ddots & \\ 0 & & t_{nn}^2 \end{pmatrix}$  donc, pour tout  $i \in \{1, \dots, n\}$ ,  $t_{ii}^2 = u_{ii}$  i.e.  $t_{ii} = \sqrt{u_{ii}}$  (car  $t_{ii} > 0$ ), et donc  $D' = D$ . D'où, comme  $TD^{-1} = T'D'^{-1}$ , l'égalité  $T = T'$ .  $\square$

*Exemple 8.5.4.* Considérons la matrice symétrique

$$S := \begin{pmatrix} 6 & 2 & -2 \\ 2 & 6 & -2 \\ -2 & -2 & 10 \end{pmatrix} \in S_3(\mathbb{R}).$$

Son polynôme caractéristique est  $\chi_S = (6 - X)(4 - X)(12 - X)$  donc la matrice symétrique  $S$  est définie positive. On cherche  $T = \begin{pmatrix} a & 0 & 0 \\ b & c & 0 \\ d & e & f \end{pmatrix} \in M_3(\mathbb{R})$  triangulaire inférieure de coefficients

diagonaux strictement positifs telle que

$$S = T^t T = \begin{pmatrix} a & 0 & 0 \\ b & c & 0 \\ d & e & f \end{pmatrix} \begin{pmatrix} a & b & d \\ 0 & c & e \\ 0 & 0 & f \end{pmatrix}.$$

On a alors

1.  $a^2 = 6$  donc  $a = \sqrt{6}$  (car  $a > 0$ ),
2.  $ba = 2$  donc  $b = \frac{2}{\sqrt{6}}$ ,
3.  $da = -2$  donc  $d = -\frac{2}{\sqrt{6}}$ ,
4.  $b^2 + c^2 = 6$  donc  $c = \sqrt{6 - b^2} = \sqrt{\frac{16}{3}}$  ( $c > 0$ ),
5.  $db + ec = -2$  donc  $e = \frac{1}{c}(-2 - db) = \sqrt{\frac{3}{16}}(-2 + \frac{2}{3}) = -\frac{4}{3}\sqrt{\frac{3}{16}}$ ,
6.  $d^2 + e^2 + f^2 = 10$  donc  $f = \sqrt{10 - \frac{2}{3} - \frac{1}{3}} = 3$  ( $f > 0$ ).

Ainsi

$$S = \begin{pmatrix} \sqrt{6} & 0 & 0 \\ \frac{2}{\sqrt{6}} & \sqrt{\frac{16}{3}} & 0 \\ -\frac{2}{\sqrt{6}} & -\frac{4}{3}\sqrt{\frac{3}{16}} & 3 \end{pmatrix} \begin{pmatrix} \sqrt{6} & \frac{2}{\sqrt{6}} & -\frac{2}{\sqrt{6}} \\ 0 & \sqrt{\frac{16}{3}} & -\frac{4}{3}\sqrt{\frac{3}{16}} \\ 0 & 0 & 3 \end{pmatrix}$$

est la décomposition de Cholesky de la matrice symétrique définie positive  $S$ .

*Remarque 8.5.5.* • Comme illustré dans l'exemple ci-dessus, le calcul de la décomposition de Cholesky de  $S$  est plus avantageux que le calcul de la décomposition  $LU$  de  $S$  (il y a moins de coefficients à déterminer).

- Si  $B$  est un vecteur colonne de  $M_{n,1}(\mathbb{R})$ , la décomposition de Cholesky de la matrice  $S$  permet de résoudre efficacement le système  $SX = B$  de vecteur inconnu  $X \in M_{n,1}(\mathbb{R})$  : résoudre ce système revient à résoudre successivement les deux systèmes triangulaires inversibles  $TY = B$ , de vecteur inconnu  $Y \in M_{n,1}(\mathbb{R})$ , et  ${}^tTX = Y$ .