UNIVERSITÉ CLAUDE BERNARD LYON 1
Licence Sciences, Technologies, Santé

**Enseignement de mathématiques
des parcours Informatique**

ANALYSE MATRICIELLE ET ALGÈBRE LINÉAIRE APPLIQUÉE

- Notes de cours et de travaux dirigés -

PHILIPPE MALBOS
malbos@math.univ-lyon1.fr

Table des matières

CHAPITRE

0

Préliminaires algébriques

Sommaire

Ce chapitre contient peu de démonstrations, son rôle est de fixer les notations et de rappeler les structures algébriques fondamentales, ainsi que les principaux résultats algébriques que nous utiliserons dans ce cours. Nous renvoyons le lecteur au cours de première année pour tout approfondissement.

§ 1 Ensembles et applications

0.1.1. Applications.— Soient A et B deux ensembles. Une *application* f de A dans B est un procédé qui à tout élement x de A associe un élément unique de B, noté $f(x)$. On note $f : A \longrightarrow B$, ou $A \xrightarrow{f} B$, ou encore

$$\begin{aligned} f : A &\longrightarrow B \\ x &\longrightarrow f(x). \end{aligned}$$

On note $f(A)$ l'image de l'ensemble A, définie par

$$f(A) = \{y \mid y \in B,\ \exists\, x \in A,\ \text{tel que } y = f(x)\}.$$

L'image inverse d'un sous-ensemble $Y \subset B$ est définie par

$$f^{-1}(Y) = \{x \mid x \in A,\ f(x) \in Y\}.$$

Une application $f : A \longrightarrow B$ est dite *injective* si, $f(x) = f(y)$ implique $x = y$. Elle est dite *surjective* si $f(A) = B$, *i.e.*, pour tout $y \in B$, il existe un $x \in A$ tel que $y = f(x)$. Une application est dite *bijective* si elle est à la fois injective et surjective.

Si $f : A \longrightarrow B$ et $g : B \longrightarrow C$ sont deux applications, on note $g \circ f$, ou encore gf, l'application, dite *composée*, définie par

$$\begin{aligned} g \circ f : A &\longrightarrow C \\ x &\longrightarrow g(f(x)). \end{aligned}$$

La composée des applications est une opération associative, i.e., étant données trois applications $A \xrightarrow{f} B \xrightarrow{g} C \xrightarrow{h} D$, on a

$$h \circ (g \circ f) = (h \circ g) \circ f.$$

0.1.2. Quelques ensembles fondamentaux de nombres.— Dans tout ce cours, nous supposons connus les ensembles de nombres suivants et les opérations d'addition, de soustraction, de multiplication et de division sur ces ensembles :

- l'ensemble des entiers naturels, 0, 1, 2, ..., noté $\mathbb{N}$,
- l'ensemble des entiers relatifs, noté $\mathbb{Z}$, formé des entiers naturels et de leurs opposés,
- l'ensemble des rationnels, noté $\mathbb{Q}$, formé des quotients $\frac{p}{q}$, où p et q sont des entiers relatifs, avec q non nul,
- l'ensemble des réels, noté $\mathbb{R}$, qui contient les nombres rationnels et les irrationnels,
- l'ensemble des complexes, noté $\mathbb{C}$, formé des nombres $a + ib$, où a et b sont des réels et i un complexe vérifiant $i^2 = -1$.

Si p et q sont deux entiers relatifs, on notera

$$[\![p, q]\!] = \{a \in \mathbb{Z} \mid p \leqslant a \leqslant q\}.$$

§ 2 Les corps

Un *corps* est un objet algébrique constitué d'un ensemble et de deux opérations sur cet ensemble, une addition et une multiplication, qui satisfont à certaines relations. Intuitivement, cette structure est proche de notre intuition de nombres et des opérations que l'on peut leur appliquer. Avant d'énoncer les relations des deux opérations de la structure de corps, rappelons la structure de groupe.

0.2.1. Les groupes.— Un *groupe* est un ensemble G muni d'une opération $\star$, associant à deux éléments a et b de G un troisième élément de G, noté $a \star b$, satisfaisant les assertions suivantes

i) l'opération est *associative*, *i.e.*, pour tous éléments a, b et c de G,

$$a \star (b \star c) = (a \star b) \star c,$$

ii) il existe un élément e dans G, appelé *neutre*, tel que, pour tout élément a de G,

$$a \star e = e \star a = a,$$

iii) pour tout élément a de G, il existe un élément *inverse*, que nous noterons a^{-1}, tel que

$$a \star a^{-1} = e = a^{-1} \star a.$$

Exercice 1.— On définit sur l'ensemble des nombres réels l'opération $\star$ en posant

$$a \star b = 2a + 2b.$$

1. Cette opération est-elle associative ?
2. L'opération

$$a \star b = 2a + b$$

est-elle associative ?

Exercice 2.—

1. Montrer qu'un groupe possède un unique élément neutre.
2. Montrer que dans un groupe, l'inverse d'un élément est unique.

0.2.2. Exemples.—

1) Le groupe *trivial* est le groupe à un seul élément, l'élément neutre.

2) L'ensemble des entiers $\mathbb{Z}$ forme un groupe pour l'addition usuelle. Il ne forme pas un groupe pour la multiplication.

3) L'ensemble des nombres rationnels $\mathbb{Q}$ forme un groupe pour l'addition. L'ensemble $\mathbb{Q} - \{0\}$ des nombres rationnels non nul est un groupe pour la multiplication.

4) L'ensemble des complexes non nuls $\mathbb{C} - \{0\}$, muni de la multiplication usuelle des complexes.

5) L'ensemble $\mathbb{R}^n$ des n-uplets ordonnées

$$(x_1, \ldots, x_n)$$

de nombres réels, muni de l'opération

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

forme un groupe.

Exercice 3.— Justifier toutes les propriétés précédentes. Dans le cas de $\mathbb{R}^n$, déterminer l'élément neutre du groupe et l'inverse d'un n-uplet $(x_1, \ldots, x_n)$.

0.2.3. Les groupes abéliens.— Un groupe est dit *abélien*, ou *commutatif*, si tous éléments a et b vérifient

$$a \star b = b \star a.$$

Les groupes des exemples 0.2.2 sont abéliens.

Exercice 4.— Les opérations de l'exercice 1 sont-elles commutatives ?

Exercice 5.— Soit X un ensemble.

1. Montrer que l'ensemble des permutations de X, i.e. des bijections de X dans lui-même, forment un groupe.
2. Montrer que ce groupe n'est pas commutatif lorsque X possède au moins trois éléments.

0.2.4. Les corps.— Un *corps* (commutatif) est un ensemble $\mathbb{K}$ sur lequel une opération d'addition $(a,b) \to a+b$ et une opération de multiplication $(a,b) \to ab$ sont définies et satisfont aux assertions suivantes :

i) $\mathbb{K}$ est un groupe abélien pour l'addition,

ii) $\mathbb{K} - \{0\}$ est un groupe abélien pour la multiplication,

iii) la multiplication est distributive par rapport à l'addition, i.e., pour tous éléments a, b et c, on a

$$a(b+c) = ab + ac.$$

L'élément neutre pour l'addition, appelé *zero*, est noté 0, l'inverse de a est appelé l'*opposé* de a et noté $-a$, l'élement neutre pour la multiplication est appelé *unité* et noté 1, l'*inverse* de a pour la multiplication est noté a^{-1}.

0.2.5. Exemples.—

1) L'ensemble des nombres rationnels $\mathbb{Q}$, l'ensemble des nombres réels $\mathbb{R}$ et l'ensemble des nombres complexes $\mathbb{C}$, munis des opérations d'addition et de multiplication usuelles sont des corps.

2) L'ensemble $\mathbb{Z}$ des entiers relatifs n'est pas un corps.

3) Un exemple de corps fini, i.e., avec un nombre fini d'éléments, est donné par l'ensemble, noté $\mathbb{Z}/p\mathbb{Z}$, des entiers modulo un entier premier p, muni des opérations d'addition et de multiplication induites de celles de $\mathbb{Z}$.

Exercice 6.— Montrer que $\mathbb{Z}/4\mathbb{Z}$ n'est pas un corps.

Exercice 7.— Montrer que dans un corps, l'élément neutre de l'addition joue le rôle d'*annulateur*, i.e., pour tout élément a, on a :

$$a0 = 0.$$

Par définition, un groupe ne peut être vide, il contient au moins un élément. Un corps contient donc au moins deux éléments 0 et 1 qui sont nécessairement distincts.

Exercice 8.— Montrer qu'un corps ne contient pas de diviseur de zero, c'est-à-dire que si a et b sont deux éléments non nul d'un corps $\mathbb{K}$, alors leur produit ab est non nul.

Il n'existe qu'un seul corps à deux éléments.

Exercice 9.— Établir les tables d'addition et de multiplication du corps à deux éléments.

0.2.6. Extension de corps.— Un sous-ensemble $\mathbb{L}$ d'un corps $\mathbb{K}$ est un *sous-corps* de $\mathbb{K}$ si les opérations du corps $\mathbb{K}$ munissent $\mathbb{L}$ d'une structure de corps. On dit alors que $\mathbb{K}$ est une *extension* du corps $\mathbb{L}$. Par exemple, le corps des réels $\mathbb{R}$ est une extension du corps des rationnels $\mathbb{Q}$ et le corps des complexes $\mathbb{C}$ est une extension du corps $\mathbb{R}$.

§ 3 Les anneaux

La structure d'anneau généralise celle de corps. Un ensemble muni d'une opération d'addition et d'une opération de multiplication qui satisfont à tous les axiomes de corps, excepté l'existence d'un élément inverse a^{-1}, pour tout élément a non nul, est appelé un *anneau commutatif*. Pour que notre définition soit complète, on convient, qu'il existe un anneau qui possède un seul élément.

Par exemple, l'ensemble des entiers relatifs $\mathbb{Z}$, muni de l'addition et de la multiplication, n'est pas un corps - les éléments non nuls ne sont pas tous inversibles - mais il forme un anneau commutatif. Nous verrons que l'ensemble $A[x]$ des polynômes à une indéterminée à coefficients dans un anneau ou un corps A forme un anneau ; les principales constructions sur les anneaux de polynômes sont rappelées dans la section suivante.

0.3.1. Les anneaux.— Un *anneau* est un ensemble A muni d'une opération d'*addition* $(a,b) \to a+b$ et d'une opération de *multiplication* $(a,b) \to ab$ qui satisfont aux assertions suivantes

i) A est un groupe abélien pour l'addition,

ii) la multiplication est associative, i.e., pour tous éléments a, b et c de A,

$$(ab)c = a(bc).$$

iii) la multiplication possède un élément neutre dans A, appelé *unité* et noté 1, vérifiant pour tout élément a de A,

$$1a = a1 = a.$$

iv) la multiplication est *distributive* par rapport à l'addition, i.e., pour tous éléments a,b,c de A, on a :

$$a(b+c) = ab+ac, \qquad (b+c)a = ba+ca.$$

Un anneau est dit *commutatif* si sa multiplication est commutative.

Exercice 10.— Montrer que dans un anneau A, on a, pour tous éléments a et b,

1. $0a = a0 = 0$,
2. $(-1)a = -a$,
3. $-(ab) = (-a)b = a(-b)$,
4. $(-a)(-b) = ab$.

0.3.2. Exemples.—

1) L'ensemble des entiers relatifs $\mathbb{Z}$, muni de l'addition et de la multiplication usuelles, forme un anneau commutatif.
2) Un corps (commutatif) est un anneau $\mathbb{K}$ non réduit à $\{0\}$, tel que la multiplication muni $\mathbb{K} - \{0\}$ d'une structure de groupe abélien.
3) Si $1 = 0$ dans un anneau A, alors A est réduit à $\{0\}$, car pour tout élément a de A, $a = 1a = 0a = 0$.

0.3.3. Endomorphismes d'un groupe abélien.— Rappelons qu'un *endomorphisme* d'un groupe $(G, \star)$ est un morphisme de groupes de G dans lui-même, c'est-à-dire, une application $f : G \longrightarrow G$ vérifiant, pour tous $a, b \in G$,

$$f(a \star b) = f(a) \star f(b).$$

L'ensemble des endomorphismes d'un groupe abélien $(G, +)$, muni de l'addition induite de celle sur G et de la composition, est un anneau non commutatif en général.

0.3.4. Formule du binôme.— Dans un anneau, si deux éléments a et b commutent, i.e., $ab = ba$, alors on a la formule dite du *binôme de Newton*, pour tout entier naturel n,

$$(a+b)^n = \sum_{p=0}^{n} \binom{n}{p} a^p b^{n-p}.$$

Exercice 11.— Démontrer la formule du binôme de Newton.

0.3.5. Caractéristique d'un anneau commutatif.— Soit A un anneau commutatif. La *caractéristique* de A est le plus petit entier naturel non nul q, tel que l'addition de q fois l'unité soit égale à zero :

$$q.1 = \underbrace{1 + 1 + \ldots + 1}_{q \text{ fois}} = 0.$$

Si un tel entier n'existe pas, on dit que l'anneau est de caractéristique nulle.

Exercice 12.—

1. Montrer qu'un anneau commutatif fini est de caractéristique non nulle.
2. Montrer que la caractéristique d'un corps fini est un nombre premier.

Exercice 13.— Construire un corps de caractéristique 3.

Exercice 14.— Montrer que dans un anneau commutatif de caractéristique un nombre premier p, alors, pour tous éléments a et b de A, on a

$$(a+b)^p = a^p + b^p.$$

0.3.6. Division euclidienne dans l'anneau $\mathbb{Z}$.— La *division euclidienne* est un résultat fondamental de l'arithmétique élémentaire sur les entiers ou les polynômes. Avant d'identifier les anneaux dans lesquels, un tel algorithme est disponible, rappelons la division euclidienne sur les entiers.

0.1 Théorème (division euclidienne).— Soient $a, b \in \mathbb{Z}$, avec $b > 0$. Il existe un couple unique (q, r) d'entiers dans $\mathbb{Z}$ tel que :

$$a = bq + r, \quad \text{avec } 0 \leqslant r < b.$$

L'entier q est appelé le *quotient* de la division euclidenne de a par b et l'entier r est appelé le *reste* de la division euclidenne de a par b.

Preuve. Montrons dans un premier temps l'unicité du couple. Supposons que (q, r) et (q', r') soient deux couples vérifiant la condition, alors

$$a = bq + r = bq' + r'.$$

D'où $r' - r = b(q - q')$, par suite b divise $r' - r$. Comme, par hypothèse, $0 \leqslant r < b$ et $0 \leqslant r' < b$, on a

$$b|q - q'| = |r' - r| < b.$$

Par suite, $|q - q'| = 0$, d'où $q = q'$ et $r = r'$.

Montrons l'existence du couple. Considérons l'ensemble $A = \{k \in \mathbb{Z} \mid bk \leqslant a\}$. C'est une partie non vide et majorée de $\mathbb{Z}$. En effet, si $a \geqslant 0$, alors $0 \in A$, d'où A est non vide, et comme $1 \leqslant b$, l'entier a majore A. Si $a < 0$, alors $a \in A$, d'où A est non vide et 0 majore A. Par suite, l'ensemble A admet un plus grand élément q. On a

$$bq \leqslant a < b(q+1).$$

En posant $r = a - bq$, on a $0 \leqslant r < b$. □

De l'unicité du quotient et du reste de la division eulcidienne, on déduit qu'un entier b divise un entier a si, et seulement si, le reste de division euclidienne de a par b est nul.

Exercice 15.— Soit n un entier naturel. Calculer la division euclidienne de

1. l'entier $n^3 + n^2 + 2n + 1$ par $n + 1$,
2. l'entier $n^4 + 4n^3 + 6n^2$ par $n^2 + 2$.

FIGURE 0.1.: Euclide de Samos (325 - 265 av. J.-C.)

Euclide est un mathématicien de la Grèce antique né vers 325 av. J.C. et mort vers 265 av. J.C.. Nous n'avons que très peu d'information sur la vie d'Euclide. L'article de Fabio Acerbi du site Image des mathématiques [1] *présente ce que nous savons à ce jour sur le personnage d'Euclide. Euclide est l'auteur des* Éléments *qui est un texte fondateur de la géométrie.*

FIGURE 0.2.: Un fragment des éléments d'Euclide, papyrus daté d'entre 75 et 125 de notre ère.

0.3.7. Les anneaux euclidiens.— Soit A un anneau commutatif. On appelle *algorithme euclidien* sur A toute application

$$\varphi : A - \{0\} \longrightarrow \mathbb{N},$$

1. Fabio Acerbi, « Euclide » - Images des Mathématiques, CNRS, 2010. En ligne, URL : `http://images.math.cnrs.fr/Euclide.html`

telle que, pour tout $a \in A$ et tout $b \in A - \{0\}$, il existe $q \in A$ et $r \in A$, tels que

$$a = bq + r, \quad \text{avec } \varphi(r) < \varphi(b) \ \text{ ou } \ r = 0.$$

Un anneau commutatif A est dit *euclidien*, s'il vérifie les deux propriétés suivantes

i) A est *intègre*, i.e., pour tous éléments a et b de A,

$$ab = 0 \ \Rightarrow (\ a = 0 \ \text{ ou } \ b = 0\ ).$$

ii) il existe sur A un algorithme euclidien.

0.3.8. Exemple.— L'anneau $\mathbb{Z}$ est euclidien. Il est en effet intègre et l'application valeur absolue $|\ | : \mathbb{Z} - \{0\} \longrightarrow \mathbb{N}$ est un algorithme euclidien, car, pour tout $a \in \mathbb{Z}$ et tout $b \in \mathbb{Z} - \{0\}$, il existe $q, r \in \mathbb{Z}$ tels que

$$a = bq + r, \quad \text{avec } |r| < |b| \ \text{ ou } \ r = 0.$$

Attention, le couple (q, r) n'est pas ici unique, par exemple, on a

$$5 = (-3)(-2) + (-1) \text{ avec } |-1| < |-3|,$$

et

$$5 = (-3)(-1) + 2 \text{ avec } |2| < |-3|.$$

Dans la suite, nous montrerons que si $\mathbb{K}$ est un corps, l'anneau $\mathbb{K}[x]$ est euclidien.

Exercice 16.— Montrer que l'anneau $\mathbb{D}$ des nombres décimaux, i.e., le sous-anneau de $\mathbb{Q}$, engendré par $1/10$, est euclidien.

§ 4 Les polynômes à une indéterminée

0.4.1. Polynômes sur un corps.— Avant d'aborder la notion de polynôme, rappelons qu'il est important de distinguer les polynômes des fonctions polynomiales. En effet, considérons le polynôme $f = x^2 - x$ à coefficients dans le corps $\mathbb{Z}/2\mathbb{Z}$. La fonction polynomiale associée $\widetilde{f} : \mathbb{Z}/2\mathbb{Z} \longrightarrow \mathbb{Z}/2\mathbb{Z}$, définie par

$$\widetilde{f}(a) = a^2 - a, \ \text{ pour tout } a \in \mathbb{Z}/2\mathbb{Z},$$

est nulle, car $\widetilde{f}(0) = 0$ et $\widetilde{f}(1) = 0$, alors que le polynôme f n'est pas nul.

Exercice 17.— Montrer qu'il n'existe que quatre fonctions polynomiales à coefficients dans le corps $\mathbb{Z}/2\mathbb{Z}$ et une infinité de polynômes à coefficients dans ce corps.

La situation est différente pour les polynômes à coefficients dans les corps infinis, dans ce cas, il existe une correspondance biunivoque entre les polynômes et les fonctions polynomiales, cf. section 0.4.5.

0.4.2. Les polynômes.— Soit $\mathbb{K}$ un corps. On appelle *polynôme* à coefficients dans $\mathbb{K}$, toute suite $f = (a_n)_{n\in\mathbb{N}}$ d'éléments de $\mathbb{K}$, nulle à partir d'un certain rang. On note $\mathbb{K}^{(\mathbb{N})}$ l'ensemble de ces suites.

On définit sur l'ensemble $\mathbb{K}^{(\mathbb{N})}$ une addition et un produit externe par un scalaire en posant, pour tous $f = (a_n)_{n\in\mathbb{N}}$, $g = (b_n)_{n\in\mathbb{N}}$ et $\lambda \in \mathbb{K}$,

$$f + g = (a_n + b_n)_{n\in\mathbb{N}}, \qquad \lambda f = (\lambda a_n)_{n\in\mathbb{N}}.$$

En outre, on définit une multiplication en posant, pour tous $f = (a_n)_{n\in\mathbb{N}}$, $g = (b_n)_{n\in\mathbb{N}}$,

$$fg = (c_n)_{n\in\mathbb{N}}, \qquad \text{avec} \qquad c_n = \sum_{i=0}^{n} a_i b_{n-i}.$$

0.4.3. L'algèbre des polynômes.— Ces trois opérations munissent l'ensemble $\mathbb{K}^{(\mathbb{N})}$ d'une structure de $\mathbb{K}$-*algèbre* associative, commutative et unitaire, c'est-à-dire,

i) $\mathbb{K}^{(\mathbb{N})}$ muni de l'addition et du produit par un scalaire est un $\mathbb{K}$-espace vectoriel,

ii) $\mathbb{K}^{(\mathbb{N})}$ muni de l'addition et de la multiplication est un anneau commutatif,

iii) pour tous $f, g \in \mathbb{K}^{(\mathbb{N})}$ et tous scalaires $\lambda, \mu \in \mathbb{K}$, on a

$$(\lambda f)(\mu g) = (\lambda\mu)(fg).$$

0.4.4. Notion d'indéterminée.— L'écriture des polynômes sous forme de suite est peu manipulable, aussi, on préfère la notation basée sur la notion d'*indéterminée*. Notons x le polynôme de $\mathbb{K}^{(\mathbb{N})}$ dont tous les termes sont nuls, sauf celui de degré 1 :

$$x = (0, 1, 0, \ldots).$$

Par convention, on pose $x^0 = 1$. On définit les puissances de x par récurrence, pour tout entier k, $x^{k+1} = xx^k$. Ainsi, si $f = (a_n)_{n\in\mathbb{N}}$, on montre que

$$f = \sum_{i=0}^{+\infty} a_i x^i.$$

Les scalaires a_i sont appelés les *coefficients* du polynôme f. On montre que deux polynômes sont égaux si, et seulement si, ils ont les mêmes coefficients :

$$\sum_{k=0}^{+\infty} a_k x^k = \sum_{k=0}^{+\infty} b_k x^k \quad \text{si, et seulement si,} \quad a_k = b_k, \text{ pour tout } k \in \mathbb{N}.$$

On notera alors $\mathbb{K}[x]$ l'ensemble des *polynômes à une indéterminée* à coefficients dans le corps $\mathbb{K}$. Avec ces notations, l'addition des polynômes est définie de la façon suivante, pour $f = \sum_{i=0}^{m} a_i x^i$ et $g = \sum_{j=0}^{n} b_j x^j$, alors

$$f + g = \sum_{k=0}^{\max\{m,n\}} (a_k + b_k) x^k,$$

avec $a_k = 0$, pour $k > m$ et $b_k = 0$ pour $k > n$. Par ailleurs, pour la multiplication, on a

$$fg = \sum_{k=0}^{m+n} \left(\sum_{\substack{i,j\geqslant 0\\ i+j=k}} (a_i b_j) \right) x^k$$

0.4.5. Fonction polynomiale.— Étant donné un polynôme $f = \sum_{i=0}^{n} a_i x^i$ de $\mathbb{K}[x]$, on définit la *fonction polynomiale* associée comme l'application

$$\widetilde{f} : \mathbb{K} \longrightarrow \mathbb{K},$$

qui, à tout $a \in \mathbb{K}$, associe le scalaire $f(a) \in \mathbb{K}$, obtenu en remplaçant dans l'expression de f l'indéterminée x par a.

Nous avons vu en 0.4.1 que sur un corps fini, les notions de polynômes et de fonction polynomiale ne coïncident pas. Nous allons voir que c'est le cas lorsque le corps est infini, par exemple lorsque $\mathbb{K}$ est $\mathbb{R}$ ou $\mathbb{C}$.

Exercice 18.— Supposons que $\mathbb{K}$ est le corps $\mathbb{R}$ ou $\mathbb{C}$.

1. Montrer que l'application
$$\varphi : \mathbb{K}[x] \longrightarrow \mathbb{K}^{\mathbb{K}}$$
définie par $\varphi(f) = \widetilde{f}$ est injective.

2. Montrer que deux polynômes à coefficients dans $\mathbb{K}$ sont égaux si, et seulement si leurs fonctions polynomiales associées sont égales.

0.4.6. Degré d'un polynôme.— Soit f un polynôme de $\mathbb{K}[x]$. Si $f = 0$, on pose $\deg f = -\infty$, si $f = \sum_{i=0}^{m} a_i x^i$ est non nul, on note $\deg f$ le plus grand entier naturel n tel que a_n soit non nul. L'entier $\deg f$ est appelé le *degré* du polynôme f.

Un polynôme non nul f de degré $n \geqslant 0$ s'écrit de façon unique sous la forme

$$f = a_0 + a_1 x + \ldots + a_n x^n,$$

où a_n est non nul. Le degré de f est le plus grand exposant de x apparaissant dans f.

0.4.7. Les monômes.— On appelera *monôme* un polynôme de la forme x^k, où k est un entier naturel. La famille de monômes $(x^n)_{n\in N}$ forme une base du $\mathbb{K}$-espace vectoriel $\mathbb{K}[x]$. On l'appelle base canonique de $\mathbb{K}[x]$.

0.4.8. Terme de plus haut degré.— Le *coefficient de plus haut degré* (*leading coefficient*) d'un polynôme f de $\mathbb{K}[x]$, noté $\mathrm{lc}(f)$, est le coefficient du monôme de plus grand exposant. Le *terme de plus haut degré* d'un polynôme f (*leading term*), noté $\mathrm{lt}(f)$, est le terme de plus haut degré de f. Par exemple, pour le polynôme $f = a_0 + a_1 x + \ldots + a_n x^n$, on a

$$\deg f = n, \quad \mathrm{lt}(f) = a_n x^n, \quad \mathrm{lc}(f) = a_n.$$

Un polynôme est dit *unitaire*, si le coefficient de son terme de plus haut degré est égal à 1.

Exercice 19.— Montrer que pour tous polynômes f et g de $\mathbb{K}[x]$, on a

1. $\mathrm{lc}(fg) = \mathrm{lc}(f)\mathrm{lc}(g)$,
2. $\mathrm{lt}(fg) = \mathrm{lt}(f)\,\mathrm{lt}(g)$.

§ 5 Arithmétique des polynômes

0.5.1. Divisibilité.— Soient f et g deux polynômes de $\mathbb{K}[x]$. On dit que g *divise* f, ou que f est *divisible* par g, ou encore que f est un multiple de g, s'il existe un polynôme q de $\mathbb{K}[x]$ tel que $f = gq$. On note alors $g|f$.

Exercice 20.— Montrer que le polynôme $x+3$ divise le polynôme x^3+27.

Exercice 21.— Montrer que pour deux polynômes f et g non nuls de $\mathbb{K}[x]$, $\deg f \leqslant \deg g$ si, et seulement si, $\mathrm{lt}(f) \mid \mathrm{lt}(g)$.

Exercice 22.— Soient f et g deux polynômes de $\mathbb{K}[x]$. Montrer que $f|g$ et $g|f$ si, et seulement si, il existe un scalaire non nul λ de $\mathbb{K}$ te que $f = \lambda g$.

0.5.2. La division euclidienne dans $\mathbb{K}[x]$.— Il existe sur l'anneau $\mathbb{K}[x]$ une division euclidienne comme celle que nous avons vue sur l'anneau $\mathbb{Z}$. Par exemple, considérons deux polynomes

$$f = x^3 - 3x^2 + 4x + 7, \qquad g = 2x^2 + 4x + 6,$$

de $\mathbb{Q}[x]$. La division de f par g a pour quotient $\frac{1}{2}x - \frac{5}{2}$ et pour reste $11x+22$. On les obtient en procédant de la même façon que la division des entiers :

$$\begin{array}{rrrr|rrr}
x^3 & -3x^2 & +4x & +7 & 2x^2 & +4x & +6 \\ \cline{5-7}
x^3 & +2x^2 & +3x & & \frac{1}{2}x & -\frac{5}{2} & \\ \cline{1-4}
 & -5x^2 & +x & +7 & & & \\
 & -5x^2 & -10x & -15 & & & \\ \cline{2-4}
 & & 11x & +22 & & &
\end{array}$$

La première étape consiste à multiplier le polynôme g par $\frac{1}{2}x$, puis à soustraire le résultat à f, soit

$$f - \frac{x^3}{2x^2}g = f - \frac{1}{2}xg = -5x^2 + x + 7.$$

L'idée consiste à multiplier g par un terme, ici $\frac{x^3}{2x^2}$, de telle façon que le terme de plus haut degré de g multiplié par ce terme annule le terme de plus haut degré de f. On obtient ainsi un nouveau polynôme $h = -5x^2 + x + 7$, on dit que h est une *reduction* de f par g, on note

$$f \xrightarrow{g} h.$$

On répète alors ce processus, jusqu'à obtenir le reste

$$r = h - (-\frac{5}{2})g = 11x + 22.$$

La division se compose ainsi d'une suite de réductions par g :

$$f \xrightarrow{g} h \xrightarrow{g} r.$$

0.5.3. Le cas général.— Plus généralement, considérons deux polynômes

$$f = a_n x^n + \ldots + a_1 x + a_0, \quad g = b_m x^m + \ldots + b_1 x + b_0,$$

avec $\deg f = n \geqslant \deg g = m$. La première étape dans la division de f par g consiste à soustraire à f le produit

$$\frac{a_n}{b_m} x^{n-m} g,$$

qui, avec les notations définies en 0.4.8, s'écrit

$$\frac{\mathrm{lt}\,(f)}{\mathrm{lt}\,(g)} g.$$

On obtient ainsi comme premier reste le polynôme

$$h = f - \frac{\mathrm{lt}\,(f)}{\mathrm{lt}\,(g)} g.$$

On dit que f se *réduit* en h par g, on note

$$f \xrightarrow{g} h.$$

On répète alors l'opération de réduction par g, pour obtenir un nouveau reste :

$$h' = h - \frac{\mathrm{lt}\,(h)}{\mathrm{lt}\,(g)} g.$$

Dans la réduction $f \xrightarrow{g} h$, on notera que le reste h a un degré strictement inférieur au degré de f. On peut alors poursuivre le processus de réduction, jusqu'à obtenir un reste, dont le degré est strictement inférieur au degré du polynôme g. On obtient ainsi une suite de réductions par g qui termine sur un reste r, tel que $\deg r < \deg g$:

$$f \xrightarrow{g} h \xrightarrow{g} h' \xrightarrow{g} \ldots \xrightarrow{g} r.$$

On montre ainsi l'existence du quotient et du reste dans le théorème de la division euclidienne :

0.2 Théorème (de la division euclidienne).— Soient f et g deux polynômes de $\mathbb{K}[x]$, avec $g \neq 0$. Il existe un couple unique (q,r) de polynômes de $\mathbb{K}[x]$ tel que :

$$f = gq + r,$$

avec $\deg r < \deg g$.

Le polynôme q est appelé le *quotient* de la division euclidenne de f par g et le polynôme r est appelé le *reste* de la division euclidenne de f par g. Si le reste de la division euclidienne de f par g est nul, alors le polynôme g divise f.

Exercice 23.— Montrer l'unicité des polynômes q et r.

Exercice 24.— Étant donnés deux élements distincts a et b d'un corps $\mathbb{K}$. Calculer le reste de la division euclidienne d'un polynôme f de $\mathbb{K}[x]$ par le polynôme $(x-a)(x-b)$ en fonction de $f(a)$ et $f(b)$.

Exercice 25.— Calculer le reste de la division de f par g avec
1. $f = x^3 + x^2 + x + 1$, $g = x + 1$,
2. $f = x^3 + x^2 + x + 1$, $g = x - 1$,
3. $f = x^3 + 3x^2 - 7x + 5$, $g = x^2 - 3$,
4. $f = x^4 + 2x^3 - 4x^2 + 5$, $g = x^3 + 5$.

0.3 Théorème.— Si $\mathbb{K}$ est un corps, l'anneau $\mathbb{K}[x]$ est euclidien.

Preuve. D'après le théorème 0.2, l'application $\deg : \mathbb{K}[x] - \{0\} \longrightarrow \mathbb{N}$ est un algorithme euclidien sur $\mathbb{K}[x]$. $\square$

ENTRÉE : $f, g \in \mathbb{K}[x]$ avec $g \neq 0$,

SORTIE : $q, r \in \mathbb{K}[x]$ tels que $f = gq + r$ avec $\big(r = 0$ ou $\deg r < \deg g\big)$.

INITIALISATION : $q := 0\,;\, r := f$

TANT QUE : $r \neq 0$ **ET** $\deg g \leqslant \deg r$ **FAIRE**

$$q := q + \frac{\operatorname{lt}(r)}{\operatorname{lt}(g)}$$
$$r := r - \frac{\operatorname{lt}(r)}{\operatorname{lt}(g)} g.$$

Algorithme de la division des polynômes d'une indéterminée.

Nous reviendrons sur cet algorithme de la division dans un prochain chapitre, en particulier pour ses nombreuses applications.

0.5.4. Polynômes premiers entre eux.— Deux polynômes sont dits *premiers entre eux*, si leurs seuls diviseurs communs sont les polynômes de degré nul. Plus généralement, des polynômes $f_1,\ldots,f_s$ de $\mathbb{K}[x]$ sont dits

- *premiers entre eux dans leur ensemble*, si les seuls polynômes qui divisent simultanément les polynômes $f_1,\ldots,f_s$ sont de degré nul,

- *premiers entre eux deux à deux*, si, pour tout i différent de j, les polynômes f_i et f_j sont premiers entre eux.

Si f_i et f_j sont premiers entre eux, alors les polynômes f_1, …, f_i, …, f_j, …, f_s sont premier entre eux dans leur ensemble. Attention, les polynômes $f_1 = x-1$, $f_2 = (x-1)(x-2)$ et $f_3 = x-3$ sont premiers entre eux dans leur ensemble, alors que les polynômes f_1 et f_2 ne sont pas premiers entre eux.

0.4 Théorème (Identité de Bézout).— Les polynômes $f_1,\ldots,f_s \in \mathbb{K}[x]$ sont premiers entre eux dans leur ensemble si, et seulement si, il existe des polynômes $u_1,\ldots,u_s$ de $\mathbb{K}[x]$, tels que

$$f_1h_1+\cdots+f_sh_s = 1.$$

L'égalité $f_1h_1+\cdots+f_sh_s = 1$ s'appelle une *identité de Bézout*.

0.5.5. Exemples.— Les polynômes $x-1$ et $x+2$ sont premiers entre eux, on a l'identité de Bézout

$$-\frac{1}{3}(x-1)+\frac{1}{3}(x+2) = 1.$$

Les polynômes x^2-1 et $x+2$ sont premiers entre eux, une identité de Bézout est donnée par

$$\frac{1}{3}(x^2-1)+(-\frac{1}{3}x+\frac{2}{3})(x+2) = 1.$$

0.5.6. Calculer une identité de Bézout.— L'algorithme d'Euclide permet de calculer une identité de Bézout. Étant donnés deux polynômes f_1 et f_2, premiers entre eux, l'algorithme suivant permet de calculer une identité de Bézout

$$f_1h_1+f_2h_2 = 1.$$

Soient $f_1, f_2 \in \mathbb{K}[x] - \{0\}$ deux polynômes premiers entre eux. On pose $r_0 = f_1$, $r_1 = f_2$. On calcule les divisions euclidiennes

$$\begin{aligned}
r_0 &= r_1 q_1 + r_2, \quad \deg r_2 < \deg r_1,\\
r_1 &= r_2 q_2 + r_3, \quad \deg r_3 < \deg r_2,\\
&\vdots\\
r_{n-2} &= r_{n-1} q_{n-1} + r_n, \quad \deg r_n < \deg r_{n-1},
\end{aligned}$$

Alors, il existe $n_0 \geqslant 0$ tel que, pour tout $n \geqslant n_0$, $r_n = 0$. Les polynômes f_1 et f_2 sont premiers entre eux, par suite le dernier reste non nul r_{n_0-1} est une constante $b \in \mathbb{K} - \{0\}$. Pour déterminer les polynômes h_1 et h_2 dans l'identité de Bézout, il suffit de partir de

$$r_{n_0-1} = b = r_{n_0-3} - r_{n_0-2} q_{n_0-2},$$

et en utilisant toutes les relations entre les restes, obtenir une relation de Bézout entre r_0 et r_1, comme dans l'exemple suivant.

0.5.7. Exemple.— Les polynômes x^4+1 et x^3+1 sont premiers entre eux. On calcule la division euclidienne de x^4+1 par x^3+1 :

$$x^4+1 = (x^3+1)(x) + (-x+1),$$

on calcule alors la division euclidienne de x^3+1 par $-x+1$:

$$x^3+1 = (-x+1)(-x^2-x-1) + 2.$$

Le dernier reste non nul est 2, on a alors

$$\begin{aligned}
2 &= (x^3+1) - (-x+1)(-x^2-x-1),\\
&= (x^3+1) - ((x^4+1) - (x^3+1)(x))(-x^2-x-1),\\
&= (x^3+1) - (x^4+1)(-x^2-x-1) + (x^3+1)x(-x^2-x-1),\\
&= (x^3+1)(1-x-x^2-x^3) + (x^4+1)(1+x+x^2).
\end{aligned}$$

On obtient ainsi une relation de Bézout :

$$1 = \frac{1}{2}(1-x-x^2-x^3)(x^3+1) + \frac{1}{2}(1+x+x^2)(x^4+1).$$

Exercice 26.— Trouver une relation de Bézout entre les polynômes f_1 et f_2, avec

1. $f_1 = x^2+2x-1, \quad f_2 = x+2,$
2. $f_1 = x^4+2x^3-x, \quad f_2 = x^3+5,$
3. $f_1 = x^2+2x-1, \quad f_2 = x+2.$

Exercice 27 (Lemme de Gauss).— Soient f, g, h des polynômes de $\mathbb{K}[x]$. Montrer que si f et g sont premiers entre eux et que f divise gh, alors f divise h.

FIGURE 0.3.: Étienne Bézout (1730 - 1783)

Étienne Bézout est un mathématicien français, auteur d'une Théorie générale des équations algébriques sur la théorie de l'élimination et des fonctions symétriques sur les racines d'une équation. Examinateur des élèves du corps de l'artillerie, il rédige un cours de mathématiques à l'usage de la marine et de l'artillerie, qui deviendra un ouvrage de référence pour les candidats au concours d'entrée à l'École polytechnique.

0.5.8. Racine d'un polynôme.— Soit f un polynôme de $\mathbb{K}[x]$. Un scalaire $a \in \mathbb{K}$ est dit *racine* de f si $f(a) = 0$, c'est-à-dire, lorsque a est un zéro de la fonction polynomiale $\widetilde{f}$. On peut dire aussi que a est racine de f si, et seulement si, $x - a$ divise f.

Si $a_1, \ldots, a_p$ sont p racines distinctes de f, alors f est divisible par le polynôme

$$(x - a_1) \ldots (x - a_p).$$

Un polynôme non nul f de degré n admet au plus n racines distinctes. Si f admet n racines distinctes $a_1, \ldots, a_n$, alors, il se décompose sous la forme

$$f = \mathrm{lc}(f)(x - a_1) \ldots (x - a_n).$$

0.5.9. Racines multiples.— Soient f un polynôme de $\mathbb{K}[x]$ et a une racine de f. On appelle *ordre de multiplicité* de la racine a l'exposant de la plus grande puissance de $x - a$ qui divise f. Autrement dit, c'est l'entier h tel que $(x-a)^h$ divise f et $(x-a)^{h+1}$ ne divise pas f. Soit f un polynôme tel que

$$f = (x-a)^h q.$$

Alors a est racine d'ordre de multiplicité h si, et seulement si, a n'est pas racine du polynôme q.

0.5.10. Polynômes scindés.— Soit $f \in \mathbb{K}[x]$ un polynôme. On dit que f est *scindé* sur $\mathbb{K}$ s'il admet des racines $a_1, \ldots, a_p$ dans $\mathbb{K}$ d'ordre de multiplicité respectifs $m_1, \ldots, m_p$

telles que $m_1 + \ldots + m_p = \deg f$. On a alors

$$f = \mathrm{lc}(f)(x-a_1)^{m_1} \ldots (x-a_p)^{m_p},$$

avec $a_i \neq a_j$ si $i \neq j$ et $m_1 + \cdots + m_p = \deg f$.

0.5 Théorème (de D'Alembert-Gauss).— Tout polynôme non constant à coefficients dans $\mathbb{C}$ possède au moins une racine dans $\mathbb{C}$.

Ce théorème est appelé aussi théorème de D'Alembert-Gauss, ou encore *théorème fondamental de l'algèbre*.

0.6 Corollaire.— Tout polynôme non nul de $\mathbb{C}[x]$ est scindé sur $\mathbb{C}$.

On dit qu'un corps $\mathbb{K}$ est *algébriquement clos*, si tout polynôme non constant de $\mathbb{K}[x]$ possède une racine dans $\mathbb{K}$.

Exercice 28.— Montrer que le corps $\mathbb{R}$ n'est pas algébriquement clos.

Le théorème fondamental de l'algèbre entraîne que le corps $\mathbb{C}$ est algébriquement clos. Plus généralement, on peut montrer que, pour tout corps $\mathbb{K}$, il existe une extension $\mathbb{L}$ telle que tout polynôme de $\mathbb{K}[x]$ soit scindé sur $\mathbb{L}$. Le corps $\mathbb{L}$ est appelé la *clôture algébrique* de $\mathbb{K}$. Par exemple $\mathbb{C}$ est la clôture algébrique de $\mathbb{R}$.

FIGURE 0.4.: Jean Le Rond D'Alembert (1717 - 1783)

Jean Le Rond D'Alembert est un mathématicien, philosophe et encyclopédiste français. Il énonce le théorème de D'Alembert-Gauss dans le Traité de dynamique, qui ne sera démontré qu'un siècle après par Carl Friedrich Gauss.

D'Alembert est célèbre pour ses nombreux travaux mathématiques, notamment sur les équations différentielles et les équations aux dérivées partielles. Il aborda des problèmes difficiles en physique avec le Traité de dynamique des systèmes, en astronomie avec le problème des trois corps, ou encore en musique avec la vibration des cordes.

§ 6 Les fractions rationnelles

Nous connaissons la construction du corps $\mathbb{Q}$ des nombres rationnels à partir de l'anneau des entiers relatifs $\mathbb{Z}$. Le corps des fractions rationnelles $\mathbb{K}(x)$ se construit de façon similaire à partir de l'anneau des polynômes $\mathbb{K}[x]$.

0.6.1. Le corps des fractions rationnelles.— Soit $\mathscr{E}$ l'ensemble $\mathbb{K}[x]\times(\mathbb{K}[x]\backslash\{0\})$. On définit sur $\mathscr{E}$ une relation $\approx$ définie par, pour tous $(f,g),(f',g')\in\mathscr{E}$,

$$(f,g)\approx(f',g') \quad \text{si, et seulement si,} \quad fg'=f'g.$$

La relation $\approx$ forme une relation d'équivalence sur $\mathscr{E}$. L'ensemble $\mathscr{E}/\approx$ des classes d'équivalence est noté $\mathbb{K}(x)$. Un élément de $\mathbb{K}(x)$ représenté par $(f,g)\in\mathscr{E}$ est noté $\dfrac{f}{g}$. et appelé *fraction rationnelle*.

On définit sur $\mathbb{K}(x)$ une addition et une multiplication définies de la façon suivante, pour toutes fractions rationnelles $\dfrac{f}{g}$ et $\dfrac{f'}{g'}$,

$$\frac{f}{g}+\frac{f'}{g'}=\frac{fg'+f'g}{gg'}, \qquad \frac{f}{g}.\frac{f'}{g'}=\frac{ff'}{gg'}.$$

0.7 Proposition.— Ces opérations sont bien définies et munissent l'ensemble $\mathbb{K}(x)$ des fractions rationnelles d'une structure de corps.

Le corps $\mathbb{K}(x)$ est appelé *corps des fractions rationnelles*.

Exercice 29.— Montrer que la relation $\approx$ est une relation d'équivalence.

Exercice 30.— Montrer la proposition 0.7.

On appelle *représentant irréductibe* d'une fraction rationnelle F, tout représentant (f,g) de F où les polynômes f et g sont premiers entre eux.

On appelle *degré* d'une fraction rationnelle F, la quantité $\deg f-\deg g\in\mathbb{Z}\cup\{-\infty\}$, où (f,g) est un représentant de F. Cette quantité ne dépend pas du représentant et est notée $\deg F$. On a $\deg 0=-\infty$.

Soit F une fraction rationnelle de forme irréductible $\frac{f}{g}$. On appelle *racine* de F toute racine de f. On appelle *pôle* de F toute racine de g. L'*ordre de multiplicité* d'une racine ou d'un pôle est l'ordre de multiplicité de la racine dans le polynôme correspondant.

0.6.2. Décomposition en éléments simples.—

0.8 Proposition.— Toute fraction rationnelle F s'écrit de façon unique comme la somme d'un polynôme, appelé *partie entière* de F, et d'une fraction rationnelle de degré strictement négatif.

On considère la fraction rationnelle $\frac{f}{g}$. En notant E le quotient et R le reste de la division euclidienne de f par g, alors

$$\frac{f}{g} = E + \frac{r}{q}, \quad \text{avec } \deg r < \deg q.$$

Par exemple, on a

$$\frac{x^5+1}{x(x-1)^2} = x^2+2x+3+\frac{4x^2-3x+1}{x(x-1)^2}.$$

Le polynôme x^2+2x+3 est la partie entière de la fraction rationnelle $\frac{x^5+1}{x(x-1)^2}$.

0.9 Proposition.— Soit F un fraction rationnelle admettant un pôle λ d'ordre k. Il existe un unique k-uplet de scalaires $(\beta_1,\ldots,\beta_k)$ et une unique fraction F_0 n'admettant pas λ pour pôle, tels que :

$$F = F_0 + \sum_{i=1}^{k} \frac{\beta_i}{(x-\lambda)^i}.$$

0.6.3. Exemple.— La fraction $\frac{x^5+1}{x(x-1)^2}$ admet pour pôles 0 et 1 d'ordre 1 et 2 respectivement. On a une *décomposition en éléments simples* :

$$\frac{x^5+1}{x(x-1)^2} = x^2+2x+3+\frac{1}{x}+\frac{2}{(x-1)^2}+\frac{3}{x-1}.$$

0.10 Proposition (décomposition en éléments simples dans $\mathbb{C}(x)$).— Soit F une fraction rationnelle de $\mathbb{C}(x)$ de pôles $\lambda_1,\ldots,\lambda_p$ distincts deux à deux et d'ordre de multiplicité respectifs $k_1, \dots, k_p$. Il existe un polynôme E de $\mathbb{C}[x]$ et une unique famille de scalaires $(\beta_{i,j})$ tels que :

$$F = E + \sum_{i=1}^{p} \sum_{j=1}^{k_i} \frac{\beta_{i,j}}{(x-\lambda_i)^j}.$$

La décomposition précédente est appelée la *décomposition en éléments simples* dans $\mathbb{C}(x)$ de la fraction rationnelle F.

Exercice 31.— Déterminer une identité de Bézout entre les polynômes suivants

1. $x-1$ et $x+2$,
2. $x+1$ et x^2-2x+1,
3. x^4-1 et x^3+2.

Exercice 32.— Décomposer en éléments simples dans $\mathbb{C}(x)$ les fractions suivantes

1. $\dfrac{10x^3}{(x^2+1)(x^2-4)}$,
2. $\dfrac{x^3-1}{(x-1)(x-2)(x-3)}$,
3. $\dfrac{x^2}{(x^2+x+1)^2}$,
4. $\dfrac{(x^2+4)^2}{(x^2+1)(x^2-2)^2}$.

Première partie .

Les matrices et abrégé d'algèbre linéaire

CHAPITRE
1

Les espaces vectoriels

Sommaire

Nous rappelons dans ce chapitre les notions d'algèbre linéaire abordées en première année. Pour une bonne compréhension des prochains chapitres, il est important que ces notions soient bien assimilées. Ce chapitre constitue un recueil de résultats sans démonstration. Le lecteur est renvoyé à son cours de première année, ou aux ouvrages suivants pour tout approfondissement de François Liret et Dominique Martinais [LM03], ainsi que de celui de Joseph Grifone [Gri11] pour tout approfondissement.

Ces deux références proposent un cours complété d'exercices avec solutions, la seconde référence couvre une partie des notions abordées dans ce cours.

§ 1 La structure d'espace vectoriel

1.1.1. Un prototype.— L'espace vectorie réel $\mathbb{R}^n$ est le prototype d'espace vectoriel réel de dimension finie. Ses éléments sont les n-uplets ordonnées de nombres réels $(x_1, x_2, \ldots, x_n)$, que nous écrirons sous la forme de *vecteur colonne* :

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

Les opérations d'addition et de multiplication de nombre réels permettent de définir deux opérations sur $\mathbb{R}^n$, l'addition de vecteurs :

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{bmatrix}$$

et la multiplication d'un vecteur par un nombre réel, pour tout réel λ,

$$\lambda \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \lambda x_1 \\ \lambda x_2 \\ \vdots \\ \lambda x_n \end{bmatrix}.$$

Ces deux opérations ont des propriétés héritées de celles satisfaites par les opérations d'addition et de multiplication sur les nombres réels. On peut vérifier ainsi que, l'addition de n-uplets satisfait les relations suivantes :

i) pour tous $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n$, on a $(\mathbf{x}+\mathbf{y})+\mathbf{z} = \mathbf{x}+(\mathbf{y}+\mathbf{z})$,

ii) pour tous $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, on a $\mathbf{x}+\mathbf{y} = \mathbf{y}+\mathbf{x}$,

iii) si $\mathbf{0}$ désigne le n-uplet $(0,0,\ldots,0)$, on a $\mathbf{x}+\mathbf{0} = \mathbf{x}$,

iv) en notant $-(x_1,x_2,\ldots,x_n)$, le n-uplet $(-x_1,-x_2,\ldots,-x_n)$, on a $\mathbf{x}+(-\mathbf{x}) = \mathbf{0}$.

D'autre part, la mupliplication par un nombre réel satisfait les relations suivantes :

v) pour tous $\lambda, \mu \in \mathbb{R}$ et tout $\mathbf{x} \in \mathbb{R}^n$, on a $\lambda(\mu\mathbf{x}) = (\lambda\mu)\mathbf{x}$,

vi) pour tout $\mathbf{x} \in \mathbb{R}^n$, on a $1\mathbf{x} = \mathbf{x}$.

Enfin, les opérations d'addition et de multiplication par un réel vérifient des relations de compatibilités entre elles :

vii) pour tout $\lambda \in \mathbb{R}$ et tous $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, on a $\lambda(\mathbf{x}+\mathbf{y}) = \lambda\mathbf{x}+\lambda\mathbf{y}$,

viii) pour tous $\lambda, \mu \in \mathbb{R}$ et tout $\mathbf{x} \in \mathbb{R}^n$, on a $(\lambda+\mu)\mathbf{x} = \lambda\mathbf{x}+\mu\mathbf{x}$.

1.1.2. D'autres exemples.— Il existe d'autres ensembles qui, munis d'une opération d'addition et de multiplication par un réel, satisfont les mêmes relations.

1) Considérons l'ensemble $\mathscr{C}([0,1],\mathbb{R})$ des fonctions continues sur l'intervale réel $[0,1]$ et à valeurs réelles. Si f et g sont deux telles fonctions, on définit leur addition en posant, pour tout réel $x \in [0,1]$,

$$(f+g)(x) = f(x)+g(x),$$

et la multiplication d'une fonction f par un réel λ, en posant

$$(\lambda f)(x) = \lambda f(x).$$

Ces opérations ainsi définies statisfont les propriétés **i)-viii)** ci-dessus, en prenant pour zéro dans **iii)** la fonction nulle et en notant dans **iv)** $-f$, la fonction définie par

$$(-f)(x) = -f(x).$$

2) On vérifiera que les mêmes opérations sur l'ensemble des fonctions deux fois différentiables qui satisfont

$$\frac{d^2f}{dt^2} + \lambda\frac{df}{dt} + \mu f = 0, \qquad \lambda, \mu \in \mathbb{R},$$

satisfont les propriétés **i)-viii)**.

3) De la même façon, les suites réelles $(x_n)_{n\in\mathbb{N}}$, satisfaisant la relation de récurrence

$$x_{n+2} = \lambda x_{n+1} + \mu x_n, \qquad \lambda, \mu \in \mathbb{R},$$

satisfont les propriétés **i)-viii)**.

En remplaçant le corps des réels par un corps quelconque $\mathbb{K}$, on obtient la struture d'espace vectoriel sur un corps $\mathbb{K}$, tel que $\mathbb{K}$ lui-même ou bien $\mathbb{K}^n$.

1.1.3. Définition.— Un *espace vectoriel* sur un corps $\mathbb{K}$, ou $\mathbb{K}$-espace vectoriel, est un ensemble E muni d'une addition $+ : E \times E \longrightarrow E$ et d'une application

$$\begin{aligned} . : \mathbb{K} \times E &\longrightarrow E \\ (\lambda, \mathbf{x}) &\longmapsto \lambda.\mathbf{x} \end{aligned}$$

appelée loi externe, on notera aussi $\lambda\mathbf{x}$ pour $\lambda.\mathbf{x}$, qui vérifient les propriétés suivantes,

i) l'ensemble E muni de l'addition est un groupe abélien,

ii) pour tout $\mathbf{x} \in E$, $1\mathbf{x} = \mathbf{x}$,

iii) pour tous $\lambda, \mu \in \mathbb{K}$ et tous $\mathbf{x}, \mathbf{y} \in E$:

$$\begin{aligned} \lambda(\mu\mathbf{x}) &= (\lambda\mu)\mathbf{x}, \\ (\lambda+\mu)\mathbf{x} &= \lambda\mathbf{x} + \mu\mathbf{x}, \\ \lambda(\mathbf{x}+\mathbf{y}) &= \lambda\mathbf{x} + \lambda\mathbf{y}. \end{aligned}$$

Les éléments d'un espace vectoriel E sont appelés *vecteurs* et les éléments du corps $\mathbb{K}$ sont appelés *scalaires*.

1.1.4. L'espace nul.— Le $\mathbb{K}$-espace vectoriel qui comporte un seul élément, l'élément nul $\mathbf{0}$, est appelé l'*espace vectoriel nul*.

Exercice 1.— Soit E un $\mathbb{K}$-espace vectoriel. Montrer que, pour tout vecteur $\mathbf{x}$ et scalaire λ, on a les relations :

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

b) $\lambda\mathbf{0} = \mathbf{0}$,

c) $0\mathbf{x} = \mathbf{0}$,

d) $\lambda\mathbf{x} = \mathbf{0}$ si, et seulement si, $\lambda = 0$ ou $\mathbf{x} = \mathbf{0}$,

e) $-(\lambda\mathbf{x}) = \lambda(-\mathbf{x}) = (-\lambda)\mathbf{x}$.

1.1.5. Combinaisons linéaires.— Soit E un $\mathbb{K}$-espace vectoriel. Soient $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_p$ des vecteurs de E. On appelle *combinaison linéaire* des vecteurs $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_p$ tout vecteur de la forme

$$\lambda_1\mathbf{x}_1 + \lambda_2\mathbf{x}_2 + \ldots + \lambda_p\mathbf{x}_p,$$

où $\lambda_1, \ldots, \lambda_p \in \mathbb{K}$.

Plus généralement, si $(\mathbf{x}_i)_{i\in I}$ est une famille de vecteurs de E, on appele *combinaison linéaire* des $(\mathbf{x}_i)_{i\in I}$ toute somme

$$\sum_{i\in I} \lambda_i\mathbf{x}_i,$$

dans laquelle, pour tout $i \in I$, $\lambda_i \in \mathbb{K}$ et où les λ_i sont tous nuls sauf un nombre fini.

1.1.6. Sous-espaces vectoriels.— Soit E un $\mathbb{K}$-espace vectoriel. Un *sous-espace vectoriel* de E est un sous ensemble F de E tel que les opérations de E induisent sur F une structure de $\mathbb{K}$-espace vectoriel.

Un sous-ensemble F de E est donc un sous-espace vectoriel si les assertions suivantes sont vérifiées :

i) $\mathbf{0} \in F$,

ii) pour tous $\mathbf{x}, \mathbf{y} \in F$, $\mathbf{x} + \mathbf{y} \in F$,

iii) pour tout $\mathbf{x} \in F$ et tout $\lambda \in \mathbb{K}$, $\lambda\mathbf{x} \in F$.

1.1.7. Exemples.—

a) Le sous-ensemble $\{\mathbf{0}\}$ réduit au vecteur nul et E sont des sous-espaces vectoriel de E.

b) Soit $\mathbb{K}[x]$ l'ensemble des polynômes à coefficients dans $\mathbb{K}$. Le sous-ensemble $\mathbb{K}_n[x]$ de $\mathbb{K}[x]$ formé des polynômes de degré au plus n est un sous-espace vectoriel de $\mathbb{K}[x]$.

Exercice 2.— Montrer qu'une intersection de sous-espaces vectoriels de E est un sous-espace vectoriel de E.

1.1.8. Faits.— Un sous-ensemble F de E est un sous-espace vectoriel si, et seulement si, toute combinaison linéaire de vecteurs de F est un vecteur de F. Pour montrer qu'un sous-ensemble F de E est un sous-espace vectoriel, il suffit alors de montrer que $\mathbf{0} \in F$ et que $\lambda\mathbf{x} + \mu\mathbf{y}$, pour tous vecteurs $\mathbf{x}, \mathbf{y} \in F$ et scalaires $\lambda, \mu \in \mathbb{K}$.

Exercice 3.— Soit α un réel. On considère le sous-ensemble suivant de $\mathbb{R}^3$:

$$F_\alpha = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_1 + x_2 + x_3 = \alpha\}$$

Montrer que F_α est un sous-espace vectoriel de $\mathbb{R}^3$, si et seulement si, $\alpha = 0$.

§ 2 Bases et dimension d'un espace vectoriel

1.2.1. Sous-espace vectoriel engendré par une partie.— Une intersection de sous-espaces vectoriels de E est un sous-espace vectoriel de E.

Soient E un $\mathbb{K}$-espace vectoriel et A une partie de E. L'intersection de tous les sous-espaces vectoriels de E qui contiennent A est le plus petit sous-espace vectoriel de E, au sens de l'inclusion, contenant A. On l'appelle *sous-espace vectoriel engendré* par A et on le note $\mathrm{Vect}(A)$.

Le sous-espace vectoriel $\mathrm{Vect}(A)$ est formé de toutes les combinaisons linéaires de vecteurs de A.

1.2.2. Exemples.— Le sous-espace vectoriel engendré par l'ensemble vide est l'espace vectoriel nul :

$$\mathrm{Vect}(\{\mathbf{0}\}) = \{\mathbf{0}\}.$$

Pour tous vecteurs $\mathbf{x}, \mathbf{y} \in E$,

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

Exercice 4.—

1. Montrer que l'ensemble $A = \{(1,0),(0,1)\}$ engendre $\mathbb{R}^2$.
2. Montrer que l'ensemble $A = \{(x,x^2,x^3) \mid x \in \mathbb{R}\}$ n'engendre pas $\mathbb{R}^3$.

1.2.3. Famille génératrice.— Une famille $(\mathbf{x}_i)_{i\in I}$ de vecteurs d'un $\mathbb{K}$-espace vectoriel E est dite *génératrice* si $E = \mathrm{Vect}((\mathbf{x}_i)_{i\in I})$.

1.2.4. Famille libre.— Une famille $(\mathbf{x}_i)_{i\in I}$ de vecteurs d'un $\mathbb{K}$-espace vectoriel E est dite *libre*, ou *linéairement indépendante*, si pour toute combinaison linéaire vérifiant

$$\sum_{i\in I} \lambda_i \mathbf{x}_i = \mathbf{0},$$

alors, pour tout $i \in I$, $\lambda_i = 0$. Autrement dit, aucun des vecteurs de la famille $(\mathbf{x}_i)_{i\in I}$ n'est combinaison linéaire des autres. Un famille qui n'est pas libre est dite *liée*.

Les éléments d'une famille libre sont non nuls et tous distincts. Toute sous-famille d'une famille libre est libre

Exercice 5.— Montrer cette propriété.

1.2.5. Base d'un espace vectoriel.— Une famille libre et génératrice d'un $\mathbb{K}$-espace vectoriel E est appelée un *base* de E. Si $(\mathbf{e}_i)_{i\in I}$ une base de E, tout vecteur $\mathbf{x}$ de E s'écrit de façon unique

$$\mathbf{x} = \sum_{i\in I} \lambda_i \mathbf{e}_i.$$

Les scalaires λ_i s'appellent les *coordonnées* de $\mathbf{x}$ dans la base $(\mathbf{e}_i)_{i\in I}$.

Tout élément non nul de $\mathbb{K}$ forme une base du $\mathbb{K}$-espace vectoriel $\mathbb{K}$. Le couple $(1,i)$ forme une base du $\mathbb{R}$-espace vectoriel $\mathbb{C}$.

1.2.6. Base canonique.— Lorsqu'un espace est muni naturellement d'une base particulière, elle est dite *canonique*. Par exemple, la famille $(x^n)_{n\in\mathbb{N}}$ est la base canonique de l'espace vectoriel $\mathbb{K}[x]$ des polynômes en l'indéterminée x.

La famille $(0,\ldots,1,\ldots,0)$, $1 \leqslant i \leqslant n$, est la base canonique de l'espace vectoriel $\mathbb{K}^n$.

1.2.7. Le concept de dimension.— Un espace vectoriel est dit de *dimension finie*, s'il admet une famille génératrice finie. Dans le cas contraire, on dit qu'il est de *dimension infinie*.

On a le théorème fondamental d'existence de base dans les espaces vectoriels de dimension finie.

1.1 Théorème.— Soit E un $\mathbb{K}$-espace vectoriel non nul de dimension finie. Soient $\mathscr{G}$ une famille génératrice de E et $\mathscr{L} \subset \mathscr{G}$ une famille libre. Alors, il existe une base $\mathscr{B}$ de E vérifiant $\mathscr{L} \subset \mathscr{B} \subset \mathscr{G}$.

De ce résultat on déduit que, pour tout espace vectoriel de dimension finie E

i) de toute famille génératrice de E, on peut extraire une base de E,

ii) (théorème de la base incomplète) toute famille libre peut être complétée en une base.

Enfin, le théorème de la dimension d'un espace vectoriel de dimension finie.

1.2 Théorème.— Dans un $\mathbb{K}$-espace vectoriel E de dimension finie, toutes les bases ont le même nombre d'éléments.

Ce nombre est appelé la *dimension* de E et est noté $\dim_{\mathbb{K}} E$, ou $\dim E$, s'il n'y a pas de confusion.

1.3 Proposition.— Soit E un $\mathbb{K}$-espace vectoriel de dimension finie et F un sous-espace vectoriel de E. Alors F est de dimension finie et $\dim F \leqslant \dim E$. De plus, F est égal à E si, et seulement si, $\dim F = \dim E$.

Par exemple, l'espace $\mathbb{K}^n$ est de dimension n. L'espace $\mathbb{K}_n[x]$ des polynômes à coefficients dans $\mathbb{K}$ et de degré inférieur ou égal à n admet comme base canonique $(1, x, \ldots, x^n)$ et est donc de dimension $n+1$. L'espace vectoriel $\mathbb{K}[x]$ est de dimension infinie.

§ 3 Somme de sous-espaces vectoriels

1.3.1. Somme de sous-espaces.— Soit E un $\mathbb{K}$-espace vectoriel et soient F et G deux sous-espaces vectoriels de E. La réunion de F et G n'est pas en général un sous-espace vectoriel de E

Exercice 6.— Donner des contre-exemples.

Le plus petit sous-espace vectoriel contenant F et G est le sous-espace vectoriel $\mathrm{Vect}(F \cup G)$. Ce sous-espace vectoriel est décrit par

$$\mathrm{Vect}(F \cup G) = \{\mathbf{x} + \mathbf{y} \mid (\mathbf{x}, \mathbf{y}) \in F \times G\}.$$

Ce sous-espace vectoriel est appelé la *somme de F et G*, on le note $F + G$.

Plus généralement, si $(E_i)_{i \in I}$ une famille de sous-espaces vectoriel de E, on note $\sum\limits_{i \in I} E_i$ le sous-espace vectoriel de E formé des combinaisons linéaires

$$\sum_{i \in I} \mathbf{x}_i,$$

où les $\mathbf{x}_i \in E_i$, $i \in I$, sont tous nuls sauf une nombre fini. L'espace $\sum\limits_{i \in I} E_i$ est appelé la *somme* des sous-espaces vectoriels E_i, $i \in I$.

Exercice 7.— Soient E un espace vectoriel et F, G deux sous-espaces vectoriels de E.

1. Montrer que $F \cap G = F + G$ si, et seulement si, $F = G$.
2. Montrer que la réunion $F \cup G$ est un sous-espace vectoriel de E si, et seulement si, $F \subset G$ ou $G \subset F$.

1.3.2. Somme de sous-espaces supplémentaires.— Soient E un $\mathbb{K}$-espace vectoriel et F et G deux sous-espaces vectoriels de E. Les deux assertions suivantes sont équivalentes :

i) pour tout vecteur $\mathbf{x} \in E$, il existe un unique couple $(\mathbf{y}, \mathbf{z}) \in F \times G$ tel que

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

ii) E est égal à la somme $F + G$ des sous-espaces F et G et $F \cap G = \{\mathbf{0}\}$.

On dit que F et G sont deux *sous-espaces vectoriels supplémentaires* de E lorsqu'ils satisfont ces propriétés. On dit aussi que E est la *somme directe* des sous-espaces F et G et on écrit :

$$E = F \oplus G.$$

Exercice 8.— Montrer l'équivalence des assertions **i)** et **ii)**.

En résumé :

1.4 Proposition.— Soient F et G deux sous-espaces vectoriels de E. Alors $E = F \oplus G$ si, et seulement si, les deux propriétés suivantes sont vérifiées :

i) $E = F + G$,

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

1.5 Proposition.— Tout sous-espace vectoriel d'un espace vectoriel de dimension finie admet au moins un supplémentaire.

Exercice 9.— On considère les trois sous-espaces vectoriels suivants de $\mathbb{R}^2$:

$$F = \mathbb{R} \times \{\mathbf{0}\}, \qquad G = \{\mathbf{0}\} \times \mathbb{R}, \qquad H = \{(\mathbf{x}, \mathbf{x}) \mid \mathbf{x} \in \mathbb{R}\}.$$

Montrer que $F \cap G = F \cap H = G \cap H = \{\mathbf{0}\}$ et que la somme $F + G + H$ n'est pas directe.

1.3.3. Somme directe d'une famille de sous-espaces.— Plus généralement, soient $E_1, \ldots, E_p$ des sous-espaces vectoriel de E. On dit que E est *somme directe* des sous-espaces vectoriels $E_1, \ldots, E_p$, si tout vecteur $\mathbf{x}$ de E s'écrit de façon unique sous la forme

$$\mathbf{x} = \mathbf{x}_1 + \ldots + \mathbf{x}_p,$$

où, pour tout $i \in [\![1, p]\!]$, $\mathbf{x}_i \in E_i$. On note alors $E = E_1 \oplus \ldots \oplus E_p$.

Pour montrer que E se décompose en la somme directe de p espaces vectoriels :

$$E = E_1 \oplus \ldots \oplus E_p,$$

il suffit de montrer que $E = E_1 + \ldots + E_p$ et que pour tout $i \in [\![2, p]\!]$, on a

$$E_i \cap (E_1 + \ldots + E_{i-1}) = \{\mathbf{0}\}.$$

En pratique, on peut aussi utiliser la caractérisation suivante. Les sous-espaces $E_1, \ldots, E_p$ sont en somme directe si, et seulement si, l'égalité

$$\mathbf{x}_1 + \ldots + \mathbf{x}_p = \mathbf{0}, \quad \text{où} \quad \mathbf{x}_i \in E_i,$$

implique que $\mathbf{x}_i = \mathbf{0}$, pour tout $i \in [\![1, p]\!]$.

1.6 Proposition.— Soit E un $\mathbb{K}$-espace vectoriel de dimension finie et $E_1, \ldots, E_p$ des sous-espaces de E de bases respectives $\mathscr{B}_1, \ldots, \mathscr{B}_p$. Alors $E = E_1 \oplus \ldots \oplus E_p$ si, et seulement si, la réunion $\mathscr{B}_1 \cup \ldots \cup \mathscr{B}_p$ est une base de E.

Considérons le $\mathbb{R}$-espace vectoriel $\mathbb{C}$ des nombres complexes. L'ensemble $\mathbb{R}$ des nombres réels et l'ensemble $i\mathbb{R}$ des nombres imaginaires purs, auquel on ajoute 0, sont deux sous-espaces vectoriel de $\mathbb{C}$. On a $\mathbb{C} = \mathbb{R} \oplus i\mathbb{R}$.

Exercice 10.— Montrer que l'ensemble $\mathscr{F}(\mathbb{R},\mathbb{R})$ des fonctions réelles à valeurs réelles se décompose en la somme directe de l'ensemble des fonctions paires et de l'ensemble des fonctions impaires.

1.7 Proposition (une formule de Grassmann).— Soient F et G deux sous-espaces vectoriels de E. On a

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

On en déduit que si $E_1, \ldots, E_p$ sont des sous-espaces vectoriels de E, alors

$$\dim(E_1 \oplus \ldots \oplus E_p) = \dim E_1 + \ldots + \dim E_p.$$

§ 4 Les applications linéaires

1.4.1. Définition.— Soient E et E' deux $\mathbb{K}$-espaces vectoriels. Une application

$$u : E \longrightarrow E'$$

est dite *linéaire* si, pour tous vecteurs $\mathbf{x}, \mathbf{y} \in E$ et scalaires $\lambda, \mu \in \mathbb{K}$,

$$u(\lambda \mathbf{x} + \mu \mathbf{y}) = \lambda u(\mathbf{x}) + \mu u(\mathbf{y}).$$

Il découle de cette définition que $u(\mathbf{0}) = \mathbf{0}$.

1.4.2. Espace vectoriel des applications linéaires.— On notera $\mathscr{L}(E,E')$ l'ensemble des applications linéaires de E dans E'. Si u et v sont deux applications linéaires de E dans E' et $\lambda \in \mathbb{K}$ un scalaire, alors les applications $u+v$ et λu définies, pour tout $\mathbf{x} \in E$, par

$$\begin{aligned}(u+v)(\mathbf{x}) &= u(\mathbf{x}) + v(\mathbf{x}),\\ (\lambda u)(\mathbf{x}) &= \lambda u(\mathbf{x}),\end{aligned}$$

sont aussi des applications linéaires de E dans E'. L'addition et la multplication par un scalaire sur l'espace vectoriel E' induisent ainsi une structure de $\mathbb{K}$-espace vectoriel sur $\mathscr{L}(E,E')$.

Une application linéaire de E dans E est appelée un *endomorphisme* de E. On notera $\mathscr{L}(E)$ l'espace vectoriel des endomorphismes de E.

Une application linéaire de E dans $\mathbb{K}$ est appelée une *forme linéaire*.

Une application linéaire et bijective est appelée un *isomorphisme*. Les endomorphismes bijectifs de E sont encore appelés les *automorphismes* de E. Les automorphismes de E, muni de la composition, forment un groupe, noté $\mathrm{GL}(E)$ et appelé le *groupe linéaire* de E.

1.4.3. Exemples.—

1) L'homothétie $h : E \longrightarrow E$, de rapport λ, définie par $h(\mathbf{x}) = \lambda\mathbf{x}$ est un endomorphisme de E.

2) Soit $E = \mathscr{C}([a,b],\mathbb{R})$ l'espace vectoriel réel formé des fonctions continues de l'intervalle réel $[a,b]$ à valeurs réelles. L'application $\varphi : E \longrightarrow \mathbb{R}$ définie par

$$\varphi(f) = \int_a^b f(t)dt$$

est une forme linéaire.

3) Soit p l'application de $\mathbb{R}^3$ dans $\mathbb{R}^3$ qui envoie un vecteur sur sa projection orthognonale sur le plan xy :

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} \longmapsto \begin{bmatrix} x \\ y \\ 0 \end{bmatrix}$$

est une application linéaire.

1.4.4. Endomorphismes nilpotents.— Un endomorphisme u de $\mathbb{K}^n$ est dit *nilpotent* s'il existe un entier naturel q tel que $u^q = 0$. Le plus petit entier non nul r tel que $u^r = 0$ est appelé l'*indice de nilpotence* de u.

1.4.5. Image d'une application linéaire.— Soient $u : E \longrightarrow E'$ une application linéaire et F un sous-espace vectoriel de E. Alors l'image directe par u du sous-espace F, donnée par

$$u(F) = \{\mathbf{y} \in E' \mid \exists \mathbf{x} \in F,\ \mathbf{y} = u(\mathbf{x})\}$$

est un sous-espace vectoriel de E'

Exercice 11.— Montrer que $u(F)$ est un sous-espace vectoriel.

On appelle *image* de u, et on note $\mathrm{Im}\, u$, le sous-espace vectoriel $u(E)$.

L'application u est surjective si, et seulement si, $\mathrm{Im}\, u = E'$.

1.4.6. Sous-espace vectoriel stable.— Soient u un endomorphisme de E et F un sous-espace vectoriel de E. On dit que u laisse *stable* F si $u(F)$ est un sous-espace vectoriel de F.

Par exemple, étant donné $\mathbf{x} \in E$, un endomorphisme u de E laisse stable la droite vectorielle $\mathrm{Vect}(\mathbf{x})$ si, et seulement si, les vecteurs $u(\mathbf{x})$ et $\mathbf{x}$ sont colinéaires. Autrement dit, si u est une homothétie.

1.4.7. Noyau d'une application linéaire.— Soient $u : E \longrightarrow E'$ une application linéaire et F' un sous-espace vectoriel de E'. Alors, l'image réciproque par u de F'

$$u^{-1}(F') = \{\mathbf{x} \in E \mid u(\mathbf{x}) \in F'\}$$

est un sous-espace vectoriel de E.

Exercice 12.— Montrer que $u^{-1}(F')$ est un sous-espace vectoriel de E.

On appelle *noyau* de u, et on note $\operatorname{Ker} u$, le sous-espace vectoriel $u^{-1}(\mathbf{0})$.

1.8 Proposition.— Une application linéaire u est injective si, et seulement si, $\operatorname{Ker} u = \{\mathbf{0}\}$.

Exercice 13.— Montrer la proposition 1.8.

1.4.8. Rang d'une application linéaire.— Soit E un $\mathbb{K}$-espace vectoriel. Le rang d'une famille finie $\mathscr{V}$ de vecteurs de E est la dimension du sous-espace vectoriel $\operatorname{Vect}(\mathscr{V})$ engendré par les vecteurs de $\mathscr{V}$, on le note $\operatorname{rang} \mathscr{V}$.

Par exemple, si

$$\mathscr{V} = \left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$$

on a $\operatorname{rang} \mathscr{V} = 2$.

Soit $u : E \longrightarrow E'$ une application linéaire. L'application u est dite de rang fini si le sous-espace vectoriel $\operatorname{Im} u$ est de dimension finie. L'entier $\dim(\operatorname{Im} u)$ est alors appelé le *rang* de u et est noté $\operatorname{rang} u$.

1.4.9. Le théorème du rang.— Étant donnée une application linéaire $u : E \longrightarrow E'$, on considère les ensembles $\operatorname{Ker} u$ et $\operatorname{Im} u$ qui sont respectivement sous-espaces vectoriels de E et E'. Soit $\mathscr{B} = (\mathbf{e}_1, \ldots, \mathbf{e}_k)$ une base de $\operatorname{Ker} u$. D'après le théorème de la base incomplète, on peut compléter cette base en une base $(\mathbf{e}_1, \ldots, \mathbf{e}_k, \mathbf{e}_{k+1}, \ldots, \mathbf{e}_n)$ de E. Alors la famille $(u(\mathbf{e}_{k+1}), \ldots, u(\mathbf{e}_n))$ est une base du sous-espace $\operatorname{Im} u$, cf. exercice 14. On en déduit une formule très utile reliant les dimensions des sous-espaces $\operatorname{Ker} u$, $\operatorname{Im} u$ avec celle de E :

1.9 Théorème (La formule du rang).— Soient E un espace vectoriel de dimension finie et $u : E \longrightarrow E'$ une application linéaire à valeurs dans un espace vectoriel E'. Alors

$$\dim E = \dim(\operatorname{Ker} u) + \operatorname{rang} u.$$

Exercice 14.— Soient E et E' des espaces vectoriels et $u : E \longrightarrow E'$ une application linéaire. On suppose que E est de dimension finie. D'après la proposition 1.5, il existe alors un supplémentaire F du sous-espace $\mathrm{Ker}\, u$ dans E.

Montrer que la restriction de l'application u à F

$$\begin{aligned} u_{|_F} : F &\longrightarrow \mathrm{Im}\, u \\ \mathbf{x} &\longmapsto u(\mathbf{x}) \end{aligned}$$

est un isomorphisme de F sur $\mathrm{Im}\, u$.

On déduit du théorème 1.9, le résultat suivant :

1.10 Proposition.— Soit $u : E \longrightarrow E'$ une application linéaire entre deux $\mathbb{K}$-espaces vectoriel de même dimension finie. Les assertions suivantes sont équivalentes :

i) u est injective,

ii) u est surjective,

iii) u est bijective

iv) $\mathrm{Ker}\, u = \{\mathbf{0}\}$,

v) $\mathrm{Im}\, u = E'$,

vi) $\mathrm{rang}\, u = \dim E$.

1.4.10. Remarque.— Attention, ce résultat est faux en dimension infinie. En effet, l'application

$$\begin{aligned} u : \mathbb{R}[x] &\longrightarrow \mathbb{R}[x] \\ f(x) &\longmapsto f'(x) \end{aligned}$$

est un endomorphisme de $\mathbb{R}[x]$. Son noyau est le sous-espace vectoriel de $\mathbb{R}[x]$ formé des polynômes constants. L'endomorphisme u n'est donc pas injectif, alors qu'il est surjective. En effet $\mathrm{Im}\, u = \mathbb{R}[x]$, car, pour tout polynôme f de $\mathbb{R}[x]$, il existe un polynôme g défini par

$$g(x) = \int_0^x f(t)dt$$

tel que $u(g) = f$. Par exemple, si

$$f(x) = \sum_{i=0}^{n} a_i x^i,$$

on pose

$$g(x) = \sum_{i=0}^{n} \frac{a_i}{i+1} x^{i+1}.$$

On a alors $u(g) = f$.

Exercice 15.— Montrer que l'application

$$u : \mathbb{R}[x] \longrightarrow \mathbb{R}[x]$$
$$f(x) \longmapsto xf(x)$$

est un endomorphisme injectif de $\mathbb{R}[x]$, mais pas surjectif.

1.4.11. Les projecteurs.— Soit E un $\mathbb{K}$-espace vectoriel. Un endomorphisme p de E tel que $p^2 = p$ est appelé *projecteur* de E.

Soient F et G deux sous-espaces vectoriels de E tels que $E = F \oplus G$. Alors tout vecteur $\mathbf{x}$ de E s'écrit de façon unique en $\mathbf{x} = \mathbf{y} + \mathbf{z}$, avec $\mathbf{y} \in F$ et $\mathbf{z} \in G$. On considère l'endomorphisme p_F de E défini par $p_F(\mathbf{x}) = \mathbf{y}$. On dit que p_F est la *projection sur F parallèlement à G*. On définit de la même façon, la *projection sur G parallèlement à F*, comme l'endomorphisme p_G de E défini par $p_G(\mathbf{x}) = \mathbf{z}$. Pour tout vecteur $\mathbf{x}$ de E, on a

$$\mathbf{x} = p_F(\mathbf{x}) + p_G(\mathbf{x}).$$

Ainsi définis, les endomorphismes p_F et p_G vérifient les assertions suivantes

i) p_F et p_G sont des projecteurs, i.e., $p_F^2 = p_F$ et $p_G^2 = p_G$,
ii) $\mathrm{Im}\, p_F = F$ et $\mathrm{Im}\, p_G = F$,
iii) $\mathrm{Ker}\, p_F = G$ et $\mathrm{Ker}\, p_G = F$.
iv) $p_F + p_G = \mathrm{id}_E$,
v) $p_F p_G = p_G p_F = 0$.

L'assertion **i)** découle de la définition de p_F et p_G. Si $\mathbf{x} \in E$, alors $p_F(\mathbf{x})$ est un vecteur de F d'où $p_F(p_F(\mathbf{x})) = p_F(\mathbf{x}) + \mathbf{0}$. On montre de la même façon que $p_G^2 = p_G$. L'assertion **ii)** est une conséquence immédiate de la définition de p_F et p_G.

Pour montrer **iii)**, considérons $\mathbf{x} \in \mathrm{Ker}\, p_F$, alors $\mathbf{x} = p_G(\mathbf{x})$ donc $\mathbf{x} \in G$. Inversement, si $\mathbf{x} \in G$, on a $p_F(\mathbf{x}) = \mathbf{0}$. On montre de la même façon que $\mathrm{Ker}\, p_G = F$.

L'assertion **iv)** correspond à la décomposition $\mathbf{x} = p_F(\mathbf{x}) + p_G(\mathbf{x})$. L'assertion **v)** est obtenue en composant la relation $\mathrm{id}_E = p_F + p_G$ par p_F et p_G.

Inversement, supposons qu'il existe deux applications linéaires

$$p : E \longrightarrow F, \quad q : E \longrightarrow G,$$

telles que $\mathrm{id}_E = p + q$ et $pq = qp = 0$. On montre alors que $E = F \oplus G$. Plus généralement, on a :

1.11 Proposition.— Soit E un $\mathbb{K}$-espace vectoriel et soient $E_1, \ldots, E_k$ des sous-espaces vectoriel de E. Une condition nécessaire et suffisante pour que $E = E_1 \oplus \ldots \oplus E_k$ est qu'il existe des applications linéaires $p_i : E \longrightarrow E_i$, $i \in [\![1, k]\!]$ vérifiant les deux assertions suivantes :

i) $\mathrm{id}_E = p_1 + \ldots + p_k$,
ii) $p_i p_j = 0$, pour tout $i \neq j$.

Pour tout $i \in [\![1,k]\!]$, p_i est appelé la projection sur le sous-espace E_i parallèlement au sous-espace $\underset{\substack{j=1\\j\neq i}}{\overset{k}{\oplus}} E_j$.

Les projection p_i sont des projecteurs, $p_i^2 = p_i$, et satisfont

$$\mathrm{Ker}\ p_i = \bigoplus_{\substack{j=1\\j\neq i}}^{k} E_j, \qquad \mathrm{Im}\ p_i = E_i.$$

Exercice 16.— Établir les deux propriétés précédentes.

Exercice 17.— Soit E un $\mathbb{K}$-espace vectoriel et soient p et q deux projecteurs de E. On suppose que le corps $\mathbb{K}$ n'est pas de caractéristique 2.

1. Montrer que $p+q$ est un projecteur si, et seulement si, $p \circ q = q \circ p = 0$.
2. Supposons que $p+q$ est un projecteur de E. Montrer l'égalité :

$$\mathrm{Im}\ (p+q) = \mathrm{Im}\ (p) + \mathrm{Im}\ (q).$$

3. Supposons que $p+q$ est un projecteur de E. Montrer l'égalité :

$$\mathrm{Ker}\ (p+q) = \mathrm{Ker}\ (p) \cap \mathrm{Ker}\ (q).$$

Exercice 18.— Soit E un $\mathbb{K}$-espace vectoriel de dimension finie n et soient p et q deux endomorphismes de E vérifiant

$$p+q = \mathrm{id}_E, \quad \text{et} \quad \mathrm{rang}\,(p) + \mathrm{rang}\,(q) \leqslant n.$$

1. Montrer que $\mathrm{Im}\ p$ et $\mathrm{Im}\ q$ sont supplémentaires dans E.
2. Montrer que $\mathrm{rang}\,(p) + \mathrm{rang}\,(q) = n$.
3. Montrer que p et q sont des projecteurs.

Exercice 19.— Soit p un projecteur d'un $\mathbb{K}$-espace vectoriel E.

1. Montrer que $E = \mathrm{Ker}\ p \oplus \mathrm{Im}\ p$.
2. La réciproque est-elle vraie ?

Exercice 20.— Soit E un $\mathbb{K}$-espace vectoriel de dimension finie. Montrer que si p est un projecteur de E, alors

$$\mathrm{rang}(p) = \mathrm{trace}(p).$$

§ 5 Exercices

Exercice 21.— Soient E un $\mathbb{K}$-espace vectoriel et u un endomorphisme de E. Montrer que u est une homothétie si, et seulement si, pour tout vecteur $\mathbf{x}$ de E, la famille $(\mathbf{x}, u(\mathbf{x}))$ est liée.

Exercice 22.— Soit E un $\mathbb{K}$-espace vectoriel, pas nécessairement de dimension finie et soit u un endomorphisme de E.

1. Montrer que $\mathrm{Ker}\, u \subseteq \mathrm{Ker}\, u^2$ et $\mathrm{Im}\, u^2 \subseteq \mathrm{Im}\, u$.
2. Montrer que $\mathrm{Ker}\, u = \mathrm{Ker}\, u^2$ si, et seulement si, $\mathrm{Ker}\, u \cap \mathrm{Im}\, u = \{\mathbf{0}\}$.
3. Montrer que $\mathrm{Im}\, u = \mathrm{Im}\, u^2$ si, et seulement si, $E = \mathrm{Ker}\, u + \mathrm{Im}\, u$.

Exercice 23.— Soit E un $\mathbb{K}$-espace vectoriel de dimension finie et soit u un endomorphisme de E. D'après le théorème du rang, on a l'égalité

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

Ce qui n'entraîne pas en général que $E = \mathrm{Ker}\, u \oplus \mathrm{Im}\, u$.

1. Soit u l'endomorphisme de $\mathbb{R}^3$ dont la matrice dans la base canonique est

$$[u]_{can} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

Déterminer $\mathrm{Ker}\, u$, $\mathrm{Ker}\, u^2$, $\mathrm{Im}\, u$, $\mathrm{Im}\, u^2$.

2. Montrer que les assertions suivantes sont équivalentes

- **a)** $E = \mathrm{Ker}\, u \oplus \mathrm{Im}\, u$,
- **b)** $E = \mathrm{Ker}\, u + \mathrm{Im}\, u$
- **c)** $\mathrm{Im}\, u = \mathrm{Im}\, u^2$,
- **d)** $\mathrm{rang}\, u = \mathrm{rang}\, u^2$
- **e)** $\mathrm{Ker}\, u = \mathrm{Ker}\, u^2$,
- **f)** $\mathrm{Ker}\, u \cap \mathrm{Im}\, u = \{\mathbf{0}\}$

Exercice 24.— Soit u un endomorphisme d'un espace vectoriel E de dimension finie n. Montrer que les deux assertions suivantes sont équivalentes

- **a)** $\mathrm{Ker}\, u = \mathrm{Im}\, u$,
- **b)** $u^2 = 0$ et $n = 2.\mathrm{rang}\,(u)$.

Exercice 25.— Soient E, F et G trois $\mathbb{K}$-espaces vectoriels et $u : E \longrightarrow F$, $v : F \longrightarrow G$ deux applications linéaires.

1. Montrer que $u(\mathrm{Ker}\, (v \circ u)) = \mathrm{Ker}\, v \cap \mathrm{Im}\, u$.
2. Montrer que $v^{-1}(\mathrm{Im}\, (v \circ u)) = \mathrm{Ker}\, v + \mathrm{Im}\, u$.

CHAPITRE
2

Les matrices

Sommaire

§ 1 Définitions

2.1.1. Définitions.— Soient m et n deux entiers naturels non nuls. Une famille $(a^i_j)^{1\leqslant i\leqslant n}_{1\leqslant j\leqslant m}$, où, pour tous entiers i et j, a^i_j est un scalaire dans $\mathbb{K}$, est appelée une *matrice* de type (m,n) à coefficients dans $\mathbb{K}$.

S'il n'y a pas de confusion, une telle matrice sera notée $[a^i_j]$ et représentée par un tableau de scalaires à n colonnes et m lignes :

$$\begin{bmatrix} a^1_1 & a^2_1 & \cdots & a^n_1 \\ a^1_2 & a^2_2 & \cdots & a^n_2 \\ \vdots & \vdots & & \vdots \\ a^1_m & a^2_m & \cdots & a^n_m \end{bmatrix}$$

On note $\mathscr{M}_{m,n}(\mathbb{K})$ l'ensemble des matrices carrées de type (m,n) à coefficients dans $\mathbb{K}$. Si $\mathbf{A} \in \mathscr{M}_{m,n}(\mathbb{K})$, on notera $\mathbf{A}^i_j$, le coefficient de $\mathbf{A}$ de la j-ième ligne de la i-ième colonne :

$$\mathbf{A} = \begin{array}{cc} \begin{array}{ccc} & i & \end{array} & \\ \begin{bmatrix} & \vdots & \\ & \mathbf{A}^i_j & \cdots \\ & & \end{bmatrix} & j \end{array}$$

Une matrice de type (n,n) est dite *carrée*. On note $\mathscr{M}_n(\mathbb{K})$ l'ensemble des matrices carrées de type (n,n). Une matrice de type $(m,1)$ est dite *colonne* et une matrice de type $(1,n)$ est dite *ligne*.

La *diagonale* d'une matrice carrée $\mathbf{A} = [a^i_j]$ de type (n,n) est formée des coefficients a^i_i, pour $i \in [\![1,n]\!]$; on note

$$\operatorname{diag}(\mathbf{A}) = (a^1_1, a^2_2, \ldots, a^n_n).$$

Une matrice est dite *diagonale* , si tous ses coefficients non diagonaux sont nuls, i.e., $a^j_i = 0$, pour tout $i \neq j$. On notera

$$\mathbf{Diag}(a_1,\ldots,a_n) = \begin{bmatrix} a_1 & 0 & \ldots & 0 \\ 0 & a_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & a_n \end{bmatrix}.$$

La *matrice identité* de $\mathscr{M}_n(\mathbb{K})$ est la matrice diagonale $\mathbf{1}_n$, où tous les éléments de la diagonale sont égaux à 1 :

$$\mathbf{1}_n = \mathbf{Diag}(1,\ldots,1).$$

2.1.2. Matrices triangulaires.— Une matrice est dite *triangulaire supérieure* (resp. *triangulaire inférieure*), si tous ses coefficients en dessous (resp. au dessus) de la diagonale sont nuls, i.e., $a^i_j = 0$, pour tout $j > i$ (resp. $i < j$). Une matrice triangulaire supérieure (resp. inférieure) sera notée

$$\begin{bmatrix} * & * & \ldots & * \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \ldots & 0 & * \end{bmatrix} \quad \left(\text{resp.} \begin{bmatrix} * & 0 & \ldots & 0 \\ * & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ * & \ldots & * & * \end{bmatrix}\right).$$

2.1.3. Matrices élémentaires.— Les *matrices élémentaires* de $\mathscr{M}_{m,n}(\mathbb{K})$ sont les matrices $\mathbf{E}_{(i,j)}$, où $(\mathbf{E}_{(i,j)})^l_k$ égale 1 si $k = j$ et $l = i$ et 0 sinon :

$$\mathbf{E}_{(i,j)} = \begin{array}{cc} \begin{array}{ccccc} & & i & & \end{array} & \\ \begin{bmatrix} 0 & \cdots & 0 & \cdots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & 1 & \cdots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & 0 & \cdots & 0 \end{bmatrix} & \begin{array}{c} \\ \\ j \\ \\ \\ \end{array} \end{array}$$

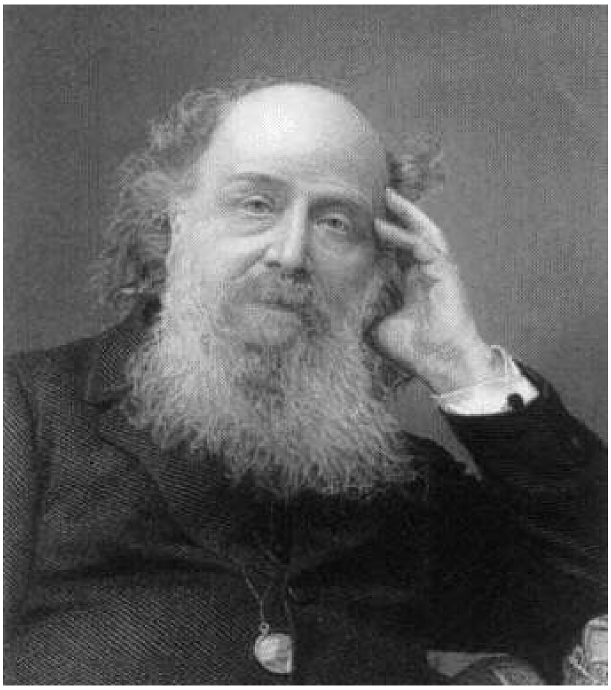


FIGURE 2.1.: James Joseph Sylvester (1814 - 1897)

Joseph Sylvester est un mathématicien anglais. Il introduit en 1951 le terme de matrice dans ses travaux géométriques. Il publie plusieurs mémoires dans lesquels il traduit des propriétés géométriques en terme de calcul de déterminant. Sylvester entretiendra une collaboration mathématique fructueuse avec le mathématicien Arthur Cayley. Ils travaillent notamment sur les invariants des formes quadratiques et les déterminants.

2.1.4. Espace vectoriel des matrices.— Soient m et n deux entiers naturels non nuls. On définit sur l'ensemble $\mathscr{M}_{m,n}(\mathbb{K})$ les opération

i) d'*addition*, pour toutes matrices $\mathbf{A} = [a^i_j]$ et $\mathbf{B} = [b^i_j]$ de $\mathscr{M}_{m,n}(\mathbb{K})$,

$$\mathbf{A} + \mathbf{B} = [a^i_j + b^i_j].$$

ii) la multiplication par un scalaire, pour toute matrice $\mathbf{A} = [a_i^j]$ de $\mathscr{M}_{m,n}(\mathbb{K})$ et tout scalaire α de $\mathbb{K}$,

$$\alpha\mathbf{A} = [\alpha a_j^i].$$

2.1 Proposition.— Muni de l'addition et de la multiplication par un scalaire l'ensemble $\mathscr{M}_{m,n}(\mathbb{K})$ des matrices de type (m,n) forme un $\mathbb{K}$-espace vectoriel de dimension mn.

La famille des mn matrices élémentaires

$$\{\mathbf{E}_{(i,j)} \quad | \quad i \in [\![1,n]\!],\ j \in [\![1,m]\!]\}$$

de $\mathscr{M}_{m,n}(\mathbb{K})$ forme une base canonique du $\mathbb{K}$-espace vectoriel $\mathscr{M}_{m,n}(\mathbb{K})$.

2.1.5. Sous-espaces vectoriels remarquables de $\mathscr{M}_{m,n}(\mathbb{K})$.— L'ensemble des matrices diagonales de $\mathscr{M}_n(\mathbb{K})$ forme un sous-espace vectoriel de dimension n.

L'ensemble des matrices triangulaires supérieures (resp. inférieures) de $\mathscr{M}_n(\mathbb{K})$ forme un sous-espace vectoriel de dimension $n(n+1)/2$.

2.1.6. Transposition.— La matrice *transposée* d'une matrice $\mathbf{A}$ de $\mathscr{M}_{m,n}(\mathbb{K})$ est la matrice de $\mathscr{M}_{n,m}(\mathbb{K})$, notée $\mathbf{A}^\top$ définie par

$$(\mathbf{A}^\top)_j^i = \mathbf{A}_i^j,$$

pour tous $i \in [\![1,n]\!]$ et $j \in [\![1,m]\!]$. Pour toute matrice $\mathbf{A}$, on a

$$(\mathbf{A}^\top)^\top = \mathbf{A}.$$

La transposition est une application linéaire

$$\begin{aligned} {}^\top : \mathscr{M}_{m,n}(\mathbb{K}) &\longrightarrow \mathscr{M}_{m,n}(\mathbb{K}) \\ \mathbf{A} &\longmapsto \mathbf{A}^\top \end{aligned}$$

On vérifie que pour toutes matrices $\mathbf{A}$, $\mathbf{B}$ et scalaire λ, on a

i) $(\mathbf{A}+\mathbf{B})^\top = \mathbf{A}^\top + \mathbf{B}^\top$,

ii) $(\lambda\mathbf{A})^\top = \lambda\mathbf{A}^\top$.

Exercice 1.— Établir ces relations.

2.1.7. Matrices symétriques et antisymétrique.— Une matrice carrée $\mathbf{S}$ est dite *symétrique* si $\mathbf{S}^\top = \mathbf{S}$. Une matrice carrée $\mathbf{A}$ est dite *antisymétrique* si $\mathbf{A}^\top = -\mathbf{A}$.

On note $\mathscr{S}_n(\mathbb{K})$ (resp. $\mathscr{A}_n(\mathbb{K})$) le sous-ensemble de $\mathscr{M}_n(\mathbb{K})$ formé des matrices symétriques (resp. antisymétriques).

2.2 Proposition.— Les sous-ensembles $\mathscr{S}_n(\mathbb{K})$ et $\mathscr{A}_n(\mathbb{K})$ forment des sous-espaces vectoriels de $\mathscr{M}_n(\mathbb{K})$, avec

$$\dim \mathscr{A}_n(\mathbb{K}) = \frac{n^2-n}{2}, \quad \dim \mathscr{S}_n(\mathbb{K}) = \frac{n^2+n}{2}.$$

De plus, on a

$$\mathscr{M}_n(\mathbb{K}) = \mathscr{A}_n(\mathbb{K}) \oplus \mathscr{S}_n(\mathbb{K})$$

Exercice 2.— Soit $u : \mathscr{M}_n(\mathbb{R}) \longrightarrow \mathscr{M}_n(\mathbb{R})$ l'application définie par

$$u(\mathbf{A}) = \mathbf{A} + \mathbf{A}^\top,$$

pour tout $\mathbf{A} \in \mathscr{M}_n(\mathbb{R})$.

1. Montrer que l'application u est un endomorphisme de $\mathscr{M}_n(\mathbb{R})$.
2. Déterminer le noyau de u.
3. Déterminer l'image de u.
4. Montrer que $\mathscr{M}_n(\mathbb{R}) = \operatorname{Ker} u \oplus \operatorname{Im} u$.

§ 2 Produit de matrices

2.2.1. Produit de matrices.— On définit le *produit* (ou *multiplication*) d'une matrice $\mathbf{A}$ de type (m,n) par une matrice $\mathbf{B}$ de type (n,p), comme la matrice

$$\mathbf{AB} = [c_i^j], \quad \text{avec} \quad c_i^j = \sum_{k=1}^{n} a_i^k b_k^j.$$

Le produit de matrices vérifie les propriétés suivantes :

i) (associativité) pour toute matrices compatibles, $\mathbf{A}$, $\mathbf{B}$ et $\mathbf{C}$,

$$\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}.$$

ii) (matrices unitées) pour toute matrice $\mathbf{A}$ de $\mathscr{M}_{m,n}(\mathbb{K})$,

$$\mathbf{1}_m\mathbf{A} = \mathbf{A} = \mathbf{A}\mathbf{1}_n.$$

iii) (distributivité), pour toutes matrices compatibles $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ et $\mathbf{D}$, on a

$$\mathbf{A}(\mathbf{B}+\mathbf{C}) = \mathbf{AB} + \mathbf{AC}, \qquad (\mathbf{B}+\mathbf{C})\mathbf{D} = \mathbf{BD} + \mathbf{CD}.$$

On en déduit que

2.3 Proposition.— L'ensemble $\mathscr{M}_n(\mathbb{K})$ des matrices carrées de type (n,n), muni de l'addition et de la multiplication, forme un anneau non commutatif.

L'unité de l'anneau est la matrice identité $\mathbf{1}_n$.

Exercice 3.— Montrer que l'anneau $\mathscr{M}_n(\mathbb{K})$ n'est pas commutatif pour $n \geqslant 2$.

Exercice 4.— Montrer que, pour toutes matrices compatibles **A** et **B**, on a

$$(\mathbf{AB})^\top = \mathbf{B}^\top\mathbf{A}^\top.$$

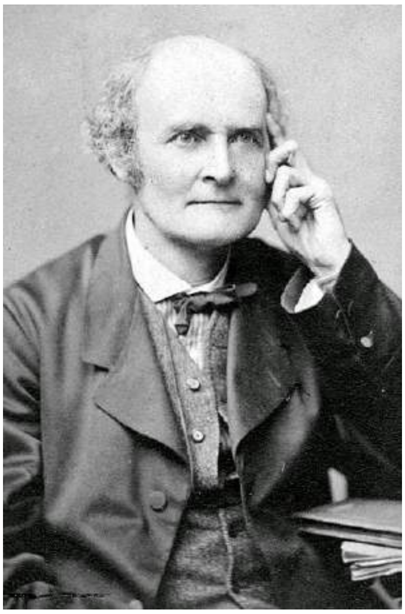


FIGURE 2.2.: Arthur Cayley (1821 - 1895)

Arthur Cayley est un mathématicien anglais auteur de nombreux travaux. Dans un article de 1855 publié dans le Journal de Crelle, il utilise la notion de matrice pour représenter des systèmes d'équations linéaires et des formes quadratiques. Les problèmes géométriques, comme les types d'intersection de coniques ou quadriques, exprimés en terme de déterminant, sont à l'origine de l'intérêt de Cayley pour la notion de matrice. Ses travaux sur les matrices conduisent à un mémoire intitulé « A memoir on the Theory of Matrices », publié en 1858 dans Philosophical Transactions of the Royal Society of London *dans lequel la notion de matrice fait l'objet d'une étude théorique indépendamment du contexte géométrique dans lequels elles apparaissent jusqu'alors.*

Ce mémoire conporte des résultats importants, en particulier, il formule le théorème dit de Cayley-Hamilton pour les matrices carrées d'ordre 3 *sans toutefois en publier de preuve*

Exercice 5.— Soient

$$\mathbf{B} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 5 & 0 & 1 \end{bmatrix}$$

et $\mathbf{A}$ des matrices de $\mathscr{M}_3(\mathbb{R})$. Décrire les lignes de la matrice $\mathbf{BA}$ en terme des lignes de $\mathbf{A}$ et les colonnes de la matrice $\mathbf{AB}$ en terme des colonnes de $\mathbf{A}$.

Exercice 6.— Soit $\mathbf{e}_i$ le vecteur colonne de $\mathscr{M}_{n,1}(\mathbb{K})$ contenant 1 à la i-ième ligne et 0 sur les autres lignes. Étant donnée une matrice $\mathbf{A}$ de $\mathscr{M}_n(\mathbb{K})$, décrire les produits suivants :

$$\mathbf{A}\mathbf{e}_i, \qquad \mathbf{e}_i^\top \mathbf{A}, \qquad \mathbf{e}_i^\top \mathbf{A}\mathbf{e}_j.$$

Exercice 7.— Soient $\mathbf{A}$ et $\mathbf{B}$ deux matrices de $\mathscr{M}_{n,m}(\mathbb{K})$. Montrer que si $\mathbf{Ax} = \mathbf{Bx}$, pour tout vecteur colonne $\mathbf{x}$, alors $\mathbf{A} = \mathbf{B}$.

Exercice 8.— Soit

$$\mathbf{A} = \begin{bmatrix} 1/2 & a \\ 0 & 1/2 \end{bmatrix}$$

une matrice à coefficients réels.

1. Calculer $\mathbf{A}^2$, $\mathbf{A}^3$ et $\mathbf{A}^4$.
2. Donner l'expression de $\mathbf{A}^k$, pour tout entier naturel k.

Exercice 9.— Soient $\mathbf{T}$ et $\mathbf{T}'$ deux matrices triangulaires supérieures.

1. Montrer que le produit $\mathbf{TT}'$ est une matrice triangulaire supérieure.
2. Déterminer la diagonale de la matrice $\mathbf{TT}'$.

Exercice 10.— Soient $\mathbf{A}$ et $\mathbf{B}$ deux matrices symétriques de $\mathscr{M}_n(\mathbb{K})$ qui commutent, i.e., $\mathbf{AB} = \mathbf{BA}$. Montrer que le produit $\mathbf{AB}$ est une matrice symétrique.

2.2.2. Matrices nilpotentes.— Une matrice $\mathbf{A}$ de $\mathscr{M}_n(\mathbb{K})$ est dite *nilpotente*, s'il existe un entier naturel q tel que $\mathbf{A}^q = 0$. Le plus petit entier non nul r tel que $\mathbf{A}^r = 0$ est appelé l'*indice de nilpotence* de $\mathbf{A}$.

2.2.3. Exemples.— Toute matrice nulle est nilpotente d'indice de nilpotence 1. Les matrices suivantes

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

sont nilpotentes d'indice de nilpotence 2, 3, 3, 3 et 4 respectivement.

2.2.4. Matrices inversibles.— Une matrice carrée $\mathbf{A}$ de $\mathscr{M}_n(\mathbb{K})$ est dite *inversible*, s'il existe une matrice $\mathbf{B}$ de $\mathscr{M}_n(\mathbb{K})$ telle que

$$\mathbf{AB} = \mathbf{1}_n \quad \text{et} \quad \mathbf{BA} = \mathbf{1}_n.$$

La matrice $\mathbf{B}$ est alors appelée la *matrice inverse* de $\mathbf{A}$, on note alors $\mathbf{B} = \mathbf{A}^{-1}$. On déduit immédiatement de cette définition que l'inverse d'une matrice est unique.

L'opération d'inversion vérifie les propriétés suivantes :

i) si $\mathbf{A}$ est inversible, $(\mathbf{A}^{-1})^{-1} = \mathbf{A}$,

ii) si $\mathbf{A}$ et $\mathbf{B}$ sont inversibles, alors $\mathbf{AB}$ est inversible et

$$(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1},$$

iii) si $\mathbf{A}$ est inversible, alors sa transposée est inversible et

$$(\mathbf{A}^\top)^{-1} = (\mathbf{A}^{-1})^\top.$$

Exercice 11.— Montrer ces trois propriétés.

On désigne par $\mathrm{GL}_n(\mathbb{K})$ l'ensemble des matrices inversibles de $\mathscr{M}_n(\mathbb{K})$:

$$\mathrm{GL}_n(\mathbb{K}) = \{\mathbf{A} \in \mathscr{M}_n(\mathbb{K}) \mid \mathbf{A} \text{ est inversible}\}.$$

2.4 Proposition.— La multiplication des matrices munit l'ensemble $\mathrm{GL}_n(\mathbb{K})$ des matrices inversibles de $\mathscr{M}_n(\mathbb{K})$ d'une structure de groupe.

Le groupe $\mathrm{GL}_n(\mathbb{K})$, est appelé le *groupe linéaire* des matrices d'ordre n.

2.2.5. Calcul de l'inverse.— Déterminer si une matrice est inversible et le calcul des inverses sont des problèmes importants d'algèbre linéaire. Par exemple, considérons la matrice suivante de $\mathscr{M}_3(\mathbb{R})$

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 3 \\ 2 & 0 & 1 \\ 1 & 1 & 2 \end{bmatrix}.$$

La matrice $\mathbf{A}$ est-elle inversible ? L'équation matricielle suivante dans $\mathbb{R}^3$:

$$\mathbf{Ax} = \mathbf{b} \ \text{avec} \ \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \ \text{et}\ \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

peut s'écrire sous la forme du système d'équations suivant

$$\left|\begin{array}{rcl} x_1 + x_2 + 3x_3 & = & b_1 \\ 2x_1 + x_3 & = & b_2 \\ x_1 + x_2 + 2x_3 & = & b_3 \end{array}\right.$$

Pour résoudre ce systèmes, i.e., exprimer les coefficients du vecteur $\mathbf{x}$ en terme de ceux du vecteur $\mathbf{b}$ on procède en appliquant des opérations sur les lignes. Retranchons 3 fois la seconde équation à la première et 2 fois à la dernière, le système devient :

$$\left|\begin{array}{lcl} -5x_1+x_2 &=& b_1-3b_2 \\ 2x_1+x_3 &=& b_2 \\ -3x_1+x_2 &=& -2b_2+b_3 \end{array}\right.$$

Retranchons la dernière équation à la première, on obtient :

$$\left|\begin{array}{lcl} -2x_1 &=& b_1-b_2-b_3 \\ 2x_1+x_3 &=& b_2 \\ -3x_1+x_2 &=& -2b_2+b_3 \end{array}\right.$$

On additionne la première équation à la seconde et on retranche $\frac{3}{2}$ de la première à la troisième, il reste :

$$\left|\begin{array}{lcl} -2x_1 &=& b_1-b_2-b_3 \\ x_3 &=& b_1-b_3 \\ x_2 &=& -\frac{3}{2}b_1-\frac{1}{2}b_2+\frac{5}{2}b_3 \end{array}\right.$$

On obtient ainsi le système

$$\left|\begin{array}{lcl} x_1 &=& -\frac{1}{2}b_1+\frac{1}{2}b_2+\frac{1}{2}b_3 \\ x_2 &=& -\frac{3}{2}b_1-\frac{1}{2}b_2+\frac{5}{2}b_3 \\ x_3 &=& b_1-b_3 \end{array}\right.$$

L'équation $\mathbf{Ax}=\mathbf{b}$ admet une unique solution $\mathbf{x}$ donnée par le système précédent, ce qui s'écrit matriciellement sous la forme

$$\mathbf{x}=\mathbf{A}^{-1}\mathbf{b},$$

avec

$$\mathbf{A}^{-1}=\begin{bmatrix} -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ -\frac{3}{2} & -\frac{1}{2} & \frac{5}{2} \\ 1 & 0 & -1 \end{bmatrix}.$$

On vérifie que $\mathbf{AA}^{-1}=\mathbf{A}^{-1}\mathbf{A}=\mathbf{1}_n$, la matrice $\mathbf{A}$ est donc inversible. Cet exemple illustre le calcul de l'inverse d'une matrice en effectuant la résolution d'un système linéaire à seconds membres. On notera cependant, que les matrices carrées ne sont pas toujours inversibles. Nous verrons plus loin d'autres méthodes permettant de déterminer si une matrice est inversible et de calculer l'inverse d'une matrice.

Exercice 12.— Montrer que les matrices suivantes ne sont pas inversibles

$$\mathbf{A}=\begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix},\qquad \mathbf{B}=\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix},\qquad \mathbf{C}=\begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 7 \\ 1 & 1 & 2 \end{bmatrix}.$$

Exercice 13.— Soit $\mathbf{A}$ un matrice de $\mathscr{M}_n(\mathbb{K})$ telle que $\mathbf{A}^2$ est inversible. Montrer que $\mathbf{A}$ est inversible.

II. *A Memoir on the Theory of Matrices.* By ARTHUR CAYLEY, *Esq., F.R.S.*

Received December 10, 1857,—Read January 14, 1858.

THE term matrix might be used in a more general sense, but in the present memoir I consider only square and rectangular matrices, and the term matrix used without qualification is to be understood as meaning a square matrix; in this restricted sense, a set of quantities arranged in the form of a square, *e. g.*

$$\begin{pmatrix} a, & b, & c \\ a', & b', & c' \\ a'', & b'', & c'' \end{pmatrix}$$

is said to be a matrix. The notion of such a matrix arises naturally from an abbreviated notation for a set of linear equations, viz. the equations

$$\begin{aligned} X &= ax + by + cz, \\ Y &= a'x + b'y + c'z, \\ Z &= a''x + b''y + c''z, \end{aligned}$$

may be more simply represented by

$$(X, Y, Z) = \begin{pmatrix} a, & b, & c \\ a', & b', & c' \\ a'', & b'', & c'' \end{pmatrix}(x, y, z),$$

and the consideration of such a system of equations leads to most of the fundamental notions in the theory of matrices. It will be seen that matrices (attending only to those of the same order) comport themselves as single quantities; they may be added, multiplied or compounded together, &c.: the law of the addition of matrices is precisely similar to that for the addition of ordinary algebraical quantities; as regards their multiplication (or composition), there is the peculiarity that matrices are not in general convertible; it is nevertheless possible to form the powers (positive or negative, integral or fractional) of a matrix, and thence to arrive at the notion of a rational and integral function, or generally of any algebraical function, of a matrix. I obtain the

FIGURE 2.3.: A Memoir on the Theory of Matrices

§ 3 Matrice d'une application linéaire

Soient E et E' deux $\mathbb{K}$-espaces vectoriel de dimensions n et m respectivement. Considérons une base $\mathscr{B} = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ de E et une base $\mathscr{B}' = (\mathbf{f}_1, \ldots, \mathbf{f}_m)$ de E'.

Soit $u : E \longrightarrow E'$ une application linéaire. On peut exprimer l'image par u de chaque vecteur de la base $\mathscr{B}$ dans la base $\mathscr{B}'$, on a :

$$\begin{aligned}
u(\mathbf{e}_1) &= a_1^1\mathbf{f}_1 + a_2^1\mathbf{f}_2 + \cdots + a_m^1\mathbf{f}_m,\\
u(\mathbf{e}_2) &= a_1^2\mathbf{f}_1 + a_2^2\mathbf{f}_2 + \cdots + a_m^2\mathbf{f}_m,\\
&\dots\\
u(\mathbf{e}_n) &= a_1^n\mathbf{f}_1 + a_2^n\mathbf{f}_2 + \cdots + a_m^n\mathbf{f}_m,
\end{aligned}$$

avec $a_j^i \in \mathbb{K}$. La matrice

$$[a_j^i]_{1\leqslant j\leqslant m}^{1\leqslant i\leqslant n}$$

est appelée la *matrice de l'application linéaire u* exprimée dans les bases $\mathscr{B}$ et $\mathscr{B}'$, on la notera

$$[u]_{\mathscr{B}}^{\mathscr{B}'}.$$

En résumé :

$$[u]_{\mathscr{B}}^{\mathscr{B}'} = \begin{array}{c} \begin{array}{cccc} u(\mathbf{e}_1) & u(\mathbf{e}_2) & \cdots & u(\mathbf{e}_n) \end{array} \\ \begin{bmatrix} a_1^1 & a_1^2 & \cdots & a_1^n \\ a_2^1 & a_2^2 & \cdots & a_2^n \\ \vdots & \vdots & & \vdots \\ a_m^1 & a_m^2 & \cdots & a_m^n \end{bmatrix} \end{array} \begin{array}{c} \\ \mathbf{f}_1 \\ \mathbf{f}_2 \\ \vdots \\ \mathbf{f}_m \end{array}.$$

Pour un endomorphisme $u : E \longrightarrow E$ de E et $\mathscr{B}$ une base de E, nous noterons $[u]_{\mathscr{B}}$ la matrice $[u]_{\mathscr{B}}^{\mathscr{B}}$.

2.3.1. Exemple.— Soit p l'application linéaire de $\mathbb{R}^3$ dans $\mathbb{R}^3$ qui envoie un vecteur $\mathbf{x}$, sur sa projection orthogonale $p(\mathbf{x})$ sur le plan (Oxy) :

$$\begin{aligned} p : \mathbb{R}^3 &\longrightarrow \mathbb{R}^3 \\ \begin{bmatrix} x \\ y \\ z \end{bmatrix} &\longmapsto \begin{bmatrix} x \\ y \\ 0 \end{bmatrix}. \end{aligned}$$

On considère la base de $\mathbb{R}^3$ donnée par

$$\mathscr{B} = \left\{ \mathbf{e}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{e}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{e}_3 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \right\}$$

On a

$$p(\mathbf{e}_1) = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \mathbf{e}_2, \quad p(\mathbf{e}_2) = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \mathbf{e}_2, \quad p(\mathbf{e}_3) = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} = -3\mathbf{e}_1 + 3\mathbf{e}_2 + \mathbf{e}_3.$$

Ainsi, la matrice de l'application p, exprimée dans la base $\mathscr{B}$ de $\mathbb{R}^3$, est

$$[p]_{\mathscr{B}} = \begin{bmatrix} 0 & 0 & -3 \\ 1 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix}.$$

2.5 Proposition.— Soient E et E' deux $\mathbb{K}$-espaces vectoriels de dimension finie respectives n et m. L'application $\Phi : \mathscr{L}(E,E') \longrightarrow \mathscr{M}_{m,n}(\mathbb{K})$ définie par

$$\Phi(u) = [u]_{\mathscr{B}}^{\mathscr{B}'}$$

est un isomorphisme de $\mathbb{K}$-espaces vectoriels. En particulier, les espaces vectoriels $\mathscr{L}(E,E')$ et $\mathscr{M}_{m,n}(\mathbb{K})$ ont la même dimension, égale à mn.

Soient $u : E \longrightarrow E'$ et $v : E' \longrightarrow E''$ deux applications linéaires et $\mathscr{B}, \mathscr{B}', \mathscr{B}''$ des bases des $\mathbb{K}$-espaces vectoriels E, E' et E'' respectivement. Alors

$$[vu]_{\mathscr{B}}^{\mathscr{B}''} = [v]_{\mathscr{B}'}^{\mathscr{B}''} [u]_{\mathscr{B}}^{\mathscr{B}'}.$$

Soit $\mathbf{x}$ un vecteur de E de coordonnées $x_1, \ldots, x_n$ dans la base $\mathscr{B}$, on note $[\mathbf{x}]_{\mathscr{B}}$ le vecteur colonne correspondant :

$$[\mathbf{x}]_{\mathscr{B}} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

Avec les notations précédentes, on a :

$$[u(\mathbf{x})]_{\mathscr{B}'} = [u]_{\mathscr{B}}^{\mathscr{B}'} [\mathbf{x}]_{\mathscr{B}}.$$

2.3.2. Matrices de passage.— Soient E un $\mathbb{K}$-espace vectoriel de dimension n. Considérons $\mathscr{B} = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ et $\mathscr{B}' = (\mathbf{f}_1, \ldots, \mathbf{f}_n)$ deux bases de E. Tout vecteur $\mathbf{f}_i$, pour $i \in [\![1,n]\!]$, se décompose de façon unique dans la base $\mathscr{B}$:

$$\mathbf{f}_i = p_1^i \mathbf{e}_1 + \ldots + p_n^i \mathbf{e}_n, \quad \text{où } p_j^i \in \mathbb{K}.$$

La *matrice de passage* de la base $\mathscr{B}$ à la base $\mathscr{B}'$, est la matrice, notée $\mathbf{P}_{\mathscr{B}}^{\mathscr{B}'}$, dont les colonnes sont les composantes des vecteurs $\mathbf{f}_i$ exprimés dans la base $\mathscr{B}$:

$$\mathbf{P}_{\mathscr{B}}^{\mathscr{B}'} = \begin{bmatrix} p_1^1 & \cdots & p_1^n \\ p_2^1 & & p_2^n \\ \vdots & & \vdots \\ p_n^1 & \cdots & p_n^n \end{bmatrix}$$

La matrice de passage $\mathbf{P}_{\mathscr{B}}^{\mathscr{B}'}$ vérifie les propriétés suivantes :

i) $\mathbf{P}_{\mathscr{B}}^{\mathscr{B}'} = [\mathrm{id}_E]_{\mathscr{B}'}^{\mathscr{B}}$,

ii) $\mathbf{P}_{\mathscr{B}}^{\mathscr{B}'}$ est inversible et $(\mathbf{P}_{\mathscr{B}}^{\mathscr{B}'})^{-1} = \mathbf{P}_{\mathscr{B}'}^{\mathscr{B}}$.

Soit $\mathbf{x}$ un vecteur de E de coordonnées $x_1, \ldots, x_n$ dans la base $\mathscr{B}$ et de coordonnées $x'_1, \ldots, x'_n$ dans la base $\mathscr{B}'$. En notant

$$[\mathbf{x}]_{\mathscr{B}} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad \text{et} \quad [\mathbf{x}]_{\mathscr{B}'} = \begin{bmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_n \end{bmatrix},$$

on a :

$$[\mathbf{x}]_{\mathscr{B}} = \mathbf{P}_{\mathscr{B}}^{\mathscr{B}'} [\mathbf{x}]_{\mathscr{B}'},$$

ou de façon équivalente :

$$[\mathbf{x}]_{\mathscr{B}'} = (\mathbf{P}_{\mathscr{B}}^{\mathscr{B}'})^{-1} [\mathbf{x}]_{\mathscr{B}} = \mathbf{P}_{\mathscr{B}'}^{\mathscr{B}} [\mathbf{x}]_{\mathscr{B}}.$$

2.6 Proposition.— Soient E et E' deux $\mathbb{K}$-espaces vectoriels de dimension finie et $u : E \longrightarrow E'$ une application linéaire. Soient $\mathscr{B}, \mathscr{B}'$ deux bases de E, $\mathscr{C}, \mathscr{C}'$ deux bases de E' et $\mathbf{P}_{\mathscr{B}}^{\mathscr{B}'}$, $\mathbf{P}_{\mathscr{C}}^{\mathscr{C}'}$ les matrices de changements de bases associées. Alors,

$$[u]_{\mathscr{B}'}^{\mathscr{C}'} = (\mathbf{P}_{\mathscr{C}}^{\mathscr{C}'})^{-1} [u]_{\mathscr{B}}^{\mathscr{C}} \mathbf{P}_{\mathscr{B}}^{\mathscr{B}'}.$$

En particulier, si u est un endomorphisme de E :

$$[u]_{\mathscr{B}'} = (\mathbf{P}_{\mathscr{B}}^{\mathscr{B}'})^{-1} [u]_{\mathscr{B}} \mathbf{P}_{\mathscr{B}}^{\mathscr{B}'}.$$

2.3.3. Exemple.— Soit $u : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ l'endomorphisme défini par

$$u \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x - z \\ y - z \\ 0 \end{bmatrix}.$$

La matrice de u dans la base canonique $\mathrm{can} = (\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3)$ de $\mathbb{R}^3$ est

$$[u]_{\mathrm{can}} = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}$$

On considère une nouvelle base de $\mathbb{R}^3$:

$$\mathscr{B} = \left\{ \mathbf{e}_1 = \mathbf{c}_1,\ \mathbf{e}_2 = \mathbf{c}_2,\ \mathbf{e}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}.$$

La matrice de passage de la base canonique à la base $\mathscr{B}$ est

$$\mathbf{P}_{\text{can}}^{\mathscr{B}} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}.$$

On calcule l'inverse de la matrice de $\mathbf{P}_{\text{can}}^{\mathscr{B}}$:

$$\mathbf{P}_{\text{can}}^{\mathscr{B}\ -1} = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}.$$

L'expression de la matrice de u exprimée dans la base $\mathscr{B}$ est

$$\begin{aligned} [u]_{\mathscr{B}} &= \mathbf{P}_{\text{can}}^{\mathscr{B}\ -1}[u]_{\text{can}}\mathbf{P}_{\text{can}}^{\mathscr{B}} \\ &= \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}. \end{aligned}$$

D'où

$$[u]_{\mathscr{B}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

On constate que cette nouvelle base permet d'exprimer plus simplement l'endomorphisme u. Nous retrouverons cet exemple en 4.3.4.

Exercice 14.— Écrire la matrice de passage de la base $\mathscr{B}$ de l'espace $\mathbb{R}^3$ donnée dans l'exemple 2.3.1 à la base canonique de $\mathbb{R}^3$. Soit p la projection donnée en 2.3.1. À l'aide de cette matrice de passage et de la matrice de l'endomorphisme p exprimée dans la base $\mathscr{B}$, calculer la matrice de p exprimée dans la base canonique de $\mathbb{R}^3$.

2.3.4. Matrices semblables.— Deux matrices $\mathbf{A}$ et $\mathbf{B}$ de $\mathscr{M}_n(\mathbb{K})$ sont dites *semblables*, s'il existe une matrice inversible $\mathbf{P}$ de $\mathscr{M}_n(\mathbb{K})$ telle que

$$\mathbf{B} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}.$$

La relation « être semblable à », dite relation de *similitude*, est une relation d'équivalence sur l'ensemble $\mathscr{M}_n(\mathbb{K})$.

Exercice 15.— Montrer cette propriété.

Deux matrices semblables représentent le même endomorphisme dans des bases différentes.

2.7 Proposition.— Deux matrices semblables ont même trace, même déterminant, même rang.

2.3.5. Matrices équivalentes.— Deux matrices $\mathbf{A}$ et $\mathbf{B}$ de $\mathscr{M}_n(\mathbb{K})$ sont dites *équivalentes*, s'il existe deux matrice inversibles $\mathbf{P}$ et $\mathbf{Q}$ de $\mathscr{M}_n(\mathbb{K})$ telles que

$$\mathbf{B} = \mathbf{PAQ}.$$

La relation d'*équivalence* est une relation d'équivalence sur $\mathscr{M}_n(\mathbb{K})$. Deux matrices semblables sont équivalentes.

§ 4 Trace d'une matrice

2.4.1. Définition.— La *trace* d'une matrice carrée $\mathbf{A} = [a^i_j]$ de $\mathscr{M}_n(\mathbb{K})$ est la somme des coefficients de sa diagonale :

$$\mathrm{trace}\,(\mathbf{A}) = \sum_{i=1}^{n} a^i_i.$$

2.8 Proposition.— L'application trace : $\mathscr{M}_n(\mathbb{K}) \longrightarrow \mathbb{K}$ est une forme linéaire vérifiant, pour toutes matrices $\mathbf{A}$ et $\mathbf{B}$ de $\mathscr{M}_n(\mathbb{K})$,

$$\mathrm{trace}\,\mathbf{AB} = \mathrm{trace}\,\mathbf{BA}.$$

Exercice 16.— Étant donné deux matrices $\mathbf{A}$ et $\mathbf{B}$ de $\mathscr{M}_n(\mathbb{K})$, déterminer toutes les matrices $\mathbf{X}$ de $\mathscr{M}_n(\mathbb{K})$ telles que

$$\mathbf{X} + \mathrm{tr}(\mathbf{X})\mathbf{A} = \mathbf{B}.$$

Exercice 17.— Montrer qu'il n'existe pas de matrices $\mathbf{A}$ et $\mathbf{B}$ de $\mathscr{M}_n(\mathbb{K})$ telles que

$$\mathbf{AB} - \mathbf{BA} = \mathbf{1}_n.$$

2.4.2. Trace d'un endomorphisme.— De la proposition 2.8, on déduit que

2.9 Proposition.— Deux matrices semblables ont la même trace.

Exercice 18.— Soient $\mathbf{A}$ et $\mathbf{B}$ deux matrices carrées.

1. Montrer la relation $\mathrm{trace}\,\mathbf{AB} = \mathrm{trace}\,\mathbf{BA}$.
2. En déduire la proposition 2.9 : deux matrices semblables ont la même trace.

§ 5 Noyau et image d'une matrice

Nous avons vu dans le chapitre précédent les notions de noyau et d'image d'une application linéaire. On peut définir de la même façon, le noyau et l'image d'une matrice.

2.5.1. Image d'une matrice.— Si $u : \mathbb{K}^n \longrightarrow \mathbb{K}^m$ est une application linéaire, l'image de u est le sous-espace vectoriel de $\mathbb{K}^m$ défini par

$$\operatorname{Im} u = \{u(\mathbf{x}) \mid \mathbf{x} \in \mathbb{K}^n\}.$$

De la même façon, si $\mathbf{A}$ est une matrice de $\mathscr{M}_{m,n}(\mathbb{K})$, on définit l'*image* de $\mathbf{A}$ comme le sous-espace vectoriel de $\mathbb{K}^m$ engendré par les vecteurs $\mathbf{Ax}$, où $\mathbf{x} \in \mathbb{K}^n$. Soit

$$\operatorname{Im} \mathbf{A} = \{\mathbf{Ax} \mid \mathbf{x} \in \mathbb{K}^n\}.$$

2.10 Proposition.— L'image d'une matrice $\mathbf{A}$ de $\mathscr{M}_{m,n}(\mathbb{K})$ est le sous-espace vectoriel de $\mathbb{K}^m$ engendré par les vecteurs colonnes de $\mathbf{A}$.

Preuve. Écrivons $\mathbf{A}$ selon ses vecteurs colonnes :

$$\mathbf{A} = \left[\ \mathbf{a}_*^1 \mid \mathbf{a}_*^2 \mid \ldots \mid \mathbf{a}_*^n\ \right].$$

Pour tout vecteur

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

de $\mathbb{K}^n$, on a

$$\mathbf{Ax} = \left[\ \mathbf{a}_*^1 \mid \mathbf{a}_*^2 \mid \ldots \mid \mathbf{a}_*^n\ \right] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1\mathbf{a}_*^1 + x_2\mathbf{a}_*^2 + \ldots + x_n\mathbf{a}_*^n.$$

Ainsi, $\mathbf{Ax}$ est une combinaison linéaire des vecteurs colonnes de $\mathbf{A}$, autrement dit $\operatorname{Im} \mathbf{A}$ est le sous-espace vectoriel de $\mathbb{K}^m$ engendré par les vecteurs colonnes de $\mathbf{A}$. □

Exercice 19.— Soit $\mathbf{A}$ une matrice de $\mathscr{M}_n(\mathbb{K})$ telle que $\operatorname{Im}(\mathbf{A})$ soit égal à $\mathbb{K}^n$. Justifier que la matrice $\mathbf{A}$ est inversible.

Exercice 20.— Soit $\mathbf{A}$ une matrice de $\mathscr{M}_{n,m}(\mathbb{K})$. Soit $\mathscr{E}$ un sous-ensemble de vecteurs colonnes de $\mathscr{M}_{m,1}(\mathbb{K})$, identifié à $\mathbb{K}^m$. L'*image de* $\mathscr{E}$ *par* $\mathbf{A}$ est l'ensemble, noté $\mathbf{A}(\mathscr{E})$, formé de tous les produits de $\mathbf{A}$ par un vecteur de $\mathscr{E}$, i.e.,

$$\mathbf{A}(\mathscr{E}) = \{\mathbf{Ax} \mid \mathbf{x} \in \mathscr{E}\}.$$

1. Montrer que si $\mathscr{E}$ est un sous-espace vectoriel de $\mathbb{K}^m$, alors $\mathbf{A}(\mathscr{E})$ est un sous-espace vectoriel de $\mathbb{K}^n$.
2. Montrer que si $\mathscr{E} = \operatorname{Vect}(\mathbf{x}_1, \ldots, \mathbf{x}_p)$, alors $\mathbf{A}(\mathscr{E}) = \operatorname{Vect}(\mathbf{Ax}_1, \ldots, \mathbf{Ax}_p)$.

Exercice 21.— Soient $\mathbf{A}$ et $\mathbf{B}$ deux matrices compatibles. Montrer que $\operatorname{Im}(\mathbf{AB})$ est un sous-espace vectoriel de $\operatorname{Im}(\mathbf{A})$.

Exercice 22.— Soit $\mathbf{A} \in \mathscr{M}_{n,m}(\mathbb{K})$ et $\mathbf{B} \in \mathscr{M}_{n,p}(\mathbb{K})$ deux matrices. Montrer que

$$\mathrm{Im}\,([\mathbf{A}\ \ \mathbf{B}]) = \mathrm{Im}\,(\mathbf{A}) + \mathrm{Im}\,(\mathbf{B}),$$

où $[\mathbf{A}\ \ \mathbf{B}]$ désigne la matrice constituée des blocs $\mathbf{A}$ et $\mathbf{B}$.

2.5.2. Noyau d'une matrice.— Si $u : \mathbb{K}^n \longrightarrow \mathbb{K}^m$ est une application linéaire, le noyau de u est le sous-espace vectoriel de $\mathbb{K}^n$ défini par

$$\mathrm{Ker}\, u = \{\mathbf{x} \in \mathbb{K}^n \mid u(\mathbf{x}) = \mathbf{0}\}.$$

De la même façon, si $\mathbf{A}$ est une matrice de $\mathscr{M}_{m,n}(\mathbb{K})$, on définit le *noyau* de $\mathbf{A}$ comme le sous-espace vectoriel suivant de $\mathbb{K}^n$

$$\mathrm{Ker}\, \mathbf{A} = \{\mathbf{x} \in \mathbb{K}^n \mid \mathbf{A}\mathbf{x} = \mathbf{0}\}.$$

Le noyau Ker $\mathbf{A}$ est formé des solutions de l'équation

$$\mathbf{A}\mathbf{x} = \mathbf{0}.$$

Exercice 23.— Soit $\mathbf{A}$ une matrice inversible de $\mathscr{M}_n(\mathbb{K})$. Décrire les sous-espaces

$$\mathrm{Ker}\,(\mathbf{A}), \qquad \mathrm{Ker}\,(\mathbf{A}^\top), \qquad \mathrm{Im}\,(\mathbf{A}), \qquad \mathrm{Im}\,(\mathbf{A}^\top).$$

Exercice 24.— Soient $\mathbf{A}$ et $\mathbf{B}$ deux matrices compatibles. Montrer que Ker $(\mathbf{B})$ est un sous-espace vectoriel de Ker $(\mathbf{AB})$.

§ 6 Le rang d'une matrice

2.6.1. Définition.— Soient E un $\mathbb{K}$-espace vectoriel. On appelle *rang* d'une famille de vecteurs $(\mathbf{x}_i)_i$ de E, la dimension du sous-$\mathbb{K}$-espace vectoriel de E engendré par $(\mathbf{x}_i)_i$. En d'autres termes, le rang de la famille $(\mathbf{x}_i)_i$ est le nombre maximal de vecteurs linéairement indépendants que l'on peut extraire de $(\mathbf{x}_i)_i$.

2.6.2. Rang d'une matrice.— Le *rang* d'une matrice $\mathbf{A}$ de $\mathscr{M}_{m,n}(\mathbb{K})$ est le rang de la famille de ses vecteurs colonnes dans $\mathbb{K}^m$. On le note rang $\mathbf{A}$. Autrement dit,

$$\mathrm{rang}\,\mathbf{A} = \dim(\mathrm{Im}\,\mathbf{A}).$$

2.11 Proposition.— Soit $\mathbf{A}$ une matrice de $\mathscr{M}_{m,n}(\mathbb{K})$. Le rang de $\mathbf{A}$ vérifie les propriétés suivantes :

- **i)** $\mathrm{rang}\,\mathbf{A} \leqslant \inf\{m,n\}$,
- **ii)** $\mathrm{rang}\,\mathbf{A} = \mathrm{rang}\,(\mathbf{A}^\top)$ (en particulier le rang de $\mathbf{A}$ est aussi le rang des vecteurs lignes),

iii) Ker $\mathbf{A} = \{\mathbf{0}\}$ si, et seulement si, $\operatorname{rang}\mathbf{A} = n$,
iv) Ker $(\mathbf{A}^\top) = \{\mathbf{0}\}$ si, et seulement si, $\operatorname{rang}\mathbf{A} = m$,
v) si $m = n$, alors $\mathbf{A}$ est inversible si, et seulement si, $\operatorname{rang}\mathbf{A} = n$.

Soit $\mathbf{A} \in \mathscr{M}_{p,q}(\mathbb{K})$ et $\mathbf{Q} \in \mathscr{M}_p(\mathbb{K})$, $\mathbf{P} \in \mathscr{M}_p(\mathbb{K})$ deux matrices inversibles, on a

$$\operatorname{rang}\mathbf{A} = \operatorname{rang}(\mathbf{QAP}).$$

En particulier si $\operatorname{rang}\mathbf{A} = r$, il existe deux matrices inversibles $\mathbf{P}$ et $\mathbf{Q}$ telles que

$$\mathbf{A} = \mathbf{Q}\begin{bmatrix} \mathbf{1}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}\mathbf{P}.$$

2.12 Proposition.— Soit $\mathbf{A}$ une matrice de $\mathscr{M}_{p,q}(\mathbb{K})$ de rang r. Alors la matric $\mathbf{A}$ est équivalente à la matrice

$$\mathbf{J}_r = \begin{bmatrix} \mathbf{1}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}.$$

En particulier, deux matrices de $\mathscr{M}_{p,q}(\mathbb{K})$ sont équivalentes si, et seulement si, elles ont même rang.

Soit $u : E \longrightarrow E'$ une application linéaire. On a défini le rang de u comme la dimension du sous-espace Im u. Soient $\mathscr{B}$ et $\mathscr{B}'$ des bases de E et E' respectivement. On a

$$\operatorname{rang} u = \dim \operatorname{Im} u = \dim \operatorname{Vect}(u(\mathscr{B})).$$

Autrement dit le rang de u est le rang de la matrice $[u]_{\mathscr{B}}^{\mathscr{B}'}$.

Exercice 25.— Soient $\mathbf{A}$ et $\mathbf{B}$ deux matrices de $\mathscr{M}_{n,m}(\mathbb{K})$.
1. Montrer que Im $(\mathbf{A}+\mathbf{B})$ est un sous-espace de Im $(\mathbf{A})$ + Im $(\mathbf{B})$.
2. En déduire que $\operatorname{rang}(\mathbf{A}+\mathbf{B}) \leqslant \operatorname{rang}(\mathbf{A}) + \operatorname{rang}(\mathbf{B})$.

Exercice 26.— Soit $\mathbf{A} \in \mathscr{M}_{n,m}(\mathbb{K})$ et $\mathbf{B} \in \mathscr{M}_{n,p}(\mathbb{K})$ deux matrices. Montrer que

$$\operatorname{rang}([\mathbf{A}\ \ \mathbf{B}]) \leqslant \operatorname{rang}(\mathbf{A}) + \operatorname{rang}(\mathbf{B}) - \dim(\operatorname{Im}(\mathbf{A}) \cap \operatorname{Im}(\mathbf{B})).$$

§ 7 Opérations matricielles par blocs

Dans certaines situations, nous rencontrerons des matrices partionnées par blocs. Les opérations sur les matrices définies sur les scalaires, comme l'addition et la multiplication, se généralisent aux matrices partitionnées par blocs.

2.7.1. Matrices par blocs.— Exprimer une matrice par colonnes ou par lignes et un cas particulier de partitionnement de matrice. Plus généralement, soit $\mathbf{A}$ une matrice de $\mathscr{M}_{m,n}(\mathbb{K})$, une partitionner $\mathbf{A}$ consiste à l'écire sous la forme

$$
\begin{array}{cc}
\begin{array}{ccc} n_1 & \cdots & n_k \end{array} & \\
\begin{bmatrix} \mathbf{A}_1^1 & \cdots & \mathbf{A}_1^k \\ \vdots & & \vdots \\ \mathbf{A}_l^1 & \cdots & \mathbf{A}_l^k \end{bmatrix} & \begin{array}{c} m_1 \\ \vdots \\ m_l \end{array}
\end{array}
$$

où $n_1+n_2+\ldots+n_k = n$, $m_1+m_2+\ldots+m_l = m$ et où $\mathbf{A}_i^j$ est une matrice de $\mathscr{M}_{n_j,m_i}(\mathbb{K})$, appelée *bloc* (i,j) ou *sous-matrice* (i,j) de la matrice $\mathbf{A}$.

Les opérations sur les matrices par blocs se définissent comme dans le cas des matrices définies par des scalaires ; il faut cependant tenir compte des compatibilités des tailles des blocs.

2.7.2. Addition par blocs.— Considérons deux matrices par blocs, partitionnées de la même façon :

$$
\mathbf{A} = \begin{array}{cc}
\begin{array}{ccc} n_1 & \cdots & n_k \end{array} & \\
\begin{bmatrix} \mathbf{A}_1^1 & \cdots & \mathbf{A}_1^k \\ \vdots & & \vdots \\ \mathbf{A}_l^1 & \cdots & \mathbf{A}_l^k \end{bmatrix} & \begin{array}{c} m_1 \\ \vdots \\ m_l \end{array}
\end{array}, \qquad
\mathbf{B} = \begin{array}{cc}
\begin{array}{ccc} n_1 & \cdots & n_k \end{array} & \\
\begin{bmatrix} \mathbf{B}_1^1 & \cdots & \mathbf{B}_1^k \\ \vdots & & \vdots \\ \mathbf{B}_l^1 & \cdots & \mathbf{B}_l^k \end{bmatrix} & \begin{array}{c} m_1 \\ \vdots \\ m_l \end{array}
\end{array}.
$$

La somme des matrices $\mathbf{A}$ et $\mathbf{B}$ est une matrice $\mathbf{C}$ de même partion :

$$
\mathbf{C} = \mathbf{A}+\mathbf{B} = \begin{array}{cc}
\begin{array}{ccc} n_1 & \cdots & n_k \end{array} & \\
\begin{bmatrix} \mathbf{C}_1^1 & \cdots & \mathbf{C}_1^k \\ \vdots & & \vdots \\ \mathbf{C}_l^1 & \cdots & \mathbf{C}_l^k \end{bmatrix} & \begin{array}{c} m_1 \\ \vdots \\ m_l \end{array}
\end{array} = \begin{bmatrix} \mathbf{A}_1^1+\mathbf{B}_1^1 & \cdots & \mathbf{A}_1^k+\mathbf{B}_1^k \\ \vdots & & \vdots \\ \mathbf{A}_l^1+\mathbf{B}_l^1 & \cdots & \mathbf{A}_l^k+\mathbf{B}_l^k \end{bmatrix}.
$$

2.7.3. Multiplication par blocs.— La multiplication par blocs est moins évidente. Considérons deux matrices par blocs

$$
\mathbf{A} = \begin{array}{cc}
\begin{array}{ccc} n_1 & \cdots & n_i \end{array} & \\
\begin{bmatrix} \mathbf{A}_1^1 & \cdots & \mathbf{A}_1^i \\ \vdots & & \vdots \\ \mathbf{A}_j^1 & \cdots & \mathbf{A}_j^i \end{bmatrix} & \begin{array}{c} m_1 \\ \vdots \\ m_l \end{array}
\end{array} \quad \text{et} \quad
\mathbf{B} = \begin{array}{cc}
\begin{array}{ccc} p_1 & \cdots & p_k \end{array} & \\
\begin{bmatrix} \mathbf{B}_1^1 & \cdots & \mathbf{B}_1^k \\ \vdots & & \vdots \\ \mathbf{B}_i^1 & \cdots & \mathbf{B}_i^k \end{bmatrix} & \begin{array}{c} n_1 \\ \vdots \\ n_i \end{array}
\end{array}.
$$

Avec ces notations, on montre que (ce n'est pas immédiat)

2.13 Proposition.— Si la matrice produit $\mathbf{C} = \mathbf{AB}$ est partionnée de la façon suivante

$$\mathbf{C} = \begin{array}{c c} \begin{array}{ccc} p_1 & \cdots & p_k \end{array} & \\ \left[\begin{array}{ccc} \mathbf{C}_1^1 & \cdots & \mathbf{C}_1^k \\ \vdots & & \vdots \\ \mathbf{C}_j^1 & \cdots & \mathbf{C}_j^k \end{array}\right] & \begin{array}{c} m_1 \\ \vdots \\ m_l \end{array} \end{array}$$

alors, pour tous $r \in [\![1, l]\!]$ et $s \in [\![1, k]\!]$,

$$\mathbf{C}_r^s = \sum_{t=1}^{i} \mathbf{A}_r^t \mathbf{B}_t^s.$$

En particulier, on a par exemple,

$$\mathbf{C}_1^1 = \mathbf{A}_1^1\mathbf{B}_1^1 + \mathbf{A}_1^2\mathbf{B}_2^1 + \ldots + \mathbf{A}_1^i\mathbf{B}_i^1.$$

2.7.4. Exemples.—

a) Soient $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ dans $\mathscr{M}_{m,n_1}(\mathbb{K})$, $\mathscr{M}_{m,n_2}(\mathbb{K})$, $\mathscr{M}_{m,n_3}(\mathbb{K})$ respectivement et $\mathbf{D}$, $\mathbf{E}$, $\mathbf{F}$ dans $\mathscr{M}_{n_1,p}(\mathbb{K})$, $\mathscr{M}_{n_2,p}(\mathbb{K})$, $\mathscr{M}_{n_3,p}(\mathbb{K})$ respectivement, alors

$$\left[\begin{array}{ccc} \mathbf{A} & \mathbf{B} & \mathbf{C} \end{array}\right] \left[\begin{array}{c} \mathbf{D} \\ \mathbf{E} \\ \mathbf{F} \end{array}\right] = \mathbf{AD} + \mathbf{BE} + \mathbf{CF}.$$

b) Soient $\mathbf{A} \in \mathscr{M}_{m_1,n_1}(\mathbb{K})$, $\mathbf{B} \in \mathscr{M}_{m_1,n_2}(\mathbb{K})$, $\mathbf{C} \in \mathscr{M}_{m_2,n_1}(\mathbb{K})$, $\mathbf{D} \in \mathscr{M}_{m_2,n_2}(\mathbb{K})$ et $\mathbf{E} \in \mathscr{M}_{n_1,p_1}(\mathbb{K})$, $\mathbf{F} \in \mathscr{M}_{n_1,p_2}(\mathbb{K})$, $\mathbf{G} \in \mathscr{M}_{n_2,p_1}(\mathbb{K})$, $\mathbf{H} \in \mathscr{M}_{n_2,p_2}(\mathbb{K})$, alors

$$\left[\begin{array}{cc} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{array}\right] \left[\begin{array}{cc} \mathbf{E} & \mathbf{F} \\ \mathbf{G} & \mathbf{H} \end{array}\right] = \left[\begin{array}{cc} \mathbf{AE} + \mathbf{BG} & \mathbf{AF} + \mathbf{BH} \\ \mathbf{CE} + \mathbf{DG} & \mathbf{CF} + \mathbf{DH} \end{array}\right]$$

c) Soient

$$\mathbf{A} = \begin{array}{cc} \begin{array}{cc} p_1 & p_2 \end{array} & \\ \left[\begin{array}{cc} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{array}\right] & \begin{array}{c} n_1 \\ n_2 \end{array} \end{array} \quad \text{et} \quad \mathbf{x} = \begin{array}{cc} 1 & \\ \left[\begin{array}{c} \mathbf{x}_1 \\ \mathbf{x}_2 \end{array}\right] & \begin{array}{c} p_1 \\ p_2 \end{array} \end{array}$$

une matrice par blocs de $\mathscr{M}_{n_1+n_2,p_1+p_2}(\mathbb{K})$ et un vecteur de $\mathbb{K}^{p_1+p_2}$. On a

$$\left[\begin{array}{cc} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{array}\right] \left[\begin{array}{c} \mathbf{x}_1 \\ \mathbf{x}_2 \end{array}\right] = \left[\begin{array}{c} \mathbf{A}\mathbf{x}_1 + \mathbf{B}\mathbf{x}_2 \\ \mathbf{C}\mathbf{x}_1 + \mathbf{D}\mathbf{x}_2 \end{array}\right].$$

Exercice 27.— Soient $\mathbf{A} \in \mathscr{M}_{n,m}(\mathbb{K})$ et $\mathbf{B} \in \mathscr{M}_{m,p}(\mathbb{K})$ deux matrices partionnées en blocs de la façon suivante

$$\mathbf{A} = \left[\begin{array}{cc} \mathbf{C} & \mathbf{1}_2 \\ \mathbf{1}_2 & \mathbf{0} \end{array}\right] \qquad \mathbf{B} = \left[\begin{array}{cc} \mathbf{1}_2 & \mathbf{0} \\ \mathbf{C} & \mathbf{C} \end{array}\right],$$

où

$$\mathbf{C} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}.$$

Exprimer le produit **AB** par blocs, puis calculer **AB**.

Exercice 28.— Soient $\mathbf{A} \in \mathscr{M}_{n,n}(\mathbb{K})$, $\mathbf{B} \in \mathscr{M}_{m,m}(\mathbb{K})$ et $\mathbf{C} \in \mathscr{M}_{m,n}(\mathbb{K})$, telles que les matrices **A** et **B** sont inversibles. Montrer que les matrices suivantes sont inversibles et calculer leur inverse

$$\begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{B} \end{bmatrix}, \qquad \begin{bmatrix} \mathbf{A} & \mathbf{C} \\ \mathbf{0} & \mathbf{B} \end{bmatrix}.$$

—

§ 8 Exercices

Exercice 29.— Parmi les sous-ensembles de matrices de $\mathscr{M}_n(\mathbb{K})$ suivants, lesquels sont des sous-espaces vectoriels :

a) les matrices symétriques,
b) les matrices diagonales,
c) les matrices inversibles,
d) les matrices non inversibles,
e) les matrices triangulaires,
f) les matrices triangulaires supérieures,
g) les matrices qui commutent avec une matrice donnée **A**,
h) les matrices **A** telles que $\mathbf{A}^2 = \mathbf{A}$,
i) les matrices de trace nulle.

Exercice 30.— Donner les dimensions des sous-espaces vectoriels de $\mathscr{M}_n(\mathbb{K})$ suivants

a) le sous-espace vectoriel des matrices à coefficients constants,
b) le sous-espace vectoriel des matrices diagonales,
c) le sous-espace vectoriel des matrices symétriques,
d) le sous-espace vectoriel des matrices antisymétriques.

Exercice 31.— Trouver des bases de l'espace des lignes et de l'espace des colonnes de la matrice

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & -1 & 3 \\ 5 & 3 & 8 \end{bmatrix}.$$

Quelles sont les dimensions des noyaux de l'application linéaire représentée par **A** et de celle représentée par sa transposée. Déterminer des bases de ces noyaux.

Exercice 32.— Soit u l'application de $\mathbb{R}^3$ à valeurs dans $\mathbb{R}^3$ définie par

$$u\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x-y \\ y-x \\ x-z \end{bmatrix}.$$

1. Montrer que u est une application linéaire.
2. Exprimer la matrice de u dans la base suivante de $\mathbb{R}^3$:

$$\mathscr{B} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \right\}$$

3. Écrire la matrice de passage de la base $\mathscr{B}$ à la base canonique de $\mathbb{R}^3$.
4. Donner la matrice de u exprimée dans la base canonique.

Exercice 33.— Mêmes questions avec l'application u définie par

$$u\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x+y \\ y+x \\ x+z \end{bmatrix}.$$

Exercice 34.— Pour chacune des applications de $\mathbb{R}^3$ à valeurs dans $\mathbb{R}^3$ définies par

$$\mathbf{a})\ u\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2x-y \\ -x+2y-z \\ z-y \end{bmatrix}, \qquad \mathbf{b})\ u\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x-2y \\ x-2z \\ 2y+z \end{bmatrix},$$

$$\mathbf{c})\ u\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x-y \\ y-z \\ x+z \end{bmatrix}, \qquad \mathbf{d})\ u\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x-y+z \\ x+y-z \\ x+y+z \end{bmatrix}.$$

1. Montrer que u est une application linéaire.
2. Exprimer la matrice de u dans la base canonique de $\mathbb{R}^3$.
3. Montrer que l'application linéaire u est inversible.
4. Déterminer $u^{-1}(x,y,z)$, pour tout vecteur (x,y,z) de $\mathbb{R}^3$.

Exercice 35.— Soit $\mathscr{T}^s$ (resp. $\mathscr{T}^i$) l'ensemble des matrices triangulaires supérieures (resp. inférieures) de $\mathscr{M}_n(\mathbb{K})$.
1. Montrer que $\mathscr{T}^s$ et $\mathscr{T}^i$ sont des sous-espaces vectoriels de $\mathscr{M}_n(\mathbb{K})$.
2. Montrer que $\mathscr{M}_n(\mathbb{K}) = \mathscr{T}^s + \mathscr{T}^i$. Cette somme est-elle directe ?
3. Quels sont les dimensions des sous-espaces $\mathscr{T}^s$ et $\mathscr{T}^i$?

Exercice 36.— Montrer que toute matrice triangulaire supérieure est semblable à une matrice triangulaire inférieure.

Exercice 37.— Soit u l'endomorphisme de $\mathbb{R}^2$ représenté dans une base $\mathscr{B}$ de $\mathbb{R}^2$ par la matrice

$$[u]_{\mathscr{B}} = \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix}.$$

1. Montrer que $u^2 = 0$.
2. Déterminer le rang, l'image et le noyau de u.
3. Montrer qu'il existe une base de $\mathbb{R}^2$ dans laquelle la matrice de u est

$$\begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix}.$$

Exercice 38.— Soit $\mathbf{A} \in \mathscr{M}_2(\mathbb{K})$.

1. Montrer que

$$\mathbf{A}^2 - \operatorname{tr}(\mathbf{A})\mathbf{A} + \det(\mathbf{A})\mathbf{1}_2 = \mathbf{0}.$$

2. On suppose que le déterminant de $\mathbf{A}$ est non nul. Calculer l'inverse de $\mathbf{A}$.

Exercice 39.— Soient $\mathbf{A} \in \mathscr{M}_n(\mathbb{K})$ et $\lambda, \mu \in \mathbb{K}$ tels que

$$\mathbf{A}^2 = \lambda\mathbf{A} + \mu\mathbf{1}_n.$$

1. Montrer que si μ est non nul, la matrice $\mathbf{A}$ est inversible et que

$$\mathbf{A}^{-1} = \mu^{-1}(\mathbf{A} - \lambda\mathbf{1}_n).$$

2. Montrer que pour tout entier k, la matrice $\mathbf{A}^k$ s'écrit comme une combinaison linéaire des matrices $\mathbf{A}$ et $\mathbf{1}_n$.

Exercice 40.— Soient E un $\mathbb{K}$-espace vectoriel et u un endomorphisme non nul de E tel que $u^2 = 0$.

1. Montrer que $\operatorname{Im} u \subset \operatorname{Ker} u$.
2. Déterminer la dimension du noyau de u et de l'image de u.
3. Montrer qu'il existe une base de $\mathbb{R}^3$ dans laquelle l'endomorphisme u a pour matrice

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

CHAPITRE
3

Les déterminants

Sommaire

Dans tout ce chapitre, $\mathbb{K}$ désignera un corps commutatif.

§ 1 Définition récursive du déterminant

3.1.1. Notation.— Soit $\mathbf{A} = [a_i^j]$ une matrice de $\mathscr{M}_n(\mathbb{K})$ et soient $i, j \in [\![1, n]\!]$. On notera $\mathbf{A}_{\check{i}\check{j}}$ la matrice obtenue en supprimant la i-ème ligne et j-ème colonne de la matrice $\mathbf{A}$:

$$\mathbf{A}_{\check{i}\check{j}} = \begin{bmatrix} a_1^1 & \cdots & a_1^j & \cdots & a_1^n \\ \vdots & & \vdots & & \vdots \\ a_i^1 & \cdots & a_i^j & \ldots & a_i^n \\ \vdots & & \vdots & & \vdots \\ a_n^1 & \cdots & a_n^j & \cdots & a_n^n \end{bmatrix} \in \mathscr{M}_{n-1}(\mathbb{K})$$

3.1.2. Définition.— Soit $\mathbf{A} = [a_i^j]$ une matrice de $\mathscr{M}_n(\mathbb{K})$. On définit une application

$$\det : \mathscr{M}_n(\mathbb{K}) \longrightarrow \mathbb{K}$$

par récurrence sur l'entier n en posant, pour $n = 1$,

$$\det \mathbf{A} = a_1^1,$$

et, pour tout $n > 1$,

$$\det \mathbf{A} = \sum_{k=1}^{n} (-1)^{k+1} a_1^k \det \mathbf{A}_{\check{1}\check{k}}.$$

Le scalaire $\det \mathbf{A}$ ainsi défini est appelé le *déterminant* de la matrice $\mathbf{A}$

3.1.3. Notation.— Le déterminant d'une matrice $\begin{bmatrix} a_1^1 & \dots & a_1^n \\ \vdots & & \vdots \\ a_n^1 & \dots & a_n^n \end{bmatrix}$ sera aussi noté

$$\begin{vmatrix} a_1^1 & \dots & a_1^n \\ \vdots & & \vdots \\ a_n^1 & \dots & a_n^n \end{vmatrix}.$$

3.1.4. Exemples.— Pour une matrice de $\mathscr{M}_2(\mathbb{K})$, on a

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc.$$

Pour une matrice de $\mathscr{M}_3(\mathbb{K})$, on a

$$\begin{aligned}
\begin{vmatrix} a_1^1 & a_1^2 & a_1^3 \\ a_2^1 & a_2^2 & a_2^3 \\ a_3^1 & a_3^2 & a_3^3 \end{vmatrix} &= a_1^1 \begin{vmatrix} a_2^2 & a_2^3 \\ a_3^2 & a_3^3 \end{vmatrix} - a_1^2 \begin{vmatrix} a_2^1 & a_2^3 \\ a_3^1 & a_3^3 \end{vmatrix} + a_1^3 \begin{vmatrix} a_2^1 & a_2^2 \\ a_3^1 & a_3^2 \end{vmatrix} \\
&= a_1^1(a_2^2a_3^3 - a_3^2a_2^3) - a_1^2(a_2^1a_3^3 - a_3^1a_2^3) + a_1^3(a_2^1a_3^2 - a_3^1a_2^2) \\
&= a_1^1a_2^2a_3^3 - a_1^1a_3^2a_2^3 - a_1^2a_2^1a_3^3 + a_1^2a_3^1a_2^3 + a_1^3a_2^1a_3^2 - a_1^3a_3^1a_2^2.
\end{aligned}$$

Exercice 1.— Montrer que le déterminant d'une matrice triangulaire est égal au produit des coefficients de la diagonale.

3.1.5. Déterminant d'une famille de vecteurs.— Soient E un $\mathbb{K}$-espace vectoriel de dimension n et $\mathscr{B} = (\mathbf{e}_1, \dots, \mathbf{e}_n)$ une base de E.

Étant donnée une famille $(\mathbf{x}_1, \dots, \mathbf{x}_n)$ de vecteurs de E, on forme la matrice de $\mathscr{M}_n(\mathbb{K})$, notée

$$[\mathbf{x}_1 \,|\, \mathbf{x}_2 \,|\, \dots \,|\, \mathbf{x}_n]_{\mathscr{B}},$$

dont la j-ème colonne est constituée des coordonnées du vecteur $\mathbf{x}_j$ exprimé dans la base $\mathscr{B}$.

Le scalaire $\det[\,\mathbf{x}_1\,|\,\mathbf{x}_2\,|\,\ldots\,|\,\mathbf{x}_n\,]_{\mathscr{B}}$ est appelé le *déterminant de la famille de vecteurs par rapport à la base* $\mathscr{B}$. On le note $\det_{\mathscr{B}}(\mathbf{x}_1,\ldots,\mathbf{x}_n)$.

§ 2 Premières propriétés du déterminant

Soit $\mathbf{A} = [a_i^j]$ une matrice de $\mathscr{M}_n(\mathbb{K})$. On notera $\mathbf{a}_*^j$ la j-ème colonne de la matrice $\mathbf{A}$. Avec cette notation, on peut écrire la matrice $\mathbf{A}$ par colonnes sous la forme

$$\mathbf{A} = \left[\ \mathbf{a}_*^1 \mid \mathbf{a}_*^2 \mid \ldots \mid \mathbf{a}_*^n\ \right].$$

FIGURE 3.1.: Gottfried Wilhelm Leibniz (1646 - 1716)

Gottfried Wilhelm Leibniz est un mathématicien, philosophe, juriste, diplomate allemand. En mathématiques, l'influence de Leibniz est considérable, dans le domaine de l'algébrisation de la géométrie, en calcul différentiel, calcul infinitésimal. Leibniz travaille aussi sur la résolution des systèmes d'équations. Il introduit la notion de déterminant d'un système d'équation. Ses travaux sur les déterminants ne sont pas publiés, mais se retrouvent dans une lettre au mathématicien Guillaume François Antoine de L'Hôpital datée du 28 avril 1693. Dans cette lettre, il écrit la condition de compatibilité d'un système de trois équations linéaires à deux inconnues sous forme de déter-

minant. Le terme de « déterminant »ne sera introduit que plus tard, par Carl Friedrich Gauss en 1801 dans l'ouvrage « Disquisitiones Arithmeticae ».

3.1 Proposition.— Le déterminant est une application linéaire par rapport à chaque colonne, i.e., pour toute matrice $\mathbf{A} = [a_i^j]$ de $\mathscr{M}_n(\mathbb{K})$, on a, pour tout $j \in [\![1,n]\!]$,

i) si $\mathbf{a}_*^j = \mathbf{b}_*^j + \mathbf{c}_*^j$,

$$\det \left[\begin{array}{c|c|c|c|c} \mathbf{a}_*^1 & \ldots & \mathbf{b}_*^j + \mathbf{c}_*^j & \ldots & \mathbf{a}_*^n \end{array} \right] = \det \left[\begin{array}{c|c|c|c|c} \mathbf{a}_*^1 & \ldots & \mathbf{b}_*^j & \ldots & \mathbf{a}_*^n \end{array} \right] + \det \left[\begin{array}{c|c|c|c|c} \mathbf{a}_*^1 & \ldots & \mathbf{c}_*^j & \ldots & \mathbf{a}_*^n \end{array} \right]$$

ii) pout tout scalaire $\alpha \in \mathbb{K}$,

$$\det \left[\begin{array}{c|c|c|c|c} \mathbf{a}_*^1 & \ldots & \alpha\, \mathbf{a}_*^j & \ldots & \mathbf{a}_*^n \end{array} \right] = \alpha \det \left[\begin{array}{c|c|c|c|c} \mathbf{a}_*^1 & \ldots & \mathbf{a}_*^j & \ldots & \mathbf{a}_*^n \end{array} \right]$$

Preuve. On montre les assertions **i)** et **ii)** par récurrence sur n. Si $n = 1$, les deux assertions sont immédiates. Si $n = 2$, on vérifie que

$$\begin{aligned} \begin{vmatrix} b_1^1 + c_1^1 & a_1^2 \\ b_2^1 + c_2^1 & a_2^2 \end{vmatrix} &= (b_1^1 + c_1^1)a_2^2 - (b_2^1 + c_2^1)a_1^2 \\ &= (b_1^1 a_2^2 - b_2^1 a_1^2) + (c_1^1 a_2^2 - c_2^1 a_1^2) \\ &= \begin{vmatrix} b_1^1 & a_1^2 \\ b_2^1 & a_2^2 \end{vmatrix} + \begin{vmatrix} c_1^1 & a_1^2 \\ c_2^1 & a_2^2 \end{vmatrix}. \end{aligned}$$

Pour tout $\alpha \in \mathbb{K}$, on a

$$\begin{aligned} \begin{vmatrix} \alpha a_1^1 & a_1^2 \\ \alpha a_2^1 & a_2^2 \end{vmatrix} &= (\alpha a_1^1)a_2^2 - (\alpha a_2^1)a_1^2 \\ &= \alpha(a_1^1 a_2^2 - a_2^1 a_1^2) \\ &= \alpha \begin{vmatrix} a_1^1 & a_1^2 \\ a_2^1 & a_2^2 \end{vmatrix}. \end{aligned}$$

On montre de façon analogue la linéarité du déterminant par rapport à la seconde colonne.

Supposons les assertions **i)** et **ii)** vraies pour toute matrice de $\mathscr{M}_{n-1}(\mathbb{K})$ et considérons une matrice $\mathbf{A} = \left[\begin{array}{c|c|c|c} \mathbf{a}_*^1 & \mathbf{a}_*^2 & \ldots & \mathbf{a}_*^n \end{array} \right]$ de $\mathscr{M}_n(\mathbb{K})$. Soit $j \in [\![1,n]\!]$, montrons que $\det \mathbf{A}$ est linéaire par rapport à la j-ème colonne.

Par définition, on a

$$\det \mathbf{A} = \sum_{k=1}^{n} (-1)^{k+1} a_1^k \det \mathbf{A}_{\check{1}\check{k}}. \tag{3.1}$$

Lorsque k est différent de j, alors la matrice $\mathbf{A}_{\check{1}\check{k}}$ contient la j-ème colonne de $\mathbf{A}$. Comme $\mathbf{A}_{\check{1}\check{k}} \in \mathscr{M}_{n-1}(\mathbb{K})$, par hypothèse de récurrence, $\det \mathbf{A}_{\check{1}\check{k}}$ dépend linéairement de la j-ème colonne de $\mathbf{A}$. Comme a_1^k ne dépend pas de la j-ème colonne, on en déduit que $a_1^k \det \mathbf{A}_{\check{1}\check{k}}$ dépend linéairement de la j-ème colonne.

Dans le cas où k est égal à j. Alors, la mice $\mathbf{A}_{\check{1}\check{k}}$ ne contient pas la j-ème colonne et $a_1^k = a_1^j$ dépend linéairement de la j-ème colonne de $\mathbf{A}$. Ainsi, $a_1^k \det \mathbf{A}_{\check{1}\check{k}}$ dépend linéairement de la j-ème colonne de $\mathbf{A}$.

On en conclue que tous les termes de la somme (3.1) dépendent linéairement de la j-ème colonne de $\mathbf{A}$, par suite $\det \mathbf{A}$ dépend linéairement de la j-ème colonne de $\mathbf{A}$. □

3.2 Proposition.— Le déterminant d'une matrice dont deux colonnes sont égales est nul.

Preuve. Nous allons montrer ce résultat à partir des deux assertions suivantes

i) *u*ne matrice qui possède deux colonnes adjacentes égales est de déterminant nul,

ii) *l*e déterminant change de signe, lorsque l'on échange deux colonnes adjacentes.

De ces deux assertions, nous déduisons la proposition. En effet, si une matrice $\mathbf{A}$ possède deux colonnes quelconques égales. D'après l'assertion **ii)**, par échange successifs, on peut faire en sorte que les deux colonnes égales soient adjacentes, quitte à changer le signe du déterminant de la matrice ainsi obtenue. De l'assertion **i)**, on en déduit que le déterminant de $\mathbf{A}$ est alors nul.

Reste à montrer les assertions **i)** et **ii)**. On procède par récurrence sur la dimension de la matrice pour l'assertion **i)**. Si $n = 2$, alors

$$\begin{vmatrix} a & a \\ c & c \end{vmatrix} = ac - ac = 0.$$

Supposons l'assertions **i)** vraie pour toute matrice de $\mathscr{M}_{n-1}(\mathbb{K})$ et considérons une matrice $\mathbf{A} = [a_i^j]$ de $\mathscr{M}_n(\mathbb{K})$ ayant deux colonnes adjacentes égales, i.e., $\mathbf{A}_*^j = \mathbf{A}_*^{j+1}$ pour un $j \in [\![1,n]\!]$. Par définition, on a

$$\det \mathbf{A} = \sum_{k=1}^{n} (-1)^{k+1} a_1^k \det \mathbf{A}_{\check{1}\check{k}}.$$

Or, si k est différent de j et de $j+1$, alors la matrice $\mathbf{A}_{\check{1}\check{k}}$ possède deux colonnes identiques. Par hypothèse de récurrence, on a donc $\det \mathbf{A}_{\check{1}\check{k}} = 0$. Ainsi,

$$\det \mathbf{A} = (-1)^{j+1} a_1^j \det \mathbf{A}_{\check{1}\check{j}} + (-1)^{j+2} a_1^{j+1} \det \mathbf{A}_{\check{1}(\check{j+1})}.$$

Or $\mathbf{A}_*^j = \mathbf{A}_*^{j+1}$, donc $a_1^j = a_1^{j+1}$ et les matrice $\mathbf{A}_{\check{1}\check{j}}$ et $\mathbf{A}_{\check{1}(\check{j+1})}$ sont égales. D'où $\det \mathbf{A} = 0$.

Pour montrer l'assertion **ii)**, considérons une matrice $\mathbf{A} = \left[\ \mathbf{a}_*^1 \mid \mathbf{a}_*^2 \mid \dots \mid \mathbf{a}_*^n\ \right]$ de $\mathscr{M}_n(\mathbb{K})$. D'après **i)**, on a pour tout $j \in [\![1,n]\!]$,

$$\det \left[\ \mathbf{a}_*^1 \mid \dots \mid \mathbf{a}_*^j + \mathbf{a}_*^{j+1} \mid \mathbf{a}_*^j + \mathbf{a}_*^{j+1} \mid \dots \mid \mathbf{a}_*^n\ \right] = 0.$$

Par linéarité, on en déduit que

$$\det\left[\ \mathbf{a}^1_* \mid \ldots \mid \mathbf{a}^j_* \mid \mathbf{a}^j_* \mid \ldots \mid \mathbf{a}^n_*\ \right] + \det\left[\ \mathbf{a}^1_* \mid \ldots \mid \mathbf{a}^j_* \mid \mathbf{a}^{j+1}_* \mid \ldots \mid \mathbf{a}^n_*\ \right]$$
$$+\det\left[\ \mathbf{a}^1_* \mid \ldots \mid \mathbf{a}^{j+1}_* \mid \mathbf{a}^j_* \mid \ldots \mid \mathbf{a}^n_*\ \right] + \det\left[\ \mathbf{a}^1_* \mid \ldots \mid \mathbf{a}^{j+1}_* \mid \mathbf{a}^{j+1}_* \mid \ldots \mid \mathbf{a}^n_*\ \right] = 0.$$

Toujours, d'après l'assertion **i)**, on en déduit que

$$\det\left[\ \mathbf{a}^1_* \mid \ldots \mid \mathbf{a}^j_* \mid \mathbf{a}^{j+1}_* \mid \ldots \mid \mathbf{a}^n_*\ \right] = -\det\left[\ \mathbf{a}^1_* \mid \ldots \mid \mathbf{a}^{j+1}_* \mid \mathbf{a}^j_* \mid \ldots \mid \mathbf{a}^n_*\ \right],$$

ce qui montre **ii)**. □

3.3 Proposition.— Le déterminant d'une matrice change de signe, lorsque l'on échange deux colonnes.

Preuve. Ce résultat est une conséquence immédiate du résultat précédent. Considérons une matrice $\mathbf{A} = [\ \mathbf{a}^1_* \mid \mathbf{a}^2_* \mid \ldots \mid \mathbf{a}^n_*\ ]$ de $\mathscr{M}_n(\mathbb{K})$. Montrons que le déterminant de la matrice $\mathbf{A}$ change de signe lorsque l'on échange les j-ème colonne et k-ème colonne. D'après la proposition 3.2, on a

$$\det\left[\ \mathbf{a}^1_* \mid \ldots \mid \mathbf{a}^j_* + \mathbf{a}^k_* \mid \ldots \mid \mathbf{a}^j_* + \mathbf{a}^k_* \mid \ldots \mid \mathbf{a}^n_*\ \right] = 0.$$

Par linéarité, on en déduit que

$$\det\left[\ \mathbf{a}^1_* \mid \ldots \mid \mathbf{a}^j_* \mid \ldots \mid \mathbf{a}^j_* \mid \ldots \mid \mathbf{a}^n_*\ \right] + \det\left[\ \mathbf{a}^1_* \mid \ldots \mid \mathbf{a}^j_* \mid \ldots \mid \mathbf{a}^k_* \mid \ldots \mid \mathbf{a}^n_*\ \right]$$
$$+\det\left[\ \mathbf{a}^1_* \mid \ldots \mid \mathbf{a}^k_* \mid \ldots \mid \mathbf{a}^j_* \mid \ldots \mid \mathbf{a}^n_*\ \right] + \det[\ \mathbf{a}^1_* \mid \ldots \mid \mathbf{a}^k_* \mid \ldots \mid \mathbf{a}^k_* \mid \ldots \mid \mathbf{a}^n_*\ ] = 0.$$

D'où le résultat attendu :

$$\det\left[\ \mathbf{a}^1_* \mid \ldots \mid \mathbf{a}^j_* \mid \ldots \mid \mathbf{a}^k_* \mid \ldots \mid \mathbf{a}^n_*\ \right] = -\det\left[\ \mathbf{a}^1_* \mid \ldots \mid \mathbf{a}^k_* \mid \ldots \mid \mathbf{a}^j_* \mid \ldots \mid \mathbf{a}^n_*\ \right].$$

□

3.4 Théorème.— Soit $\mathbf{A} = [\ \mathbf{a}^1_* \mid \mathbf{a}^2_* \mid \ldots \mid \mathbf{a}^n_*\ ]$ une matrice de $\mathscr{M}_n(\mathbb{K})$. La famille de vecteurs $(\mathbf{a}^1_*, \mathbf{a}^2_*, \ldots, \mathbf{a}^n_*)$ forme une base de $\mathbb{K}^n$ si, et seulement si, $\det \mathbf{A}$ est non nul.

Preuve. Supposons que la famille $(\mathbf{a}^1_*, \mathbf{a}^2_*, \ldots, \mathbf{a}^n_*)$ soit liée et montrons que le déterminant $\det \mathbf{A}$ est nul. Par exemple, supposons que la première colonne soit combinaison linéaire des autres colonnes :

$$\mathbf{a}_*^1 = \sum_{k=2}^{n} \alpha_k \, \mathbf{a}_*^k.$$

Alors,

$$\begin{aligned}
\det \mathbf{A} &= \det \left[\sum_{k=2}^{n} \alpha_k \, \mathbf{a}_*^k \;\middle|\; \mathbf{a}_*^2 \;\middle|\; \dots \;\middle|\; \mathbf{a}_*^n \right] \\
&= \sum_{k=2}^{n} \alpha_k \det \left[\, \mathbf{a}_*^k \mid \mathbf{a}_*^2 \mid \dots \mid \mathbf{a}_*^n \, \right]
\end{aligned}$$

Or pour tout $k \in [\![2,n]\!]$, la matrice $\left[\, \mathbf{a}_*^k \mid \mathbf{a}_*^2 \mid \dots \mid \mathbf{a}_*^n \, \right]$ possède deux colonnes égales. De la proposition 3.2, on déduit que $\det \mathbf{A} = 0$.

Inversement, supposons que la famille $\mathscr{B} = (\mathbf{a}_*^1, \mathbf{a}_*^2, \dots, \mathbf{a}_*^n)$ forme une base de l'espace $\mathbb{K}^n$, montrons par l'absurde que le déterminant de $\mathbf{A}$ est non nul.

Supposons que $\det \mathbf{A} = 0$. Considérons une famille quelconque de vecteurs $(\mathbf{x}_1, \dots, \mathbf{x}_n)$ de $\mathbb{K}^n$. Pour tout $j \in [\![1,n]\!]$, le vecteur $\mathbf{x}_j$ se décompose de façon unique dans la base $\mathscr{B}$ en une combinaison linéaire de colonnes de $\mathbf{A}$:

$$\mathbf{x}_j = \sum_{k=1}^{n} \alpha_j^k \mathbf{a}_*^k.$$

On a

$$\begin{aligned}
\det \left[\, \mathbf{x}_1 \mid \mathbf{x}_2 \mid \dots \mid \mathbf{x}_n \, \right] &= \det \left[\sum_{k_1=1}^{n} \alpha_1^{k_1} \mathbf{a}_*^{k_1} \;\middle|\; \sum_{k_2=1}^{n} \alpha_2^{k_2} \mathbf{a}_*^{k_2} \;\middle|\; \dots \;\middle|\; \sum_{k_n=1}^{n} \alpha_n^{k_n} \mathbf{a}_*^{k_n} \right] \\
&= \sum_{k_1=1}^{n} \sum_{k_2=1}^{n} \dots \sum_{k_n=1}^{n} \alpha_1^{k_1} \alpha_2^{k_2} \dots \alpha_n^{k_n} \det \left[\, \mathbf{a}_*^{k_1} \mid \mathbf{a}_*^{k_2} \mid \dots \mid \mathbf{a}_*^{k_n} \, \right]
\end{aligned}$$

Si un terme de la somme possède deux indices k_i égaux, alors

$$\det \left[\, \mathbf{a}_{k_1} \mid \mathbf{a}_{k_2} \mid \dots \mid \mathbf{a}_{k_n} \, \right] = 0$$

d'après la proposition 3.2. Si un terme ne possède que des indices différents, alors le déterminant $\det \left[\, \mathbf{a}_{k_1} \mid \mathbf{a}_{k_2} \mid \dots \mid \mathbf{a}_{k_n} \, \right]$ est égal, au signe près, au déterminant $\det \mathbf{A}$, supposé nul. Ainsi, pour toute famille de vecteur de $\mathbb{K}^n$, on a

$$\det \left[\, \mathbf{x}_1 \mid \mathbf{x}_2 \mid \dots \mid \mathbf{x}_n \, \right] = 0.$$

Ce qui est absurde, car si $\mathscr{C}an = (\mathbf{c}_1, \dots, \mathbf{c}_n)$ désigne la base canonique de $\mathbb{K}^n$, on a

$$\det \left[\, \mathbf{c}_1 \mid \dots \mid \mathbf{c}_n \, \right] = \det \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{bmatrix} = 1.$$

Le déterminant $\det \mathbf{A}$ est donc non nul. $\square$

FIGURE 3.2.: Gabriel Cramer (1704 - 1752)

Gabriel Cramer est un mathématicien suisse, il travaille notamment sur les courbes algébriques. Il publie en 1750, « Introduction à l'analyse des lignes courbes algébriques ». C'est la première publication en Europe sur les déterminants. Cramer donne un formulation basée sur les déterminants pour la solution au problème de la construction de la conique passant par cinq points donnés, problème qu'il réduit à un système d'équations linéaires.

§ 3 Les formules de Cramer

Soit $\mathbf{A} = \left[\ \mathbf{a}_*^1 \mid \mathbf{a}_*^2 \mid \ldots \mid \mathbf{a}_*^n\ \right]$ une matrice de $\mathscr{M}_n(\mathbb{K})$. Si son déterminant $\det \mathbf{A}$ est non nul, nous allons montrer qu'il existe une formule explicite basée sur le déterminant de $\mathbf{A}$ pour les solutions de l'équation

$$\mathbf{A}\mathbf{x} = \mathbf{b}.$$

Posons

$$\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \quad \text{et} \quad \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$$

Le produit $\mathbf{A}\mathbf{x}$ est une combinaison linéaire de colonnes de $\mathbf{A}$, à coefficients les coefficients du vecteur $\mathbf{x}$:

$$\mathbf{A}\mathbf{x} = x_1\mathbf{a}_*^1 + x_2\mathbf{a}_*^2 + \ldots + x_n\mathbf{a}_*^n.$$

Ainsi, l'équation $\mathbf{Ax} = \mathbf{B}$ s'écrit sous la forme

$$x_1 \begin{bmatrix} a_1^1 \\ \vdots \\ a_n^1 \end{bmatrix} + x_2 \begin{bmatrix} a_1^2 \\ \vdots \\ a_n^2 \end{bmatrix} + \ldots + x_n \begin{bmatrix} a_1^n \\ \vdots \\ a_n^n \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$$

Nous souhaitons déterminer les inconnues $x_1, \ldots, x_n$. Soit x_i l'une d'entre elles. Dans l'équation précédente, on soustrait le vecteur $\mathbf{b}$ au i-ème vecteur colonne, on obtient

$$x_1 \begin{bmatrix} a_1^1 \\ \vdots \\ a_n^1 \end{bmatrix} + x_2 \begin{bmatrix} a_1^2 \\ \vdots \\ a_n^2 \end{bmatrix} + \ldots + x_i \begin{bmatrix} a_1^i - b_1 \\ \vdots \\ a_n^i - b_n \end{bmatrix} + \ldots + x_n \begin{bmatrix} a_1^n \\ \vdots \\ a_n^n \end{bmatrix} = \mathbf{0}.$$

Par suite, les colonnes de la matrice suivante sont linéairement dépendantes :

$$\begin{bmatrix} a_1^1 & \cdots & (x_i a_1^i - b_1) & \cdots & a_1^n \\ \vdots & & & & \\ a_n^1 & \cdots & (x_i a_n^i - b_n) & \cdots & a_n^n \end{bmatrix}$$

Du théorème 3.4, on déduit alors que le déterminant de cette matrice est nul, soit

$$x_i \det \begin{bmatrix} a_1^1 & \cdots & a_1^i & \cdots & a_1^n \\ \vdots & & & & \\ a_n^1 & \cdots & a_n^i & \cdots & a_n^n \end{bmatrix} - \det \begin{bmatrix} a_1^1 & \cdots & b_1 & \cdots & a_1^n \\ \vdots & & & & \\ a_n^1 & \cdots & b_n & \cdots & a_n^n \end{bmatrix} = 0.$$

La première matrice est la matrice $\mathbf{A}$, la seconde est la matrice $\mathbf{A}$ dans laquelle on a remplacé la i-ème colonne par le vecteur colonne $\mathbf{b}$. On a ainsi montré

3.5 Théorème (Formules de Cramer).— Si $\det \mathbf{A}$ est non nul, alors l'équation

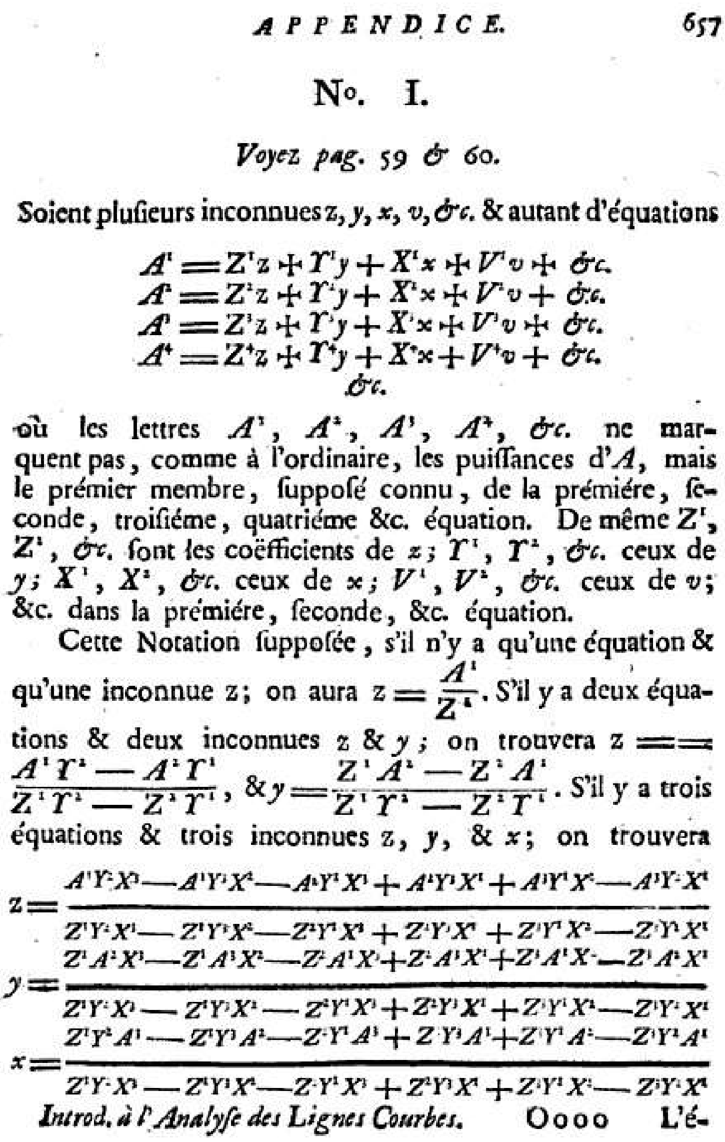
$$\mathbf{A} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$$

admet pour solutions

$$x_i = \frac{1}{\det \mathbf{A}} \det \begin{bmatrix} a_1^1 & \cdots & b_1 & \cdots & a_1^n \\ \vdots & & & & \\ a_n^1 & \cdots & b_n & \cdots & a_n^n \end{bmatrix},$$

pour tout $i \in [\![1, n]\!]$.

Cette formule n'est pas toujours très utilisable en pratique, les méthodes de résolution de systèmes linéaires par élimination gaussienne sont plus efficaces. Les formules de Cramer gardent un intérêt théorique qui sera exploité plus loin.



APPENDICE. 657

No. I.

Voyez pag. 59 & 60.

Soient plusieurs inconnues z, y, x, v, &c. & autant d'équations

$$\begin{aligned} A^1 &= Z^1z + Y^1y + X^1x + V^1v + \&c. \\ A^2 &= Z^2z + Y^2y + X^2x + V^2v + \&c. \\ A^3 &= Z^3z + Y^3y + X^3x + V^3v + \&c. \\ A^4 &= Z^4z + Y^4y + X^4x + V^4v + \&c. \\ &\&c. \end{aligned}$$

où les lettres A^1, A^2, A^3, A^4, &c. ne marquent pas, comme à l'ordinaire, les puissances d'A, mais le prémier membre, supposé connu, de la prémiére, seconde, troisiéme, quatriéme &c. équation. De même Z^1, Z^2, &c. sont les coëfficients de z; Y^1, Y^2, &c. ceux de y; X^1, X^2, &c. ceux de x; V^1, V^2, &c. ceux de v; &c. dans la prémiére, seconde, &c. équation.

Cette Notation supposée, s'il n'y a qu'une équation & qu'une inconnue z; on aura $z = \frac{A^1}{Z^1}$. S'il y a deux équations & deux inconnues z & y; on trouvera $z = \frac{A^1Y^2 - A^2Y^1}{Z^1Y^2 - Z^2Y^1}$, & $y = \frac{Z^1A^2 - Z^2A^1}{Z^1Y^2 - Z^2Y^1}$. S'il y a trois équations & trois inconnues z, y, & x; on trouvera

$$z = \frac{A^1Y^2X^3 - A^1Y^3X^2 - A^2Y^1X^3 + A^2Y^3X^1 + A^3Y^1X^2 - A^3Y^2X^1}{Z^1Y^2X^3 - Z^1Y^3X^2 - Z^2Y^1X^3 + Z^2Y^3X^1 + Z^3Y^1X^2 - Z^3Y^2X^1}$$

$$y = \frac{Z^1A^2X^3 - Z^1A^3X^2 - Z^2A^1X^3 + Z^2A^3X^1 + Z^3A^1X^2 - Z^3A^2X^1}{Z^1Y^2X^3 - Z^1Y^3X^2 - Z^2Y^1X^3 + Z^2Y^3X^1 + Z^3Y^1X^2 - Z^3Y^2X^1}$$

$$x = \frac{Z^1Y^2A^3 - Z^1Y^3A^2 - Z^2Y^1A^3 + Z^2Y^3A^1 + Z^3Y^1A^2 - Z^3Y^2A^1}{Z^1Y^2X^3 - Z^1Y^3X^2 - Z^2Y^1X^3 + Z^2Y^3X^1 + Z^3Y^1X^2 - Z^3Y^2X^1}$$

Introd. à l'Analyse des Lignes Courbes. Oooo L'é-

FIGURE 3.3.: « *Introduction à l'analyse des lignes courbes algébriques* » de Gabriel Cramer 1750.

§ 4 Formulation explicite du déterminant

Nous avons vu dans le paragraphe 1, une définition récursive du déterminant. Dans cette section, nous établissons une formulation explicite qui sera utile dans la suite pour établir certaines propriétés des déterminants. Nous renvoyons le lecteur à une annexe de ce chapitre, section 8, pour un rappel sur les notions du groupe symétrique utilisées pour établir cette formulation.

Considérons le cas $n = 2$. Soit

$$\mathbf{A} = \begin{bmatrix} a_1^1 & a_1^2 \\ a_2^1 & a_2^2 \end{bmatrix}$$

une matrice de $\mathscr{M}_2(\mathbb{K})$. La matrice $\mathbf{A}$ s'exprime par colonne sous la forme $\mathbf{A} = \left[\ \mathbf{a}_*^1 \mid \mathbf{a}_*^2\ \right]$.

Notons $\mathscr{C}\text{an} = (\mathbf{c}_1, \mathbf{c}_2)$ la base canonique de $\mathbb{K}^2$. Les colonnes $\mathbf{a}_*^1$ et $\mathbf{a}_*^2$ qui s'identifient à des vecteurs de $\mathbb{K}^2$ se décomposent dans la base canonique en

$$\begin{aligned}\mathbf{a}_*^1 &= a_1^1\mathbf{c}_1 + a_2^1\mathbf{c}_2\\ \mathbf{a}_*^2 &= a_1^2\mathbf{c}_1 + a_2^2\mathbf{c}_2\end{aligned}$$

On a

$$\begin{aligned}\det \mathbf{A} &= \det \left[\ \mathbf{a}_*^1 \mid \mathbf{a}_*^2\ \right]\\ &= \det \left[\ a_1^1\mathbf{c}_1 + a_2^1\mathbf{c}_2 \mid a_1^2\mathbf{c}_1 + a_2^2\mathbf{c}_2\ \right]\\ &= (a_1^1a_2^2 - a_2^1a_1^2)\det \left[\ \mathbf{c}_1 \mid \mathbf{c}_2\ \right]\end{aligned}$$

Or $\det \left[\ \mathbf{c}_1 \mid \mathbf{c}_2\ \right] = 1$, on retrouve ainsi la formule donnée précédemment : $\det \mathbf{A} = a_1^1a_2^2 - a_2^1a_1^2$. Ce qui s'écrit aussi sous la forme

$$\det \mathbf{A} = \operatorname{sgn}(\mathrm{id})a^1_{\mathrm{id}(1)}a^2_{\mathrm{id}(2)} + \operatorname{sgn}(\tau)a^1_{\tau(1)}a^2_{\tau(2)},$$

avec $\tau = (1,2)$. Plus généralement, montrons la formulation suivante

3.6 Théorème (Formule de Leibniz).— Soit $\mathbf{A} = [a_i^j]$ une matrice de $\mathscr{M}_n(\mathbb{K})$. On a

$$\det \mathbf{A} = \sum_{\sigma \in \mathscr{S}_n} \operatorname{sgn}(\sigma)a^1_{\sigma(1)}a^2_{\sigma(2)}\ldots a^n_{\sigma(n)}.$$

Preuve. Soit $\mathbf{A} = \left[\ \mathbf{a}_*^1 \mid \mathbf{a}_*^2 \mid \ldots \mid \mathbf{a}_*^n\ \right]$ une matrice de $\mathscr{M}_n(\mathbb{K})$. Considérons la base canonique $\mathscr{C}\text{an} = (\mathbf{c}_1, \ldots, \mathbf{c}_n)$ de $\mathbb{K}^n$. Tout vecteur colonne $\mathbf{a}_*^j$ de $\mathbf{A}$ se décompose dans la base canonique en

$$\mathbf{a}_*^j = \sum_{k=1}^n a_k^j \mathbf{c}_k.$$

On a

$$\begin{aligned}\det \mathbf{A} &= \det \left[\ \mathbf{a}_*^1 \mid \mathbf{a}_*^2 \mid \ldots \mid \mathbf{a}_*^n\ \right]\\ &= \det \left[\ \sum_{k_1=1}^n a^1_{k_1}\mathbf{c}_{k_1} \,\middle|\, \sum_{k_2=1}^n a^2_{k_2}\mathbf{c}_{k_2} \,\middle|\, \ldots \,\middle|\, \sum_{k_n=1}^n a^n_{k_n}\mathbf{c}_{k_n}\ \right]\\ &= \sum_{k_1=1}^n\sum_{k_2=1}^n \ldots \sum_{k_n=1}^n a^1_{k_1}a^2_{k_2}\ldots a^n_{k_n}\det \left[\ \mathbf{c}_{k_1} \mid \mathbf{c}_{k_2} \mid \ldots \mid \mathbf{c}_{k_n}\ \right]\end{aligned}$$

Si un terme de la somme précédente possède deux indices k_i égaux, alors

$$\det \left[\ \mathbf{c}_{k_1} \mid \mathbf{c}_{k_2} \mid \ldots \mid \mathbf{c}_{k_n}\ \right] = 0.$$

Sinon, un terme ne possède que des indices différents $k_1, \ldots, k_n$ de l'intervalle $[\![1,n]\!]$. La donnée de ces n indices correspond à la donnée d'une permutation de $\mathscr{S}_n$, telle que

$$\sigma(i) = k_i, \quad \text{pour tout } i \in [\![1,n]\!].$$

On en déduit que

$$\det \mathbf{A} = \sum_{\sigma \in \mathscr{S}_n} a^1_{\sigma(1)} a^2_{\sigma(2)} \ldots a^n_{\sigma(n)} \det \left[\ \mathbf{c}_{\sigma(1)} \mid \mathbf{c}_{\sigma(2)} \mid \ldots \mid \mathbf{c}_{\sigma(n)}\ \right]$$

Or, d'après la proposition 3.16, toute permutation de S_n se décompose en un produit de transpositions. Ainsi, d'après la proposition 3.3, on a

$$\det \left[\ \mathbf{c}_{\sigma(1)} \mid \mathbf{c}_{\sigma(2)} \mid \ldots \mid \mathbf{c}_{\sigma(n)}\ \right] = \operatorname{sgn}(\sigma) \det \left[\ \mathbf{c}_1 \mid \mathbf{c}_2 \mid \ldots \mid \mathbf{c}_n\ \right].$$

Comme $\det \left[\ \mathbf{c}_1 \mid \mathbf{c}_2 \mid \ldots \mid \mathbf{c}_n\ \right] = 1$, on en déduit que

$$\det \mathbf{A} = \sum_{\sigma \in \mathscr{S}_n} \operatorname{sgn}(\sigma) a^1_{\sigma(1)} a^2_{\sigma(2)} \ldots a^n_{\sigma(n)}.$$

□

§ 5 Calcul des déterminants

3.7 Proposition.— Pour toute matrice $\mathbf{A}$ de $\mathscr{M}_n(\mathbb{K})$, on a

$$\det(\mathbf{A}^\top) = \det \mathbf{A}.$$

3.5.1. Déterminant de la transposée d'une matrice.—

Preuve. Soit $\mathbf{A} = [a^j_i]$ une matrice de $\mathscr{M}_n(\mathbb{K})$. On a montré que

$$\det \mathbf{A} = \sum_{\sigma \in \mathscr{S}_n} \operatorname{sgn}(\sigma) a^1_{\sigma(1)} a^2_{\sigma(2)} \ldots a^n_{\sigma(n)}.$$

Pour toute permutation σ de $\mathscr{S}_n$, le corps $\mathbb{K}$ étant commutatif, on a, pour toute permutation τ de $\mathscr{S}_n$,

$$a^1_{\sigma(1)} a^2_{\sigma(2)} \ldots a^n_{\sigma(n)} = a^{\tau(1)}_{\sigma\circ\tau(1)} a^{\tau(2)}_{\sigma\circ\tau(2)} \ldots a^{\tau(n)}_{\sigma\circ\tau(n)}.$$

En particulier, lorsque $\tau = \sigma^{-1}$, on a

$$a^1_{\sigma(1)} a^2_{\sigma(2)} \ldots a^n_{\sigma(n)} = a^{\sigma^{-1}(1)}_1 a^{\sigma^{-1}(2)}_2 \ldots a^{\sigma^{-1}(n)}_n.$$

Par ailleurs, $\mathrm{sgn}(\sigma) = \mathrm{sgn}(\sigma^{-1})$, donc

$$\mathrm{sgn}(\sigma) a^1_{\sigma(1)} a^2_{\sigma(2)} \ldots a^n_{\sigma(n)} = \mathrm{sgn}(\sigma^{-1}) a_1^{\sigma^{-1}(1)} a_2^{\sigma^{-1}(2)} \ldots a_n^{\sigma^{-1}(n)}.$$

On a ainsi,

$$\begin{aligned}
\det \mathbf{A} &= \sum_{\sigma \in \mathscr{S}_n} \mathrm{sgn}(\sigma) a^1_{\sigma(1)} a^2_{\sigma(2)} \ldots a^n_{\sigma(n)} \\
&= \sum_{\sigma \in \mathscr{S}_n} \mathrm{sgn}(\sigma^{-1}) a_1^{\sigma^{-1}(1)} a_2^{\sigma^{-1}(2)} \ldots a_n^{\sigma^{-1}(n)} \\
&= \sum_{\sigma^{-1} \in \mathscr{S}_n} \mathrm{sgn}(\sigma^{-1}) a_1^{\sigma^{-1}(1)} a_2^{\sigma^{-1}(2)} \ldots a_n^{\sigma^{-1}(n)}
\end{aligned}$$

Soit encore

$$\det \mathbf{A} = \sum_{\sigma \in \mathscr{S}_n} \mathrm{sgn}(\sigma) a_1^{\sigma(1)} a_2^{\sigma(2)} \ldots a_n^{\sigma(n)}.$$

Cette dernière expression correspond au déterminant de la transposée $\mathbf{A}^\top$, car si $\mathbf{A}^\top = (b^j_i)$, on a $b^j_i = a^i_j$ et

$$\begin{aligned}
\det \mathbf{A}^\top &= \sum_{\sigma \in \mathscr{S}_n} \mathrm{sgn}(\sigma) b^1_{\sigma(1)} b^2_{\sigma(2)} \ldots b^n_{\sigma(n)} \\
&= \sum_{\sigma \in \mathscr{S}_n} \mathrm{sgn}(\sigma) a_1^{\sigma(1)} a_2^{\sigma(2)} \ldots a_n^{\sigma(n)}.
\end{aligned}$$

□

Comme conséquence de la proposition précédente, on a

3.8 Proposition.—

i) le déterminant d'une matrice est une application linéaire par rapport à chaque ligne,

ii) le déterminant d'une matrice dont deux lignes sont égales est nul,

iii) le déterminant d'une matrice change de signe, lorsque l'on échange deux lignes,

iv) le déterminant d'une matrice est non nul si, et seulement si, les vecteurs lignes de la matrice sont linéairement indépendants.

3.5.2. Déterminant d'un produit.— La propriété suivante exprime que le déterminant est compatible au produit de matrices.

3.9 Théorème.— Pour toutes matrices $\mathbf{A}$ et $\mathbf{B}$ de $\mathscr{M}_n(\mathbb{K})$, on a

$$\det(\mathbf{AB}) = \det(\mathbf{A})\det(\mathbf{B}).$$

Preuve. Soit $\mathbf{A} = [a_i^j]$ et $\mathbf{B} = [b_i^j]$ deux matrices de $\mathscr{M}_n(\mathbb{K})$. Notons

$$\mathbf{C} = \mathbf{AB} = (c_i^j) = [\ \mathbf{c}_*^1 \mid \mathbf{c}_*^2 \mid \ldots \mid \mathbf{c}_*^n\ ]$$

le produit de $\mathbf{A}$ et $\mathbf{B}$. Par définition du produit matriciel, on a

$$c_i^j = \sum_{k=1}^{n} a_i^k b_k^j.$$

Par suite,

$$\mathbf{c}_*^j = \sum_{k=1}^{n} b_k^j \mathbf{a}_*^k.$$

Ainsi

$$\begin{aligned}
\det(\mathbf{AB}) &= \det[\ \mathbf{c}_*^1 \mid \mathbf{c}_*^2 \mid \ldots \mid \mathbf{c}_*^n\ ] \\
&= \det\left[\ \sum_{k_1=1}^{n} b_{k_1}^1 \mathbf{a}_*^{k_1} \ \middle|\ \sum_{k_2=1}^{n} b_{k_2}^2 \mathbf{a}_*^{k_2} \ \middle|\ \ldots \ \middle|\ \sum_{k_n=1}^{n} b_{k_n}^n \mathbf{a}_*^{k_n}\ \right] \\
&= \sum_{k_1=1}^{n}\sum_{k_2=1}^{n}\cdots\sum_{k_n=1}^{n} b_{k_1}^1 b_{k_2}^2 \ldots b_{k_n}^n \det\left[\ \mathbf{a}_*^{k_1} \mid \mathbf{a}_*^{k_2} \mid \ldots \mid \mathbf{a}_*^{k_n}\ \right]
\end{aligned}$$

Si un des terme de la somme précédente possède deux indices k_i égaux, alors

$$\det\left[\ \mathbf{a}_*^{k_1} \mid \mathbf{a}_*^{k_2} \mid \ldots \mid \mathbf{a}_*^{k_n}\ \right] = 0.$$

Ne reste que les termes avec des indices $k_1, k_2, \ldots, k_n$ de $[\![1,n]\!]$ tous différents entre eux. Cette somme peut donc s'écrire

$$\det(\mathbf{AB}) = \sum_{\sigma\in\mathscr{S}_n} b_{\sigma(1)}^1 b_{\sigma(2)}^2 \ldots b_{\sigma(n)}^n \det\left[\ \mathbf{a}_*^{\sigma(1)} \mid \mathbf{a}_*^{\sigma(2)} \mid \ldots \mid \mathbf{a}_*^{\sigma(n)}\ \right] \tag{3.2}$$

Or

$$\det\left[\ \mathbf{a}_*^{\sigma(1)} \mid \mathbf{a}_*^{\sigma(2)} \mid \ldots \mid \mathbf{a}_*^{\sigma(n)}\ \right] = \operatorname{sgn}(\sigma)\det[\ \mathbf{a}_*^1 \mid \mathbf{a}_*^2 \mid \ldots \mid \mathbf{a}_*^n\ ].$$

On déduit donc de l'équation (3.2) que

$$\det(\mathbf{AB}) = \sum_{\sigma\in\mathscr{S}_n} \operatorname{sgn}(\sigma) b_{\sigma(1)}^1 b_{\sigma(2)}^2 \ldots b_{\sigma(n)}^n \det[\ \mathbf{a}_*^1 \mid \mathbf{a}_*^2 \mid \ldots \mid \mathbf{a}_*^n\ ]$$

D'où $\det(\mathbf{AB}) = \det(\mathbf{A})\det(\mathbf{B})$. $\square$

On en déduit la caractérisation importante suivante des matrices inversibles

3.10 Proposition.— Une matrice $\mathbf{A}$ de $\mathscr{M}_n(\mathbb{K})$ est inversible si, et seulement si, $\det(\mathbf{A})$ est non nul. On a alors

$$\det(\mathbf{A}^{-1}) = \frac{1}{\det(\mathbf{A})}.$$

Preuve. Soit $\mathbf{A}$ une matrice inversible de $\mathscr{M}_n(\mathbb{K})$. Il existe alors une matrice $\mathbf{A}^{-1}$ de $\mathscr{M}_n(\mathbb{K})$ telle que $\mathbf{A}\mathbf{A}^{-1} = \mathbf{1}_n$. D'après le théorème 3.9, on en déduit que

$$\det(\mathbf{A})\det(\mathbf{A}^{-1}) = \det(\mathbf{A}\mathbf{A}^{-1}) = \det(\mathbf{1}_n) = 1.$$

D'où $\det(\mathbf{A})$ est non nul et

$$\det(\mathbf{A}^{-1}) = \frac{1}{\det(\mathbf{A})}.$$

Inversement, si $\mathbf{A} = \left[\ \mathbf{a}_*^1 \mid \mathbf{a}_*^2 \mid \ldots \mid \mathbf{a}_*^n\ \right]$ est une matrice de $\mathscr{M}_n(\mathbb{K})$ de déterminant non nul. D'après le théorème 3.4, la famille $\mathscr{B} = (\mathbf{a}_*^1, \mathbf{a}_*^2, \ldots, \mathbf{a}_*^n)$ forme une base de $\mathbb{K}^n$. La matrice $\mathbf{A}$ est ainsi la matrice de passage de la base canonique de $\mathbb{K}^n$ à la base $\mathscr{B}$; elle est donc inversible. □

§ 6 Calcul de l'inverse d'une matrice

3.6.1. Matrice des cofacteurs.— Soit $\mathbf{A} = [a_i^j]$ une matrice de $\mathscr{M}_n(\mathbb{K})$. On appelle *cofacteur* du coefficient a_i^j le scalaire

$$\mathrm{cof}(a_i^j) = (-1)^{i+j}\det(\mathbf{A}_{\check{i}\check{j}}),$$

où $\mathbf{A}_{\check{i}\check{j}}$ est la matrice obtenue en supprimant la i-ème ligne et la j-ème colonne de $\mathbf{A}$.

On appelle *comatrice de* $\mathbf{A}$, la matrice notée $\mathbf{Com}(\mathbf{A}) = (c_i^j)$, définie par

$$c_i^j = \mathrm{cof}(a_i^j).$$

3.11 Proposition.— Pour tout $i \in [\![1,n]\!]$, on a la relation du *développement du déterminant selon la i-ème ligne*

$$\det(\mathbf{A}) = \sum_{j=1}^{n} a_i^j \,\mathrm{cof}(a_i^j). \tag{3.3}$$

De la même façon, pour tout $j \in [\![1,n]\!]$, on a le *développement du déterminant selon* *la j-ème colonne*

$$\det(\mathbf{A}) = \sum_{i=1}^{n} a_i^j \operatorname{cof}(a_i^j). \tag{3.4}$$

Preuve. Exercice □

De la proposition précédente, on déduit une formule utile pour le calcul de l'inverse d'une matrice.

3.12 Théorème.— Pour toute matrice $\mathbf{A}$ de $\mathscr{M}_n(\mathbb{K})$, on a :

$$\mathbf{A}\mathbf{Com}(\mathbf{A})^\top = \mathbf{Com}(\mathbf{A})^\top\mathbf{A} = \det(\mathbf{A})\mathbf{1}_n.$$

Si la matrice $\mathbf{A}$ est inversible, on a la formule d'inversion

$$\mathbf{A}^{-1} = \frac{1}{\det(\mathbf{A})}\mathbf{Com}(\mathbf{A})^\top.$$

Cette formule d'inversion est aussi connue sous le nom de *règle de Cramer*.

Preuve. Soient $\mathbf{A} = [a_i^j]$ une matrice de $\mathscr{M}_n(\mathbb{K})$ et $\mathbf{Com}(\mathbf{A})$ sa matrice des cofacteurs. Exprimons le produit

$$(\mathbf{A}\mathbf{Com}(\mathbf{A})^\top)_i^j = \sum_{k=1}^{n} a_i^k \operatorname{cof}(a_j^k).$$

D'après la relation (3.3), on a, pour tout $i \in [\![1,n]\!]$,

$$(\mathbf{A}\mathbf{Com}(\mathbf{A})^\top)_i^i = \det(\mathbf{A}).$$

Par ailleurs, si $i \neq j$, alors $(\mathbf{A}\mathbf{Com}(\mathbf{A})^\top)_i^j = 0$.

Par suite, on a

$$\mathbf{A}\mathbf{Com}(\mathbf{A})^\top = \det(\mathbf{A})\mathbf{1}_n.$$

De la même façon, on montre que $\mathbf{Com}(\mathbf{A})^\top\mathbf{A} = \det(\mathbf{A})\mathbf{1}_n$. □

Exercice 2.— Calculer l'inverse des matrices inversibles suivantes

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \quad \mathbf{C} = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix} \quad \mathbf{D} = \begin{bmatrix} 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 \end{bmatrix}$$

$$\mathbf{E} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 2 & 1 \end{bmatrix} \quad \mathbf{F} = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \\ 2 & 3 & 1 \end{bmatrix} \quad \mathbf{G} = \begin{bmatrix} 0 & 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 & 0 \\ 2 & 3 & 4 & 0 & 1 \\ 3 & 4 & 0 & 1 & 2 \\ 4 & 0 & 1 & 2 & 3 \end{bmatrix}$$

Exercice 3.— Trouver la matrice $\mathbf{X}$, telle que $\mathbf{X} = \mathbf{AX} + \mathbf{B}$, avec

$$\mathbf{A} = \begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{bmatrix} \qquad \text{et} \qquad \mathbf{B} = \begin{bmatrix} 1 & 2 \\ 2 & 1 \\ 3 & 3 \end{bmatrix}.$$

Exercice 4.— Soit $\mathbf{A}$ une matrice inversible symétrique de $\mathscr{M}_n(\mathbb{K})$. Montrer que la matrice $\mathbf{A}^{-1}$ est symétrique.

§ 7 Déterminant d'un endomorphisme

3.13 Proposition.— Deux matrices semblables ont le même déterminant.

Preuve. Soient $\mathbf{A}$ et $\mathbf{B}$ deux matrices semblables de $\mathscr{M}_n(\mathbb{K})$. Il existe donc une matrice inversible $\mathbf{P}$ de $\mathscr{M}_n(\mathbb{K})$ telle que

$$\mathbf{A} = \mathbf{PBP}^{-1}.$$

Du théorème 3.9, on déduit que

$$\det(\mathbf{A}) = \det(\mathbf{PBP}^{-1}) = \det(\mathbf{P})\det(\mathbf{B})\det(\mathbf{P}^{-1}).$$

D'après la proposition 3.10, on a $\det(\mathbf{P})\det(\mathbf{P}^{-1}) = 1$. Ainsi, $\det\mathbf{A} = \det\mathbf{B}$. □

Deux matrices semblables, représentent le même endomorphisme exprimé dans des bases différentes. La proposition 3.13 permet de définir le déterminant d'un endomorphisme de la façon suivante.

3.7.1. Définition.— Soit E un $\mathbb{K}$-espace vecetoriel de dimension finie n et soit u un endomorphisme de E. On appelle *déterminant de u* le déterminant de la matrice $[u]_{\mathscr{B}}$ qui représente u dans une base quelconque $\mathscr{B}$ de E :

$$\det u = \det[u]_{\mathscr{B}}.$$

On peut compléter la caractérisation des applications linéaires inversibles en dimension finie donnée dans la proposition 1.10 :

3.14 Proposition.— Soit $u : E \longrightarrow E'$ une application linéaire entre deux $\mathbb{K}$-espaces vectoriel de même dimension finie. Les assertions suivantes sont équivalentes :

i) $\det u$ est non nul,

ii) u est injective,

iii) u est surjective,

iv) u est bijective

v) $\operatorname{Ker} u = \{\mathbf{0}\}$,

vi) $\operatorname{Im} u = E'$,

vii) $\operatorname{rang} u = \dim E'$.

§ 8 Annexe : rappels sur les groupes de symétries

Dans le paragraphe suivant, nous donnons une formulation explicite du déterminant d'une matrice en terme de permutations des colonnes de la matrice. Nous rappelons ici quelques notions autour des groupes symétriques.

3.8.1. Définitions.— Soit n un entier naturel non nul. On note $\mathscr{S}_n$ l'ensemble des *permutations* de l'ensemble $[\![1,n]\!]$, i.e., des bijections de l'ensemble $[\![1,n]\!]$ dans lui-même.

Pour $n = 1$, on a $\mathscr{S}_1 = \{\mathrm{id}\}$. Dans la suite de ce paragraphe section, nous supposerons $n > 1$.

3.15 Proposition.— L'ensemble $\mathscr{S}_n$, muni de la composition des applications, forme un groupe de cardinal $n!$.

Le groupe $\mathscr{S}_n$ est appelé le *groupe symétrique d'indice n*. La permutation composée de deux permutation $\sigma \circ \sigma'$ est appelée le produit des permutations et notée $\sigma\sigma'$.

Soient i et j deux éléments distincts de $[\![1,n]\!]$. La permutation τ de $\mathscr{S}_n$ définie par

$$\tau(i) = j, \quad \tau(j) = i, \quad \tau(k) = k, \text{ si } k \notin \{i, j\}$$

est appelée *transposition* des éléments i et j et est notée (i, j).

3.16 Proposition.— Le groupe symétrique $\mathscr{S}_n$ est engendré par les transpositions, i.e., toute permutation de $\mathscr{S}_n$ se décompose en un produit de transpositions.

Preuve. On procède par récurrence. Pour $n = 2$, le groupe $\mathscr{S}_2$ est composé de l'identité et de la permutation $(1,2)$; l'énoncé est donc clair dans ce cas.

Supposons que toute permutation de $\mathscr{S}_{n-1}$ se décompose en un produit de transpositions. Soit $\sigma \in \mathscr{S}_n$. Si $\sigma(n) = n$, alors σ permute $[\![1,n-1]\!]$. Par hypothèse de récurrence, σ s'écrit comme un produit de transpositions de $[\![1,n-1]\!]$.

Si $\sigma(n) = i$ avec $i \neq n$, la permutation $\sigma' = \sigma \circ (n,i)$ vérifie $\sigma'(n) = n$. D'après la remarque précédente, σ' s'écrit comme un produit de transpositions, on en déduit qu'il en est de même pour $\sigma = (n,i) \circ \sigma'$. □

Pour tout $i \in [\![1,n]\!]$, notons σ_i la transposition $(i,i+1)$. On notera que, pour tout $i < j$ dans $[\![1,n]\!]$, on a

$$\begin{aligned}(i,j) &= \sigma_i\sigma_{i+1}\ldots\sigma_{j-2}\sigma_{j-1}\sigma_{j-2}\ldots\sigma_{i+1}\sigma_i\\ &= \sigma_{j-1}\sigma_{j-2}\ldots\sigma_{i+1}\sigma_i\sigma_{i+1}\ldots\sigma_{j-2}\sigma_{j-1}.\end{aligned}$$

En particulier, on en déduit que le groupe $\mathscr{S}_n$ est engendré par les transpositions $\sigma_1,\ldots,\sigma_{n-1}$.

3.8.2. Signature.— Soit σ une permutation de $\mathscr{S}_n$. On appelle *inversion* de σ, un couple (i,j) d'éléments de $[\![1,n]\!]$ tel que $i < j$ et $\sigma(i) > \sigma(j)$. Notons $I(\sigma)$ le nombre d'inversions de σ.

On appelle *signature* de la permutation σ, le réel

$$\operatorname{sgn}(\sigma) = (-1)^{I(\sigma)}.$$

On peut montrer que la signature est donnée par la relation suivante

$$\operatorname{sgn}(\sigma) = \prod_{1\leqslant i<j\leqslant n} \frac{\sigma(j)-\sigma(i)}{j-i}.$$

Une permutation est dite *paire* (resp. *impaire*), si elle est de signature 1 (resp. -1).

3.17 Proposition.— L'application signature sgn est un morphisme de groupes de $\mathscr{S}_n$ dans le groupe multiplicatif $\{-1,1\}$. En particulier, pour toutes permutations σ,τ, on a

$$\operatorname{sgn}(\sigma\tau) = \operatorname{sgn}(\sigma)\operatorname{sgn}(\tau).$$

Le noyau de ce morphisme est formé des permutations paires et est appelé *groupe alterné d'indice n* . On le note $\mathscr{A}_n$.

§ 9 Annexe : déterminants et formes multilinéaires alternées

3.9.1. Applications multilinéaires alternées.— Soient E et F deux $\mathbb{K}$-espaces vectoriels et p un entier naturel non nul. Une *application p-linéaire* de E dans F est une application

$$\varphi : E^p \longrightarrow F$$

telle que, pour toute famille $(\mathbf{x}_1, \ldots, \mathbf{x}_p)$ de vecteurs de E et tout $i \in [\![1, n]\!]$, l'application

$$\begin{aligned} E &\longrightarrow F \\ \mathbf{x} &\longmapsto \varphi(\mathbf{x}_1, \ldots, \mathbf{x}_{i-1}, \mathbf{x}, \mathbf{x}_{i+1}, \ldots, \mathbf{x}_p) \end{aligned}$$

est linéaire.

Une application p-linéaire de E à valeurs dans $\mathbb{K}$ est appelée une *forme p-linéaire* .

Une application p-linéaire φ de E dans F est dite *alternée* si, pour toute famille $(\mathbf{x}_1, \ldots, \mathbf{x}_p)$ de vecteurs de E et pour tout $i \neq j$, on a :

$$\mathbf{x}_i = \mathbf{x}_j \Rightarrow \varphi(\mathbf{x}_1, \ldots, \mathbf{x}_p) = 0.$$

3.18 Proposition.— Soient E un $\mathbb{K}$-espace vectoriel de dimension n et F un $\mathbb{K}$-espace vectoriel. Soit φ une application n-linéaire alternée de E dans F. Pour toute permutation $\sigma \in \mathscr{S}_n$ et toute famille de vecteurs $(\mathbf{x}_1, \ldots, \mathbf{x}_n)$ de E, on a

$$\varphi(\mathbf{x}_{\sigma(1)}, \mathbf{x}_{\sigma(2)}, \ldots, \mathbf{x}_{\sigma(n)}) = \operatorname{sgn}(\sigma)\varphi(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n).$$

Preuve. Exercice. □

3.19 Théorème.— Soient E un $\mathbb{K}$-espace vectoriel de dimension n et $\mathscr{B} = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ une base de E. Soit φ une forme n-linéaire alternée sur E. Alors, pour toute famille $(\mathbf{x}_1, \ldots, \mathbf{x}_n)$ de vecteurs de E, on a

$$\varphi(\mathbf{x}_1, \ldots, \mathbf{x}_n) = \det\,[\mathbf{x}_1 \,|\, \ldots \,|\, \mathbf{x}_n]_{\mathscr{B}}\, \varphi(\mathbf{e}_1, \ldots, \mathbf{e}_n),$$

où $\det\,[\mathbf{x}_1 \,|\, \ldots \,|\, \mathbf{x}_n]_{\mathscr{B}}$ est le déterminant de la matrice $[\mathbf{x}_1 \,|\, \ldots \,|\, \mathbf{x}_n]_{\mathscr{B}}$ dont les coefficients de la j-ème colonne sont les coordonnées du vecteur $\mathbf{x}_j$ exprimé dans la base $\mathscr{B}$.

Preuve. Pour tout $i \in [\![1, n]\!]$, considérons la décomposition du vecteur $\mathbf{x}_i$ dans la base $\mathscr{B}$:

$$\mathbf{x}_i = \sum_{k=1}^{n} a_k^i \mathbf{e}_k.$$

On a

$$[\mathbf{x}_1 \,|\, \ldots \,|\, \mathbf{x}_n]_{\mathscr{B}} = \begin{bmatrix} a_1^1 & \ldots & a_1^n \\ \vdots & & \vdots \\ a_n^1 & \ldots & a_n^n \end{bmatrix}.$$

L'application φ étant n-linéaire, on a

$$\varphi(\mathbf{x}_1,\ldots,\mathbf{x}_n) = \sum_{k_1=1}^{n}\sum_{k_2=1}^{n}\cdots\sum_{k_n=1}^{n} a_{k_1}^1 a_{k_2}^2 \ldots a_{k_n}^n \,\varphi(\mathbf{e}_{k_1},\mathbf{e}_{k_2},\ldots,\mathbf{e}_{k_n}).$$

Si un terme de cette somme possède deux indices k_i égaux, alors le scalaire $\varphi(\mathbf{x}_{k_1},\mathbf{x}_{k_2},\ldots,\mathbf{x}_{k_n})$ est nul. Sinon, un terme ne possède que des indices différents $k_1,\ldots,k_n$ de l'intervalle $[\![1,n]\!]$. La donnée de ces n indices correspond à la donnée d'une permutation de $\mathscr{S}_n$, telle que

$$\sigma(i) = k_i, \quad \text{pour tout } i \in [\![1,n]\!].$$

On en déduit que

$$\varphi(\mathbf{x}_1,\ldots,\mathbf{x}_n) = \sum_{\sigma\in\mathscr{S}_n} a_{\sigma(1)}^1 a_{\sigma(2)}^2 \ldots a_{\sigma(n)}^n \,\varphi(\mathbf{e}_{\sigma(1)},\mathbf{e}_{\sigma(2)},\ldots,\mathbf{e}_{\sigma(n)}).$$

D'après la proposition précédente, on en déduit que

$$\varphi(\mathbf{x}_1,\ldots,\mathbf{x}_n) = \mathrm{sgn}(\sigma) \sum_{\sigma\in\mathscr{S}_n} a_{\sigma(1)}^1 a_{\sigma(2)}^2 \ldots a_{\sigma(n)}^n \,\varphi(\mathbf{e}_1,\mathbf{e}_2,\ldots,\mathbf{e}_n),$$

D'où la formule à établir :

$$\varphi(\mathbf{x}_1,\ldots,\mathbf{x}_n) = \det\,[\mathbf{x}_1 \,|\, \ldots \,|\, \mathbf{x}_n]_{\mathscr{B}}\, \varphi(\mathbf{e}_1,\ldots,\mathbf{e}_n),$$

□

Une conséquence immédiate du théorème précédent est

3.20 Proposition.— Soient E un $\mathbb{K}$-espace vectoriel de dimension n et $\mathscr{B} = (\mathbf{e}_1,\ldots,\mathbf{e}_n)$ une base de E. Il existe une unique forme n-linéaire alternée φ_0 sur E telle que

$$\varphi_0(\mathbf{e}_1,\ldots,\mathbf{e}_n) = 1.$$

De plus, toute forme n-linéaire alternée sur E est proportionnelle à φ_0.

Preuve. Si on pose, pour toute famille $(\mathbf{x}_1,\ldots,\mathbf{x}_n)$ de vecteurs de E :

$$\varphi_0(\mathbf{x}_1,\ldots,\mathbf{x}_n) = \det[\mathbf{x}_1 \,|\, \ldots \,|\, \mathbf{x}_n]_{\mathscr{B}}. \tag{3.5}$$

Ainsi définie, on déduit des propriétés de l'application det que φ_0 est une forme n-linéaire alternée.

D'après le théorème précédent, pour toute forme n-linéaire φ sur E, on a

$$\varphi = \varphi(\mathbf{e}_1, \ldots, \mathbf{e}_n)\varphi_0.$$

Il existe ainsi une unique forme n-linéaire φ sur E vérifiant $\varphi(\mathbf{e}_1, \ldots, \mathbf{e}_n) = 1$, c'est la forme φ_0 définie par (3.5). $\square$

3.9.2. Remarque.— On aurait pu définir le déterminant d'une matrice en utilisant le point de vue des formes multiliéaires alternées.

D'après la proposition précédente, il existe une unique forme n-linéaire alternée sur $\mathbb{K}^n$ telle que $\varphi_0(\mathbf{c}_1, \ldots, \mathbf{c}_n) = 1$, où $(\mathbf{c}_1, \ldots, \mathbf{c}_n)$ est la base canonique de $\mathbb{K}^n$. Le déterminant d'une matrice $\mathbf{A} = \left[\ \mathbf{a}_*^1 \mid \mathbf{a}_*^2 \mid \ldots \mid \mathbf{a}_*^n\ \right]$ de $\mathscr{M}_n(\mathbb{K})$ est alors donné par

$$\det \mathbf{A} = \varphi_0(\mathbf{a}_*^1, \mathbf{a}_*^2, \ldots, \mathbf{a}_*^n).$$

Deuxième partie .

La réduction des matrices

CHAPITRE
4

Pour se mettre en appétit

Sommaire

§ 1 Équations d'évolution linéaire couplées

4.1.1. Évolution discrète du temps.— On s'intéresse à l'évolution d'une population de renards dans une forêt au cours des années. Supposons que le nombre de renards l'année n est proportionnel au nombre de renards l'année précédente $n-1$. En désignant par $x(n)$ le nombre de renards vivant dans la forêt l'année n, l'évolution du nombre de renards se modélise par l'équation

$$x(n) = ax(n-1), \tag{4.1}$$

où a est une constante positive réelle. Cet exemple est un modèle d'évolution *linéaire*. La population est croissante si $a > 1$ et en extinction si $a < 1$. Si $a = 1$ la population reste constante à sa valeur initiale. Connaissant la population initiale $x(0)$ de renards l'année 0, de l'équation de récurrence 4.1, on déduit le nombre de renards en fonction de l'année n :

$$x(n) = a^n x(0).$$

De nombreux problèmes d'évolution discrète linéaire peuvent être décrits par une telle équation.

4.1.2. Évolution continue du temps.— Afin d'illustrer les problèmes d'évolution continue, considérons une cuve pleine d'eau de volume V litre et contenant p litre de polluant à l'instant $t = 0$. Un robinet d'eau non polluée verse dans la cuve à raison de r litre d'eau par seconde. Afin de maintenir la quantité d'eau constante dans la cuve, un robinet vide la cuve à raison de r litre d'eau par seconde, comme indiqué sur la figure 4.1.

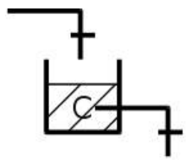


FIGURE 4.1.: Évolution de la concentration de polluant dans une cuve

On souhaite déterminer la quantité de polluant dans la cuve à tout instant t. Notons $x(t)$ la quantité de polluant dans la cuve à l'instant t. Le taux de variation de la quantité de polluant par rapport au temps s'exprime par

$$\frac{dx(t)}{dt} = \text{concentration qui arrive} - \text{concentration évacuée.}$$

L'eau versée dans la cuve étant non polluée, la concentration de poluant qui arrive est nulle et la concentration évacuée est $\frac{r}{V}x(t)$, on a donc

$$\frac{dx(t)}{dt} = 0 - \frac{r}{V}x(t), \tag{4.2}$$

avec la condition initiale $x(0) = p_0$. Cette équation différentielle admet pour solution

$$x(t) = e^{-\frac{r}{V}t}p_0.$$

4.1.3. Équations couplées, la cas discret.— Dans de nombreuses situations, on étudie l'évolution de plusieurs espèces cohabitant dans un même écosystème. Le système est alors décrit par plusieurs variables du temps $x_1(n), x_2(n), \ldots, x_k(n)$ qui représentent le nombre d'individus de chaque espèce. Ces variables dépendent les unes des autres, on dit qu'elles sont *« couplées »*. Il existe plusieurs cas d'évolution :

- chaque population a un effet négatif sur la croissance des autres populations, il s'agit de la *compétition interspécifique*,
- on parle de relations *proie-prédateur* ou *hôte-parasite*, lorsqu'une population a un effet positif sur une autre, mais l'effet est négatif dans l'autre sens,
- chaque population a un effet positif sur la croissance des autres populations, on parle alors d'évolution en *symbiose*.

4.1.4. Un modèle proie-prédateur.— Comme illustration du modèle proie-prédateur, considérons les populations de renards et de mulots dans une forêt. L'évolution de ces deux populations est liée du fait que l'on suppose que les renards mangent les mulots. En conséquence, plus il y aura de mulots, plus les renards auront à manger ; on peut supposer alors que cela encourage leur reproduction, par suite le nombre de renards augmente. D'autre part, plus il y aura de renards, plus de mulots seront mangés, ce qui va réduire la population de mulots l'année suivante.

Ces deux espèces sont ainsi en interaction, on peut modéliser cette interaction de la façon suivante. Notons respectivement par $x_1(n)$ et $x_2(n)$ le nombre de renards et de mulots l'année n. Les équations qui décrivent l'évolution des deux populations forment le système :

$$\left| \begin{array}{ll} x_1(n) & = ax_1(n-1) + bx_2(n-1), \\ x_2(n) & = -cx_1(n-1) + dx_2(n-1), \end{array} \right. \tag{4.3}$$

où a, b, c, d sont quatre constantes réelles positives. On regroupe les deux quantités $x_1(n)$ et $x_2(n)$ pour former un vecteur

$$\mathbf{x}(n) = \begin{bmatrix} x_1(n) \\ x_2(n) \end{bmatrix}.$$

Le système d'équation (4.3) s'écrit alors sous la forme d'une seule équation

$$\mathbf{x}(n) = \mathbf{A}\mathbf{x}(n-1), \tag{4.4}$$

où $\mathbf{A}$ désigne la matrice

$$\begin{bmatrix} a & b \\ -c & d \end{bmatrix}$$

L'équation (4.4) est une généralisation au cas de plusieurs variables de l'équation (4.1). De la même façon que dans le cas à une seule variable, connaissant les deux populations l'année 0, i.e., étant donné le vecteur $\mathbf{x}(0)$, on montre par récurrence que le nombre de renards et de mulots l'année n est donné par

$$\mathbf{x}(n) = \mathbf{A}^n\mathbf{x}(0).$$

4.1.5. Équations couplées, la cas continu.— On considère le problème de concentration de polluant de l'exemple 4.1.2, mais maintenant avec trois cuves. On dispose trois cuves pleines d'eau polluée, d'un volume V chacune, de telle façon que l'eau circule entre elles selon le schéma de la figure 4.2.

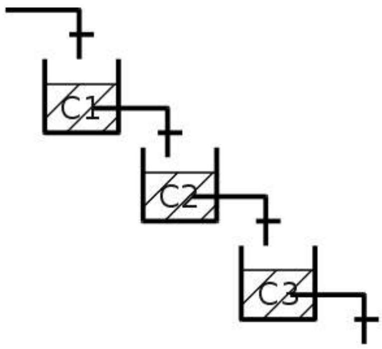


FIGURE 4.2.: Évolution de la concentration de polluant dans trois cuves

Il y a r litre d'eau par seconde qui passe de la cuve C_1 à la cuve C_2, de la cuve C_2 à la cuve C_3 et de la cuve C_3 à l'égout. De plus r litre d'eau pur par seconde est versé dans la cuve C_1. Les cuves contiennent initialement p_1, p_2 et p_3 litre de polluant respectivement. On souhaite déterminer la quantité de polluant qui se trouve dans chacune des cuves à chaque instant t de l'évolution du système.

Notons $x_i(t)$ la quantité de polluant dans la cuve i, à l'instant t. On a

$$\frac{dx_i(t)}{dt} = \text{concentration qui arrive} - \text{concentration évacuée.}$$

La concentration du polluant évacuée de la cuve i est donc $\frac{r}{V}x_i(t)$. Les quantités de polluant satisfont donc le système différentiel suivant :

$$\left|\begin{array}{ll} \dfrac{dx_1(t)}{dt} & = -\dfrac{r}{V}x_1(t) \\ \dfrac{dx_2(t)}{dt} & = \dfrac{r}{V}x_1(t) - \dfrac{r}{V}x_2(t) \\ \dfrac{dx_3(t)}{dt} & = \dfrac{r}{V}x_2(t) - \dfrac{r}{V}x_3(t) \end{array}\right.$$

Ce système s'écrit matriciellement sous la forme

$$\frac{d}{dt}\begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix} = \frac{r}{V}\begin{bmatrix} -1 & 0 & 0 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix}\begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix}.$$

Ou encore,

$$\frac{d\mathbf{x}(t)}{dt} = \mathbf{A}\mathbf{x}(t), \tag{4.5}$$

en posant

$$\mathbf{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix} \quad \text{et} \quad \mathbf{A} = \frac{r}{V}\begin{bmatrix} -1 & 0 & 0 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix}$$

et où $\frac{d\mathbf{x}(t)}{dt}$ désigne la dérivée du vecteur $\mathbf{x}(t)$:

$$\frac{d\mathbf{x}(t)}{dt} = \begin{bmatrix} \frac{dx_1(t)}{dt} \\ \frac{dx_2(t)}{dt} \\ \frac{dx_3(t)}{dt} \end{bmatrix}.$$

Nous montrerons plus loin dans le cours que l'équation (4.5) admet pour solution

$$\mathbf{x}(t) = e^{t\mathbf{A}}\mathbf{x}(0), \tag{4.6}$$

où $\mathbf{x}(0) = \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix}$ désigne le vecteur représentant les concentrations de polluant dans les trois cuves à l'instant $t = 0$ et où $e^{\frac{rt}{V}\mathbf{A}}$ est la matrice définie par une série de matrices :

$$e^{\frac{rt}{V}\mathbf{A}} = \sum_{k=0}^{+\infty} \frac{1}{k!}\left(\frac{rt}{V}\mathbf{A}\right)^k.$$

Nous justifierons plus loin cette définition et nous présenterons des méthodes de calcul des exponetielles de matrices. On remarquera l'analogie entre la solution (4.6) et la solution de l'équation différentielle (4.2).

Comme dans le cas discret, le calcul de la solution passe ainsi par celui des puissances de la matrice $\mathbf{A}$. Nous allons voir comment l'on peut réduire le calcul des puissances de $\mathbf{A}$ à celui des puissances d'une matrice diagonale semblable à $\mathbf{A}$. On appelle cette opération le *découplage*.

§ 2 Le découplage de système d'équations

Nous allons montrer que pour résoudre un système d'équations couplées, il est judicieux de transformer ce système en un système découplé. Dans un premier temps, on montre que la résolution d'un système découplé est immédiate.

4.2.1. Les équations découplées.— Supposons que nous ayons une équation d'évolution discrète

$$\mathbf{x}(n) = \mathbf{A}\mathbf{x}(n-1),$$

avec par exemple

$$\mathbf{A} = \begin{bmatrix} 3 & 0 \\ 0 & 5 \end{bmatrix}.$$

L'équation est équivalente au système de deux équations linéaires :

$$\left| \begin{array}{l} x_1(n) = 3x_1(n-1) \\ x_2(n) = 5x_2(n-1) \end{array} \right.$$

Ces équations sont dites *non couplées*, car $x_1(n)$ ne dépend que des valeurs des $x_1(k)$, pour $k \leqslant n$, et est indépendant des valeurs de $x_2(k)$. De même pour $x_2(n)$ qui ne dépend que des valeurs de $x_2(k)$, pour $k \leqslant n$.

Les solutions du système se déduisent par une simple récurrence :

$$x_1(n) = 3^n x_1(0) \quad \text{et} \quad x_2(n) = 5^n x_2(0).$$

4.2.2. Le découplage.— Considérons le système suivant de deux équations couplées :

$$\left| \begin{array}{ll} x_1(n) & = 2x_1(n-1) + x_2(n-1), \\ x_2(n) & = x_1(n-1) + 2x_2(n-1). \end{array} \right. \tag{4.7}$$

Afin de résoudre ce système, nous effectuons le changement de variables suivant :

$$y_1(n) = \frac{x_1(n) + x_2(n)}{2}, \qquad y_2(n) = \frac{x_1(n) - x_2(n)}{2}. \tag{4.8}$$

On a alors

$$x_1(n) = y_1(n) + y_2(n), \qquad x_2(n) = y_1(n) - y_2(n). \tag{4.9}$$

En additionnant et soustrayant les deux équations du système précédent, on obtient

$$\left| \begin{array}{ll} y_1(n) & = 3y_1(n-1), \\ y_2(n) & = y_2(n-1). \end{array} \right. \tag{4.10}$$

Ce dernier système est découplé, il admet pour solutions :

$$y_1(n) = 3^n y_1(0), \qquad y_2(n) = y_2(0).$$

On peut alors résoudre notre problème initial. On a

$$y_1(0) = \frac{x_1(0) + x_2(0)}{2}, \qquad y_2(0) = \frac{x_1(0) - x_2(0)}{2}.$$

D'où

$$y_1(n) = 3^n \frac{x_1(0) + x_2(0)}{2}, \qquad y_2(n) = \frac{x_1(0) - x_2(0)}{2}.$$

Et en utilisant les équations (4.9), on obtient

$$x_1(n) = \frac{3^n + 1}{2} x_1(0) + \frac{3^n - 1}{2} x_2(0), \qquad x_2(n) = \frac{3^n - 1}{2} x_1(0) + \frac{3^n + 1}{2} x_2(0). \tag{4.11}$$

4.2.3. Exégèse matricielle.— Le système (4.7) s'écrit sous la forme matricielle suivante

$$\mathbf{x}(n) = \mathbf{A}\mathbf{x}(n-1),$$

avec

$$\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \qquad \text{et} \qquad \mathbf{x}(n) = \begin{bmatrix} x_1(n) \\ x_2(n) \end{bmatrix}.$$

Le changement de variable (4.8) consiste à considérer la matrice

$$\mathbf{Q} = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

et définir le vecteur $\mathbf{y}(n)$ par

$$\mathbf{y}(n) = \mathbf{Q}\mathbf{x}(n). \tag{4.12}$$

Par ailleurs, le changement de variable (4.9) s'écrit matriciellement sous la forme :

$$\mathbf{x}(n) = \mathbf{P}\mathbf{y}(n), \quad \text{avec} \quad \mathbf{P} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}. \tag{4.13}$$

On vérifiera que $\mathbf{P} = \mathbf{Q}^{-1}$. Le système découplé (4.10) s'écrit matriciellement sous la forme

$$\mathbf{y}(n) = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix} \mathbf{y}(n-1).$$

On en déduit que

$$\begin{aligned} \mathbf{y}(n) &= \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}^n \mathbf{y}(0) \\ &= \begin{bmatrix} 3^n & 0 \\ 0 & 1 \end{bmatrix} \mathbf{y}(0). \end{aligned}$$

Donc, de (4.12), on déduit que $\mathbf{y}(0) = \mathbf{Q}\mathbf{x}(0)$, d'où :

$$\begin{aligned} \mathbf{y}(n) &= \frac{1}{2}\begin{bmatrix} 3^n & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \mathbf{x}(0) \\ &= \frac{1}{2}\begin{bmatrix} 3^n(x_1(0) - x_2(0)) \\ x_1(0) - x_2(0) \end{bmatrix}. \end{aligned}$$

D'après (4.13), on en déduit finalement que

$$\begin{aligned} \mathbf{x}(n) &= \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} 3^n(x_1(0) - x_2(0)) \\ x_1(0) - x_2(0) \end{bmatrix} \\ &= \begin{bmatrix} 3^n/2(x_1(0) + x_2(0)) + (x_1(0) - x_2(0)) \\ 3^n/2(x_1(0) + x_2(0)) - (x_1(0) - x_2(0)) \end{bmatrix}. \end{aligned}$$

Ceci correspond à la solution que nous avons obtenu en (4.11).

Le choix du changement de variable (4.13) pour obtenir un système découplé, i.e., avec une matrice diagonale, est possible car il existe deux « *vecteurs propres* »

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \text{et} \quad \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

de la matrice $\mathbf{A}$ associés respectivement aux « *valeurs propres* » 3 et 1 de la matrice $\mathbf{A}$.

C'est-à-dire que l'on a

$$\mathbf{A}\begin{bmatrix}1\\1\end{bmatrix}=3\begin{bmatrix}1\\1\end{bmatrix}\quad\text{et}\quad\mathbf{A}\begin{bmatrix}1\\-1\end{bmatrix}=\begin{bmatrix}1\\-1\end{bmatrix}.$$

La matrice $\mathbf{A}$ est semblable à une matrice diagonale :

$$\begin{bmatrix}3&0\\0&1\end{bmatrix}=\frac{1}{2}\begin{bmatrix}1&1\\1&-1\end{bmatrix}\mathbf{A}\begin{bmatrix}1&1\\1&-1\end{bmatrix}.$$

La première matrice dans le second membre de l'égalité est l'inverse de la matrice de changement de base

$$\mathbf{P}=\begin{bmatrix}1&1\\1&-1\end{bmatrix}$$

formée de deux vecteurs propres associés aux deux valeurs propres.

Nous retiendrons que « découpler » un système revient à le « diagonaliser ». On a construit deux matrices de changement de bases $\mathbf{P}$ et $\mathbf{Q}$ telles que

$$\mathbf{D}=\mathbf{QAP},$$

ou de façon équivalente

$$\mathbf{A}=\mathbf{PDQ},$$

car $\mathbf{P}$ et $\mathbf{Q}$ sont inverses l'une de l'autre. Ainsi, le calcul des puissances de $\mathbf{A}$ se réduit à celui des puissances de $\mathbf{D}$:

$$\begin{tikzcd}
\mathbf{x}(0) \arrow[r, squiggly, "\mathbf{A}"] \arrow[d, "\mathbf{Q}"'] & \mathbf{x}(1) \arrow[r, squiggly, "\mathbf{A}"] & \mathbf{x}(2) \arrow[r, squiggly, "\mathbf{A}"] & \cdots \arrow[r, squiggly, "\mathbf{A}"] & \mathbf{x}(n-1) \arrow[r, squiggly, "\mathbf{A}"] & \mathbf{x}(n) \\
\mathbf{y}(0) \arrow[r, "\mathbf{D}"'] & \mathbf{y}(1) \arrow[r, "\mathbf{D}"'] \arrow[u, "\mathbf{P}"'] & \mathbf{y}(2) \arrow[r, "\mathbf{D}"'] \arrow[u, "\mathbf{P}"'] & \cdots \arrow[r, "\mathbf{D}"'] & \mathbf{y}(n-1) \arrow[r, "\mathbf{D}"'] \arrow[u, "\mathbf{P}"'] & \mathbf{y}(n) \arrow[u, "\mathbf{P}"']
\end{tikzcd}$$

Plus généralement, nous montrerons dans ce cours que lorsque l'on peut trouver n valeurs propres différentes d'une matrice $\mathbf{A}$ de $\mathscr{M}_n(\mathbb{R})$, on peut résoudre un système couplé

$$\mathbf{x}(n)=\mathbf{A}\mathbf{x}(n-1)$$

en considérant un système découplé

$$\mathbf{y}(n)=\mathbf{D}\mathbf{y}(n-1).$$

§ 3 La diagonalisation des matrices et des endomorphismes

4.3.1. Problème de diagonalisation des matrices.— De nombreux problèmes comme celui du découplage conduisent au problème suivant :

Donnée : $\mathbf{A}$ une matrice de $\mathscr{M}_n(\mathbb{K})$.
Question : Trouver une matrice inversible $\mathbf{P}$ de $\mathscr{M}_n(\mathbb{K})$, telle que $\mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ soit diagonale.

Nous verrons qu'il existe de nombreuses autres situtations dans lesquelles ce problème intervient.

4.3.2. Problème de diagonalisation des endomorphismes.— On peut donner une autre formulation de ce problème. Soit E un $\mathbb{K}$-espace vectoriel de dimension finie n.

Donnée : u un endomorphisme de E représenté par une matrice $[u]_{\mathscr{B}}$ dans une base $\mathscr{B}$ de E.
Question : Trouver une base $\mathscr{B}'$ de E, telle que la matrice $[u]_{\mathscr{B}'}$ soit diagonale.

Deux matrices semblables représentant le même endomorphisme dans des bases différentes, la version matricielle de ce problème correspond au problème 4.3.1 : si $\mathbf{P}$ désigne la matrice de passage de la base $\mathscr{B}$ à la base $\mathscr{B}'$, on a la relation :

$$[u]_{\mathscr{B}'} = \mathbf{P}^{-1}[u]_{\mathscr{B}}\mathbf{P}.$$

Lorsque la matrice $[u]_{\mathscr{B}'}$ est une matrice diagonale, i.e., de la forme :

$$\begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix}$$

on dit que l'endomorphisme u est *diagonalisable*. Les vecteurs $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ de la base $\mathscr{B}'$ ont une propriété remarquable, ils satisfont, pour tout $i \in [\![1,n]\!]$,

$$u(\mathbf{e}_i) = \lambda_i \mathbf{e}_i.$$

Nous verrons que ce problème n'admet pas toujours de solution : un endomorphisme n'est pas toujours diagonalisable. Cependant, nous verrons qu'il est toujours possible de trouver un changement de base tel que la matrice s'exprime dans cette base comme une matrice diagonale par blocs.

4.3.3. Une reformulation géométrique.— On peut donner une autre formulation du problème de diagonalisation d'un endomorphisme. Soit u un endomorphisme de E, diagonaliser u revient à trouver une décomposition de E en une somme directe

$$E = E_1 \oplus \ldots \oplus E_p$$

de sous-espaces vectoriels telle que, pour tout $i \in [\![1,p]\!]$,

i) E_i est stable par u, i.e., $u(E_i)$ est un sous-espace de E_i,

ii) la restriction $u_{|E_i}$ de l'endomorphisme u à E_i est une homothétie, c'est-à-dire, qu'il existe un scalaire λ_i tel que
$$u(\mathbf{x}) = \lambda_i \mathbf{x},$$

pour tout vecteur $\mathbf{x}$ de E_i.

En choisissant une base $\mathscr{B}_i$ pour chaque sous-espace E_i, la réunion $\mathscr{B}_1 \cup \ldots \cup \mathscr{B}_p$ de ces bases forme alors une base $\mathscr{B}$ de E dans laquelle la matrice de u est de la forme

$$[u]_{\mathscr{B}} = \begin{bmatrix} \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_1 \end{bmatrix} & & & 0 \\ & \begin{bmatrix} \lambda_2 & & 0 \\ & \ddots & \\ 0 & & \lambda_2 \end{bmatrix} & & \\ & & \ddots & \\ 0 & & & \begin{bmatrix} \lambda_p & & 0 \\ & \ddots & \\ 0 & & \lambda_p \end{bmatrix} \end{bmatrix}$$

où chaque bloc

$$\begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_1 \end{bmatrix}$$

est la matrice de la restriction de l'endomorphisme u au sous-espace E_i, exprimé dans la base $\mathscr{B}_i$.

4.3.4. Exemple.— Soient Π le plan de $\mathbb{R}^3$ d'équation $z = 0$ et Δ la droite d'équation $x = y = z$. Soit u l'endomorphisme de $\mathbb{R}^3$ qui projette tout vecteur $\mathbf{x}$ de $\mathbb{R}^3$ sur le plan Π parallèlement à la droite Δ.

Notons $\mathscr{C}\text{an} = (\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3)$ la base canonique de $\mathbb{R}^3$:

$$\mathbf{c}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{c}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{c}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Soit $\mathbf{e} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ un vecteur directeur de Δ. L'endomorphisme u est caractérisé par

$$u(\mathbf{x}) + \lambda \mathbf{e} = \mathbf{x}, \quad u(\mathbf{x}) \in \Pi, \quad \text{pour tout } \mathbf{x} \in \mathbb{R}^3.$$

Donc si $[\mathbf{x}]_{\mathscr{C}\text{an}} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$, on a

$$[u]_{\mathscr{C}\text{an}} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 - \lambda \\ x_2 - \lambda \\ x_3 - \lambda \end{bmatrix} \quad \text{et} \quad [u]_{\mathscr{C}\text{an}} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \Pi.$$

Soit encore

$$[u]_{\mathscr{C}\mathrm{an}} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 - x_3 \\ x_2 - x_3 \\ 0 \end{bmatrix}.$$

Ainsi

$$[u]_{\mathscr{C}\mathrm{an}} = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Alors que, étant donnée une base $(\mathbf{e}_1, \mathbf{e}_2)$ du plan Π, la famille $\mathscr{B} = (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e})$ forme une base de $\mathbb{R}^3$; dans cette base, on a

$$[u]_{\mathscr{B}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Ce qui montre que l'endomorphisme u est diagonalisable. On observera que

$$\mathbb{R}^3 = \Pi \oplus \Delta = \mathrm{Vect}(\mathbf{c}_1, \mathbf{c}_2) \oplus \mathrm{Vect}(\mathbf{e}),$$

et que

i) pour tout $\mathbf{x} \in \Pi$, $u(\mathbf{x}) = \mathbf{x} = 1.\mathbf{x}$,

ii) pour tout $\mathbf{x} \in \Delta$, $u(\mathbf{x}) = 0 = 0.\mathbf{x}$.

On dit dans ce cas qu'il existe deux directions *globalement invariantes* et que u admet pour *valeurs propres* 1 et 0. Nous montrerons que tout endomorphisme de $\mathbb{R}^2$ admettant deux valeurs propres disctinctes est diagonalisable.

4.3.5. Exemple.— Soit u la rotation de $\mathbb{R}^2$ d'angle θ et de centre l'origine. La matrice de u dans la base canonique est

$$[u]_{\mathscr{C}\mathrm{an}} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$$

Si $\theta \neq k\pi$, alors il n'existe pas de direction globalement invariante et u ne possède pas de valeurs propres. Dans ce cas, il est impossible de trouver une base de diagonalisation, ainsi l'endomorphisme u n'est pas diagonalisable.

§ 4 Marches sur un graphe et diagonalisation

4.4.1. Graphes orientés.— Un *graphe (orienté fini)* $\mathscr{G}$ est la donnée d'un ensemble fini de *sommets* $\mathscr{S}$, d'un ensemble fini d'*arcs* $\mathscr{A}$, de deux applications

$$s : \mathscr{A} \longrightarrow \mathscr{S}, \quad t : \mathscr{A} \longrightarrow \mathscr{S}$$

appelées respectivement *source* et *but*. Un arc de source p et de but q est noté graphiquement de la façon suivante

$$p \longrightarrow q.$$

4.4.2. Matrices d'adjacence.— Soit $\mathscr{G}$ un graphe à n sommets. On suppose que les sommets sont numérotés de 1 à n. On associe au graphe G, une matrice $\mathbf{G}$ de $\mathscr{M}_n(\mathbb{R})$ telle que $\mathbf{G}^i_j$ est le nombre d'arcs de source j et de but i dans $\mathscr{G}$. Cette matrice est appelée la *matrice d'adjacence* du graphe $\mathscr{G}$.

Par exemple, la matrice du graphe suivant

$$\begin{tikzcd}
\mathbf{1} \arrow[rr] \arrow[dr] & & \mathbf{2} \arrow[loop above] \arrow[dl, shift right] \arrow[dl, shift left] \\
& \mathbf{3} &
\end{tikzcd}$$

est la matrice

$$\mathbf{G} = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

4.1 Proposition.— Soient $\mathbf{G}$ la matrice associée à un graphe $\mathscr{G}$ et k un entier naturel. Le nombre de chemins de longueur k du sommet j au sommet i dans le graphe $\mathscr{G}$ est donnée par le coefficient $(\mathbf{G}^k)^i_j$.

Preuve. On obtient ce résultat par récurrence sur la longueur du nombre de chemins. Le nombre de chemins de longueur 2 de j à i est

$$\sum_{l=1}^{n} \mathbf{G}^l_j \mathbf{G}^i_l,$$

car $\mathbf{G}^l_j$ est le nombre de chemins de longueur 1 de j à l et $\mathbf{G}^i_l$ est le nombre de chemins de longueur 1 de l à i. Cette somme correspond au coefficient $(\mathbf{G}^2)^i_j$.

On suppose que, pour tous i, j, $(\mathbf{G}^{k-1})^i_j$ est le nombre de chemins de longueur $k-1$ entre i et j. Alors

$$(\mathbf{G}^k)^i_j = \sum_{l=1}^{n} (\mathbf{G}^{k-1})^l_j \mathbf{G}^i_l.$$

□

4.4.3. Exemple.— Un conducteur de train réalise chaque jour un trajet reliant l'une des trois villes de la figure 4.3. Chaque jour, il ne peut faire qu'un des trajets indiqués. En deux jours, il ne pourra donc faire que les quatre trajets suivants au départ de Lyon :

$$\text{Lyon} \rightarrow \text{Paris} \rightarrow \text{Genève}$$

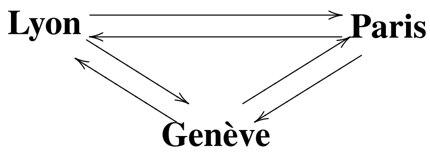


FIGURE 4.3.: Liaisons féroviaires entre trois villes

$$\text{Lyon} \to \text{Genève} \to \text{Paris}$$

$$\text{Lyon} \to \text{Genève} \to \text{Lyon}$$

$$\text{Lyon} \to \text{Paris} \to \text{Lyon}$$

Les itinéraires possibles forment les chemins d'un graphe dont les sommets sont les villes. La matrice de ce graphe est

$$\mathbf{G} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}.$$

Pour calculer le nombre d'itinéraires possibles en k jours d'une ville à une autre, on calcule les coefficients de la matrice $\mathbf{G}^k$. La matrice $\mathbf{G}$ est diagonalisable. On montre en effet

$$\mathbf{D} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{bmatrix} = \mathbf{P}^{-1}\mathbf{G}\mathbf{P},$$

avec

$$\mathbf{P} = \begin{bmatrix} -1 & -1 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix} \qquad \text{et} \qquad \mathbf{P}^{-1} = \frac{1}{3}\begin{bmatrix} -1 & -1 & 2 \\ -1 & 2 & -1 \\ 1 & 1 & 1 \end{bmatrix}.$$

Comme $\mathbf{G} = \mathbf{P}\mathbf{D}\mathbf{P}^{-1}$, on a

$$\mathbf{G}^k = \mathbf{P}\mathbf{D}\mathbf{P}^{-1}\mathbf{P}\mathbf{D}\mathbf{P}^{-1}\ldots\mathbf{P}\mathbf{D}\mathbf{P}^{-1},$$

soit

$$\mathbf{G}^k = \mathbf{P}\mathbf{D}^k\mathbf{P}^{-1}.$$

Donc

$$\mathbf{G}^k = \mathbf{P}\begin{bmatrix} (-1)^k & 0 & 0 \\ 0 & (-1)^k & 0 \\ 0 & 0 & 2^k \end{bmatrix}\mathbf{P}^{-1}.$$

Soit

$$\mathbf{G}^k = \frac{2^k}{3}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} + \frac{(-1)^k}{3}\begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}.$$

En particulier,

$$\mathbf{G}^2 = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}, \quad \mathbf{G}^3 = \begin{bmatrix} 2 & 3 & 3 \\ 3 & 2 & 3 \\ 3 & 3 & 2 \end{bmatrix}, \quad \mathbf{G}^4 = \begin{bmatrix} 6 & 5 & 5 \\ 5 & 6 & 5 \\ 5 & 5 & 6 \end{bmatrix}.$$

—

§ 5 Exercices

Exercice 1.— Soit Π le plan de $\mathbb{R}^3$ d'équation $x+3y+5z=0$ et soit Δ la droite d'équation $3x=5y=15z$. Notons u la projection de $\mathbb{R}^3$ sur le plan Π parallèlement à la droite Δ.

1. Écrire la matrice de l'endomorphisme u dans la base canonique de $\mathbb{R}^3$.
2. Sans calcul, trouver une base de $\mathbb{R}^3$ dans laquelle la matrice de u est la plus simple possible, i.e., ayant le plus grand nombre de coefficients nuls.

Exercice 2.— Soient E un $\mathbb{C}$-espace vectoriel de dimension finie et u un endomorphisme de E vérifiant $u^2 = \mathrm{id}_E$.

1. Montrer que
$$E = \mathrm{Ker}\,(u-\mathrm{id}_E) \oplus \mathrm{Ker}\,(u+\mathrm{id}_E).$$
2. Montrer que si $u \neq \mathrm{id}_E$ et $u \neq -\mathrm{id}_E$, alors il existe une base de E dans laquelle l'endomorphisme u est représenté par la matrice
$$\begin{bmatrix} \mathbf{1}_p & \mathbf{0} \\ \mathbf{0} & -\mathbf{1}_q \end{bmatrix},$$
où $\mathbf{1}_p$ et $\mathbf{1}_q$ désignent les matrices identités de $\mathscr{M}_p(\mathbb{C})$ at $\mathscr{M}_q(\mathbb{C})$ respectivement.

Exercice 3.— On considère les deux matrices de $\mathscr{M}_n(\mathbb{K})$

$$\mathbf{I} = \begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 1 & \dots & 1 & 1 \end{bmatrix}, \qquad \mathbf{A} = \begin{bmatrix} a & b & \dots & b \\ b & a & \ddots & \vdots \\ \vdots & \ddots & \ddots & b \\ b & \dots & b & a \end{bmatrix}.$$

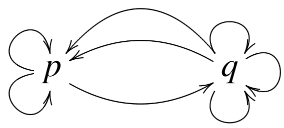
1. Écrire $\mathbf{A}$ comme combinaison linéaire de $\mathbf{I}$ et $\mathbf{1}_n$.
2. Calculer $\mathbf{I}^k$ et $\mathbf{A}^k$, pour tout entier naturel k.

Exercice 4.— Soit $\mathbf{M} = \begin{bmatrix} 2 & 1 \\ 2 & 3 \end{bmatrix}$ et $\mathbf{P} = \begin{bmatrix} -1 & \frac{1}{2} \\ 1 & 1 \end{bmatrix}$.

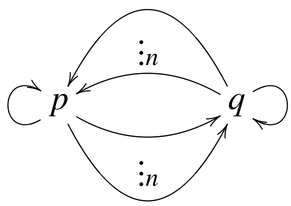
1. Montrer que $\mathbf{P}^{-1}\mathbf{MP}$ est de la forme

$$\mathbf{D} = \begin{bmatrix} \lambda & 0 \\ 0 & \mu \end{bmatrix}.$$

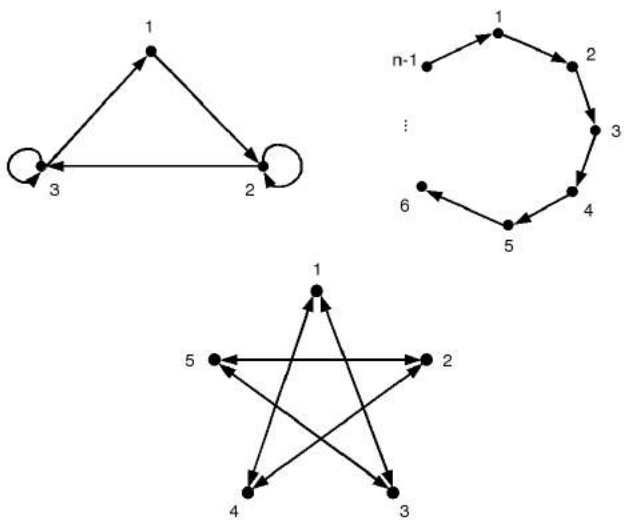
2. Combien y a-t-il de chemins de longueur 2, 3 et 4 dans le graphe ci-dessous ?



3. Comparer les résultats avec $\mathbf{M}^2$, $\mathbf{M}^3$, $\mathbf{M}^4$ que l'on calculera à l'aide de la matrice $\mathbf{D}$.
4. Déterminer, pour tout entier $k \geqslant 1$, le nombre de chemins de longueur k dans le graphe suivant



Exercice 5.— Écrire les matrices d'adjacence des graphes suivants



CHAPITRE
5

Valeurs propres et vecteurs propres

Sommaire

L'objectif de ce chapitre est de définir les notions de valeurs propres et vecteurs propres et de donner des méthodes permettant de déterminer ces valeurs propres, notamment via le calcul avec le polynôme caractéristique.

§ 1 Préliminaires

Nous avons vu dans le chapitre 4 que la notion de valeur propre d'une matrice intervient dans le découplage de systèmes d'équations. Les vecteurs propres interviennent pour décrire le changement de variable permettant d'obtenir un système découplé Nous avons illustré cela avec les systèmes d'équations différentielles linéaires et les systèmes d'équations récurrentes. Avant de définir les notions de valeur propre et vecteur propre dans la section suivante, considérons un nouvel exemple d'illustration avec le problème de migration de population suivant.

5.1.1. Migration de population.— On étudie la répartition de la population entre deux zones géographiques d'un même pays, disons entre une région nord et une région sud. Chaque année, un mouvement de population entre ces deux zones s'opère de la façon suivante :

- 80% de la population des régions du nord part vivre dans les régions du sud,

\- 70% de la population des régions du sud rejoint les régions nords.

On suppose que la population totale du pays reste constante d'une année à l'autre, aussi 20% des habitants des régions du nords restent dans le nord et 30% des habitants des régions du sud restent dans le sud. Le mouvement de population est schématisé par la figure 5.1

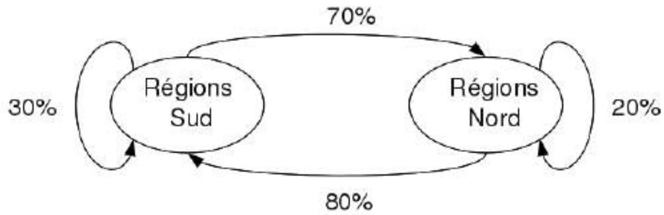


FIGURE 5.1.: Mouvement de population entre deux zones géographiques

Étant donné une répartition initiale, l'année $k = 0$, de la population entre les deux zones, quelle sera la répartition de population la k-ième année ? Est-ce qu'au bout d'un certain temps, toute la population va se retrouver en zone rurale ?

5.1.2. Modélisation du problème.— Notons n_k le nombre d'habitants des régions du nord la k-ième année et s_k le nombre d'habitants des régions du sud la k-ième année. L'évolution des deux populations est régie par le système d'équations couplées suivant

$$\left|\begin{array}{l} n_{k+1} = 0.2n_k + 0.7s_k \\ s_{k+1} = 0.8n_k + 0.3s_k \end{array}\right.$$

qui s'écrit matriciellement sous la forme :

$$\begin{bmatrix} n_{k+1} \\ s_{k+1} \end{bmatrix} = \mathbf{A} \begin{bmatrix} n_k \\ s_k \end{bmatrix},$$

avec

$$\mathbf{A} = \begin{bmatrix} 0.2 & 0.7 \\ 0.8 & 0.3 \end{bmatrix}$$

et où le vecteur $\mathbf{p}(k) = \begin{bmatrix} n_k \\ s_k \end{bmatrix}$ exprime la répartition de la population au nord et au sud l'année k. Par exemple, si l'année $k = 0$, on a un million d'habitants au nord et dix mille habitants au sud, i.e., $\mathbf{p}(0) = \begin{bmatrix} 10^6 \\ 10^4 \end{bmatrix}$, alors l'année suivante, à $k = 1$, il y aura

$$\begin{aligned} \mathbf{p}(1) = \begin{bmatrix} n_1 \\ s_1 \end{bmatrix} &= \mathbf{A} \begin{bmatrix} 10^6 \\ 10^4 \end{bmatrix} \\ &= \begin{bmatrix} 207000 \\ 803000 \end{bmatrix}. \end{aligned}$$

Pour connaître la population dans chacune des régions la k-ième année, on peut appliquer la formule suivante, obtenue par récurrence sur l'entier k :

$$\begin{bmatrix} n_k \\ s_k \end{bmatrix} = \mathbf{A}^k \begin{bmatrix} n_0 \\ s_0 \end{bmatrix}.$$

Voici les valeurs du vecteur $\mathbf{p}(k)$ pour quelques valeurs de k :

$$\begin{aligned} &\mathbf{p}(1) = \begin{bmatrix} 207000 \\ 803000 \end{bmatrix}, \quad \mathbf{p}(2) = \begin{bmatrix} 603500 \\ 406500 \end{bmatrix}, \quad \mathbf{p}(3) = \begin{bmatrix} 405250 \\ 604750 \end{bmatrix}, \\ &\mathbf{p}(4) = \begin{bmatrix} 504375 \\ 505625 \end{bmatrix}, \quad \mathbf{p}(5) = \begin{bmatrix} 454812.5 \\ 555187.5 \end{bmatrix}, \quad \mathbf{p}(6) = \begin{bmatrix} 479593.75 \\ 530406.25 \end{bmatrix}, \\ &\mathbf{p}(7) = \begin{bmatrix} 467203.12 \\ 542796.87 \end{bmatrix}, \quad \mathbf{p}(8) = \begin{bmatrix} 473398.43 \\ 536601.56 \end{bmatrix}, \quad \mathbf{p}(30) = \begin{bmatrix} 471333.33 \\ 538666.66 \end{bmatrix}. \end{aligned} \tag{5.1}$$

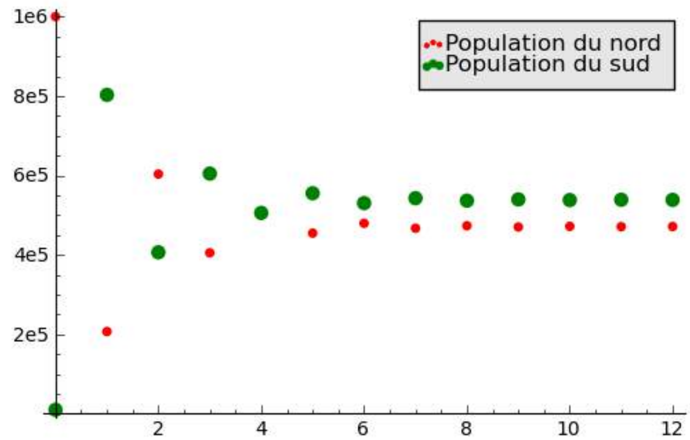


FIGURE 5.2.: Évolution de la population entre le nord et le sud les 12 premières années

Nous allons montrer que le calcul des puissances k-ième successives de la matrice $\mathbf{A}$ peut se faire à l'aide d'un simple calcul de valeurs propres de la matrice $\mathbf{A}$. Cette approche permet d'éviter les multiplications de la matrice $\mathbf{A}$ par elle-même, cette opération étant coûteuse en nombre d'opérations. Cette approche va nous permettre aussi de mieux comprendre l'évolution du système, ici de l'évolution de la population dans chaque zone géographique.

5.1.3. Calcul des valeurs propres.— Nous allons donner dans ce chapitre une méthode de calcul des valeurs propres pour une matrice : les valeurs propres de la matrice $\mathbf{A}$ sont les solutions λ de l'équation suivante :

$$\det(\mathbf{A} - \lambda \mathbf{1}_2) = 0.$$

Soit

$$\begin{vmatrix} 0.2-\lambda & 0.7 \\ 0.8 & 0.3-\lambda \end{vmatrix} = \lambda^2 - 0.5\lambda - 0.5 = (\lambda-1)(\lambda+0.5) = 0.$$

La matrice $\mathbf{A}$ possède ainsi deux valeurs propres, $\lambda = 1$ et $\lambda = -0.5$, qui sont les racines du polynôme $\lambda^2 - 0.5\lambda - 0.5$.

Si λ est valeur propre, comme $\det(\mathbf{A} - \lambda \mathbf{1}_2) = 0$, la matrice $\mathbf{A} - \lambda \mathbf{1}_2$ admet un noyau non trivial. Autrement dit, il existe des vecteurs $\mathbf{x}$ solutions de l'équation suivante

$$\mathbf{A}\mathbf{x} = \lambda \mathbf{x}.$$

Ces vecteurs sont dans la même direction que $\mathbf{A}\mathbf{x}$, on les appelles *vecteurs propres* associés à la valeur propre λ. La valeur propre indique si les vecteurs propres sont laissés inchangés, s'ils sont étirés, réduits ou encore inversés. Dans notre exemple, les vecteurs propres associés à la valeur propre $\lambda = 1$ reste inchangés, ils vérifient $\mathbf{A}\mathbf{x} = \mathbf{x}$. Ceux associés à la valeur propre $\lambda = -0.5$ sont réduits et inversés : $\mathbf{A}\mathbf{x} = -0.5\mathbf{x}$.

5.1.4. Les sous-espaces propres.— Lorsque l'on calcule les puissances de la matrice $\mathbf{A}$, les valeurs propres sont élevées à la puissance, alors que les vecteurs propres restent inchangés. En effet, pour tout vecteur propre $\mathbf{x}$ associé à la valeur propre -0.5, on a

$$\mathbf{A}^2\mathbf{x} = \mathbf{A}(\mathbf{A}\mathbf{x}) = -0.5\mathbf{A}\mathbf{x} = (-0.5)^2\mathbf{x},$$

ainsi, pour tout entier k,

$$\mathbf{A}^k\mathbf{x} = (-0.5)^k\mathbf{x}.$$

Pour un vecteur propre $\mathbf{x}$ associé à la valeur propre 1, on a $\mathbf{A}^k\mathbf{x} = \mathbf{x}$, pour tout entier k.

Par suite, en décomposant les vecteurs dans une base de vecteurs propres, on pourra alors calculer les puissances de $\mathbf{A}$ simplement en calculant les puissances des valeurs propres.

L'ensemble des vecteurs propres associés à la valeur propre -0.5 est défini par

$$E_{-0.5} = \{\mathbf{x} \in \mathbb{R}^2 \mid \mathbf{A}\mathbf{x} = -0.5\mathbf{x}\}.$$

Donc $E_{-0.5}$ est formé des vecteurs $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ dont les composantes satisfont le système d'équations suivant :

$$\left| \begin{array}{l} -0.5x_1 = 0.2x_1 + 0.7x_2 \\ -0.5x_2 = 0.8x_1 + 0.3x_2 \end{array} \right.$$

Ainsi, $E_{-0.5}$ est l'espace vectoriel de dimension 1 engendré par le vecteur $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$. De la même façon, on montre que l'espace vectoriel formé des vecteurs propres associés à la valeur propre 1, défini par

$$E_1 = \{\mathbf{x} \in \mathbb{R}^2 \mid \mathbf{A}\mathbf{x} = \mathbf{x}\}.$$

est engendré par le vecteur $\begin{bmatrix} 0.875 \\ 1 \end{bmatrix}$. Ces deux vecteurs forment une base de $\mathbb{R}^2$. En particulier, étant donné le vecteur initial $\begin{bmatrix} n_0 \\ s_0 \end{bmatrix}$, il existe deux réels uniques β_1 et β_2 tels que

$$\begin{bmatrix} n_0 \\ s_0 \end{bmatrix} = \beta_1 \begin{bmatrix} 1 \\ -1 \end{bmatrix} + \beta_2 \begin{bmatrix} 0.875 \\ 1 \end{bmatrix}.$$

De cette équation, on déduit les deux réels β_1 et β_2, on a

$$\beta_1 = \frac{1646.10^3}{3} \quad \text{et} \quad \beta_2 = \frac{1616.10^3}{3}.$$

On peut alors exprimer le vecteur $\begin{bmatrix} n_k \\ s_k \end{bmatrix}$, on a

$$\begin{aligned}
\begin{bmatrix} n_k \\ s_k \end{bmatrix} &= \mathbf{A}^k \begin{bmatrix} n_0 \\ s_0 \end{bmatrix} \\
&= \beta_1 \mathbf{A}^k \begin{bmatrix} 1 \\ -1 \end{bmatrix} + \beta_2 \mathbf{A}^k \begin{bmatrix} 0.875 \\ 1 \end{bmatrix} \\
&= \beta_1 (-0.5)^k \begin{bmatrix} 1 \\ -1 \end{bmatrix} + \beta_2 \begin{bmatrix} 0.875 \\ 1 \end{bmatrix}
\end{aligned}$$

D'où

$$\begin{bmatrix} n_k \\ s_k \end{bmatrix} = \frac{1646.10^3}{3} (-0.5)^k \begin{bmatrix} 1 \\ -1 \end{bmatrix} + \frac{1616.10^3}{3} \begin{bmatrix} 0.875 \\ 1 \end{bmatrix}$$

On a ainsi résolu le problème du calcul de la répartition de la population la k-ième année sans avoir à faire le produit de la matrice $\mathbf{A}$ par elle-même k-fois. Si k tend vers l'infini, on a de façon immédiate que

$$\lim_{k \to +\infty} \begin{bmatrix} n_k \\ s_k \end{bmatrix} = \frac{1616.10^3}{3} \begin{bmatrix} 0.875 \\ 1 \end{bmatrix} = \begin{bmatrix} 471333,33 \\ 538666,66 \end{bmatrix}$$

Nous retrouvons les valeurs numériques obtenues en (5.1).

§ 2 Valeurs propres et espaces propres

5.2.1. Définitions.— Soit $\mathbf{A}$ une matrice de $\mathscr{M}_n(\mathbb{K})$. Un scalaire λ dans $\mathbb{K}$ est appelé *valeur propre* de $\mathbf{A}$, si l'une des conditions équivalentes suivantes est vérifiée :

i) Ker $(\mathbf{A} - \lambda \mathbf{1}_n) \neq \{0\}$,

ii) $\det(\mathbf{A} - \lambda \mathbf{1}_n) = 0$,

iii) il existe un vecteur non nul $\mathbf{x}$ de $\mathbb{K}^n$, solution de l'équation

$$\mathbf{A}\mathbf{x} = \lambda \mathbf{x}.$$

Les équivalences des assertions **i)**, **ii)** sont immédiates. L'assertion **iii)** est une simple reformulation de **i)**

Un vecteur $\mathbf{x}$ vérifiant l'assertion **iv)** est appelé *vecteur propre* de $\mathbf{A}$ associé à la valeur propre λ. Le sous ensemble de $\mathbb{K}$ formé des valeurs propres de $\mathbf{A}$ est appelé le *spectre* sur $\mathbb{K}$ de la matrice $\mathbf{A}$ et sera noté dans la suite $\mathrm{Sp}_{\mathbb{K}}(\mathbf{A})$, ou $\mathrm{Sp}(\mathbf{A})$ s'il n'y a pas de confusion possible avec le corps de base.

5.2.2. Exemple.— Considérons la matrice suivante de $\mathscr{M}_2(\mathbb{R})$:

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Alors, on a

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

ainsi le vecteur $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ est un vecteur propre de $\mathbf{A}$ associé à la valeur propre 1. Par ailleurs, on a

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 \\ -1 \end{bmatrix} = -1\begin{bmatrix} 1 \\ -1 \end{bmatrix},$$

le vecteur $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ est un vecteur propre de $\mathbf{A}$ associé à la valeur propre -1.

5.2.3. Sous-espaces propres.—

5.1 Proposition.— Soit $\mathbf{A}$ une matrice de $\mathscr{M}_n(\mathbb{K})$ et soit λ une valeur propre de $\mathbf{A}$. L'ensemble des vecteurs propres associés à λ

$$E_\lambda = \{\mathbf{x} \in \mathbb{K}^n \mid \mathbf{A}\mathbf{x} = \lambda\mathbf{x}\}$$

est un sous-espace vectoriel de $\mathbb{K}^n$, stable par $\mathbf{A}$.

Preuve. On vérifie que l'ensemble E_λ contient le vecteur nul et que si $\mathbf{x}$ et $\mathbf{y}$ sont deux vecteurs de E_λ, alors

$$\begin{aligned}\mathbf{A}(\alpha\mathbf{x}+\beta\mathbf{y}) &= \alpha\mathbf{A}\mathbf{x}+\beta\mathbf{A}\mathbf{y},\\ &= \alpha\lambda\mathbf{x}+\beta\lambda\mathbf{y},\\ &= \lambda(\alpha\mathbf{x}+\beta\mathbf{y}).\end{aligned}$$

Plus simplement, on peut remarquer que $E_\lambda = \mathrm{Ker}\,(\mathbf{A}-\lambda\mathbf{1}_n)$ est un sous-espace vectoriel de $\mathbb{K}^n$, en tant que noyau d'une matrice.

Montrons la stabilité du sous-espace E_λ par $\mathbf{A}$. Si $\mathbf{x} \in E_\lambda$, on a

$$\mathbf{A}(\mathbf{A}\mathbf{x}) = \mathbf{A}(\lambda\mathbf{x}) = \lambda\mathbf{A}\mathbf{x},$$

donc $\mathbf{A}\mathbf{x} \in E_\lambda$. $\square$

5.2.4. Définition.— L'espace E_λ est appelé le *sous-espace propre* de la matrice $\mathbf{A}$ associé à la valeur propre λ.

5.2.5. Exemple.— Nous avons montré que la matrice

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},$$

admet pour valeur propres -1 et 1. Les sous-espaces propres associés sont

$$E_{-1} = \mathrm{Vect}(\begin{bmatrix} 1 \\ -1 \end{bmatrix}), \qquad E_1 = \mathrm{Vect}(\begin{bmatrix} 1 \\ 1 \end{bmatrix}).$$

5.2.6. Somme de sous-espaces propres.— Soit $\mathbf{A}$ une matrice de $\mathscr{M}_n(\mathbb{K})$ admettant deux valeurs propres distinctes λ_1 et λ_2. Les sous-espaces propres E_{λ_1} et E_{λ_2} sont alors en somme directe. En effet, si $\mathbf{x}$ est un vecteur propre de $\mathbf{A}$ associé aux deux valeurs propres λ_1 et λ_2. On a

$$\mathbf{A}\mathbf{x} = \lambda_1\mathbf{x} \quad \text{et} \quad \mathbf{A}\mathbf{x} = \lambda_2\mathbf{x}.$$

D'où

$$(\lambda_1 - \lambda_2)\mathbf{x} = \mathbf{0}.$$

Comme les valeurs propres λ_1 et λ_2 sont supposées distinctes, on en déduit que $\mathbf{x} = \mathbf{0}$. On a donc

$$E_{\lambda_1} \cap E_{\lambda_2} = \{\mathbf{0}\}.$$

C'est ainsi que E_{λ_1} et E_{λ_2} sont en somme directe.

Plus généralement on a :

5.2 Proposition.— Soient $\lambda_1, \ldots, \lambda_p$ des valeurs propres d'une matrice $\mathbf{A}$ de $\mathscr{M}_n(\mathbb{K})$, distinctes deux à deux. Les sous-espaces propres E_{λ_1} , ... , E_{λ_p} sont en somme directe.

Preuve. On procède par récurrence sur le nombre p de valeurs propres distinctes. Si $p = 1$, il n'y a rien à démontrer. Supposons que les sous-espaces $E_{\lambda_1}, \ldots, E_{\lambda_p}$ soient en somme directe. D'après 1.3.3, c'est équivalent à

$$E_{\lambda_1} \cap E_{\lambda_2} = \{\mathbf{0}\}, \quad (E_{\lambda_1} + E_{\lambda_2}) \cap E_{\lambda_3} = \{\mathbf{0}\}, \quad \ldots \quad , (E_{\lambda_1} + E_{\lambda_2} + \cdots + E_{\lambda_{p-1}}) \cap E_{\lambda_p} = \{\mathbf{0}\}.$$

Pour montrer que les sous-espaces $E_{\lambda_1}, \ldots, E_{\lambda_p}, E_{\lambda_{p+1}}$ sont en somme directe, il suffit de montrer que

$$(E_{\lambda_1} + E_{\lambda_2} + \cdots + E_{\lambda_p}) \cap E_{\lambda_{p+1}} = \{\mathbf{0}\}.$$

Soit $\mathbf{x} = \mathbf{x}_1 + \cdots + \mathbf{x}_p$ un vecteur de $(E_{\lambda_1} + E_{\lambda_2} + \cdots + E_{\lambda_p}) \cap E_{\lambda_{p+1}}$, où $\mathbf{x}_k \in E_{\lambda_k}$, pour tout $k \in [\![1, p]\!]$. Comme $\mathbf{A}\mathbf{x}_k = \lambda_k\mathbf{x}_k$, on a

$$\mathbf{A}\mathbf{x} = \lambda_1\mathbf{x}_1 + \cdots + \lambda_p\mathbf{x}_p.$$

Par ailleurs, $\mathbf{x}$ est un vecteur propre associé à λ_{p+1}, d'où

$$\mathbf{A}\mathbf{x} = \lambda_{p+1}\mathbf{x} = \lambda_{p+1}\mathbf{x}_1 + \cdots + \lambda_{p+1}\mathbf{x}_p.$$

Par suite,

$$\mathbf{0} = (\lambda_1 - \lambda_{p+1})\mathbf{x}_1 + \cdots + (\lambda_p - \lambda_{p+1})\mathbf{x}_p.$$

Les sous-espaces $E_{\lambda_1}, \ldots, E_{\lambda_p}$ étant en somme directe, $\mathbf{0} = \mathbf{0}_{E_{\lambda_1}} + \cdots + \mathbf{0}_{E_{\lambda_p}}$ est la seule décomposition de $\mathbf{0}$. D'où, pour tout $k \in [\![1,p]\!]$,

$$(\lambda_k - \lambda_{k+1})\mathbf{x}_k = \mathbf{0}_{E_{\lambda_k}},$$

Comme les valeurs propres λ_i sont deux à deux distinctes, tous les vecteurs $\mathbf{x}_k$ sont nuls, par suite le vecteur $\mathbf{x}$ est nul. Ainsi, le sous-espace $(E_{\lambda_1} + E_{\lambda_2} + \cdots + E_{\lambda_p}) \cap E_{\lambda_{p+1}}$ est trivial et les sous-espaces $E_{\lambda_1}, \ldots, E_{\lambda_p}, E_{\lambda_{p+1}}$ forment une somme directe. □

De ce résultat, on déduit que si une matrice $\mathbf{A}$ de $\mathscr{M}_n(\mathbb{K})$ admet n valeurs propres disctinctes $\lambda_1, \ldots, \lambda_n$, alors

$$\mathbb{K}^n = E_{\lambda_1} \oplus \ldots \oplus E_{\lambda_n}.$$

5.2.7. Exemple.— Considérons la matrice suivante de $\mathscr{M}_2(\mathbb{R})$

$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

La matrice $\mathbf{A}$ est de rang 1, du théorème du rang, on déduit que son noyau Ker $\mathbf{A}$, qui est le sous-espace propre associé à la valeur propre 0, est de dimension 1. On remarque que $E_0 = \mathrm{Vect}(\mathbf{e}_1)$ avec $\mathbf{e}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$. On remarque aussi que le vecteur $\mathbf{e}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ est vecteur propre de $\mathbf{A}$ associé à la valeur propre 2, car $\mathbf{A}\mathbf{e}_2 = 2\mathbf{e}_2$. Le sous-espace propre E_2 est de dimension 1 et engendré par le vecteur $\mathbf{e}_2$. Les vecteur $\mathbf{e}_1$ et $\mathbf{e}_2$ forment une base $\mathscr{B}$ de $\mathbb{R}^2$, on a :

$$\mathbb{R}^2 = E_0 \oplus E_2.$$

Considérons la matrice de passage de la base canonique de $\mathbb{R}^2$ à la base $\mathscr{B}$:

$$\mathbf{P}_{\mathscr{C}\mathrm{an}}^{\mathscr{B}} = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}.$$

Elle admet pour inverse la matrice

$${\mathbf{P}_{\mathscr{C}\mathrm{an}}^{\mathscr{B}}}^{-1} = \frac{1}{2}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}.$$

On vérifie alors que

$$\begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix} = {\mathbf{P}_{\mathscr{C}\mathrm{an}}^{\mathscr{B}}}^{-1}\mathbf{A}\mathbf{P}_{\mathscr{C}\mathrm{an}}^{\mathscr{B}}.$$

§ 3 Calcul des valeurs propres

5.3.1. Définition.— Le *polynôme caractéristique* d'une matrice $\mathbf{A}$ de $\mathscr{M}_n(\mathbb{K})$ est le polynôme de $\mathbb{K}[x]$ défini par

$$p_{\mathbf{A}} = \det(\mathbf{A} - x\mathbf{1}_n).$$

5.3.2. Exemples.— Le polynôme caractéristique de la matrice

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

est

$$p_{\mathbf{A}} = \begin{vmatrix} -x & 1 \\ 1 & -x \end{vmatrix} = x^2 - 1 = (x-1)(x+1).$$

Le polynôme caractéristique de la matrice

$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

est

$$p_{\mathbf{A}} = \begin{vmatrix} 1-x & 1 \\ 1 & 1-x \end{vmatrix} = (1-x)^2 - 1 = -x(2-x).$$

5.3 Proposition.— Deux matrices semblables ont le même polynôme caractéristique.

Preuve. Soient $\mathbf{A}$ et $\mathbf{B}$ deux matrices semblables, montrons que $p_{\mathbf{A}} = p_{\mathbf{B}}$. Par hypothèse, il existe une matrice inversible $\mathbf{P}$ telle que $\mathbf{B} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$. Il découle de la propriété du déterminant que :

$$\begin{aligned} p_{\mathbf{P}^{-1}\mathbf{A}\mathbf{P}} &= \det(\mathbf{P}^{-1}\mathbf{A}\mathbf{P} - x\mathbf{1}_n) \\ &= \det(\mathbf{P}^{-1}(\mathbf{A} - x\mathbf{1}_n)\mathbf{P}) \\ &= (\det\mathbf{P})^{-1}\det(\mathbf{A} - x\mathbf{1}_n)\det\mathbf{P}. \end{aligned}$$

Ainsi $p_{\mathbf{A}} = p_{\mathbf{B}}$. □

Par définition, un scalaire λ est valeur propre d'une matrice $\mathbf{A}$ si le déterminant de la matrice $\mathbf{A} - \lambda\mathbf{1}_n$ est nul. On a donc

5.4 Proposition.— Un scalaire λ est valeur propre d'une matrice $\mathbf{A}$ si, et seulement si, λ est racine de son polynôme caractéristique $p_{\mathbf{A}}$.

5.3.3. Exemple.— Le polynôme caractéristique d'une matrice de $\mathscr{M}_n(\mathbb{K})$ peut ne pas avoir de racines dans le corps $\mathbb{K}$. Par exemple, le polynôme caractéristique de la matrice

$$\mathbf{A} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

est $p_\mathbf{A} = x^2 + 1$, qui n'a pas de racines réelles, par suite $\mathrm{Sp}_\mathbb{R}(\mathbf{A}) = \emptyset$. Dans le chapitre suivant, nous verrons qu'alors la matrice $\mathbf{A}$ n'est pas diagonalisable dans $\mathscr{M}_2(\mathbb{R})$, mais est diagonalisable dans $\mathscr{M}_2(\mathbb{C})$. En effet, on a $\mathrm{Sp}_\mathbb{C}(\mathbf{A}) = \{-i, i\}$ et les sous-espaces propres de $\mathbf{A}$ dans $\mathbb{C}^2$ sont

$$E_i = \mathrm{Vect}(\begin{bmatrix} i \\ 1 \end{bmatrix}) \quad \text{et} \quad E_{-i} = \mathrm{Vect}(\begin{bmatrix} -i \\ 1 \end{bmatrix}).$$

5.3.4. Exemple.— L'exemple précédent correspond à la matrice de la rotation du plan vectoriel $\mathbb{R}^2$ d'angle $\pi/2$. Plus généralement, considérons la matrice de la rotation d'angle
$\theta \neq 0 \mod (\pi)$ représentée dans la base canonique

$$\mathbf{R}_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$$

Le polynôme caractéristique $p_{\mathbf{R}_\theta} = x^2 - 2\cos\theta x + 1$ est de discriminant $-4\sin^2\theta$ non nul, car $\theta \neq 0 \mod (\pi)$. Il ne possède donc pas de racine réelle, donc la matrice $\mathbf{R}_\theta$ ne possède pas de valeur propre réelle. Géométriquement, cela s'interprète en disant qu'il n'existe pas de droite globalement invariante par la rotation d'angle θ, lorsque $\theta \neq 0 \mod (\pi)$.

Si l'on se place sur le corps $\mathbb{C}$, i.e., on regarde la matrice $\mathbf{R}_\theta$ comme une matrice de $\mathscr{M}_2(\mathbb{C})$. Alors, le polynôme $p_{\mathbf{R}_\theta}$ se factorise dans $\mathbb{C}[x]$ sous la forme

$$p_{\mathbf{R}_\theta} = (x - e^{i\theta})(x - e^{-i\theta}).$$

Dans $\mathbb{C}$, la matrcie $\mathbf{R}_\theta$ admet 2 valeurs propres :

$$\mathrm{Sp}_\mathbb{C}(\mathbf{R}_\theta) = \{e^{i\theta}, e^{-i\theta}\}.$$

5.5 Proposition.— Soit une matrice de $\mathscr{M}_n(\mathbb{K})$

$$\mathbf{M} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{0} & \mathbf{C} \end{bmatrix}$$

définie par blocs, où les blocs $\mathbf{A}$ et $\mathbf{C}$ sont carrés. Alors les polynôme caractéristiques de $\mathbf{A}$ et $\mathbf{C}$ divisent le polynôme caractéristique de $\mathbf{M}$.

Preuve. Si $\mathbf{A} \in \mathscr{M}_p(\mathbb{K})$, on a

$$\begin{aligned} p_{\mathbf{M}} &= \det(\mathbf{M} - x\mathbf{1}_n) \\ &= \det \begin{bmatrix} \mathbf{A} - x\mathbf{1}_p & \mathbf{B} \\ \mathbf{0} & \mathbf{C} - x\mathbf{1}_{n-p} \end{bmatrix}, \\ &= \det(\mathbf{A} - x\mathbf{1}_p)\det(\mathbf{C} - x\mathbf{1}_{n-p}) \\ &= p_{\mathbf{A}} p_{\mathbf{C}}. \end{aligned}$$

Ainsi, les polynômes $p_{\mathbf{A}}$ et $p_{\mathbf{C}}$ divisent $p_{\mathbf{M}}$. □

§ 4 Le cas des endomorphismes

Soit E un $\mathbb{K}$-espace vectoriel de dimension finie n. Le $\mathbb{K}$-espace vectoriel $\mathscr{L}(E)$ formé des endomorphismes de E est isomorphe au $\mathbb{K}$-espace vectoriel $\mathscr{M}_n(\mathbb{K})$. Étant donnée une base $\mathscr{B}$ de E, un isomorphisme est donné par l'application

$$\begin{aligned} \Phi : \mathscr{L}(E) &\longrightarrow \mathscr{M}_n(\mathbb{K}) \\ u &\longmapsto [u]_{\mathscr{B}}. \end{aligned}$$

Par suite, les notions de valeur propre et de vecteur propre se définissent pour les endomorphismes de la même façon que pour les matrices.

5.4.1. Valeurs propres et vecteurs propres d'endomorphismes.— Soit u un endomorphisme de E.

Un scalaire λ dans $\mathbb{K}$ est appelé *valeur propre* de u si l'une des conditions équivalentes suivantes est vérifiée :

- **i)** l'application $u - \lambda \mathrm{id}_E$ est non injective,
- **ii)** Ker $(u - \lambda \mathrm{id}_E) \neq \{0\}$,
- **iii)** $\det(u - \lambda \mathrm{id}_E) = 0$,
- **iv)** il existe un vecteur non nul $\mathbf{x}$ de E tel que $u(\mathbf{x}) = \lambda \mathbf{x}$.
- **v)** λ est valeur propre de la matrice $[u]_{\mathscr{B}}$ de u, exprimé dans une base $\mathscr{B}$ de E.

Les équivalences des assertions **i)**, **ii)** et **iii)** sont données par la proposition 1.10. L'assertion **iv)** est une simple reformulation de **ii)**. Pour tout base $\mathscr{B}$ de E, λ est valeur propre de la matrice $[u]_{\mathscr{B}}$ si, et seulement si, λ est valeur propre de l'endomorphisme u. En effet, si $[u]_{\mathscr{B}}\mathbf{x} = \lambda \mathbf{x}$, alors si $\widetilde{\mathbf{x}}$ désigne le vecteur de E tel que $\mathbf{x} = [\widetilde{\mathbf{x}}]_{\mathscr{B}}$, on a $u(\widetilde{\mathbf{x}}) = \lambda \widetilde{\mathbf{x}}$. Ce qui montre l'équivalence de **iv)** avec **v)**.

Un vecteur $\mathbf{x}$ vérifiant l'assertion **iv)** est appelé *vecteur propre* de u associé à la valeur propre λ. L'ensemble des valeurs propres de u est appelé le *spectre* de u et sera noté dans la suite $\mathrm{Sp}_{\mathbb{K}}(u)$, ou Sp(u) s'il n'y a pas de confusion possible avec le corps de base.

5.4.2. Exemple.— Soit u l'endomorphisme de $\mathbb{R}^2$ défini par

$$\begin{aligned} u : \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ \begin{bmatrix} x \\ y \end{bmatrix} &\longmapsto \begin{bmatrix} y \\ x \end{bmatrix}. \end{aligned}$$

Alors la matrice u dans la base canonique de $\mathbb{R}^2$ est

$$[u]_{\mathscr{C}\mathrm{an}} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Les valeurs propres de l'endomorphime u sont les valeurs propres de la matrices $[u]_{\mathscr{C}\mathrm{an}}$, soit $\mathrm{Sp(u)} = \{-1, 1\}$.

5.4.3. Polynôme caractéristique d'un endomorphisme.— D'après la proposition 5.3, deux matrices semblables admettent le même polynôme caractéristique. Par suite, pour un endomorphisme u de E, le polynôme caractéristique de la matrice $[u]_{\mathscr{B}}$ est indépendant de la base $\mathscr{B}$. Autrement dit, si $\mathscr{B}'$ est une autre base de E, les polynômes caractéristiques des matrices $[u]_{\mathscr{B}}$ et $[u]_{\mathscr{B}'}$ sont égaux.

On appele *polynôme caractéristique* de u, noté p_u, le polynôme caractéristique d'une matrice de u exprimé dans une base de E.

On a l'analogue de la proposition 5.4. Par définition, un scalaire λ est valeur propre de u si $\det(u - \lambda \mathrm{id}_E) = 0$. Par suite, on a

5.6 Proposition.— Un scalaire λ est valeur propre d'un endomorphisme u si, et seulement si, λ est racine de son polynôme caractéristique p_u.

Le résultat suivant relie le polynôme caractéristique de la restriction d'un endomorphisme à un sous-espace stable au polynôme caractéristique de l'endomorphisme.

5.7 Proposition.— Soient F un sous-espace vectoriel non nul de E stable par u et $u_{|F}$ la restriction de u à F. Alors le polynôme $p_{u_{|F}}$ divise le polynôme p_u.

Preuve. Le sous-espace F est stable par u, donc la restriction $u_{|F}$ de u à F est un endomorphisme de F. Considérons une base $\mathscr{B}' = (e_1, \ldots, e_p)$ de F. On la complète en une base $\mathscr{B} = (e_1, \ldots, e_n)$ de E. La matrice de l'endomorphisme u, exprimée dans la base $\mathscr{B}$, se décompose par blocs de la façon suivante :

$$[u]_{\mathscr{B}} = \begin{bmatrix} [u_{|F}]_{\mathscr{B}'} & \mathbf{B} \\ \mathbf{0} & \mathbf{C} \end{bmatrix}$$

par suite, d'après la proposition 5.5, le polynôme $p_{u_{|F}}$ divise le polynôme p_u. $\square$

—

§ 5 Exercices

Exercice 1.— Déterminer les valeurs propres et espaces propres des matrices suivantes de $\mathscr{M}_3(\mathbb{R})$:

$$\textbf{a)} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \textbf{b)} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \textbf{c)} \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix},$$

$$\textbf{d)} \begin{bmatrix} 0 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \quad \textbf{e)} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}, \quad \textbf{f)} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix},$$

$$\textbf{g)} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}, \quad \textbf{h)} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}.$$

Exercice 2.— Déterminer les valeurs propres et vecteurs propres des matrices suivantes

$$\mathbf{A} = \begin{bmatrix} -10 & -7 \\ 14 & 11 \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} 2 & 16 & 8 \\ 4 & 14 & 8 \\ -8 & -32 & -18 \end{bmatrix} \quad \mathbf{C} = \begin{bmatrix} 3 & -2 & 5 \\ 0 & 1 & 4 \\ 0 & -1 & 5 \end{bmatrix}$$

$$\mathbf{D} = \begin{bmatrix} 0 & 6 & 3 \\ -1 & 5 & 1 \\ -1 & 2 & 4 \end{bmatrix} \quad \mathbf{E} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

Exercice 3.— Soit $\mathbf{T} = \begin{bmatrix} \mathbf{A} & \mathbf{C} \\ \mathbf{0} & \mathbf{B} \end{bmatrix}$ une matrice définie par blocs, où les matrices $\mathbf{A}$ et $\mathbf{B}$ sont carrées. Montrer que

$$\mathrm{Spec}(\mathbf{T}) = \mathrm{Spec}(\mathbf{A}) \cup \mathrm{Spec}(\mathbf{B}).$$

Exercice 4.— Construire des exemples de matrices $\mathbf{A}$ et $\mathbf{B}$ de $\mathscr{M}_2(\mathbb{K})$ telles que

1. λ valeur propre de $\mathbf{A}$ et β valeur propre de $\mathbf{B}$ n'implique pas $\lambda+\beta$ valeur propre de $\mathbf{A}+\mathbf{B}$.

2. λ valeur propre de $\mathbf{A}$ et β valeur propre de $\mathbf{B}$ n'implique pas $\lambda\beta$ valeur propre de $\mathbf{AB}$.

Exercice 5.— Soit $\mathbf{A}$ une matrice de $\mathscr{M}_n(\mathbb{K})$ et soient λ une valeur propre de $\mathbf{A}$ et $\mathbf{x}$ un vecteur propre associé à la valeur propre λ.

1. Montrer que λ^k est une valeur propre de $\mathbf{A}^k$ et $\mathbf{x}$ est aussi vecteur propre pour la valeur propre λ^k.

2. Soit $q(x) = a_0 + a_1x + \ldots a_kx^k$ un polynôme de $\mathbb{K}[x]$. Montrer que $q(\lambda)$ est une valeur propre de la matrice suivante

$$q(\mathbf{A}) = a_0\mathbf{1}_n + a_1\mathbf{A} + a_2\mathbf{A}^2 + \ldots + a_k\mathbf{A}^k$$

et que $\mathbf{x}$ est un vecteur propre associé.

Exercice 6.— Montrer que la trace d'une matrice nilpotente de $\mathscr{M}_n(\mathbb{K})$ est nulle.

Exercice 7.— Soient $\mathbf{x}_1, \ldots, \mathbf{x}_k$ des vecteurs propres d'une matrice $\mathbf{A}$ de $\mathscr{M}_n(\mathbb{K})$, associées à une même valeur propre λ. Montrer que toute combinaison linéaire non nulle des vecteurs $\mathbf{x}_1, \ldots, \mathbf{x}_k$ est un vecteur propre de $\mathbf{A}$ associé à la valeur propre λ.

Exercice 8.— Soit $\mathbf{A}$ une matrice de $\mathscr{M}_n(\mathbb{K})$ de spectre $\operatorname{spec}(\mathbf{A}) = \{\lambda_1, \ldots, \lambda_p\}$. On considère une valeur propre λ_k de $\mathbf{A}$ et un vecteur propre $\mathbf{x}$ associé.

1. Soit λ un scalaire qui n'est pas valeur propre de $\mathbf{A}$. Montrer que

$$(\mathbf{A} - \lambda\mathbf{1}_n)^{-1}\mathbf{x} = \frac{1}{\lambda_k - \lambda}\mathbf{x}.$$

2. Soit $\mathbf{y}$ un vecteur de $\mathbb{K}^n$, montrer que les valeurs propres de $\mathbf{A} + \mathbf{x}\mathbf{y}^\top$ coïncident avec celles de $\mathbf{A}$ sauf pour la valeur propre λ_k qui est remplacée par $\lambda_k + \mathbf{y}^\top\mathbf{x}$.

3. Soit μ un scalaire. Comment construire un vecteur $\mathbf{y}$ afin que les valeurs propres de $\mathbf{A} + \mathbf{x}\mathbf{y}^\top$ et $\mathbf{A}$ coïncident sauf pour la valeur propre λ_k qui est remplacée par μ.

Exercice 9.— Soient $\mathbf{A} \in \mathscr{M}_{m,n}(\mathbb{K})$ et $\mathbf{B} \in \mathscr{M}_{n,m}(\mathbb{K})$ deux matrices.

1. Montrer que pour tout scalaire λ,

$$\lambda^m \det(\lambda\mathbf{1}_n - \mathbf{BA}) = \lambda^n \det(\lambda\mathbf{1}_m - \mathbf{AB}).$$

2. Montrer que si $m = n$, alors les matrices $\mathbf{AB}$ et $\mathbf{BA}$ ont le même polynôme caractéristique.

3. On suppose que $m \neq n$. Montrer que les polynômes caractéristiques de $\mathbf{AB}$ et $\mathbf{BA}$ ne peuvent pas être égaux. Que peut-on dire des spectres de $\mathbf{AB}$ et $\mathbf{BA}$?

Exercice 10.— Soient $\mathbf{A} \in \mathscr{M}_n(\mathbb{K})$ et $\mathbf{B} \in \mathscr{M}_n(\mathbb{K})$ deux matrices. Montrer que si $\mathbf{AB} = \mathbf{BA}$, alors $\mathbf{A}$ et $\mathbf{B}$ ont un vecteur propre commun.

Exercice 11.— Soient $\mathbf{P} \in \mathscr{M}_m(\mathbb{K})$ et $\mathbf{Q} \in \mathscr{M}_n(\mathbb{K})$ deux matrices fixées. Soit

$$u : \mathscr{M}_{m,n}(\mathbb{K}) \longrightarrow \mathscr{M}_{m,n}(\mathbb{K})$$

une application définie par

$$u(\mathbf{A}) = \mathbf{PAQ}.$$

1. Montrer que u est linéaire.

2. Montrer que $\operatorname{trace}(u) = \operatorname{trace}(\mathbf{P})\operatorname{trace}(\mathbf{Q})$.

Exercice $^{(\textsf{Sage})}$ **12.—** Sage propose plusieurs méthodes sur les matrices permettant de calculer par exemple le polynôme caractéristique d'une matrice `charpoly()`, son polynôme minimal `minpoly()`, ses valeurs propres `eigenvalues()`, une liste de vecteurs propres `eigenvalues_right()`. On pourra regarder l'aide en ligne de ces méthodes qui fournit quantité d'exemples d'utilisation.

Calculer les valeurs propres de la matrice suivante :

$$\begin{bmatrix} a & b & b & b \\ b & a & b & b \\ b & b & a & b \\ b & b & b & a \end{bmatrix}$$

Exercice $^{(\textsf{Sage})}$ **13.—** On considère les trois matrices réelles suivantes

$$\mathbf{A} = \begin{bmatrix} a & 1 & 1 & 1 \\ 1 & b & 1 & 1 \\ 1 & 1 & c & 1 \\ d & 1 & 1 & d \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} -1 & a & d \\ b & e & -1 \\ c & 1 & f \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} a & 2 & 1 \\ 2 & b & -1 \\ -1 & 1 & c \end{bmatrix}.$$

1. Déterminer les quadruplets $(a,b,c,d) \in \mathbb{R}^4$, pour lesquels la matrice **A** admet $\begin{bmatrix} -1 \\ 2 \\ 2 \\ -1 \end{bmatrix}$ pour vecteur propre. On pourra utiliser la fonction `solve`.

2. Déterminer les réels a,b,c,d,e,f, pour lesquels la matrice **B** admet $\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$, $\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$, $\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$ pour vecteurs propres.

3. Déterminer les triplets $(a,b,c) \in \mathbb{R}^3$, pour lesquels la matrice **C** admet 1, a et b comme valeurs propres.

Exercice $^{(\textsf{Sage})}$ **14.—** Déterminer l'expression générale du polynôme minimal d'une matrice antisymétrique de taille 3×3 puis d'une matrice antisymétrique de taille 4×4.

CHAPITRE
6

Trigonalisation et diagonalisation

Sommaire

Nous abordons dans ce chapitre les problèmes de trigonalisation et diagonalisation des matrices. Nous montrons que toute matrice à coefficients complexes est trigonalisable, c'est-à-dire semblable à une matrice triangulaire supérieure. On présente quelques conséquences théoriques importantes de ce résultat.

Le problème de la diagonalisation est plus épineux. Une matrice n'est pas en général diagonalisable, c'est-à-dire semblable à une matrice diagonale. Dans ce chapitre, on s'intéressera aux obstructions au caractère diagonalisable. En particulier, nous donnerons une caractérisation de nature géométrique des matrices diagonalisables.

Nous présentons deux applications immédiates de la diagonalisation des matrices avec le calcul des puissances d'une matrice diagonalisable et la résolution des systèmes différentiels linéaires définis par une matrice diagonalisable. Nous reviendrons sur ces deux applications dans les prochains chapitres, notamment dans le cas où ils mettent en jeu des matrices non diagonalisables.

§ 1 Trigonalisation des matrices

6.1.1. Définition.— Une matrice $\mathbf{A}$ de $\mathscr{M}_n(\mathbb{K})$ est dite *trigonalisable* dans $\mathscr{M}_n(\mathbb{K})$, si $\mathbf{A}$ est semblable à une matrice triangulaire supérieure de $\mathscr{M}_n(\mathbb{K})$. C'est-à-dire, s'il existe une matrice inversible $\mathbf{P}$ de $\mathscr{M}_n(\mathbb{K})$ et une matrice triangulaire supérieure $\mathbf{T}$ à coefficients dans $\mathbb{K}$ telles que

$$\mathbf{A} = \mathbf{PTP}^{-1}. \tag{6.1}$$

On notera que toute matrice triangulaire supérieure étant semblable à une matrice triangulaire inférieure, une matrice est trigonalisable dans $\mathscr{M}_n(\mathbb{K})$ si, et seulement si, elle est semblable à une matrice triangulaire inférieure.

Exercice 1.— Soit $\mathbf{A}$ une matrice de $\mathscr{M}_n(\mathbb{K})$ et soit λ une valeur propre de $\mathbf{A}$. Montrer que la matrice $\mathbf{A}$ est semblable à une matrice de la forme

$$\begin{bmatrix} \lambda & * & \cdots & * \\ 0 & & & \\ \vdots & & \mathbf{B} & \\ 0 & & & \end{bmatrix}$$

6.1.2. Caractérisation des matrices trigonalisables.— Le résultat suivant fournit une caractérisation des matrices trigonalisables.

6.1 Théorème (Théorème de trigonalisation).— Une matrice $\mathbf{A}$ de $\mathscr{M}_n(\mathbb{K})$ est trigonalisable dans $\mathscr{M}_n(\mathbb{K})$ si, et seulement si, son polynôme caractéristique $p_{\mathbf{A}}$ est scindé sur $\mathbb{K}$.

Ce théorème est une conséquence du théorème de trigonalisation de Schur que nous démontrerons dans la deuxième partie de cours : *Analyse matricielle et algèbre linéaire appliquée II*. Dans l'attente, voici une preuve directe du théorème 6.1.

Preuve. La condition est nécessaire. Si $\mathbf{A}$ est une matrice trigonalisable, par définition, elle est semblable à une matrice triangulaire supérieure :

$$\mathbf{T} = \begin{bmatrix} \lambda_1 & * & \cdots & * \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix}$$

Le polynôme caractéristique de la matrice $\mathbf{T}$ est scindé :

$$p_{\mathbf{T}} = (-1)^n (x - \lambda_1) \dots (x - \lambda_n).$$

D'après la proposition 5.3, deux matrices semblables ont même polynôme caractéristique. Ainsi, $p_{\mathbf{A}} = p_{\mathbf{T}}$ et par suite le polynôme caractéristique de $\mathbf{A}$ est scindé sur $\mathbb{K}$.

La condition est suffisante. On procède par récurrence sur n. Toute matrice de $\mathscr{M}_1(\mathbb{K})$ est trigonalisable. On suppose que tout matrice de $\mathscr{M}_{n-1}(\mathbb{K})$, dont le polynôme caractéristique est scindé, est trigonalisable, montrons que cela est vrai pour toute matrice de $\mathscr{M}_n(\mathbb{K})$.

Soit $\mathbf{A} \in \mathscr{M}_n(\mathbb{K})$, telle que le polynôme $p_\mathbf{A}$ soit scindé sur $\mathbb{K}$. Le polynôme $p_\mathbf{A}$ admet donc au moins une racine λ dans $\mathbb{K}$. Considérons un vecteur propre $\mathbf{e}$ dans $\mathbb{K}^n$ associé à la valeur propre λ. Complétons le vecteur $\mathbf{e}$ en une base $\mathscr{B} = (\mathbf{e}, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ de $\mathbb{K}^n$. Soit $u_\mathbf{A}$ l'endomorphisme de $\mathbb{K}^n$ associé à la matrice $\mathbf{A}$, i.e., l'endomorphisme défini, pour tout vecteur $\mathbf{x}$ de $\mathbb{K}^n$, par $u_\mathbf{A}(\mathbf{x}) = \mathbf{A}\mathbf{x}$. On a

$$u_\mathbf{A}(\mathbf{e}) = \mathbf{A}\mathbf{e} = \lambda\mathbf{e},$$

par suite, la matrice de l'endomorphisme $u_\mathbf{A}$ exprimé dans la base $\mathscr{B}$ est

$$[u_\mathbf{A}]_\mathscr{B} = \begin{bmatrix} \lambda & * & \cdots & * \\ 0 & & & \\ \vdots & & \mathbf{B} & \\ 0 & & & \end{bmatrix},$$

où $\mathbf{B}$ est une matrice de $\mathscr{M}_{n-1}(\mathbb{K})$. La matrice $\mathbf{A}$ étant semblable à la matrice $[u_\mathbf{A}]_\mathscr{B}$, il existe une matrice inversible $\mathbf{P}$ de $\mathscr{M}_n(\mathbb{C})$, telle que

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \begin{bmatrix} \lambda & * & \cdots & * \\ 0 & & & \\ \vdots & & \mathbf{B} & \\ 0 & & & \end{bmatrix}.$$

De plus, d'après 5.5, le polynôme caractéristique du bloc $\mathbf{B}$ divise le polynôme caractéristique de la matrice $\mathbf{A}$, il est donc scindé comme ce dernier. Par hypothèse de récurrence, la matrice $\mathbf{B}$ est semblable à une matrice triangulaire supérieure, il existe une matrice inversible $\mathbf{Q}$ dans $\mathscr{M}_{n-1}(\mathbb{K})$, telle que $\mathbf{T}' = \mathbf{Q}^{-1}\mathbf{B}\mathbf{Q}$ soit triangulaire supérieure. En multipliant par blocs, on a :

$$\begin{aligned}
\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & \mathbf{Q} & \\ 0 & & & \end{bmatrix}^{-1} \mathbf{P}^{-1}\mathbf{A}\mathbf{P} \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & \mathbf{Q} & \\ 0 & & & \end{bmatrix} &= \begin{bmatrix} \lambda & * & \cdots & * \\ 0 & & & \\ \vdots & & \mathbf{Q}^{-1}\mathbf{B}\mathbf{Q} & \\ 0 & & & \end{bmatrix} \\
&= \begin{bmatrix} \lambda & * & \cdots & * \\ 0 & & & \\ \vdots & & \mathbf{T}' & \\ 0 & & & \end{bmatrix}.
\end{aligned}$$

En posant

$$\mathbf{R} = \mathbf{P}\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & \mathbf{Q} & \\ 0 & & & \end{bmatrix},$$

la dernière égalité s'écrit

$$\mathbf{R}^{-1}\mathbf{A}\mathbf{R} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & \mathbf{Q} & \\ 0 & & & \end{bmatrix}.$$

Ainsi, $\mathbf{A}$ est semblable à une triangulaire supérieure. □

6.1.3. Trigonalisation sur $\mathbb{C}$.— Voici une première conséquence importante du théorème de trigonalisation. D'après le théorème de D'Alembert-Gauss, théorème 0.5, tout polynôme non nul de $\mathbb{C}[x]$ est scindé sur $\mathbb{C}$. Par suite, on a

6.2 Proposition.— Toute matrice $\mathbf{A}$ de $\mathscr{M}_n(\mathbb{C})$ est trigonalisable dans $\mathscr{M}_n(\mathbb{C})$.

Notons que toute matrice $\mathbf{A}$ de $\mathscr{M}_n(\mathbb{R})$ peut toujours se trigonaliser dans $\mathscr{M}_n(\mathbb{C})$. En effet, si le polynôme caratéristique de $\mathbf{A}$ est scindé sur $\mathbb{R}$, $\mathbf{A}$ est trigonalisable dans $\mathscr{M}_n(\mathbb{R})$. Sinon, le polynôme $p_{\mathbf{A}}$ est toujours scindé dans $\mathscr{M}_n(\mathbb{C})$. Il existe alors une matrice inversible $\mathbf{P}$ et une matrice triangulaire $\mathbf{T}$ de $\mathscr{M}_n(\mathbb{C})$ telles que $\mathbf{A} = \mathbf{P}\mathbf{T}\mathbf{P}^{-1}$.

6.1.4. Exemple.— La matrice suivante de $\mathscr{M}_4(\mathbb{R})$

$$\mathbf{A} = \begin{bmatrix} 0 & -1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

admet pour polynôme caractéristique

$$p_{\mathbf{A}} = (x^2+1)^2.$$

Ce polynôme n'est pas scindé dans $\mathbb{R}[x]$, la matrice $\mathbf{A}$ n'est donc pas trigonalisable dans $\mathscr{M}_4(\mathbb{R})$. Cependant, il est scindé dans $\mathbb{C}[x]$:

$$p_{\mathbf{A}} = (x-i)^2(x+i)^2.$$

La matrice est trigonalisable. Posons

$$\mathbf{P} = \begin{bmatrix} 1 & -1 & 1 & 0 \\ -i & 0 & i & i \\ 0 & 1 & 0 & 1 \\ 0 & -i & 0 & i \end{bmatrix}.$$

Le premier et troisième vecteur colonne de la matrice $\mathbf{P}$ sont des vecteurs propres associés aux valeurs propres i et $-i$ respectivement. Les deux autres vecteurs colonnes complètent ces vecteurs en une base de trigonalisation. On a

$$\mathbf{A} = \mathbf{P} \begin{bmatrix} i & 1 & 0 & 0 \\ 0 & i & 0 & 0 \\ 0 & 0 & -i & 1 \\ 0 & 0 & 0 & -i \end{bmatrix} \mathbf{P}^{-1},$$

avec

$$\mathbf{P}^{-1} = \frac{1}{2} \begin{bmatrix} 1 & i & 1 & 0 \\ 0 & 0 & 1 & i \\ 1 & -i & 0 & i \\ 0 & 0 & 1 & -i \end{bmatrix}.$$

6.1.5. Somme et produit des valeurs propres.— Le théorème de trigonalisation nous permet de relier des invariants d'une matrice, tels que sa trace et son déterminant, à ses valeurs propres.

Si une matrice $\mathbf{A}$ est trigonalisable, semblable à une matrice triangulaire supérieure $\mathbf{T}$, alors les valeurs propres de $\mathbf{A}$ étant les racines du polynôme $p_{\mathbf{A}}$, sont aussi les coefficients de la diagonale de la matrice $\mathbf{T}$.

Étant donnée une matrice $\mathbf{A}$ de $\mathscr{M}_n(\mathbb{C})$, alors son polynôme caractéristique est scindé sur $\mathbb{C}$:

$$p_{\mathbf{A}} = (-1)^n (x-\lambda_1)\ldots(x-\lambda_n).$$

La matrice $\mathbf{A}$ est semblable à une matrice triangulaire $\mathbf{T}$, i.e., il existe une matrice inversible $\mathbf{P}$ telle que

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \begin{bmatrix} \lambda_1 & * & \cdots & * \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix}$$

Étant semblables, les matrices $\mathbf{A}$ et $\mathbf{T}$ ont même trace et même déterminant, on en déduit que la trace (resp. le déterminant) de $\mathbf{A}$ est égale à la somme (resp. le produit) des valeurs propres, comptées avec leur ordre de multiplicité. Précisément, on a

6.3 Proposition.— Soit $\mathbf{A}$ une matrice de $\mathscr{M}_n(\mathbb{C})$ de polynôme caractéristique

$$p_{\mathbf{A}} = (-1)^n (x-\lambda_1)^{n_1}\ldots(x-\lambda_p)^{n_p},$$

où n_i désigne l'ordre de multiplicité de la valeur propre λ_i dans le polynôme caractéristique. Alors,

i) $\operatorname{trace}(\mathbf{A}) = n_1\lambda_1 + \ldots + n_p\lambda_p$,

ii) $\det(\mathbf{A}) = \lambda_1^{n_1}\ldots\lambda_p^{n_p}$.

Plus généralement, pour tout entier $k \geqslant 1$, on a

iii) $\operatorname{trace}(\mathbf{A}^k) = n_1\lambda_1^k + \ldots + n_p\lambda_p^k$,

iv) $\det(\mathbf{A}^k) = \lambda_1^{k.n_1}\ldots\lambda_p^{k.n_p}$.

6.1.6. Exemples.— Dans l'exemple 5.3.3, on a montré que la matrice

$$\mathbf{A} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix},$$

possède deux valeurs propres $-i$ et i ; la somme de ces valeurs propres est égale à la trace de $\mathbf{A}$ et leur produit est le déterminant de $\mathbf{A}$.

Dans l'exemple 5.3.4, on a montré que le spectre de la matrice de la rotation du plan vectoriel

$$\mathbf{R}_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$$

est

$$\mathrm{Sp}_{\mathbb{C}}(\mathbf{R}_\theta) = \{e^{i\theta}, e^{-i\theta}\}.$$

La proposition précédente, nous permet de retrouver les relations trigonométriques bien connues :

$$\operatorname{trace}\mathbf{R}_\theta = 2\cos\theta = e^{i\theta} + e^{-i\theta},$$
$$\det\mathbf{R}_\theta = 1 = e^{i\theta}e^{-i\theta}.$$

Exercice 2.— Montrer qu'une matrice de $\mathscr{M}_n(\mathbb{R})$ est inversible si, et seulement si, elle n'admet pas de valeur propre nulle.

6.1.7. Exemple.— Dans l'exemple 6.3.3, nous avons montré que la matrice

$$\mathbf{A} = \begin{bmatrix} 0 & \cdots & 0 & 1 \\ \vdots & & \vdots & \vdots \\ 0 & \cdots & 0 & 1 \\ 1 & \cdots & 1 & 1 \end{bmatrix}$$

admet pour valeur propre 0, d'ordre de multiplicité géométrique $n-2$, par suite le polynôme caractéristique s'écrit sous la forme

$$p_{\mathbf{A}} = (-1)^n x^{n-2}(x^2 + \alpha x + \beta).$$

Déterminons les autres valeurs propres de $\mathbf{A}$. Supposons que

$$p_{\mathbf{A}} = (-1)^n x^{n-2}(x-\lambda_1)(x-\lambda_2).$$

D'après la proposition 6.3, λ_1 et λ_2 satisfont les relations

$$\left|\begin{array}{l} \operatorname{trace}\mathbf{A} = \lambda_1 + \lambda_2 \\ \operatorname{trace}\mathbf{A}^2 = \lambda_1^2 + \lambda_2^2 \end{array}\right.$$

avec

$$\mathbf{A}^2 = \begin{bmatrix} 1 & \cdots & 1 & 1 \\ \vdots & & \vdots & \vdots \\ 1 & \cdots & 1 & 1 \\ 1 & \cdots & 1 & n \end{bmatrix}.$$

Ainsi, $\operatorname{trace}\mathbf{A} = 1$ et $\operatorname{trace}\mathbf{A}^2 = 2n-1$, par suite, λ_1 et λ_2 satisfont les deux relations

$$\left|\begin{array}{l} \lambda_1 + \lambda_2 = 1 \\ \lambda_1^2 + \lambda_2^2 = 2n-1 \end{array}\right.$$

Comme $(\lambda_1+\lambda_2)^2 = \lambda_1^2 + \lambda_2^2 + 2\lambda_1\lambda_2$, le système précédent se réduit à

$$\left|\begin{array}{l} \lambda_1 + \lambda_2 = 1 \\ \lambda_1\lambda_2 = 1-n \end{array}\right.$$

Donc λ_1 et λ_2 sont solutions de l'équation

$$\lambda^2 - \lambda + (1-n) = 0.$$

D'où

$$\lambda_1 = \frac{1+\sqrt{4n-3}}{2}, \qquad \lambda_2 = \frac{1-\sqrt{4n-3}}{2}.$$

Le spectre de $\mathbf{A}$ est donc

$$\operatorname{Sp}(\mathbf{A}) = \left\{0, \frac{1-\sqrt{4n-3}}{2}, \frac{1+\sqrt{4n-3}}{2}\right\}.$$

Les sous-espaces propres sont définis par

$$E_0 = \operatorname{Vect}(\begin{bmatrix} 1 \\ -1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \\ \vdots \\ 0 \end{bmatrix}, \quad \ldots \quad, \begin{bmatrix} 0 \\ \vdots \\ 1 \\ -1 \\ 0 \end{bmatrix}),$$

$$E_{\lambda_1} = \mathrm{Vect}(\begin{bmatrix} 1 \\ \vdots \\ 1 \\ \lambda_1 \end{bmatrix}), \qquad E_{\lambda_2} = \mathrm{Vect}(\begin{bmatrix} 1 \\ \vdots \\ 1 \\ \lambda_2 \end{bmatrix}).$$

Pour ces deux derniers, on calcule en effet

$$\mathbf{A}\begin{bmatrix} 1 \\ \vdots \\ 1 \\ \lambda_i \end{bmatrix} = \begin{bmatrix} \lambda_i \\ \vdots \\ \lambda_i \\ \lambda_i + n - 1 \end{bmatrix},$$

avec $\lambda_i^2 = \lambda_i + n - 1$, pour $i = 1, 2$.

Exercice 3.— Déterminer les valeurs propres et sous-espaces propres des matrices suivantes de $\mathscr{M}_n(\mathbb{R})$:

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 1 & 0 & \cdots & 0 \end{bmatrix}, \qquad \mathbf{B} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 1 & 0 & \cdots & 0 \\ 1 & 1 & \cdots & 1 \end{bmatrix}, \qquad \mathbf{C} = \begin{bmatrix} 1 & \cdots & 1 & 1 \\ 0 & \cdots & 0 & 1 \\ \vdots & & \vdots & \vdots \\ 0 & \cdots & 0 & 1 \end{bmatrix}$$

$$\mathbf{D} = \begin{bmatrix} 1 & 1 & \cdots & 1 & 1 \\ 1 & 0 & \cdots & 0 & 1 \\ \vdots & \vdots & 0 & \vdots & \vdots \\ 1 & 0 & \cdots & 0 & 1 \\ 1 & 1 & \cdots & 1 & 1 \end{bmatrix},$$

6.1.8. Exemple.— Soit $\mathbf{A}$ la matrice de $\mathscr{M}_n(\mathbb{R})$ définie par

$$\mathbf{A} = \begin{bmatrix} n & 1 & \cdots & 1 \\ 1 & n & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 1 & \cdots & 1 & n \end{bmatrix}.$$

On remarque que

$$\mathbf{A} - (n-1)\mathbf{1}_n = \begin{bmatrix} 1 & \cdots & 1 \\ \vdots & & \vdots \\ 1 & \cdots & 1 \end{bmatrix}.$$

On a donc rang $(\mathbf{A} - (n-1)\mathbf{1}_n) = 1$. D'après le la formule du rang, théorème 1.9, on a $\dim E_{n-1} = n-1$. Donc $n-1$ est valeur propre de $\mathbf{A}$, avec $\mathrm{mult}_{\mathrm{alg}}(n-1) \geqslant n-1$. Pour déterminer l'autre éventuelle valeur propre λ, on calcule

$$\mathrm{trace}\,\mathbf{A} = n^2 = \lambda + (n-1)(n-1).$$

Par suite $\lambda = 2n-1$. On a donc $\mathrm{mult}_{\mathrm{alg}}(2n-1) \geqslant 1$. On en déduit donc que $\mathrm{mult}_{\mathrm{alg}}(n-1) = \mathrm{mult}_{\mathrm{geo}}(n-1) = n-1$ et que $\mathrm{mult}_{\mathrm{alg}}(2n-1) = \mathrm{mult}_{\mathrm{geo}}(2n-1) = 1$.

Dans cet exemple, on a
$$\mathbb{K}^n = E_{n-1} \oplus E_{2n-1},$$
on dit dans ce cas que la matrice $\mathbf{A}$ est diagonalisable.

§ 2 Diagonalisation des matrices

6.2.1. Matrices diagonalisables.— Une matrice $\mathbf{A}$ de $\mathscr{M}_n(\mathbb{K})$ est dite *diagonalisable* dans $\mathscr{M}_n(\mathbb{K})$, si elle est semblable à une matrice diagonale de $\mathscr{M}_n(\mathbb{K})$. C'est-à-dire, s'il existe une matrice inversible $\mathbf{P}$ de $\mathscr{M}_n(\mathbb{K})$ et une matrice diagonale $\mathbf{D}$ à coefficients dans $\mathbb{K}$ telles que
$$\mathbf{A} = \mathbf{PDP}^{-1}. \tag{6.2}$$

Exercice 4.— Montrer que la matrice suivante
$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$
est diagonalisable dans $\mathscr{M}_2(\mathbb{R})$.

Exercice 5.— Soit $\mathbf{A}$ la matrice définie par blocs :
$$\mathbf{A} = \begin{bmatrix} \mathbf{B} & \mathbf{0} \\ \mathbf{0} & \mathbf{C} \end{bmatrix},$$
où $\mathbf{B}$ et $\mathbf{C}$ sont deux matrices carrés de $\mathscr{M}_{n_1}(\mathbb{K})$ et $\mathscr{M}_{n_2}(\mathbb{K})$ respectivement. Montrer que si $\mathbf{B}$ et $\mathbf{C}$ sont diagonalisables, alors $\mathbf{A}$ est diagonalisable.

6.2.2. Matrices diagonalisables et vecteurs propres.— Les matrices $\mathbf{A}$ et $\mathbf{D}$ de la décomposition (6.2) étant semblables, d'après la proposition 5.3, elles ont le même polynôme caractéristique. Il s'ensuit que la diagonale de la matrice $\mathbf{D}$ est formée des valeurs propres de $\mathbf{A}$.

6.4 Proposition.— Une matrice $\mathbf{A}$ de $\mathscr{M}_n(\mathbb{K})$ est diagonalisable si, et seulement si, il existe une base de $\mathbb{K}^n$ formée de vecteurs propres de $\mathbf{A}$.

Preuve. Supposons qu'il existe une base $(\mathbf{x}_1, \ldots, \mathbf{x}_n)$ de E composée de vecteurs propres de $\mathbf{A}$. Considérons la matrice $\mathbf{P}$ dont les colonnes sont formées par les éléments de cette base :
$$\mathbf{P} = \begin{bmatrix} \vdots & \vdots & & \vdots \\ \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_n \\ \vdots & \vdots & & \vdots \end{bmatrix}.$$

Comme les vecteurs $\mathbf{x}_1, \ldots, x_n$ forment une base de $\mathbb{K}^n$, la matrice $\mathbf{P}$ est inversible et on a

$$\begin{aligned}
\mathbf{AP} &= \begin{bmatrix} \vdots & \vdots & & \vdots \\ \mathbf{Ax}_1 & \mathbf{Ax}_2 & \cdots & \mathbf{Ax}_n \\ \vdots & \vdots & & \vdots \end{bmatrix} \\
&= \begin{bmatrix} \vdots & \vdots & & \vdots \\ \lambda_1\mathbf{x}_1 & \lambda_2\mathbf{x}_2 & \cdots & \lambda_n\mathbf{x}_n \\ \vdots & \vdots & & \vdots \end{bmatrix} \\
&= \begin{bmatrix} \vdots & \vdots & & \vdots \\ \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_n \\ \vdots & \vdots & & \vdots \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix}
\end{aligned}$$

Ainsi

$$\mathbf{AP} = \mathbf{P} \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix}. \tag{6.3}$$

Par suite, la matrice $\mathbf{A}$ est diagonalisable.

Inversement, si $\mathbf{A}$ est diagonalisable, il existe une matrice inversible $\mathbf{P}$ telle que $\mathbf{P}$ satisfait (6.3). Pour les mêmes raisons, les vecteurs colonnes de $\mathbf{P}$ forment une base de vecteurs propres de $\mathbb{K}^n$. $\square$

6.2.3. Une condition suffisante de diagonalisation.— Toutes les matrices ne sont pas diagonalisables, nous l'avons vu en particulier avec l'exemple 5.3.4 des rotations. La matrice $\mathbf{R}_\theta$ qui représente la rotation d'angle θ du plan vectoriel $\mathbb{R}^2$ n'est pas diagonalisable sur $\mathbb{R}$ si $\theta \neq 0$ modulo π, car alors $\mathbf{R}_\theta$ ne possède pas de valeur propre réelle.

Cet exemple illustre le caractére non diagonalisable d'une matrice dans $\mathscr{M}_2(\mathbb{R})$, car elle n'admet pas de valeur propre réelle. La matrice suivante

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

admet 0 comme valeur propre et n'est pas diagonalisable dans $\mathscr{M}_2(\mathbb{R})$. En effet, si elle était diagonalisable, alors son unique valeur propre étant 0, car son polynôme caractéristique est $p_{\mathbf{A}} = x^2$, la matrice $\mathbf{A}$ serait semblable à une matrice nulle. Or toute matrice semblable à une matrice nulle est nulle. Ceci n'étant pas le cas de $\mathbf{A}$ la matrice $\mathbf{A}$ n'est pas diagonalisable.

Exercice 6.— Les matrices suivantes sont-elles diagonalisables dans $\mathscr{M}_2(\mathbb{R})$

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}?$$

6.5 Proposition (Condition suffisante de diagonalisation).— Soit $\mathbf{A}$ une matrice de $\mathscr{M}_n(\mathbb{K})$. Si le polynôme caractéristique de $\mathbf{A}$ est scindé sur $\mathbb{K}$ et possède toutes ses racines simples, alors $\mathbf{A}$ est diagonalisable dans $\mathscr{M}_n(\mathbb{K})$.

Preuve. Supposons que le polynôme caractéristique $p_{\mathbf{A}}$ soit scindé à racines simples :

$$p_{\mathbf{A}} = (-1)^n(x-\lambda_1)\dots(x-\lambda_n),$$

avec $\lambda_i \neq \lambda_j$, si $i \neq j$. Pour tout $i \in [\![1,n]\!]$, λ_i étant valeur propre de $\mathbf{A}$, il existe un vecteur propre $\mathbf{x}_i$ de $\mathbf{A}$ associé. D'aprés la proposition 5.2, les sous-espaces propres de $\mathbf{A}$ formant une somme directe, la famille $(\mathbf{x}_1,\dots,\mathbf{x}_n)$ est libre. Elle possède n éléments, c'est donc une base de $\mathbb{K}^n$. $\square$

On en déduit le résultat suivant :

6.6 Corollaire.— Toute matrice de $\mathscr{M}_n(\mathbb{K})$ qui admet n valeurs propres distinctes est diagonalisable.

6.2.4. Remarques.— Attention, la réciproque du corollaire 6.6 est fausse en général. Par exemple, la matrice

$$\mathbf{A} = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}$$

est diagonalisable, alors qu'elle n'admet que deux valeurs propres distinctes, son spectre est $\mathrm{Sp}(\mathbf{A}) = \{1,4\}$.

De la même façon, la proposition 6.5 est une condition suffisante de diagonalisation, mais elle n'est pas nécessaire. Par exemple, la matrice identité $\mathbf{1}_n$ est diagonalisable, son polynôme caractéristique est $p_{\mathbf{1}_n} = (-1)^n(x-1)^n$; son unique valeur propre 1 est d'ordre de multiplicité n.

6.2.5. Exemple.— On considère la matrice réelle

$$\mathbf{A} = \begin{bmatrix} a & bc^2 \\ b & a \end{bmatrix},$$

avec b,c non nul. Le polynôme caractéristique de $\mathbf{A}$ est

$$p_{\mathbf{A}} = (a-x)^2 - b^2c^2 = ((a-bc)-x)((a+bc)-x).$$

Par suite $\lambda_1 = a - bc$ et $\lambda_2 = a + bc$ sont deux valeurs propres, distinctes par hypothèses sur b et c. On en déduit que $\mathbf{A}$ est diagonalisable.

Déterminons les sous-espaces propres E_{λ_1} et E_{λ_2}. On a $\begin{bmatrix} x \\ y \end{bmatrix} \in E_{\lambda_1}$ si, et seulement si,

$$\mathbf{A}\begin{bmatrix} x \\ y \end{bmatrix} = \lambda_1 \begin{bmatrix} x \\ y \end{bmatrix}.$$

Autrement dit, x et y satisfont

$$\left| \begin{array}{l} ax + bc^2 y = ax - bcx \\ bx + ay = ay - bcy \end{array} \right. ,$$

Donc $x = -cy$. On déduit que

$$E_{\lambda_1} = \operatorname{Vect}\begin{bmatrix} -c \\ 1 \end{bmatrix}.$$

De la même façon, on montre que

$$E_{\lambda_2} = \operatorname{Vect}\begin{bmatrix} c \\ 1 \end{bmatrix}.$$

On a donc

$$\begin{bmatrix} a-bc & 0 \\ 0 & a+bc \end{bmatrix} = \begin{bmatrix} -c & c \\ 1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} a & bc^2 \\ b & a \end{bmatrix} \begin{bmatrix} -c & c \\ 1 & 1 \end{bmatrix}.$$

6.2.6. Exemple.— Une autre illustration du caractère non nécessaire de la condition de la proposition 6.5 est donnée par la matrice suivante :

$$\mathbf{G} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

vue dans l'exemple 4.4.3. Nous avons montré que $\mathbf{G}$ est diagonalisable, semblable à la matrice

$$\begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{bmatrix},$$

alors que son polynôme caractéristique $p_{\mathbf{G}} = (x+1)^2(2-x)$ n'est pas à racines simples.

§ 3 Une obstruction au caractère diagonalisable

Dans cette partie, nous présentons une caractérisation de nature géométrique des matrices diagonalisables. Nous verrons dans le chapitre suivant une caractérisation algébrique de la diagonalisation avec le polynôme minimal.

D'après la proposition 5.2, les sous-espaces propres d'une matrice $\mathbf{A}$ sont en somme directe. Il est possible que cette somme ne « remplisse » pas l'espace $\mathbb{K}^n$ tout entier, i.e., que la somme directe $E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_p}$ soit un sous-espace vectoriel strict de $\mathbb{K}^n$. C'est en particulier le cas lorsque l'on a $\dim E_{\lambda_1} + \cdots + \dim E_{\lambda_p} < n$. L'objectif de cette section est de montrer que ceci constitue une obstruction à la diagonalisation de $\mathbf{A}$.

6.3.1. Exemple.— Considérons la matrice de $\mathscr{M}_3(\mathbb{R})$ suivante

$$\mathbf{A} = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}.$$

On calcule son polynôme caractéristique $p_{\mathbf{A}} = -(x-2)^2(x-3)$ et, par ailleurs, on a

$$\mathbf{A} - 3\mathbf{1}_3 = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \qquad \mathbf{A} - 2\mathbf{1}_3 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Il est immédiat que

$$\operatorname{rang}(\mathbf{A} - 3\mathbf{1}_3) = \operatorname{rang}(\mathbf{A} - 2\mathbf{1}_3) = 2.$$

D'après la formule du rang, théorème 1.9, on en déduit que $\dim E_3 = 1$ et $\dim E_2 = 1$. Ainsi, la matrice $\mathbf{A}$ n'est pas diagonalisable, car la somme directe de tous les sous-espaces propres $E_2 \oplus E_3$ est de dimension 2. Il ne peut donc pas exister de base de $\mathbb{R}^3$ formée de vecteurs propres de $\mathbf{A}$. (Nous verrons dans un prochain chapitre, que la matrice $\mathbf{A}$ est présentée ici sous une forme canonique, dite de Jordan, et que son caractère non diagonalisable est évident.)

6.3.2. Multiplicité des valeurs propres.— Soient $\mathbf{A}$ une matrice de $\mathscr{M}_n(\mathbb{K})$ et λ une valeur propre de $\mathbf{A}$. L'ordre de multiplicité de λ en tant que racine du polynôme $p_{\mathbf{A}}$ est appelé *multiplicité algébrique* de λ, on la note $\operatorname{mult}^{\mathbf{A}}_{\mathrm{alg}}(\lambda)$, ou $\operatorname{mult}_{\mathrm{alg}}(\lambda)$ s'il n'y a pas de confusion possible.

La dimension du sous-espace propre E_λ est appelé la *multiplicité géométrique* de λ, on la note $\operatorname{mult}^{\mathbf{A}}_{\mathrm{geo}}(\lambda)$, ou $\operatorname{mult}_{\mathrm{geo}}(\lambda)$ s'il n'y a pas de confusion possible. Autrement dit

$$\operatorname{mult}^{\mathbf{A}}_{\mathrm{geo}}(\lambda) = \dim(\operatorname{Ker}(\mathbf{A} - \lambda\mathbf{1}_n)).$$

6.7 Proposition.— Soit $\mathbf{A}$ une matrice de $\mathscr{M}_n(\mathbb{C})$. Pour toute valeur propre λ de $\mathbf{A}$, on a

$$\operatorname{mult}_{\mathrm{geo}}(\lambda) \leqslant \operatorname{mult}_{\mathrm{alg}}(\lambda).$$

Preuve. Soit λ une racine de $p_{\mathbf{A}}$ d'ordre de multiplicité h. D'après le théorème de trigonalisation 6.1, la matrice $\mathbf{A}$ est semblable à une matrice triangulaire supérieure $\mathbf{T}$ de $\mathscr{M}_n(\mathbb{C})$:

$$\mathbf{T} = \begin{bmatrix} \mathbf{T}_1 & \mathbf{B} \\ \mathbf{0} & \mathbf{T}_2 \end{bmatrix},$$

où $\mathbf{T}_1$ est une matrice triangulaire supérieure de $\mathscr{M}_h(\mathbb{C})$ dont tous les coefficients diagonaux sont égaux à λ et où la matrice $\mathbf{T}_2$ est triangulaire supérieure n'ademettant pas le coefficient λ sur sa diagonale. On a

$$\operatorname{rang}(\mathbf{A}-\lambda\mathbf{1}_n) = \operatorname{rang}\begin{bmatrix} \mathbf{T}_1-\lambda\mathbf{1}_h & \mathbf{B} \\ \mathbf{0} & \mathbf{T}_2-\lambda\mathbf{1}_{n-h} \end{bmatrix}$$

Par suite, on a

$$\operatorname{rang}(\mathbf{A}-\lambda\mathbf{1}_n) \geqslant \operatorname{rang}(\mathbf{T}_2-\lambda\mathbf{1}_{n-h}) = n-h.$$

La dernière égalité provient du fait que $\mathbf{T}_2-\lambda\mathbf{1}_{n-h}$ est inversible, car λ n'est pas sur la diagonale de $\mathbf{T}_2$, elle est donc de déterminant non nul.

On en déduit que

$$h = \operatorname{mult}_{\text{alg}}(\lambda) \geqslant n - \operatorname{rang}(\mathbf{A}-\lambda\mathbf{1}_n).$$

Or, d'après le théorème du rang, on a $n = \operatorname{rang}(\mathbf{A}-\lambda\mathbf{1}_n) + \operatorname{mult}_{\text{geo}}(\lambda)$. Ainsi

$$\operatorname{mult}_{\text{geo}}(\lambda) \leqslant \operatorname{mult}_{\text{alg}}(\lambda).$$

□

6.3.3. Exemple.— Considérons la matrice de $\mathscr{M}_n(\mathbb{R})$, avec $n \geqslant 3$, suivante

$$\mathbf{A} = \begin{bmatrix} 0 & \cdots & 0 & 1 \\ \vdots & & \vdots & \vdots \\ 0 & \cdots & 0 & 1 \\ 1 & \cdots & 1 & 1 \end{bmatrix}$$

La matrice $\mathbf{A}$ est de rang 2, par conséquence, par la formule du rang, théorème 1.9, la dimension du sous-espace propre $E_0 = \operatorname{Ker} \mathbf{A}$ est $\dim E_0 = n-2$. Ainsi, d'après la proposition 6.7, la multiplicité algébrique de la valeur propre 0 satisfait

$$n-2 \leqslant \operatorname{mult}_{\text{alg}}(0).$$

Le polynôme caractéristique de $\mathbf{A}$ se factorise sous la forme

$$p_{\mathbf{A}} = (-1)^n x^{n-2}(x^2+\alpha x+\beta)$$

où α et β sont deux réels. Nous avons vu en 6.1.7 que le calcul des traces des matrices $\mathbf{A}$ et $\mathbf{A}^2$ permet de déterminer les deux réels α et β ; on obtient ainsi toutes les valeurs propres de la matrice $\mathbf{A}$.

§ 4 Caractérisation des matrices diagonalisables

6.8 Théorème (Caractérisation des matrices diagonalisables).— Soit $\mathbf{A}$ une matrice de $\mathscr{M}_n(\mathbb{K})$. Les assertions suivantes sont équivalentes :

i) $\mathbf{A}$ est diagonalisable dans $\mathscr{M}_n(\mathbb{K})$,

ii) le polynôme $p_\mathbf{A}$ est scindé sur $\mathbb{K}$ et, pour toute valeur propre λ de $\mathbf{A}$,

$$\mathrm{mult}_{\mathrm{geo}}(\lambda) = \mathrm{mult}_{\mathrm{alg}}(\lambda),$$

iii) il existe des valeurs propres $\lambda_1, \ldots, \lambda_p$ de $\mathbf{A}$, telles que

$$\mathbb{K}^n = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_p}.$$

Preuve. Montrons que **i)** implique **ii)**. Supposons que $\mathbf{A}$ soit diagonalisable. Alors $\mathbf{A}$ est semblable à une matrice diagonale dont la diagonale est formée des valeurs propres de $\mathbf{A}$. Notons $\lambda_1, \ldots, \lambda_p$ ces valeurs propres. On a

$$p_\mathbf{A} = (-1)^n (x-\lambda_1)^{n_1} \ldots (x-\lambda_p)^{n_p},$$

où $n_i = \mathrm{mult}_{\mathrm{alg}}(\lambda_i)$ est la multiplicité algébrique de la valeur propre λ_i, c'est-à-dire, le nombre de fois que λ_i apparaît sur la diagonale.

Montrons que pour tout $i \in [\![1, p]\!]$, $\mathrm{mult}_{\mathrm{alg}}(\lambda_i) = \mathrm{mult}_{\mathrm{geo}}(\lambda_i)$. D'après la proposition 6.4, il existe une base $\mathscr{B}$ de $\mathbb{K}^n$ formée de vecteurs propres de $\mathbf{A}$. Il existe n_i vecteurs de la base $\mathscr{B}$ vérifiant $u(\mathbf{x}) = \lambda_i \mathbf{x}$, c'est-à-dire, n_i vecteurs linéairement indépendants dans le sous-espace propre E_{λ_i}. Par suite,

$$\mathrm{mult}_{\mathrm{alg}}(\lambda_i) = n_i \leqslant \dim E_{\lambda_i} = \mathrm{mult}_{\mathrm{geo}}(\lambda_i).$$

Or, d'après la proposition 6.7, on a $\mathrm{mult}_{\mathrm{geo}}(\lambda_i) \leqslant \mathrm{mult}_{\mathrm{alg}}(\lambda_i)$. On obtient ainsi l'égalité $\mathrm{mult}_{\mathrm{geo}}(\lambda_i) = \mathrm{mult}_{\mathrm{alg}}(\lambda_i)$.

Montrons que **ii)** implique **iii)**. Soit $\mathbf{A}$ une matrice dont le polynôme caractéristique est scindé, soit

$$p_\mathbf{A} = (-1)^n (x-\lambda_1)^{n_1} (x-\lambda_2)^{n_2} \ldots (x-\lambda_p)^{n_p},$$

et tel que pour tout i,

$$n_i = \mathrm{mult}_{\mathrm{alg}}(\lambda_i) = \mathrm{mult}_{\mathrm{geo}}(\lambda_i).$$

D'après la proposition 5.2, les sous-espaces propres sont en somme directe. Soit F le sous-espace de $\mathbb{K}^n$ défini par

$$F = E_{\lambda_1} \oplus E_{\lambda_2} \oplus \ \ldots \ \oplus E_{\lambda_p}.$$

On a

$$\begin{aligned}\dim F &= \dim E_{\lambda_1} + \ldots + \dim E_{\lambda_p}\\ &= n_1 + \ldots + n_p\\ &= \deg(p_{\mathbf{A}}) = n.\end{aligned}$$

Ainsi le sous-espace F de $\mathbb{K}^n$ est de dimension n, par suite $F = \mathbb{K}^n$. Ce qui montre l'assertion **iii)**.

Montrons que **iii)** implique **i)**. Supposons que **A** admette des valeurs propres $\lambda_1, \ldots, \lambda_p$ telles que

$$\mathbb{K}^n = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_p}.$$

Considérons, pour tout $i \in [\![1,p]\!]$, $\mathscr{B}_i$ une base de E_{λ_i}, alors $\mathscr{B} = \mathscr{B}_1 \cup \cdots \cup \mathscr{B}_p$ forme une base de vecteurs propres de $\mathbb{K}^n$. De la proposition 6.4, on déduit alors que **A** est diagonalisable. $\square$

6.4.1. Remarque.— On peut résumer ce résultat, en disant qu'une matrice **A** est diagonalisable si, et seulement si, son polynôme caractéristique s'écrit

$$p_{\mathbf{A}} = (-1)^n (x-\lambda_1)^{\mathrm{mult_{geo}}(\lambda_1)} \ldots (x-\lambda_p)^{\mathrm{mult_{geo}}(\lambda_p)},$$

avec $\lambda_i \in \mathbb{K}$.

En particulier, on retrouve la proposition 6.5. Si **A** n'admet que des racines simples, alors, pour tout i,

$$1 \leqslant \mathrm{mult_{geo}}(\lambda_i) \leqslant \mathrm{mult_{alg}}(\lambda_i) = 1.$$

Par suite, $\mathrm{mult_{geo}}(\lambda_i) = \mathrm{mult_{alg}}(\lambda_i) = 1$ et **A** est diagonalisable.

6.4.2. Exemple.— Nous avons vu dans l'exemple 5.3.4 que la matrice $\mathbf{R}_\theta$ de la rotation du plan vectoriel $\mathbb{R}^2$ n'est pas diagonalisable sur $\mathbb{R}$. Cependant elle possède deux valeurs propres complexes distinctes : $e^{i\theta}$ et $e^{-i\theta}$. La matrice $\mathbf{R}_\theta$ est donc diagonalisable sur $\mathbb{C}$, on a

$$E_{i\theta} = \mathrm{Vect}(\begin{bmatrix} 1 \\ -i \end{bmatrix}), \qquad E_{-i\theta} = \mathrm{Vect}(\begin{bmatrix} 1 \\ i \end{bmatrix}).$$

D'où, en posant

$$\mathbf{P} = \begin{bmatrix} 1 & 1 \\ -i & -i \end{bmatrix}$$

la matrice de changement de la base canonique à la base formée des vecteurs propres $\begin{bmatrix} 1 \\ -i \end{bmatrix}$ et $\begin{bmatrix} 1 \\ i \end{bmatrix}$, on a

$$\begin{bmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{bmatrix} = \mathbf{P}^{-1} \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \mathbf{P},$$

où l'inverse de la matrice **P** est

$$\mathbf{P}^{-1} = \frac{1}{2}\begin{bmatrix} 1 & i \\ 1 & -i \end{bmatrix}.$$

Exercice 7.— Nous avons vu dans l'exercice 5, qu'une matrice **A** par blocs :

$$\mathbf{A} = \begin{bmatrix} \mathbf{B} & \mathbf{0} \\ \mathbf{0} & \mathbf{C} \end{bmatrix},$$

où **B** et **C** sont deux matrices carrées diagonalisables de $\mathscr{M}_{n_1}(\mathbb{K})$ et $\mathscr{M}_{n_2}(\mathbb{K})$ respectivement, est diagonalisable. L'objectif de cet exercice est de montrer que la réciproque est vraie.

1. Montrer que $p_{\mathbf{A}} = p_{\mathbf{B}}.p_{\mathbf{C}}$.
2. Montrer que, pour toute valeur propre λ de **A**, on a

$$\mathrm{mult}^{\mathbf{A}}_{\mathrm{geo}}(\lambda) = \mathrm{mult}^{\mathbf{B}}_{\mathrm{geo}}(\lambda) + \mathrm{mult}^{\mathbf{C}}_{\mathrm{geo}}(\lambda).$$

3. Montrer que si **B** ou **C** n'est pas diagonalisable, alors il existe une valeur propre λ de **A** telle que

$$\mathrm{mult}^{\mathbf{A}}_{\mathrm{geo}}(\lambda) < \mathrm{mult}^{\mathbf{A}}_{\mathrm{alg}}(\lambda).$$

4. En déduire, que si **A** est diagonalisable, alors **B** et **C** sont diagonalisables.

§ 5 Matrices diagonalisables : premières applications

Nous présentons ici une première application de la diagonalisation des matrices avec le calcul des puissances des matrices et la résolution de systèmes d'équations différentielles linéaires. Nous reviendrons plus longuement sur ces applications dans les prochains chapitres, en particulier, nous traiterons aussi le cas des matrices non diagonalisables.

6.5.1. Calcul des puissances d'une matrice diagonalisable.— Une première application classique de la diagonalisation est le calcul des puissance d'une matrice.

Soit **A** une matrice diagonalisable de $\mathscr{M}_n(\mathbb{K})$. Il existe une matrice diagonale

$$\mathbf{D} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_p \end{bmatrix}$$

et une matrice inversible **P** telles que : $\mathbf{D} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$. Alors $\mathbf{A}^k = \mathbf{P}\mathbf{D}^k\mathbf{P}^{-1}$, pour tout entier naturel k, d'où

$$\mathbf{A}^k = \mathbf{P}\begin{bmatrix} \lambda_1^k & 0 & \cdots & 0 \\ 0 & \lambda_2^k & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_p^k \end{bmatrix}\mathbf{P}^{-1}$$

6.5.2. Exemple.— Les puissances successives de la matrice

$$\mathbf{A} = \begin{bmatrix} 1 & -1 \\ 2 & 4 \end{bmatrix}$$

sont

$$\mathbf{A}^k = \mathbf{P}\mathbf{D}^k\mathbf{P}^{-1} = \begin{bmatrix} 2^{k+1}-3^k & 2^{k+1}-2.3^k \\ -2^k+3^k & -2^k+2.3^k \end{bmatrix}$$

avec

$$\mathbf{D} = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}, \qquad \mathbf{P} = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix}, \qquad \mathbf{P}^{-1} = \begin{bmatrix} 2 & 1 \\ -1 & -1 \end{bmatrix}.$$

6.5.3. Résolution des systèmes différentiels linéaires.— Une autre application classique de la diagonalisation d'une matrice est la résolution des systèmes différentiels linéaires.

On se propose de résoudre les systèmes différentiels linéaires de la forme :

$$\left|\begin{array}{ll} \dfrac{dx_1(t)}{dt} & = a_1^1 x_1(t) + \cdots + a_1^n x_n(t) \\ \vdots & \qquad\qquad \vdots \\ \dfrac{dx_n(t)}{dt} & = a_n^1 x_1(t) + \cdots + a_n^n x_n(t) \end{array}\right.$$

où les a_i^j sont des réels et les x_i des fonctions réelles à valeurs réelles. Un tel système différentiel prend la forme matricielle suivante :

$$\frac{d\mathbf{x}(t)}{dt} = \mathbf{A}\mathbf{x}(t), \tag{6.4}$$

avec

$$\mathbf{A} = \begin{bmatrix} a_1^1 & \ldots & a_1^n \\ \vdots & & \vdots \\ a_n^1 & \ldots & a_n^n \end{bmatrix}, \qquad \mathbf{x}(t) = \begin{bmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{bmatrix}$$

et où $\dfrac{d\mathbf{x}(t)}{dt}$ désigne la dérivé du vecteur $\mathbf{x}(t)$.

Supposons que la matrice $\mathbf{A}$ soit diagonalisable (nous aborderons le cas des systèmes différentiels avec $\mathbf{A}$ non diagonalisable plus tard dans le cours), il existe une matrice $\mathbf{D}$ diagonale et $\mathbf{P}$ inversible telles que

$$\mathbf{D} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}.$$

La matrice $\mathbf{P}$ désigne un changement de base de $\mathbb{R}^n$: si $\mathbf{x}(t)$ est le vecteur colonne exprimant un vecteur $\mathbf{x}(t)$ dans la base initiale et $\mathbf{y}(t)$ celui exprimant $\mathbf{x}(t)$ dans la nouvelle base. On fait le changement de variable $\mathbf{y}(t) = \mathbf{P}^{-1}\mathbf{x}(t)$. D'où

$$\frac{d\mathbf{y}(t)}{dt} = \mathbf{P}^{-1}\frac{d\mathbf{x}(t)}{dt},$$

donc

$$\frac{d\mathbf{y}(t)}{dt} = \mathbf{P}^{-1}\mathbf{A}\mathbf{x}(t) = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}\mathbf{y}(t) = \mathbf{D}\mathbf{y}(t).$$

Le système est donc équivalent au système :

$$\frac{d\mathbf{y}(t)}{dt} = \mathbf{D}\mathbf{y}(t),$$

qui est facile à intégrer, car $\mathbf{D}$ est diagonale. En effet, si

$$\mathbf{D} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_p \end{bmatrix}$$

les solutions de cette équation sont les vecteurs $\mathbf{y}(t)$ dont la i-ième composante est

$$y_i(t) = e^{\lambda_i t} y_i(0).$$

Il suffit alors de calculer $\mathbf{x}(t)$ en calculant $\mathbf{x}(t) = \mathbf{P}\mathbf{y}(t)$.

6.5.4. Exemple.— Soit à résoudre le système différentiel suivant :

$$\left| \begin{aligned} \frac{dx(t)}{dt} &= x(t) - y(t), \\ \frac{dy(t)}{dt} &= 2x(t) + 4y(t) \end{aligned} \right. \qquad (6.5)$$

où $x, y : \mathbb{R} \longrightarrow \mathbb{R}$ sont deux fonctions réelles. On pose

$$\mathbf{A} = \begin{bmatrix} 1 & -1 \\ 2 & 4 \end{bmatrix}, \quad \mathbf{D} = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} \quad \text{et} \quad \mathbf{P} = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix}.$$

Le système $\dfrac{d\mathbf{y}}{dt} = \mathbf{D}\mathbf{y}$, avec $\mathbf{y}(t) = \begin{bmatrix} u(t) \\ v(t) \end{bmatrix}$ s'écrit sous la forme :

$$\left| \begin{aligned} \frac{du(t)}{dt} &= 2u(t) \\ \frac{dv(t)}{dt} &= 3v(t) \end{aligned} \right.$$

Ces deux équations ont pour solution

$$\left| \begin{array}{l} u(t) = \beta_1 e^{2t} \\ v(t) = \beta_2 e^{3t} \end{array} \right.$$

où β_1 et β_2 sont deux constantes réelles. On en déduit que le système (6.5) admet pour solution

$$\left| \begin{array}{l} x(t) = \beta_1 e^{2t} + \beta_2 e^{3t} \\ y(t) = -\beta_1 e^{2t} - 2\beta_2 e^{3t} \end{array} \right.$$

Exercice 8.— On considère la matrice suivante de $\mathscr{M}_n(\mathbb{R})$, avec $n \geqslant 2$,

$$\mathbf{A} = \begin{bmatrix} n & 1 & \cdots & 1 \\ 1 & n & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 1 & \cdots & 1 & n \end{bmatrix}$$

1. La matrice $\mathbf{A}$ est-elle diagonalisable ?
2. Montrer que $n-1$ est une valeur propre de $\mathbf{A}$.
3. Déterminer le sous-espace propre associé à la valeur propre $n-1$.
4. Quel est l'ordre de multiplicité de la valeur propre $n-1$?
5. Calculer la trace de la matrice $\mathbf{A}$. En déduire la valeur de toutes les valeurs propres de $\mathbf{A}$.
6. Résoudre le système différentiel suivant

$$\left| \begin{array}{l} \dfrac{dx(t)}{dt} = 3x(t) + y(t) + z(t) \\ \dfrac{dy(t)}{dt} = x(t) + 3y(t) + z(t) \\ \dfrac{dz(t)}{dt} = x(t) + y(t) + 3z(t) \end{array} \right.$$

où $x, y, z : \mathbb{R} \longrightarrow \mathbb{R}$ sont trois fonctions dérivables vérifiant $x(0) = 1$, $y(0) = z(0) = 0$.

§ 6 Trigonalisation et diagonalisation des endomorphismes

La trigonalisation et la diagonalisation des endomorphismes sur un espace vectoriel de dimension finie se traite de la même façon que pour les matrices.

Dans la suite de cette section, E désigne un $\mathbb{K}$-espace vectoriel de dimension finie n.

6.6.1. Endomorphismes trigonalisables.— Un endomorphisme u de E est dit *trigonalisable* sur $\mathbb{K}$, s'il existe une base $\mathscr{B}$ de E, telle que la matrice $[u]_{\mathscr{B}}$ de u exprimée dans cette base soit triangulaire supérieure.

6.9 Proposition.— Un endomorphisme de E est trigonalisable si, et seulement si, sa matrice, exprimée dans une base de E, est trigonalisable.

En particulier, tout endomorphisme u d'un $\mathbb{C}$-espace vectoriel E est trigonalisable sur $\mathbb{C}$.

6.10 Proposition.— Soit u un endomorphisme trigonalisable de E. La restriction de u à un sous-espace vectoriel stable par u est trigonalisable.

Preuve. Soit F un sous-espace vectoriel E stable par u. Le polynôme caractéristique $p_{u_{|F}}$ de la restriction de u à F divise le polynôme caractéristique de u. L'endomorphisme étant trigonalisable, ce dernier est scindé, donc $p_{u_{|F}}$ est scindé et $u_{|F}$ est trigonalisable. $\square$

6.6.2. Endomorphismes diagonalisables.— Un endomorphisme u de E est dit *diagonalisable* sur $\mathbb{K}$, s'il existe une base de E formée de vecteurs propres de u.

Par exemple, l'endomorphisme u de $\mathbb{R}^2$ vu dans l'exemple 5.2.7 est diagonalisable. En effet, nous avons montré que $\mathbb{R}^2$ se décompose en une somme directe de sous-espaces propres associés à u. Il existe donc une base de $\mathbb{R}^2$ formée de vecteurs propres de u.

6.6.3. Lien avec les matrices diagonalisables.— Soit u un endomorphisme diagonalisable de E. Soit $\mathscr{B} = (\mathbf{x}_1, \ldots, \mathbf{x}_n)$ une base de diagonalisation de E, où, $\mathbf{x}_i$ désigne un vecteur propre associé à une valeur propre λ_i de u. Alors, pour tout i, on a $u(\mathbf{x}_i) = \lambda_i \mathbf{x}_i$ et la matrice de u exprimée dans la base $\mathscr{B}$ est

$$[u]_{\mathscr{B}} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix}$$

Inversement, si $\mathscr{B}' = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ est une base de E telle que la matrice $[u]_{\mathscr{B}'}$ soit diagonale, alors les éléments de la diagonale sont valeurs propres de u et les vecteurs $\mathbf{e}_i$ sont des vecteurs propres de u. Ainsi

6.11 Proposition.— Un endomorphisme de E est diagonalisable sur $\mathbb{K}$ si, et seulement si, sa matrice, exprimée dans une base de E, est diagonalisable.

On en déduit les résultats analogues au cas matriciel :

6.12 Proposition.— Soit u un endomorphisme de E.

i) L'endomorphisme u est trigonalisable si, et seulement si, son polynôme caractéristique est scindé sur $\mathbb{K}$.

ii) Si le polynôme caractéristique de u est scindé sur $\mathbb{K}$ et possède toutes ses racines simples, alors u est diagonalisable sur $\mathbb{K}$.

6.6.4. Multiplicités des valeurs propres d'un endomorphisme.— Soient u un endomorphisme de E et λ une valeur propre de u. L'ordre de multiplicité de λ en tant que racine du polynôme p_u est appelé *multiplicité algébrique* de λ, on la note $\mathrm{mult}_{\mathrm{alg}}(\lambda)$.

La dimension du sous-espace propre E_λ est appelé la *multiplicité géométrique* de λ, on la note $\mathrm{mult}_{\mathrm{geo}}(\lambda)$.

Comme dans le cas matriciel, on a l'inégalité :

$$\mathrm{mult}_{\mathrm{geo}}(\lambda) \leqslant \mathrm{mult}_{\mathrm{alg}}(\lambda).$$

On peut utiliser la preuve matricielle de la proposition 6.7. Sinon, on peut procéder de la façon suivante.

Supposons que λ soir une valeur propre de u d'ordre de multiplicité algébrique h. On a

$$p_u = (x-\lambda)^h g(x),$$

avec $g(x) \in \mathbb{K}[x]$ et $g(\lambda) \neq 0$. Le sous-espace propre E_λ est stable par u, donc d'aprés la proposition 5.7, le polynôme $p_{u|E_\lambda}$ divise le polynôme p_u. Or, la restriction de u au sous-espace propre E_λ est une homothétie :

$$u_{|E_\lambda} = \lambda \,\mathrm{id}_{E_\lambda}.$$

D'où $p_{u_{|E_\lambda}} = (-1)^{\dim E_\lambda}(x-\lambda)^{\dim E_\lambda}$. Par suite, $\dim E_\lambda \leqslant h$.

6.13 Théorème (Caractérisation des endomorphismes diagonalisables).— Soit u un endomorphisme de E. Les assertions suivantes sont équivalentes :

i) u est diagonalisable sur $\mathbb{K}$,

ii) p_u est scindé sur $\mathbb{K}$ et, pour toute valeur propre λ de u, $\mathrm{mult}_{\mathrm{geo}}(\lambda) = \mathrm{mult}_{\mathrm{alg}}(\lambda)$,

iii) il existe des valeurs propres $\lambda_1, \ldots, \lambda_p$ de u telles que

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

6.6.5. Exemple.— Soit u l'endomorphisme de $\mathbb{R}^2$ défini par

$$u(x_1\mathbf{e}_1 + x_2\mathbf{e}_2) = x_2\mathbf{e}_1.$$

La matrice de u dans la base canonique de $\mathbb{R}^2$ est

$$[u]_{\mathscr{C}\mathrm{an}} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

L'endomorphisme u est la projection sur la droite $(O\mathbf{e}_2)$ parallèlement à la droite $(O\mathbf{e}_1)$ suivie de la symétrie d'axe $x_1 = x_2$:

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

Montrons que u n'est pas diagonalisable en considérant les sous-espaces propres possibles pour u. Supposons que F soit un sous-espace propre de u. Alors F ne peut être nul, ni de dimension 2 puisque u n'est pas une homothétie. Supposons que F soit de dimension 1.

La droite $\mathrm{Vect}(\mathbf{e}_1)$ est stable par u, pour tout $\mathbf{x} \in \mathrm{Vect}(\mathbf{e}_1)$ on a

$$u(\mathbf{x}) = \mathbf{0} \in \mathrm{Vect}(\mathbf{e}_1).$$

Ainsi, $\mathrm{Vect}(\mathbf{e}_1)$ est un sous-espace propre, pour la valeur propre 0. Soit $\mathrm{Vect}(\mathbf{e}_1 + \lambda\mathbf{e}_2)$, $\lambda \neq 0$, une droite différente de $\mathrm{Vect}(\mathbf{e}_1)$. On a

$$u(\mathbf{e}_1 + \lambda\mathbf{e}_2) = \lambda\mathbf{e}_1 \notin \mathrm{Vect}(\mathbf{e}_1 + \lambda\mathbf{e}_2).$$

Donc $\mathrm{Vect}(\mathbf{e}_1 + \lambda\mathbf{e}_2)$ n'est pas stable par u et $\mathrm{Vect}(\mathbf{e}_1)$ n'admet de supplémentaire stable par u. Ainsi, l'endomorphisme u n'est pas diagonalisable.

Exercice 9.— Soit u l'endomorphisme de $\mathbb{R}^4$ représenté dans la base canonique $\mathscr{C}\mathrm{an} = (\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3, \mathbf{c}_4)$ par la matrice

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}.$$

1. Déterminer le rang de u. En déduire que 0 est valeur propre de u.
2. Montrer que $\mathbf{c}_1 + \mathbf{c}_2 + \mathbf{c}_3 + \mathbf{c}_4$ est vecteur propre de u.
3. Construire une base de $\mathbb{R}^4$ formée de vecteurs propres de u.

Exercice 10.— Soit u l'endomorphisme de $\mathbb{R}^3$ représenté dans la base canonique $(\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3)$ par la matrice

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

1. Calculer $u(\mathbf{c}_2)$, $u(\mathbf{c}_1 + \mathbf{c}_3)$ et $u(\mathbf{c}_1 - \mathbf{c}_3)$.

2. En déduire que u est diagonalisable et écrire la matrice de u dans une base de vecteurs propres.
3. Donner une interprétation géométrique de u.

Exercice 11.— On considère l'endomorphisme u de $\mathbb{C}^3$ représenté dans la base canonique par la matrice

$$\begin{bmatrix} 4 & 2 & 4 \\ 0 & 0 & -1 \\ -3 & -2 & -3 \end{bmatrix}$$

1. Déterminer les valeurs propres de l'endomorphisme u.
2. Montrer sans calcul qu'il existe une base de $\mathbb{C}^3$ formée de vecteurs propres.
3. Déterminer la matrice de passage de la base canonique à la base formée de vecteurs propres.
4. L'endomorphisme u est-il diagonalisable sur les corps $\mathbb{R}$ ou $\mathbb{Q}$?

—

§ 7 Exercices

Exercice 12.— Montrer que la matrice suivante n'est pas diagonalisable :

$$\begin{bmatrix} 0 & 1 & & & 0 \\ & 0 & 1 & & \\ & & 0 & \ddots & \\ & & & \ddots & 1 \\ 0 & & & & 0 \end{bmatrix}$$

Exercice 13.— Les matrices élémentaires de $\mathscr{M}_n(\mathbb{K})$, voir 2.1.3, sont-elles diagonalisables ?

Exercice 14.— Diagonaliser ou trigonaliser dans $\mathscr{M}_n(\mathbb{C})$, en donnant la matrice de passage, les matrices suivantes :

$$\mathbf{A} = \begin{bmatrix} 5 & 1 & 3 \\ 4 & 3 & 4 \\ -1 & -1 & 1 \end{bmatrix}, \ \mathbf{B} = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}, \ \mathbf{C} = \begin{bmatrix} -1 & 1 & 3 \\ -2 & 2 & 2 \\ -2 & 1 & 4 \end{bmatrix}.$$

Exercice 15.— Discuter en fonction de a, b et c la possibilité de diagonaliser les matrices de $\mathscr{M}_3(\mathbb{C})$ suivantes :

$$\begin{bmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & -1 \end{bmatrix}, \qquad \begin{bmatrix} 1 & a & 1 \\ 0 & 1 & b \\ 0 & 0 & c \end{bmatrix}.$$

Exercice 16.— Soit θ un réel. On considère la matrice de rotation

$$\mathbf{R}_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$$

1. Calculer dans $\mathbb{C}$ les valeurs propres de A.
2. Discuter en fonction de θ la possibilité de diagonaliser A dans $\mathscr{M}_n(\mathbb{R})$ et $\mathscr{M}_n(\mathbb{C})$.

Exercice 17.— Soit $\mathbf{A} \in \mathscr{M}_n(\mathbb{R})$.

1. Montrer que si λ est une valeur propre complexe de $\mathbf{A}$, alors $\overline{\lambda}$ est aussi valeur propre de $\mathbf{A}$, de même ordre de multiplicité.
2. Montrer que si $\mathbf{v}$ est un vecteur propre associé à λ, alors $\overline{\mathbf{v}}$ est un vecteur propre associé à $\overline{\lambda}$.
3. Diagonaliser en donnant une matrice de passage la matrice

$$\mathbf{A} = \begin{bmatrix} 0 & -2 & 0 \\ 1 & 0 & -1 \\ 0 & 2 & 0 \end{bmatrix}.$$

4. Calculer $\mathbf{A}^k$, pour tout entier naturel k.

Exercice 18.— On note $\mathbb{R}_n[x]$ le $\mathbb{R}$-espace vectoriel formé des polynômes de degré inférieur ou égal à n. Soit u l'endomorphisme de $\mathbb{R}_n[x]$ défini par

$$u(p) = (x^2-1)p'' + 3xp'.$$

1. Écrire la matrice de u dans la base canonique de $\mathbb{R}_n[x]$.
2. Montrer que u est diagonalisable.
3. Résoudre l'équation $u(p) = p$.

Exercice 19.— Soit u l'endomorphisme de $\mathscr{M}_2(\mathbb{R})$ défini par

$$u\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

Montrer que l'endomorphisme u est diagonalisable et construire une base de vecteurs propres de u.

Exercice 20.— Soit

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}.$$

En diagonalisant $\mathbf{A}$, trouver une solution dans $\mathscr{M}_3(\mathbb{C})$ à l'équation $\mathbf{X}^2 = \mathbf{A}$.

Exercice 21.— Soient E un $\mathbb{K}$-espace vectoriel de dimension finie et u un endomorphisme de E de rang un.

1. Montrer que la trace de u est une valeur propre de u.
2. En déduire que u est diagonalisable si, et seulement si, sa trace est non nulle.

Exercice 22.— On considère la matrice de $\mathscr{M}_2(\mathbb{C})$ suivante

$$\mathbf{A} = \begin{bmatrix} a & b \\ c & -a \end{bmatrix}.$$

1. Quelle est la somme des valeurs propres de **A** ?
2. Quel est le produit des valeurs propres de **A** ?
3. Montrer que, si son déterminant n'est pas nul, **A** est diagonalisable.
4. Montrer que, si son déterminant est nul, **A** n'est diagonalisable que si elle est nulle.
5. Montrer que **A** est diagonalisable sauf si elle est de rang un.
6. En supposant que la matrice **A** est réelle ; à quelle condition est-elle diagonalisable par un changement de base réel ?

Exercice 23.— Soient E un $\mathbb{K}$-espace vectoriel de dimension n et u un endomorphisme de E. On considère l'ensemble des endomorphismes de E qui commutent avec u :

$$\mathrm{Com}_u = \{v \in \mathscr{L}(E) \mid u \circ v = v \circ u\}.$$

1. Montrer que Com_u est un $\mathbb{K}$-espace vectoriel.

2. On suppose que u est diagonalisable et on note $\lambda_1, \ldots, \lambda_p$ les valeurs propres de u.

2.1. Montrer que

$$\mathrm{Com}_u = \{v \in \mathscr{L}(E) \mid E_{\lambda_i} \text{ est stable par } v, \text{ pour tout } i \in \{1, \ldots, p\}\}.$$

2.2. Ecrire la forme générale de la matrice d'un endomorphisme $v \in \mathrm{Com}_u$ dans la base $B = B_1 \cup \cdots \cup B_p$, où pour tout i, B_i est une base de E_{λ_i}. En déduire que

$$\dim \mathrm{Com}_u = \sum_{i=1}^{p} (\dim E_{\lambda_i})^2.$$

3. On suppose que u admet n valeurs propres distinctes $\lambda_1, \ldots, \lambda_n$. On veut montrer que le $\mathbb{K}$-espace vectoriel Com_u coïncide avec le $\mathbb{K}$-espace vectoriel $\mathbb{K}\langle \mathrm{id}, u, \ldots, u^{n-1}\rangle$, librement engendré par les endomorphismes $\mathrm{id}, u, \ldots, u^{n-1}$.

3.1. Montrer que $\mathbb{K}\langle \mathrm{id}, u, \ldots, u^{n-1}\rangle \subset \mathrm{Com}_u$.

3.2. Soit n scalaires $a_0, \ldots, a_{n-1}$ de $\mathbb{K}$ tels que $\sum_{i=0}^{n} a_i u^i = 0$. Montrer que les a_i sont solutions du système linéaire :

$$(*) \left| \begin{array}{ll} a_0 + a_1\lambda_1 + a_2\lambda_1^2 + \cdots + a_{n-1}\lambda_1^{n-1} & = 0 \\ a_0 + a_1\lambda_2 + a_2\lambda_2^2 + \cdots + a_{n-1}\lambda_2^{n-1} & = 0 \\ \vdots \qquad\qquad\qquad\qquad \vdots & \vdots \\ a_0 + a_1\lambda_n + a_2\lambda_n^2 + \cdots + a_{n-1}\lambda_n^{n-1} & = 0 \end{array} \right.$$

3.3. Soit

$$\mathbf{V} = \begin{bmatrix} 1 & \lambda_1 & \lambda_1^2 & \dots & \lambda_1^{n-1} \\ 1 & \lambda_2 & \lambda_2^2 & \dots & \lambda_2^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \lambda_n & \lambda_n^2 & \dots & \lambda_n^{n-1} \end{bmatrix}$$

la matrice associée au système $(*)$. Montrer par récurrence que

$$\det \mathbf{V} = \prod_{1 \leqslant i < j \leqslant n} (\lambda_j - \lambda_i).$$

3.4. Conclure.

Exercice 24.— Soient E un $\mathbb{K}$-espace vectoriel et u, v deux endomorphismes de E tels que
$u \circ v = v \circ u$.

1. Montrer que tout sous-espace propre de u est stable par v.
2. Montrer que u et v sont diagonalisables si, et seulement si, il existe une base commune de diagonalisation.
3. Soient u et v les endomorphismes de $\mathbb{R}^3$ représentés respectivement dans la base canonique de $\mathbb{R}^3$ par les matrices

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 2 \end{bmatrix}, \qquad \begin{bmatrix} 0 & 1 & 1 \\ -1 & 1 & -1 \\ 1 & 1 & 3 \end{bmatrix}$$

   Montrer que u et v commutent.
4. Déterminer les valeurs propres et sous-espaces propres de u et v.
5. Déterminer deux sous-espaces de $\mathbb{R}^3$ invariant par u et v dont l'un est de dimension 1 et l'autre de dimension 2.
6. En déduire une base de $\mathbb{R}^3$ trigonalisant les endomorphismes u et v.

Exercice 25.— Soient E un $\mathbb{K}$-espace vectoriel de dimension finie et u, v deux endomorphismes de E.

1. Montrer que si 0 est valeur propre de $u \circ v$ alors 0 est aussi valeur propre de $v \circ u$.
2. Donner un exemple d'endomorphismes u et v tels que les sous-espaces Ker $(u \circ v)$ et Ker $(v \circ u)$ ne soient pas de même dimension.
   Pour la suite de l'exercice, on suppose que $u \circ v$ possède une valeur propre non nulle λ.
3. Montrer que λ est valeur propre de $v \circ u$.
4. On note $E_\lambda^{u \circ v}$ le sous-espace propre de $u \circ v$ associé à la valeur propre λ et $E_\lambda^{v \circ u}$ le sous-espace propre de $v \circ u$ associé à la valeur propre λ. Montrer que $u(E_\lambda^{v \circ u})$ est un sous-espace vectoriel de $E_\lambda^{u \circ v}$ et que $v(E_\lambda^{u \circ v})$ est un sous-espace vectoriel de $E_\lambda^{v \circ u}$.
5. Montrer que la restriction de u à $E_\lambda^{v \circ u}$ est injective. En déduire que

$$\dim E_\lambda^{v \circ u} \leqslant \dim E_\lambda^{u \circ v}.$$

6. Déduire des questions précédentes que

$$\dim E_\lambda^{v\circ u} = \dim E_\lambda^{u\circ v}.$$

7. On suppose que u et v sont des isomorphismes. Montrer que $u\circ v$ et $v\circ u$ ne possèdent pas de valeurs propres nulles.
8. En déduire que si u et v sont des isomorphismes alors $u\circ v$ est diagonalisable si, et seulement si, $v\circ u$ est diagonalisable.

Exercice 26 (Lemme de Schur).— Soient E un $\mathbb{C}$-espace vectoriel de dimension finie et Q un sous-ensemble irréductible d'endomorphismes de E, i.e., les seuls sous-espaces de E stables par tous les éléments de Q sont $\{\mathbf{0}\}$ et E.

1. Montrer que, pour tout $u \in \mathscr{L}(E)$ commutant avec tout les éléments de Q, il existe une valeur propre λ dont le sous-espace propre est E.
2. En déduire que les seuls endomorphismes de E qui commutent avec les éléments de Q sont les homothéties.
3. Dans le cas où E est un $\mathbb{R}$-espace vectoriel de dimension finie, montrer en trouvant un contre exemple que le résultat précédent est faux.
4. Montrer que le résultat est vrai si E est un $\mathbb{R}$-espace vectoriel de dimension finie impaire.

Exercice 27.— On considère la matrice $\mathbf{A} = \begin{bmatrix} 1 & -1 \\ 2 & 4 \end{bmatrix}$ de $\mathscr{M}_2(\mathbb{R})$.

1. Montrer que $\mathbf{A}$ est diagonalisable et diagonaliser $\mathbf{A}$.
2. Soit $\mathbf{N}$ une matrice de $\mathscr{M}_n(\mathbb{R})$ et $\mathbf{M}$ la matrice de $\mathscr{M}_{2n}(\mathbb{R})$ défini par blocs :

$$\mathbf{M} = \begin{bmatrix} \mathbf{N} & -\mathbf{N} \\ 2\mathbf{N} & 4\mathbf{N} \end{bmatrix}.$$

Montrer que la matrice $\mathbf{M}$ est diagonalisable si, et seulement si, la matrice $\mathbf{N}$ est diagonalisable.

CHAPITRE
7

Le polynôme minimal

Sommaire

La caractérisation des matrices diagonalisables donnée par le théorème 6.8 porte sur la dimension des sous-espaces propres : une matrice de $\mathscr{M}_n(\mathbb{K})$ est diagonalisable si, et seulement si, l'espace $\mathbb{K}^n$ se décompose en une somme directe de sous-espaces propres de $\mathbf{A}$.

Dans ce chapitre, nous abordons une nouvelle caractérisation, de nature algébrique, qui porte uniquement sur les coefficients de la matrice. Précisément, nous allons montrer qu'une matrice $\mathbf{A}$ de $\mathscr{M}_n(\mathbb{K})$ est diagonalisable dans $\mathscr{M}_n(\mathbb{K})$ si, et seulement si, il existe un polynôme g à coefficients dans $\mathbb{K}$, non nul, scindé, n'ayant que des racines simples et tel que $\mathbf{A}$ soit racine de g, i.e., $g(\mathbf{A}) = 0$.

§ 1 Préliminaires

7.1.1. Exemple.— Considèrons la projection p sur le plan Π de $\mathbb{R}^3$ d'équation $z = 0$ parallèlement à la droite Δ d'équation $x = y = z$, vue dans l'exemple 4.3.4. La projection p vérifie, voir figure 7.1 :

$$p(\mathbf{x}) = \begin{cases} \mathbf{x} & \text{si } \mathbf{x} \in \Pi, \\ \mathbf{0} & \text{si } \mathbf{x} \in \Delta. \end{cases}$$

FIGURE 7.1.: Projection sur le plan Π parallèlement à la droite Δ

Ainsi, l'endomorphisme p satisfait, pour tout vecteur $\mathbf{x}$ de $\mathbb{R}^3$, l'équation

$$p^2(\mathbf{x}) = p(\mathbf{x}).$$

De façon équivalente, cela s'exprime en disant que l'endomorphisme $p^2 - p$ est nul, soit :

$$p(p - \mathrm{id}_{\mathbb{R}^3}) = 0. \tag{7.1}$$

On dit alors que l'endomorphisme p est racine du polynôme $x(x-1)$.

De l'équation (7.1), nous déduisons que l'endomorphisme p est diagonalisable. D'une part, il est immédiat que

$$\mathrm{Ker}\ p \cap \mathrm{Ker}\ (p - \mathrm{id}_{\mathbb{R}^3}) = \{\mathbf{0}\}.$$

En effet, si $\mathbf{x} \in \mathrm{Ker}\ p$ et $\mathbf{x} \in \mathrm{Ker}\ (p - \mathrm{id}_{\mathbb{R}^3})$, alors $p(\mathbf{x}) = \mathbf{0}$ et $p(\mathbf{x}) = \mathbf{x}$, d'où $\mathbf{x} = \mathbf{0}$. D'autre part, pour tout vecteur $\mathbf{x}$ de $\mathbb{R}^3$, on a la décomposition :

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

L'endomorphisme p satisfaisant l'équation (7.1), on a $\mathbf{x} - p(\mathbf{x}) \in \mathrm{Ker}\ p$ et $p(\mathbf{x}) \in \mathrm{Ker}\ (p - \mathrm{id}_{\mathbb{R}^3})$. On montre ainsi que l'on a une somme directe

$$\mathbb{R}^3 = \mathrm{Ker}\ p \oplus \mathrm{Ker}\ (p - \mathrm{id}_{\mathbb{R}^3}).$$

Les sous-espaces $\mathrm{Ker}\ p$ et $\mathrm{Ker}\ (p - \mathrm{id}_{\mathbb{R}^3})$ désignant respectivement les sous-espaces propres associées aux valeurs propres 0 et 1, du théorème 6.8, on déduit que l'endomorphisme p est diagonalisable. On pourra remarquer que $\mathrm{Ker}\ p$ correspond à la droite Δ et $\mathrm{Ker}\ (p - \mathrm{id}_{\mathbb{R}^3})$ correspond au plan Π. Si $\mathbf{e}$ est un vecteur de Δ et $(\mathbf{e}_1, \mathbf{e}_2)$ une base de Π, alors $\mathscr{B} = (\mathbf{e}, \mathbf{e}_1, \mathbf{e}_2)$ forme une base de $\mathbb{R}^3$ et

$$[u]_{\mathscr{B}} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

7.1.2. Exemple.— Considérons la matrice

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

La matrice $\mathbf{A}$ vérifie la relation $\mathbf{A}^2 = \mathbf{1}_3$ et donc l'équation :

$$(\mathbf{A} - \mathbf{1}_3)(\mathbf{A} + \mathbf{1}_3) = \mathbf{0}. \tag{7.2}$$

On dit alors que la matrice $\mathbf{A}$ est racine du polynôme $(x-1)(x+1)$. De la même façon que dans l'exemple précédent, de l'équation (7.2), on déduit une décomposition de l'espace $\mathbb{R}^3$ en une somme directe de noyaux :

$$\mathbb{R}^3 = \operatorname{Ker}(\mathbf{A}-\mathbf{1}_3) \oplus \operatorname{Ker}(\mathbf{A}+\mathbf{1}_3).$$

Les deux sous-espaces de cette décomposition correspondent aux sous-espaces propres de la matrice $\mathbf{A}$ associés aux valeurs propres 1 et -1 respectivement. D'après le théorème 6.8, on en déduit que la matrice $\mathbf{A}$ est diagonalisable. De plus, ayant

$$\mathbf{A}+\mathbf{1}_3 = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix},$$

la matrice $\mathbf{A}+\mathbf{1}_3$ est de rang 2, par le théorème du rang, on en déduit que $\operatorname{Ker}(\mathbf{A}+\mathbf{1}_3)$ est de dimension 1. Il en découle que le noyau $\operatorname{Ker}(\mathbf{A}-\mathbf{1}_3)$ est alors de rang 2. En d'autres termes, on a $\mathrm{mult}_{\mathrm{geo}}(-1) = 1$ et $\mathrm{mult}_{\mathrm{geo}}(1) = 2$. La matrice $\mathbf{A}$ étant diagonalisable, elle est alors semblable à la matrice

$$\begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Exercice 1.— Soit $\mathbf{A}$ une matrice de $\mathscr{M}_n(\mathbb{K})$, avec $n \geqslant 2$, vérifiant

$$(\mathbf{A}-2\mathbf{1}_n)(\mathbf{A}-3\mathbf{1}_n) = \mathbf{0}.$$

1. Montrer que $\mathbf{A}$ est diagonalisable dans $\mathscr{M}_n(\mathbb{K})$.
2. Montrer que la matrice $\mathbf{A}$ est inversible.
3. Exprimer l'inverse $\mathbf{A}^{-1}$ en fonction de la matrice $\mathbf{A}$.

Exercice 2.— Soit $\mathbf{A}$ une matrice de $\mathscr{M}_n(\mathbb{K})$, avec $n \geqslant 3$, vérifiant

$$(\mathbf{A}-2\mathbf{1}_n)(\mathbf{A}-3\mathbf{1}_n)(\mathbf{A}-5\mathbf{1}_n) = \mathbf{0}.$$

1. Montrer que $\mathbf{A}$ est diagonalisable dans $\mathscr{M}_n(\mathbb{K})$.
2. Montrer que la matrice $\mathbf{A}$ est inversible.
3. Exprimer l'inverse $\mathbf{A}^{-1}$ en fonction de la matrice $\mathbf{A}$.

§ 2 Polynômes de matrices

7.2.1. Définition.— Soit $\mathbf{A}$ une matrice de $\mathscr{M}_n(\mathbb{K})$. Étant donné un polynôme

$$p = a_m x^m + a_{m-1}x^{m-1} + \cdots + a_1 x + a_0$$

de $\mathbb{K}[x]$, on définit la matrice

$$P(\mathbf{A}) = a_m\mathbf{A}^m + a_{m-1}\mathbf{A}^{m-1} + \cdots + a_1\mathbf{A} + a_0\mathbf{1}_n.$$

Noter que si p est le polynôme constant égal à 1, alors $p(\mathbf{A}) = \mathbf{1}_n$. On associe ainsi à un polynôme p de $\mathbb{K}[x]$, un *polynôme de matrices* $p(\mathbf{A})$. Cette correspondance est compatible aux opérations d'addition et de multiplication sur les polynômes, on a :

7.1 Proposition.— Soient f et g deux polynôme de $\mathbb{K}[x]$. Alors, pour toute matrice $\mathbf{A}$, on a

i) $(f+g)(\mathbf{A}) = f(\mathbf{A}) + g(\mathbf{A})$,

ii) $(fg)(\mathbf{A}) = f(\mathbf{A})g(\mathbf{A})$.

Preuve. Considérons deux polynômes f et g définis par

$$f = \sum_{i=1}^{l} a_i x^i, \qquad g = \sum_{i=1}^{m} b_i x^i.$$

Quitte à rajouter des coefficients nuls, on peut supposer que $l = m$. On a

$$\begin{aligned}
(f+g)(\mathbf{A}) &= (\sum_{i=1}^{l}(a_i+b_i)x^i)(\mathbf{A}) \\
&= \sum_{i=1}^{l}(a_i+b_i)\mathbf{A}^i \\
&= \sum_{i=1}^{l} a_i\mathbf{A}^i + \sum_{i=1}^{l} b_i\mathbf{A}^i \\
&= f(\mathbf{A}) + g(\mathbf{A}).
\end{aligned}$$

Par distributivité du produit matriciel par rapport à l'addition, cf. 2.2.1, on a

$$\begin{aligned}
(fg)(\mathbf{A}) &= (\sum_{i=1}^{l}\sum_{j=1}^{l} a_i b_j x^{i+j})(\mathbf{A}) \\
&= (\sum_{i=1}^{l}\sum_{j=1}^{l} a_i b_j \mathbf{A}^{i+j}) \\
&= \sum_{i=1}^{l} a_i\mathbf{A}^i \sum_{j=1}^{l} b_j\mathbf{A}^j \\
&= f(\mathbf{A})g(\mathbf{A}).
\end{aligned}$$

□

De la seconde propriété, on déduit les polynômes d'une matrice $\mathbf{A}$ commutent entre eux :

7.2 Proposition.— Soit $\mathbf{A}$ une matrice de $\mathscr{M}_n(\mathbb{K})$. Pour tous polynômes f et g de $\mathbb{K}[x]$, les matrices $f(\mathbf{A})$ et $g(\mathbf{A})$ commutent :

$$f(\mathbf{A})g(\mathbf{A}) = g(\mathbf{A})f(\mathbf{A}).$$

Exercice 3.— Montrer la proposition 7.2.

7.2.2. Exemple.— Si $\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ et $p = x^2 + x + 1$, alors

$$p(\mathbf{A}) = \mathbf{A}^2 + \mathbf{A} + \mathbf{1}_2 = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}.$$

7.2.3. Polynômes annulateurs.— Un polynôme non nul q de $\mathbb{K}[x]$ est dit *annulateur* d'une matrice $\mathbf{A}$ de $\mathscr{M}_n(\mathbb{K})$, si la matrice $q(\mathbf{A})$ est nulle ; on dit aussi que $\mathbf{A}$ est racine du polynôme q.

7.2.4. Exemple : polynôme annulateur des matrices de $\mathscr{M}_2(\mathbb{K})$.— Toute matrice $\mathbf{A}$ de $\mathscr{M}_2(\mathbb{K})$ admet pour polynôme annulateur, le polynôme

$$p_{\mathbf{A}} = x^2 - \mathrm{trace}\,(\mathbf{A})x + \det(\mathbf{A}).$$

Pour montrer que $p_{\mathbf{A}}(\mathbf{A}) = \mathbf{0}$, on peut procéder par un calcul direct comme dans la preuve initiale de Cayley, voir figure 7.4. Si $\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, alors on montre que

$$\mathbf{A}^2 - \mathrm{trace}\,(\mathbf{A})\mathbf{A} + \det(\mathbf{A})\mathbf{1}_2 = \mathbf{0}.$$

7.3 Proposition.— Toute matrice de possède un polynôme annulateur.

Preuve. Soit $\mathbf{A}$ une matrice de $\mathscr{M}_n(\mathbb{K})$. Le $\mathbb{K}$-espace vectoriel $\mathscr{M}_n(\mathbb{K})$ est de dimension n^2. Par suite, toute famille de n^2+1 matrices de $\mathscr{M}_n(\mathbb{K})$ est liée ; c'est le cas en particulier de la famille

$$(\mathbf{1}_n, \mathbf{A}, \mathbf{A}^2, \ldots, \mathbf{A}^{n^2}).$$

Il existe donc des scalaires $a_0, \ldots, a_{n^2}$ non tous nuls, tels que

$$a_0\mathbf{1}_n + a_1\mathbf{A} + \cdots + a_{n^2}\mathbf{A}^{n^2} = \mathbf{0}.$$

Le polynôme $g = a_0 + a_1x + \cdots + a_{n^2}x^{n^2}$ est ainsi non nul et annulateur de la matrice $\mathbf{A}$. $\square$

7.4 Proposition.— Soient $\mathbf{A}$ une matrice de $\mathscr{M}_n(\mathbb{K})$ et g un polynôme annulateur de $\mathbf{A}$. Alors, toute valeur propre de $\mathbf{A}$ est racine du polynôme g.

Preuve. Soit $g = a_0 + a_1x + a_2x^2 + \cdots + a_mx^m$ un polynôme annulateur de $\mathbf{A}$. La matrice $g(\mathbf{A})$ est nulle :

$$a_0\mathbf{1}_n + a_1\mathbf{A} + a_2\mathbf{A}^2 + \cdots + a_m\mathbf{A}^p = \mathbf{0}. \tag{7.3}$$

Soient λ une valeur propre de $\mathbf{A}$ et $\mathbf{x}$ un vecteur propre associé, i.e., $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$. D'après l'équation (7.3), on a

$$(a_0\mathbf{1}_n + a_1\mathbf{A} + a_2\mathbf{A}^2 + \cdots + a_m\mathbf{A}^m)\mathbf{x} = \mathbf{0}.$$

Or $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$, d'où pour tout entier naturel k, $\mathbf{A}^k\mathbf{x} = \lambda^k\mathbf{x}$ et

$$(a_0 + a_1\lambda + a_2\lambda^2 + \cdots + a_m\lambda^m)\mathbf{x} = \mathbf{0}.$$

Comme le vecteur $\mathbf{x}$ est non nul, on en déduit que $g(\lambda)$ est nul. $\square$

Attention, la réciproque de ce résultat est fausse en général ; toutes les racines d'un polynôme annulateur de $\mathbf{A}$ ne sont pas toujours valeurs propres de $\mathbf{A}$. Par exemple, la matrice identité $\mathbf{1}_n$ est racine du polynôme $x(x-1)$, car $\mathbf{1}_n^2 = \mathbf{1}_n$, alors que 0 n'est pas une valeur propre de la matrice identité.

§ 3 Le lemme de décomposition en noyaux

La matrice $\mathbf{A}$ de l'exemple 7.1.2 satisfait la relation

$$(\mathbf{A} - \mathbf{1}_3)(\mathbf{A} + \mathbf{1}_3) = \mathbf{0}.$$

Nous en avons déduit que

$$\mathbb{R}^3 = \mathrm{Ker}\,(\mathbf{A} - \mathbf{1}_3) \oplus \mathrm{Ker}\,(\mathbf{A} + \mathbf{1}_3).$$

Dans cette section, nous montrons que ceci est une conséquence du résultat général suivant.

7.5 Proposition.— Soit $\mathbf{A}$ une matrice de $\mathscr{M}_n(\mathbb{K})$ et soient f_1, f_2 deux polynômes de $\mathbb{K}[x]$ premiers entre eux. Alors

$$\mathrm{Ker}\,(f_1f_2)(\mathbf{A}) = \mathrm{Ker}\,f_1(\mathbf{A}) \oplus \mathrm{Ker}\,f_2(\mathbf{A}).$$

Si de plus, le polynôme $f_1 f_2$ est annulateur de $\mathbf{A}$, on a

$$\mathbb{K}^n = \text{Ker } f_1(\mathbf{A}) \oplus \text{Ker } f_2(\mathbf{A}).$$

Preuve. Les polynômes f_1 et f_2 étant premiers entre eux, d'après l'identité de Bézout, théorème 0.4, il existe des polynômes h_1 et h_2 de $\mathbb{K}[x]$ tels que :

$$f_1 h_1 + f_2 h_2 = 1.$$

En conséquence, nous avons la relation matricielle suivante :

$$f_1(\mathbf{A})h_1(\mathbf{A}) + f_2(\mathbf{A})h_2(\mathbf{A}) = \mathbf{1}_n. \tag{7.4}$$

Montrons l'égalité Ker $f_1(\mathbf{A}) \oplus$ Ker $f_2(\mathbf{A}) =$ Ker $(f_1 f_2)(\mathbf{A})$. Le noyau Ker $f_2(\mathbf{A})$ est contenu dans Ker $(f_1(\mathbf{A})f_2(\mathbf{A}))$, de la même façon Ker $f_1(\mathbf{A})$ est contenu dans Ker $(f_1(\mathbf{A})f_2(\mathbf{A})) =$ Ker $(f_2(\mathbf{A})f_1(\mathbf{A}))$, car les polynômes $f_1(\mathbf{A})$ et $f_2(\mathbf{A})$ commutent entre eux. Ainsi

$$\text{Ker } f_1(\mathbf{A}) + \text{Ker } f_2(\mathbf{A}) \subseteq \text{Ker } (f_1 f_2)(\mathbf{A}).$$

Inversement, si $\mathbf{x} \in$ Ker $(f_1 f_2)(\mathbf{A})$, alors d'après la relation (7.4), il existe une décomposition

$$\mathbf{x} = f_1(\mathbf{A})h_1(\mathbf{A})(\mathbf{x}) + f_2(\mathbf{A})h_2(\mathbf{A})(\mathbf{x})$$

On a

$$f_2(\mathbf{A})f_1(\mathbf{A})h_1(\mathbf{A})(\mathbf{x}) = h_1(\mathbf{A})(f_1 f_2)(\mathbf{A})(\mathbf{x}) = \mathbf{0}.$$

La première égalité découle du fait que les polynômes en $\mathbf{A}$ commutent entre eux, la seconde du fait que $\mathbf{x}$ est un vecteur de Ker $(f_1 f_2)(\mathbf{A})$ par hypothèse. Ainsi $f_1(\mathbf{A})h_1(\mathbf{A})(\mathbf{x})$ est un vecteur de Ker $f_2(\mathbf{A})$. De la même façon, on montre que $f_2(\mathbf{A})h_2(\mathbf{A})(\mathbf{x})$ est un vecteur de Ker $f_1(\mathbf{A})$.

Reste à montrer que la somme

$$\text{Ker } f_1(\mathbf{A}) + \text{Ker } f_2(\mathbf{A}) = \text{Ker } (f_1 f_2)(\mathbf{A})$$

est directe. Si $\mathbf{x} \in$ Ker $f_1(\mathbf{A}) \cap$ Ker $f_2(\mathbf{A})$, d'après la relation (7.4), on a

$$\mathbf{x} = f_1(\mathbf{A})h_1(\mathbf{A})(\mathbf{x}) + f_2(\mathbf{A})h_2(\mathbf{A})(\mathbf{x}).$$

Donc

$$\mathbf{x} = h_1(\mathbf{A})f_1(\mathbf{A})(\mathbf{x}) + h_2(\mathbf{A})f_2(\mathbf{A})(\mathbf{x})$$

soit $\mathbf{x} = \mathbf{0}$.

Si l'on suppose de plus que le polynôme $f_1 f_2$ est annulateur de $\mathbf{A}$, on a $(f_1 f_2)(\mathbf{A}) = \mathbf{0}$, d'où Ker $(f_1 f_2)(\mathbf{A}) = \mathbb{K}^n$. Ce qui montre la deuxième assertion :

$$\mathbb{K}^n = \text{Ker } f_1(\mathbf{A}) \oplus \text{Ker } f_2(\mathbf{A}).$$

□

7.3.1. Exemple.— Soit **A** une matrice de $\mathscr{M}_n(\mathbb{R})$ satisfaisant

$$(\mathbf{A}-\alpha\mathbf{1}_2)(\mathbf{A}-\beta\mathbf{1}_2)=\mathbf{0} \tag{7.5}$$

où α et β sont deux réels distincts. D'après la proposition précédente, on a

$$\mathbb{R}^2 = E_\alpha \oplus E_\beta,$$

où E_α et E_β sont les deux sous-espaces propres associés à la matrice **A**. Un point important de la preuve de la proposition 7.5 est qu'elle fournit une méthode constructive pour déterminer l'expression des projections de $\mathbb{R}^2$ sur les sous-espaces propres E_α et E_β. Les polynômes $x-\alpha$ et $x-\beta$ sont premiers entre eux, il existe donc une identité de Bézout :

$$\frac{-1}{\alpha-\beta}(x-\alpha)+\frac{1}{\alpha-\beta}(x-\beta)=1. \tag{7.6}$$

Notons

$$\Pi_1=\frac{1}{\alpha-\beta}(\mathbf{A}-\beta\mathbf{1}_2), \qquad \Pi_2=\frac{-1}{\alpha-\beta}(\mathbf{A}-\alpha\mathbf{1}_2)$$

Alors Π_1 représente la projection de $\mathbb{R}^2$ sur E_α et Π_2 représente la projection de $\mathbb{R}^2$ sur E_β, exprimée dans la base canonique de $\mathbb{R}^2$. En effet, de la relation (7.6), on déduit que

$$\Pi_1+\Pi_2=\mathbf{1}_2. \tag{7.7}$$

Par ailleurs, de la relation 7.5, on déduit que

$$\Pi_1\Pi_2=\Pi_2\Pi_1=\mathbf{0} \tag{7.8}$$

Des relation 7.6 et 7.7, on déduit que Π_1 et Π_2 sont des matrices de projecteurs de $\mathbb{R}^2$, i.e.,

$$\Pi_1^2=\Pi_1, \quad \Pi_2^2=\Pi_2.$$

Comme

$$\operatorname{Im}\Pi_1=\operatorname{Ker}\Pi_2 \text{ et } \operatorname{Im}\Pi_2=\operatorname{Ker}\Pi_1,$$

on a $\operatorname{Im}\Pi_1=E_\alpha$ et $\operatorname{Im}\Pi_2=E_\beta$. Ainsi, Π_1 et Π_2 sont les matrices des projecteurs de $\mathbb{R}^2$ sur les sous-espaces E_α et E_β respectivement.

Exercice 4.— Considérons la matrice réelle

$$\mathbf{A}=\begin{bmatrix} 1 & 0 & 1 & 2\\ 0 & 1 & 3 & 4\\ 0 & 0 & 2 & 0\\ 0 & 0 & 0 & 2\end{bmatrix}$$

1. Montrer que $(\mathbf{A}-2\mathbf{1}_4)(\mathbf{A}-\mathbf{1}_4)=\mathbf{0}$.
2. Déterminer les sous-espaces propres ainsi que les matrices des projections de $\mathbb{R}^4$ sur ces sous-espaces propres, exprimées dans la base canonique de $\mathbb{R}^4$.

7.3.2. Le lemme des noyaux.— La formulation générale de cette décomposition, avec un produit fini quelconque de polynômes, est appelée le *lemme des noyaux*. La preuve se déroule sur le même principe qu'en présence de deux polynômes.

7.6 Théorème (Lemme des noyaux).— Soient $\mathbf{A}$ une matrice de $\mathscr{M}_n(\mathbb{K})$ et $f_1, \ldots, f_p$ des polynômes de $\mathbb{K}[x]$ premiers entre eux deux à deux. Alors,

$$\mathrm{Ker}\,\big((f_1 \ldots f_p)(\mathbf{A})\big) = \mathrm{Ker}\, f_1(\mathbf{A}) \oplus \ldots \oplus \mathrm{Ker}\, f_p(\mathbf{A}).$$

Si de plus, le polynôme $f_1 f_2 \ldots f_p$ est annulateur de $\mathbf{A}$, on a

$$\mathbb{K}^n = \mathrm{Ker}\, f_1(\mathbf{A}) \oplus \ldots \oplus \mathrm{Ker}\, f_p(\mathbf{A}).$$

Preuve. Posons $g = f_1 \ldots f_p$ et, pour tout entier $i \in [\![1, p]\!]$, posons

$$g_i = \prod_{\substack{j=1 \\ j \neq i}}^{p} f_j.$$

Les polynômes $f_1, \ldots, f_p$ sont premiers entre eux deux à deux, donc les polynômes $g_1, \ldots, g_p$ sont premiers entre eux dans leur ensemble. D'après l'identité de Bézout, théorème 0.4, il existe des polynômes $h_1, \ldots, h_p$ de $\mathbb{K}[x]$ tels que :

$$g_1 h_1 + \ldots + g_p h_p = 1.$$

Par suite, on a la relation matricielle :

$$g_1(\mathbf{A}) h_1(\mathbf{A}) + \ldots + g_p(\mathbf{A}) h_p(\mathbf{A}) = \mathbf{1}_n. \tag{7.9}$$

Montrons alors que $\mathrm{Ker}\, f_1(\mathbf{A}) + \ldots + \mathrm{Ker}\, f_p(\mathbf{A}) = \mathrm{Ker}\, g(\mathbf{A})$.

Pour tout $i \in [\![1, p]\!]$, les polynômes en $\mathbf{A}$ commutant entre eux, on a

$$g(\mathbf{A}) = f_1(\mathbf{A}) \;\ldots\; f_i(\mathbf{A}) \;\ldots\; f_p(\mathbf{A}) = f_1(\mathbf{A}) \ldots f_p(\mathbf{A}) f_i(\mathbf{A}).$$

Par suite, pour tout $i \in [\![1, p]\!]$, le noyau de $f_i(\mathbf{A})$ est contenu dans le noyau de $g(\mathbf{A})$. La somme des noyaux

$$\mathrm{Ker}\, f_1(\mathbf{A}) + \ldots + \mathrm{Ker}\, f_p(\mathbf{A})$$

est ainsi contenue dans le noyau $\mathrm{Ker}\, g(\mathbf{A})$.

Inversement, si $\mathbf{x} \in \mathrm{Ker}\, g(\mathbf{A})$, d'après la relation (7.9), on a :

$$\mathbf{x} = g_1(\mathbf{A}) h_1(\mathbf{A})(\mathbf{x}) + \ldots + g_p(\mathbf{A}) h_p(\mathbf{A})(\mathbf{x}). \tag{7.10}$$

Par ailleurs, pour tout $i \in [\![1, p]\!]$, on a

$$\begin{aligned} f_i(\mathbf{A})(g_i(\mathbf{A})h_i(\mathbf{A})(\mathbf{x})) &= g(\mathbf{A})h_i(\mathbf{A})(\mathbf{x}) \\ &= h_i(\mathbf{A})g(\mathbf{A})(\mathbf{x}) \\ &= \mathbf{0}. \end{aligned}$$

Donc $g_i(\mathbf{A})h_i(\mathbf{A})(\mathbf{x}) \in \operatorname{Ker} f_i(\mathbf{A})$ et d'après la décomposition (7.10), on en déduit que

$$\mathbf{x} \in \operatorname{Ker} f_1(\mathbf{A}) + \ldots + \operatorname{Ker} f_p(\mathbf{A}).$$

Reste à montrer que la somme des noyaux $\operatorname{Ker} f_1(\mathbf{A}) + \ldots + \operatorname{Ker} f_p(\mathbf{A})$ est directe. D'après 1.3.3, il suffit de montrer que pour tout $i \in [\![2, p]\!]$, l'intersection

$$\operatorname{Ker} f_i(\mathbf{A}) \cap \sum_{j=1}^{i-1} \operatorname{Ker} f_j(\mathbf{A}) \tag{7.11}$$

est le sous-espace vectoriel nul. Soit $i \in [\![2, p]\!]$ et soit $\mathbf{x}$ un vecteur de l'intersection (7.11). D'après la relation (7.9), on a $\mathbf{x} = g_1(\mathbf{A})h_1(\mathbf{A})(\mathbf{x}) + \ldots + g_p(\mathbf{A})h_p(\mathbf{A})(\mathbf{x})$, d'où

$$\mathbf{x} = g_i(\mathbf{A})h_i(\mathbf{A})(\mathbf{x}) + \sum_{\substack{j=1 \\ j \neq i}}^{p} g_j(\mathbf{A})h_j(\mathbf{A})(\mathbf{x}). \tag{7.12}$$

Si $j \neq i$, le polynôme f_i divise g_j, or $\mathbf{x} \in \operatorname{Ker} g_i(u)$ donc

$$\sum_{\substack{j=1 \\ j \neq i}}^{p} g_j(\mathbf{A})h_j(\mathbf{A})(\mathbf{x}) = \sum_{\substack{j=1 \\ j \neq i}}^{p} h_j(\mathbf{A})g_j(\mathbf{A})(\mathbf{x}) = \mathbf{0}.$$

Par ailleurs, $\mathbf{x} \in \sum_{j=1}^{i-1} \operatorname{Ker} f_j(\mathbf{A})$, donc $\mathbf{x} = \sum_{j=1}^{i-1} \mathbf{x}_j$, où $\mathbf{x}_j \in \operatorname{Ker} f_j(\mathbf{A})$. Ainsi

$$g_i(\mathbf{A})h_i(\mathbf{A})(\mathbf{x}) = h_i(\mathbf{A})g_i(\mathbf{A})(\mathbf{x}) = \mathbf{0}.$$

D'après la décomposition (7.12), on en déduit que $\mathbf{x} = \mathbf{0}$.

Si le polynôme $g = f_1 f_2 \ldots f_p$ est annulateur de $\mathbf{A}$, on a

$$\operatorname{Ker} g(\mathbf{A}) = \operatorname{Ker} f_1(\mathbf{A}) \oplus \ldots \oplus \operatorname{Ker} f_p(\mathbf{A}).$$

Or $g(\mathbf{A}) = 0$, donc $\operatorname{Ker} g(\mathbf{A}) = \mathbb{K}^n$, d'où la seconde assertion du théorème. □

7.3.3. Conséquence immédiate du lemme des noyaux.— Soit $\mathbf{A}$ une matrice de $\mathscr{M}_n(\mathbb{K})$ dont le polynôme caractéristique est

$$p_{\mathbf{A}} = (x - \lambda_1)^{n_1} \ldots (x - \lambda_p)^{n_p},$$

avec $\lambda_i \neq \lambda_j$. Les polynômes $(x-\lambda_i)^{n_i}$ sont premiers entre eux deux à deux, du lemme des noyaux, nous déduisons la décomposition

$$\mathbb{K}^n = \mathrm{Ker}\,(\mathbf{A}-\lambda_1\mathbf{1}_n)^{n_1} \oplus \ldots \oplus \mathrm{Ker}\,(\mathbf{A}-\lambda_p\mathbf{1}_n)^{n_p}.$$

Exercice 5.— Soit **A** une matrice de $\mathscr{M}_4(\mathbb{R})$ de trace nulle, admettant 1 et i comme valeurs propres.

1. Déterminer toutes les autres valeures propres de **A**.

2. Montrer que
$$\mathbb{R}^4 = \mathrm{Ker}\,(\mathbf{A}^2-\mathbf{1}_n) \oplus \mathrm{Ker}\,(\mathbf{A}^2+\mathbf{1}_n).$$

§ 4 Le polynôme minimal

7.4.1. Une autre caractérisation des matrices diagonalisables.— La principale conséquence du lemme des noyaux, théorème 7.6, que nous énoncerons ici est une nouvelle caractérisation de la diagonalisation des matrices.

7.7 Théorème.— Une matrice **A** de $\mathscr{M}_n(\mathbb{K})$ est diagonalisable dans $\mathscr{M}_n(\mathbb{K})$ si, et seulement si, il existe un polynôme de $\mathbb{K}[x]$ annulateur de **A** scindé et ne possédant que des racines simples.

Preuve. Montrons que la condition est suffisante. Soit g un polynôme de $\mathbb{K}[x]$ annulateur de **A** qui soit scindé sur $\mathbb{K}$ et avec toutes ses racines simples. Alors, le polynôme g s'écrit sous la forme
$$g = (x-\lambda_1)\ldots(x-\lambda_p),$$
avec $\lambda_i \neq \lambda_j$, si $i \neq j$. Comme par hypothèse $g(\mathbf{A}) = \mathbf{0}$ et que les facteurs dans la décomposition du polynôme g sont premiers entre eux deux à deux, d'après le lemme des noyaux, théorème 7.6, on a :

$$\mathbb{K}^n = \mathrm{Ker}\,(\mathbf{A}-\lambda_1\mathbf{1}_n) \oplus \ldots \oplus \mathrm{Ker}\,(\mathbf{A}-\lambda_p\mathbf{1}_n).$$

Le sous-espace $\mathrm{Ker}\,(\mathbf{A}-\lambda_i\mathbf{1}_n)$ étant le sous-espace propre associé à la valeur propre λ_i, l'espace $\mathbb{K}^n$ se décompose en une somme directe de sous-espaces propres. Du théorème 6.8, on en déduit que la matrice **A** est diagonalisable.

Montrons que la condition est nécessaire. Supposons que la matrice **A** soit diagonalisable. Il existe alors une base $\mathscr{B}$ de $\mathbb{K}^n$ formée de vecteurs propres de **A**. Notons $\lambda_1,\ldots,\lambda_p$ les valeurs propres de **A** et soient $E_{\lambda_1},\ldots,E_{\lambda_p}$ les sous-espaces propres associés.

Le polynôme $g = (x-\lambda_1)\ldots(x-\lambda_p)$ de $\mathbb{K}[x]$ est scindé et toutes ses racines sont simples. Pour tout vecteur $\mathbf{e}$ de la base $\mathscr{B}$, il existe au moins une valeur propre λ_i telle que $(\mathbf{A}-\lambda_i\mathbf{1}_n)\mathbf{e} = \mathbf{0}$. Par suite

$$g(\mathbf{A})(\mathbf{e}) = (\mathbf{A}-\lambda_1\mathbf{1}_n)\ldots(\mathbf{A}-\lambda_p\mathbf{1}_n)\mathbf{e} = \mathbf{0}.$$

On a $g(\mathbf{A})\mathbf{e} = \mathbf{0}$ pour tout vecteur $\mathbf{e}$ de la base $\mathscr{B}$ de $\mathbb{K}^n$, par suite $g(\mathbf{A})\mathbf{x} = \mathbf{0}$ pour tout vecteur de $\mathbb{K}^n$. Il s'ensuit que $g(\mathbf{A}) = \mathbf{0}$. □

Exercice 6.— Trouver une condition nécessaire et suffisante pour que les matrices réelles suivantes soient diagonalisables :

$$\mathbf{A} = \begin{bmatrix} a & b & c \\ 0 & a & d \\ 0 & 0 & a \end{bmatrix}, \qquad \mathbf{B} = \begin{bmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & d \end{bmatrix}.$$

7.4.2. Le polynôme minimal.— D'après le théorème 7.7, une matrice est diagonalisable si, et seulement si, elle possède un polynôme annulateur scindé dont toutes les racines sont simples. Il s'agit d'une caractérisation de nature algébrique, dans le sens où elle ne porte que sur les coefficients de la matrice. On comprend alors l'intérêt de caractériser l'ensemble des polynômes annulateurs d'une matrice. Nous abordons maintenant une méthode permettant de déterminer tous les polynômes annulateurs d'une matrice.

7.4.3. Définition.— Soit $\mathbf{A}$ une matrice de $\mathscr{M}_n(\mathbb{K})$. On appelle *polynôme minimal* de $\mathbf{A}$ et on note $m_{\mathbf{A}}$, le polynôme unitaire annulateur de $\mathbf{A}$ de plus petit degré.

7.8 Proposition.— Soit $\mathbf{A}$ une matrice de $\mathscr{M}_n(\mathbb{K})$. Le polynôme minimal $m_{\mathbf{A}}$ est unique et divise tout polynôme annulateur de $\mathbf{A}$.

Preuve. Montrons que $m_{\mathbf{A}}$ divise tout polynôme annulateur de $\mathbf{A}$. Soit g un polynôme annulateur de $\mathbf{A}$. En effectuant la division euclidienne de g par $m_{\mathbf{A}}$, théorème **??**, il existe deux polynômes :

$$g = g' m_{\mathbf{A}} + r,$$

avec $\deg r < \deg m_{\mathbf{A}}$. La matrice $\mathbf{A}$ étant racine du polynôme g et de son polynôme minimal $m_{\mathbf{A}}$, on a

$$\mathbf{0} = g(\mathbf{A}) = g'(\mathbf{A})\, m_{\mathbf{A}}(\mathbf{A}) + r(\mathbf{A}).$$

D'où $\mathbf{A}$ est racine du polynôme r. Or, par définition, $m_{\mathbf{A}}$ est le polynôme annulateur de $\mathbf{A}$ de plus petit degré, par suite le polynôme r ne peut pas être nul, car sinon on aurait R annulateur de $\mathbf{A}$ et $\deg r < \deg m_{\mathbf{A}}$. Ainsi, le polynôme reste r est nul, et par conséquent le polynôme $m_{\mathbf{A}}$ divise g.

Montrons l'unicité du polynôme minimal $m_{\mathbf{A}}$. Supposons que la matrice $\mathbf{A}$ admette deux polynômes minimaux m et m'. Ils sont tous deux annulateurs de $\mathbf{A}$, donc m divise m' et m' divise m. Par suite, il existe un scalaire $\alpha \in \mathbb{K}$ tel que

$$m = \alpha m'.$$

Les polynômes m et m' étant unitaires, on en déduit que $m = m'$. Ce qui montre l'unicité. □

L'ensemble des polynômes annulateur d'une matrice **A** de $\mathscr{M}_n(\mathbb{K})$ est ainsi déterminé par son polynôme minimal $m_{\mathbf{A}}$. En effet, tout polynôme annulateur g de **A** s'écrit :

$$g = g' m_{\mathbf{A}},$$

où g' est un polynôme de $\mathbb{K}[x]$. Autrement dit, l'ensemble des polynômes annulateurs de **A** s'écrit $m_{\mathbf{A}} \cdot \mathbb{K}[x]$.

Nous pouvons alors reformuler le théorème 7.7 de caractérisation algébrique des matrices diagonalisables de la façon suivante.

7.9 Théorème.— Une matrice **A** de $\mathscr{M}_n(\mathbb{K})$ est diagonalisable dans $\mathscr{M}_n(\mathbb{K})$ si, et seulement si, son polynôme minimal $m_{\mathbf{A}}$ est scindé sur $\mathbb{K}$ et possède toutes ses racines simples.

Preuve. D'après le théorème 7.7, la condition est suffisante. Supposons que **A** soit diagonalisable, d'après le théorème 7.7, il admet un polynôme annulateur g scindé à racines simples. Or d'après la proposition 7.8 le polynôme $m_{\mathbf{A}}$ divise g, donc $m_{\mathbf{A}}$ est scindé. □

7.4.4. Exemple : matrices nilpotentes.— Soit **A** une matrice nilpotente, il existe un entier q tel que $\mathbf{A}^q = 0$. Le polynôme x^q est alors annulateur de **A** et le polynôme minimal de **A** est de la forme x^k avec $1 \leqslant k \leqslant q$. D'après le théorème 7.9, une matrice nilpotente est donc diagonalisable si, et seulement si, elle est nulle.

Exercice 7.— Montrer qu'une matrice triangulaire n'ayant que des 0 sur la diagonale est diagonalisable si, et seulement si, elle est nulle.

FIGURE 7.2.: William Rowan Hamilton (1805 - 1865)

Sir William Rowan Hamilton est un mathématicien, astronome et physicien irlandais. Il contribua au développement de l'optique et de la dynamique, en mathématiques ses travaux portent sur l'algèbre, sa contribution majeure reste la découverte des quaternions. Il découvrent ces nombres en 1843 en essayant d'étendre les nombres complexes à des espaces de dimension supérieure à 2*. L'objectif était de construire des nombres dans l'espace qui ont des propriétés analogues à celles des nombres complexes dans le plan. Hamilton aurait découvert la structure multiplicative des quaternions, en se promenant avec son épouse le long du Royal Canal à Dublin. Une plaque sur le pont indique « Ici, le 16 octobre 1843, alors qu'il se promenait, Sir William Rowan Hamilton découvrit dans un éclair de génie la formule fondamentale sur la multiplication des quaternions* $i = j = k = ijk = -1$ *et la grava sur une pierre du pont. »*

FIGURE 7.3.: Broom Bridge sur le Royal Canal à Dublin

Le **théorème de Cayley-Hamilton** *: toute matrice est racine de son polynôme caractéristique a été montré par Hamilton en 1853 pour la matrice d'une rotation de l'espace de dimension* 3*. Cayley formule le théorème pour toutes les matrices carrées d'ordre* 3 *en 1958. Cependant, ni Hamilton, ni Cayley ne publieront de preuve de ce résultat. Il faudra attendre 1878 pour une première preuve par Ferdinand Georg Frobenius. Nous connaissons aujourd'hui un nombre important de preuves différentes de ce résultat.*

§ 5 Le théorème de Cayley-Hamilton

Le résultat suivant montre que le polynôme caractéristique d'une matrice est annulateur d'une matrice. Ainsi, du polynôme caractéristique on pourra déduire le polynôme minimal.

7.10 Théorème (Théorème de Cayley-Hamilton).— Toute matrice est racine de son polynôme caractéristique.

Preuve. Soit $\mathbf{A}$ une matrice de $\mathscr{M}_n(\mathbb{R})$ ou $\mathscr{M}_n(\mathbb{C})$, montrons que $\mathbf{A}$ est racine de son polynôme caractéristique $p_{\mathbf{A}}$. Le polynôme caractéristique $p_{\mathbf{A}}$ est scindé sur $\mathbb{C}$:

$$p_{\mathbf{A}} = (x-\lambda_1)^{n_1}\dots(x-\lambda_p)^{n_p}. \tag{7.13}$$

La matrice $\mathbf{A}$ est donc trigonalisable dans $\mathscr{M}_n(\mathbb{C})$: il existe une matrice inversible $\mathbf{P}$ de $\mathscr{M}_n(\mathbb{C})$ telle que

$$\mathbf{T} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P},$$

où $\mathbf{T}$ est une matrice triangulaire de la forme

$$\begin{bmatrix} \mathbf{T}_1 & * & \cdots & * \\ 0 & \mathbf{T}_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \cdots & 0 & \mathbf{T}_p \end{bmatrix} \in \mathscr{M}_n(\mathbb{C}), \quad \text{avec} \quad \mathbf{T}_i = \begin{bmatrix} \lambda_i & * & \cdots & * \\ 0 & \lambda_i & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \cdots & 0 & \lambda_i \end{bmatrix} \in \mathscr{M}_{n_i}(\mathbb{C})$$

La matrice $\mathbf{T}_i - \lambda_i \mathbf{1}_{n_i}$ est nilpotente d'indice de nilpotence n_i, soit

$$(\mathbf{T}_i - \lambda_i \mathbf{1}_{n_i})^{n_i} = \mathbf{0}.$$

Par suite,

$$(\mathbf{T} - \lambda_i \mathbf{1}_n)^{n_i} = \begin{bmatrix} (\mathbf{T}_1 - \lambda_i \mathbf{1}_{n_1})^{n_i} & * & \cdots & & * \\ 0 & \ddots & & & \vdots \\ \vdots & & (\mathbf{T}_i - \lambda_i \mathbf{1}_{n_i})^{n_i} & & \vdots \\ \vdots & \ddots & & \ddots & * \\ 0 & \cdots & 0 & & (\mathbf{T}_p - \lambda_i \mathbf{1}_{n_p})^{n_i} \end{bmatrix}$$

Donc

$$(\mathbf{T} - \lambda_i \mathbf{1}_n)^{n_i} = \begin{bmatrix} (\mathbf{T}_1 - \lambda_i \mathbf{1}_{n_1})^{n_i} & * & \cdots & & * \\ 0 & \ddots & & & \vdots \\ \vdots & & \boxed{\mathbf{0}} & & \vdots \\ \vdots & \ddots & & \ddots & * \\ 0 & \cdots & 0 & & (\mathbf{T}_p - \lambda_i \mathbf{1}_{n_p})^{n_i} \end{bmatrix}$$

Le i-ième bloc $\boxed{\mathbf{0}}$ est nul, par conséquent, on a

$$(\mathbf{T} - \lambda_1 \mathbf{1}_n)^{n_1}(\mathbf{T} - \lambda_2 \mathbf{1}_n)^{n_2}\dots(\mathbf{T} - \lambda_p \mathbf{1}_n)^{n_p} = \mathbf{0}.$$

Par ailleurs, vu l'expression (7.13) du polynôme caractéristique de $\mathbf{A}$, on a

$$p_{\mathbf{A}}(\mathbf{A}) = (\mathbf{A} - \lambda_1 \mathbf{1}_n)^{n_1}\dots(\mathbf{A} - \lambda_p \mathbf{1}_n)^{n_p}.$$

30. The equation satisfied by the matrix may be of the form $\mathrm{M}^n=1$; the matrix is in this case said to be periodic of the nth order. The preceding considerations apply to the theory of periodic matrices; thus, for instance, suppose it is required to find a matrix of the order 2, which is periodic of the second order. Writing

$$\mathrm{M}=\left(\begin{array}{cc} a, & b \\ c, & d \end{array}\right),$$

we have

$$\mathrm{M}^2-(a+d)\mathrm{M}+ad-bc=0,$$

and the assumed equation is

$$\mathrm{M}^2-1=0.$$

These equations will be identical if

$$a+d=0,\ ad-bc=-1,$$

that is, these conditions being satisfied, the equation $\mathrm{M}^2-1=0$ required to be satisfied, will be identical with the equation which is always satisfied, and will therefore itself be satisfied. And in like manner the matrix M of the order 2 will satisfy the condition $\mathrm{M}^3-1=0$, or will be periodic of the third order, if only M^3-1 contains as a factor

$$\mathrm{M}^2-(a+d)\mathrm{M}+ad-bc,$$

and so on.

31. But suppose it is required to find a matrix of the order 3,

$$\mathrm{M}=\left(\begin{array}{ccc} a, & b, & c \\ d, & e, & f \\ g, & h, & i \end{array}\right)$$

which shall be periodic of the second order. Writing for shortness

$$\left|\begin{array}{ccc} a-\mathrm{M}, & b & , c \\ d & , e-\mathrm{M}, & f \\ g & , h & , i-\mathrm{M} \end{array}\right|=-(\mathrm{M}^3-\mathrm{A}\mathrm{M}^2+\mathrm{B}\mathrm{M}-\mathrm{C}),$$

the matrix here satisfies

$$\mathrm{M}^3-\mathrm{A}\mathrm{M}^2+\mathrm{B}\mathrm{M}-\mathrm{C}=0,$$

FIGURE 7.4.: Formulation du théorème de Cayley-Hamilton par Cayley en 1858 dans [?].

D'où

$$\begin{aligned}
\mathbf{P}^{-1}p_{\mathbf{A}}(\mathbf{A})\mathbf{P} &= \mathbf{P}^{-1}(\mathbf{A}-\lambda_1\mathbf{1}_n)^{n_1}\ldots(\mathbf{A}-\lambda_p\mathbf{1}_n)^{n_p}\mathbf{P}\\
&= (\mathbf{P}^{-1}\mathbf{A}\mathbf{P}-\lambda_1\mathbf{1}_n)^{n_1}\ldots(\mathbf{P}^{-1}\mathbf{A}\mathbf{P}-\lambda_p\mathbf{1}_n)^{n_p}\\
&= (\mathbf{T}-\lambda_1\mathbf{1}_n)^{n_1}(\mathbf{T}-\lambda_2\mathbf{1}_n)^{n_2}\ldots(\mathbf{T}-\lambda_p\mathbf{1}_n)^{n_p}\\
&= \mathbf{0}.
\end{aligned}$$

Ainsi $p_{\mathbf{A}}(\mathbf{A})=\mathbf{0}$. $\square$

7.5.1. Calcul du polynôme minimal.— Une méthode permettant de déterminer le polynôme minimal d'une matrice $\mathbf{A}$ de $\mathscr{M}_n(\mathbb{K})$, consiste à considérer un polynôme annulateur de $\mathbf{A}$ et à chercher dans l'ensemble de ses diviseurs, le polynôme unitaire annulateur de plus petit degré. D'après le théorème de Cayley-Hamilton, théorème 7.10, un candidat naturel pour le polynôme annulateur est le polynôme caractéristique. Nous allons voir comment mettre en oeuvre cette méthode dans ce cas.

Montrons dans un premier temps que le spectre d'une matrice coïncide avec les racines de son polynôme minimal.

7.11 Proposition.— Soit $\mathbf{A}$ une matrice de $\mathscr{M}_n(\mathbb{K})$. Un scalaire λ est valeur propre de $\mathbf{A}$ si, et seulement si, il est racine du polynôme minimal $m_{\mathbf{A}}$.

Preuve. Le polynôme $p_{\mathbf{A}}$ est annulateur de $\mathbf{A}$, il admet donc le polynôme $m_{\mathbf{A}}$ comme diviseur. Il existe un polynôme g de $\mathbb{K}[x]$ tel que $p_{\mathbf{A}} = g m_{\mathbf{A}}$. Par suite, toute racine du polynôme $m_{\mathbf{A}}$ est racine de $p_{\mathbf{A}}$, donc est valeur propre de $\mathbf{A}$.

Inversement, le polynôme $m_{\mathbf{A}}$ est annulateur de $\mathbf{A}$, donc, d'après la proposition 7.4, toute valeur propre de $\mathbf{A}$ est racine de $m_{\mathbf{A}}$. □

De ce résultat on déduit

7.12 Proposition.— Soit $\mathbf{A}$ une matrice de $\mathscr{M}_n(\mathbb{K})$ dont le polynôme $p_{\mathbf{A}}$ est scindé :

$$p_{\mathbf{A}} = (-1)^n(x-\lambda_1)^{n_1}\dots(x-\lambda_p)^{n_p},$$

avec $\lambda_i \neq \lambda_j$, pour tout $i \neq j$ et $n_1 + \ldots + n_p = n$, alors

$$m_{\mathbf{A}} = (x-\lambda_1)^{k_1}\dots(x-\lambda_p)^{k_p},$$

avec $1 \leqslant k_i \leqslant n_i$.

Preuve. D'après le théorème 7.10, le polynôme $m_{\mathbf{A}}$ divise le polynôme $p_{\mathbf{A}}$ et, d'après la proposition 7.11, les polynômes $m_{\mathbf{A}}$ et $p_{\mathbf{A}}$ possèdent les mêmes racines. □

7.5.2. Exemple.— On considère la matrice suivante de de $\mathscr{M}_8(\mathbb{R})$:

$$\mathbf{A} = \begin{bmatrix} 3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 3 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 5 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 7 \end{bmatrix}.$$

Le polynôme caractéristique de $\mathbf{A}$ est

$$p_{\mathbf{A}} = (x-3)^4(x-5)^3(x-7).$$

La matrice $\mathbf{A}$ est diagonale par blocs :

$$\mathbf{A} = \begin{bmatrix} \mathbf{B} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{C} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{D} \end{bmatrix},$$

avec

$$\mathbf{B} = \begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 3 & 1 & 0 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 3 \end{bmatrix}, \qquad \mathbf{C} = \begin{bmatrix} 5 & 1 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 5 \end{bmatrix}, \qquad \mathbf{D} = \begin{bmatrix} 7 \end{bmatrix}.$$

On note que $(\mathbf{B}-3\mathbf{1}_4)^3 = \mathbf{0}$, $(\mathbf{C}-5\mathbf{1}_3)^2 = \mathbf{0}$ et $\mathbf{D}-7\mathbf{1}_1 = \mathbf{0}$. On obtient alors,

$$\begin{aligned}
&(\mathbf{A}-3\mathbf{1}_8)^3(\mathbf{A}-5\mathbf{1}_8)^2(\mathbf{A}-7\mathbf{1}_8) \\
&= \begin{bmatrix} (\mathbf{B}-3\mathbf{1}_4)^3 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & (\mathbf{C}-3\mathbf{1}_3)^3 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & (\mathbf{D}-3\mathbf{1}_1)^3 \end{bmatrix} \begin{bmatrix} (\mathbf{B}-5\mathbf{1}_4)^2 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & (\mathbf{C}-5\mathbf{1}_3)^2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & (\mathbf{D}-5\mathbf{1}_1)^2 \end{bmatrix} (\mathbf{A}-7\mathbf{1}_8) \\
&= \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & (\mathbf{C}-3\mathbf{1}_3)^3 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & (\mathbf{D}-3\mathbf{1}_1)^3 \end{bmatrix} \begin{bmatrix} (\mathbf{B}-5\mathbf{1}_4)^2 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & (\mathbf{D}-5\mathbf{1}_1)^2 \end{bmatrix} \begin{bmatrix} \mathbf{B}-7\mathbf{1}_4 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{C}-7\mathbf{1}_3 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & (\mathbf{D}-5\mathbf{1}_1)^5 \end{bmatrix} \begin{bmatrix} \mathbf{B}-7\mathbf{1}_4 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{C}-7\mathbf{1}_3 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}
\end{aligned}$$

Or 3, 2 et 1 sont les indices de nilpotence des matrices $\mathbf{A}-3\mathbf{1}_8$, $\mathbf{A}-5\mathbf{1}_8$ et $\mathbf{A}-7\mathbf{1}_8$ respectivement. Par suite le polynôme minimal de $\mathbf{A}$ est

$$m_{\mathbf{A}} = (x-3)^3(x-5)^3(x-7).$$

7.5.3. Exemple.— Le polynôme minimal de la rotation d'angle θ du plan vectoriel, représentée dans la base canonique par la matrice

$$\mathbf{R}_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$$

est $p_{\mathbf{R}_\theta} = x^2 - 2\cos\theta x + 1$.

7.5.4. Exemples.— Soit

$$\mathbf{A} = \begin{bmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{bmatrix}.$$

Alors $p_{\mathbf{A}} = -(x-1)(x+2)^2$. Les deux valeurs possibles pour le polynôme minimal de $\mathbf{A}$ sont donc soit $(x-1)(x+2)^2$, soit $(x-1)(x+2)$. Or

$$(\mathbf{A}-\mathbf{1}_3)(\mathbf{A}+2\mathbf{1}_3) = \begin{bmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Par suite $m_{\mathbf{A}} = (x-1)(x+2)$ est le polynôme minimal. On en déduit que la matrice $\mathbf{A}$ est diagonalisable et que $\mathrm{Sp}(\mathbf{A}) = \{1,-2\}$, où -2 est valeur propre double.

Soit

$$\mathbf{B} = \begin{bmatrix} 3 & 2 & -2 \\ -1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}.$$

On a $p_{\mathbf{B}} = -(x-1)^3$. Donc $m_{\mathbf{A}} = (x-1)^k$, où $k = 1,2$ ou 3. Par ailleurs, $\mathbf{A}$ est diagonalisable si, et seulement si, $m_{\mathbf{A}} = x-1$. Or $m_{\mathbf{A}}(\mathbf{A}) = \mathbf{A}-\mathbf{1}_3 \neq \mathbf{0}$, donc $\mathbf{A}$ n'est pas diagonalisable.

Exercice 8.— Construire deux matrices qui ont même polynôme caractéristique et même polynôme minimal et qui ne sont pas semblables.

7.5.5. Exemple : calcul de l'inverse d'une matrice admettant deux valeurs propres.— Soit $\mathbf{A}$ une matrice de $\mathscr{M}_n(\mathbb{K})$. On suppose que son polynôme minimal est scindé, de degré 2 :

$$m_{\mathbf{A}} = (x-\lambda_1)(x-\lambda_2).$$

On suppose de plus que les deux valeurs propres de $\mathbf{A}$ sont non nulles. On peut alors en déduire les puissances de $\mathbf{A}$ et son inverse. La matrice $\mathbf{A}$ est inversible, car λ_1 et λ_2 sont non nuls. On a

$$m_{\mathbf{A}} = (x-\lambda_1)(x-\lambda_2) = x(x-(\lambda_1+\lambda_2)) + \lambda_1\lambda_2.$$

Par suite, la matrice $\mathbf{A}$ vérifie

$$\mathbf{A}(\mathbf{A}-(\lambda_1+\lambda_2)\mathbf{1}_n) = -\lambda_1\lambda_2\mathbf{1}_n.$$

On en déduit l'inverse de $\mathbf{A}$:

$$\mathbf{A}^{-1} = -\frac{1}{\lambda_1\lambda_2}(\mathbf{A}-(\lambda_1+\lambda_2)\mathbf{1}_n). \tag{7.14}$$

Par exemple, dans l'exemple 6.1.8, nous avons montré que la matrice

$$\mathbf{A} = \begin{bmatrix} n & 1 & \cdots & 1 \\ 1 & n & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 1 & \cdots & 1 & n \end{bmatrix},$$

est diagonalisable et qu'elle possède deux valeurs propres distinctes $n-1$ et $2n-1$ (on suppose que $n > 1$). Son polynôme minimal est donc $m_{\mathbf{A}} = (x-(n-1))(x-(2n-1))$.

D'après ce qui précéde, on en déduit que

$$\mathbf{A}^{-1} = \frac{1}{(1-n)(2n-1)}(\mathbf{A}-(3n-2)\mathbf{1}_n).$$

Soit

$$\mathbf{A} = \frac{1}{(1-n)(2n-1)}\begin{bmatrix} 2(1-n) & 1 & \cdots & 1 \\ 1 & 2(1-n) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 1 & \cdots & 1 & 2(1-n) \end{bmatrix}.$$

Exercice 9.— Utiliser la formule (7.14) pour calculer l'inverse de la matrice $\mathbf{R}_\theta$ de l'exemple 5.3.4.

7.5.6. Exemple.— Plus généralement, on peut utiliser le théorème de Cayley-Hamilton pour calculer l'inverse d'une matrice. Soit $\mathbf{A}$ une matrice inversible de $\mathscr{M}_n(\mathbb{K})$. Son polynôme caractéristique est de degré n, il s'écrit sous la forme

$$p_\mathbf{A} = (-1)^n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0.$$

D'après le théorème de Cayley-Hamilton, $\mathbf{A}$ est racine de $p_\mathbf{A}$ et on a :

$$(-1)^n\mathbf{A}^n + a_{n-1}\mathbf{A}^{n-1} + \ldots + a_1\mathbf{A} + a_0\mathbf{1}_n = \mathbf{0}.$$

D'où

$$\mathbf{A}\left[(-1)^n\mathbf{A}^{n-1} + a_{n-1}\mathbf{A}^{n-2} + \ldots + a_1\mathbf{1}_n\right] = -a_0\mathbf{1}_n.$$

Le coefficient a_0 est non nul, car $a_0 = p_\mathbf{A}(0) = \det\mathbf{A}$ et $\mathbf{A}$ est inversible. Ainsi,

$$\mathbf{A}^{-1} = \frac{-1}{a_0}\left[(-1)^n\mathbf{A}^{n-1} + a_{n-1}\mathbf{A}^{n-2} + \ldots + a_1\mathbf{1}_n\right].$$

Exercice 10 (Une autre preuve du lemme des noyaux).— L'objectif de cet exercice est d'établir une autre preuve du lemme de décomposition en noyaux. Soient $\mathbf{A}$ une matrice de $\mathscr{M}_n(\mathbb{K})$ et g un polynôme de $\mathbb{K}[x]$, annulateur de $\mathbf{A}$ et tel que

$$g = g_1 \ldots g_p,$$

où les g_i sont des polynômes premiers entre eux deux à deux. L'objectif est de montrer par récurrence sur l'entier p, que

$$\mathbb{K}^n = \operatorname{Ker} g_1(\mathbf{A}) \oplus \cdots \oplus \operatorname{Ker} g_p(\mathbf{A}).$$

1. Soit g_1 un polynôme de $\mathbb{K}[x]$ tel que $g_1(\mathbf{A}) = \mathbf{0}$. Montrer que $\mathbb{K}^n = \operatorname{Ker} g(\mathbf{A})$.

2. Soient g_1 et g_2 deux polynômes de $\mathbb{K}[x]$ premiers entre eux, tels que $g_1(\mathbf{A})g_2(\mathbf{A}) = \mathbf{0}$. Montrer que $\mathbb{K}^n = \operatorname{Ker} g_1(\mathbf{A}) \oplus \operatorname{Ker} g_2(\mathbf{A})$.

3. Soient $g_1, \ldots, g_p, g_{p+1}$ des polynômes de $\mathbb{K}[x]$ premiers entre eux deux à deux. tels que

$$\mathbb{K}^n = \operatorname{Ker} g_1(\mathbf{A}) \oplus \cdots \oplus \operatorname{Ker} g_p(\mathbf{A}).$$

En utilisant le fait que les polynômes $g_1 \ldots g_p$ et g_{p+1} sont premiers entre eux, montrer que

$$\mathbb{K}^n = \text{Ker } g_1(\mathbf{A}) \oplus \cdots \oplus \text{Ker } g_p(\mathbf{A}) \oplus \text{Ker } g_{p+1}(\mathbf{A}).$$

§ 6 Le cas des endomorphismes

Les résultats exposés dans ce chapitre s'énoncent à l'identique pour les endomorphismes sur un espace vectoriel de dimension finie. Dans cette section, nous reformulons quelques résultats dans ce contexte. C'est aussi l'occasion de présenter une autre preuve du théorème de Cayley-Hamilton, par voie de trigonalisation.

Dans toute cette section, E désigne un $\mathbb{K}$-espace vectoriel de dimension finie $n > 0$ et u un endomorphisme de de E.

7.6.1. Polynômes d'endomorphismes.— Étant donné un polynôme de $\mathbb{K}[x]$ donné par

$$f = a_m x^m + a_{m-1} x^{m-1} + \cdots + a_1 x + a_0.$$

On note $f(u)$ le *polynôme d'endomorphismes* de E défini par :

$$f(u) = a_m u^m + a_{m-1} u^{m-1} + \cdots + a_1 u + a_0 \mathrm{id}_E,$$

où u^n désigne la composée de l'endomorphisme u avec lui-même n fois :

$$u^n = u \circ \underset{n\text{-fois}}{\ldots} \circ u,$$

et en convenant que u^0 est l'endomorphisme identité de E.

7.6.2. Polynôme annulateur d'un endomorphisme.— Un polynôme non nul g de $\mathbb{K}[x]$ est dit *annulateur* de l'endomorphisme u si $g(u)$ est l'endomorphisme nul. On dit aussi que u est *racine* du polynôme g.

De la même façon que pour les matrices, proposition 7.3, on montre que tout endomorphisme de E possède un polynôme annulateur. Cependant, on notera que cette propriété n'est pas vraie en général en dimension infinie. Prenons par exemple l'espace vectoriel réel E formé des suites réelles nulles à partir d'un certain rang. L'espace E admet pour base $(\mathbf{e}_k)_{k\in\mathbb{N}}$, où $\mathbf{e}_k$ désigne la suite $(x_n)_{n\in\mathbb{N}}$ avec

$$x_n = \begin{cases} 0 & \text{si } k \neq n, \\ 1 & \text{si } k = n. \end{cases}$$

Soit u l'endomorphisme de E défini par

$$u(\mathbf{e}_k) = \mathbf{e}_{k+1}, \quad \text{pour tout } k \in \mathbb{N},$$

On montre que le polynôme nul est le seul polynôme annulateur de u.

7.6.3. Caractérisation algébrique des endomorphismes diagonalisables.— De façon analogue au cas matriciel, si g est un polynôme annulateur de l'endomorphisme u, alors toute valeur propre de u est racine du polynôme g. On montre que l'endomorphisme u est diagonalisable sur $\mathbb{K}$ si, et seulement si, il existe un polynôme de $\mathbb{K}[x]$ annulateur de u scindé et ne possédant que des racines simples.

Le polynôme minimal d'un endomorphisme se définit comme celui d'une matrice. On obtient la même caractérisation pour les endomorphismes diagonalisable : l'endomorphisme u est diagonalisable si, et seulement si, son polynôme minimal est scindé et possède toutes ses racines simples.

7.6.4. Théorème de Cayley-Hamilton.— Nous avons montré en 7.10 que toute matrice est racine de son polynôme caractéristique. Il en est de même pour un endomorphisme sur un espace de dimension finie. Nous donnons ici une autre preuve du théorème de Cayley-Hamilton.

7.13 Théorème.— Tout endomorphisme est racine de son polynôme caractéristique.

Preuve. On suppose que $\mathbb{K}$ est le corps $\mathbb{R}$ ou le corps $\mathbb{C}$. Soit u un endomorphisme du $\mathbb{K}$-espace vectoriel E. D'après le théorème de D'Alembert-Gauss, théorème 0.5, le polynôme caractéristique p_u est scindé sur $\mathbb{C}$, par suite l'endomorphisme u est trigonalisable sur $\mathbb{C}$. Il existe donc une base $\mathscr{B} = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ de E telle que la matrice dans $\mathscr{M}_n(\mathbb{C})$ de l'endomorphisme u exprimé dans la base $\mathscr{B}$ soit triangulaire supérieure :

$$[u]_{\mathscr{B}} = \begin{bmatrix} \lambda_1 & * & \ldots & * \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix}.$$

On a

$$p_u = (\lambda_1 - x)\ldots(\lambda_n - x).$$

Montrons que

$$(\lambda_1 \mathrm{id}_E - u) \circ \cdots \circ (\lambda_n \mathrm{id}_E - u) = 0. \tag{7.15}$$

Posons, pour tout $i \in [\![1, n]\!]$,

$$v_i = (\lambda_1 \mathrm{id}_E - u) \circ \cdots \circ (\lambda_i \mathrm{id}_E - u).$$

Nous établissons l'équation (7.15) en montrant par récurrence sur i que, pour tout $i \in [\![1, n]\!]$, on a

$$v_i(\mathbf{e}_1) = v_i(\mathbf{e}_2) = \ldots v_i(\mathbf{e}_i) = \mathbf{0}.$$

Pour $i = 1$, on a

$$v_1(\mathbf{e}_1) = (\lambda_1 \mathbf{e}_1 - u(\mathbf{e}_1)) = (\lambda_1 \mathbf{e}_1 - \lambda_1 \mathbf{e}_1) = \mathbf{0}.$$

Supposons que $v_{i-1}(\mathbf{e}_1) = \cdots = v_{i-1}(\mathbf{e}_{i-1}) = \mathbf{0}$ et montrons que $v_i(\mathbf{e}_1) = \ldots = v_i(\mathbf{e}_i) = \mathbf{0}$. On a

$$v_i = v_{i-1} \circ (\lambda_i \mathrm{id}_E - u) = (\lambda_i \mathrm{id}_E - u) \circ v_{i-1}.$$

Donc

$$v_i(\mathbf{e}_1) = \cdots = v_i(\mathbf{e}_{i-1}) = \mathbf{0}, \quad \text{et} \quad v_i(\mathbf{e}_i) = v_{i-1}(\lambda_i \mathbf{e}_i - u(\mathbf{e}_i)).$$

Or $u(\mathbf{e}_i)$ peut s'écrire :

$$u(\mathbf{e}_i) = a_1\mathbf{e}_1 + a_2\mathbf{e}_2 + \cdots + a_{i-1}\mathbf{e}_{i-1} + \lambda_i \mathbf{e}_i,$$

car $[u]_{\mathscr{B}}$ est triangulaire. Donc

$$v_i(\mathbf{e}_i) = v_{i-1}(\lambda_i \mathbf{e}_i - a_1\mathbf{e}_1 - \cdots - a_{i-1}\mathbf{e}_{i-1} - \lambda_i \mathbf{e}_i) = \mathbf{0}$$

Par suite $v_i = \mathbf{0}$, pour tout $i \in [\![1,n]\!]$, or $v_n = p_u$, d'où $v_n = p_u(u) = 0$. $\square$

7.6.5. Exemple.— Un projecteur de E, cf., 1.4.11, est un endomorphisme p de E tel que $p^2 = p$. Autrement dit, un endomorphisme est un projecteur si, et seulement si, il est racine du polynôme $x^2 - x$. Si p n'est pas l'identité de E, son polynôme minimal est donc $m_p = x(x-1)$. Ainsi tout projecteur de E est diagonalisable et on a :

$$E = \mathrm{Ker}\, p \oplus \mathrm{Ker}\, (p - \mathrm{id}_E).$$

Dans l'exemple 7.1.1, nous avions établi cette relation pour une projection sur un plan de $\mathbb{R}^3$ parallèlement à une droite, sans faire appel au polynôme minimal. On comprend ici, l'intérêts du polynôme minimal.

Exercice 11.— Soient u un endomorphisme diagonalisable d'un espace vectoriel E et F un sous-espace vectoriel de E stable par u. Montrer que la restriction de u au sous-espace F est un endomorphisme diagonalisable de F.

Exercice 12.— Soit $\mathbb{R}_n[x]$ le $\mathbb{R}$-espace vectoriel formé des polynômes de degré inférieur ou égal à n. Soit $u : \mathbb{R}_n[x] \to \mathbb{R}_n[x]$ l'application qui à un polynôme p associe le reste de la division euclidienne de p par $x^2 - 1$.

1. Montrer que u est linéaire.
2. Calculer u^2 et en déduire que l'endomorphisme u est diagonalisable.

Exercice 13.— Soit n un entier supérieur ou égal à 2. On considère l'application linéaire

$$\mathrm{trace} : \mathscr{M}_n(\mathbb{R}) \longrightarrow \mathbb{R}$$

qui à toute matrice associe sa trace, cf. 2.4.1.

1. Déterminer l'image de l'application trace et la dimension de son noyau.
2. Montrer qu'on a une somme directe :

$$\mathscr{M}_n(\mathbb{R}) = \mathrm{Ker}\,(\mathrm{trace}) \oplus \mathrm{Vect}(\mathbf{1}_n).$$

Soit u l'application de $\mathscr{M}_n(\mathbb{R})$ dans $\mathscr{M}_n(\mathbb{R})$ définie par

$$u(\mathbf{A}) = \mathbf{A} + \operatorname{trace}(\mathbf{A})\mathbf{1}_n.$$

3. Montrer que u est un endomorphisme de $\mathscr{M}_n(\mathbb{R})$.
4. L'endomorphisme u est-il diagonalisable ? Est-il inversible ?

Exercice 14.— Soient E un $\mathbb{K}$-espace vectoriel de dimension finie n et u un endomorphisme de E. On considère deux sous-espaces vectoriel F et G de E stables par u tels que

$$E = F \oplus G.$$

Si m_F et m_G désignent les polynômes minimaux des restrictions de u à F et G respectivement, montrer que le polynôme minimal m_u de u est donné par

$$m_u = \operatorname{ppcm}(m_F, m_G).$$

—

§ 7 Exercices

Exercice 15.— Déterminer le polynôme minimal des matrices réelles suivantes, où a, b et c sont trois réels distincts deux à deux :

$$\mathbf{A}_1 = \begin{bmatrix} a & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & a \end{bmatrix}, \quad \mathbf{A}_2 = \begin{bmatrix} a & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & b \end{bmatrix}, \quad \mathbf{A}_3 = \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}, \quad \mathbf{A}_4 = \begin{bmatrix} a & 1 & 0 \\ 0 & a & 1 \\ 0 & 0 & a \end{bmatrix},$$

$$\mathbf{A}_5 = \begin{bmatrix} a & 1 & 0 \\ 0 & a & 0 \\ 0 & 0 & a \end{bmatrix}, \quad \mathbf{A}_6 = \begin{bmatrix} a & 1 & 0 \\ 0 & a & 0 \\ 0 & 0 & b \end{bmatrix}, \quad \mathbf{A}_7 = \begin{bmatrix} a & 1 & 0 \\ 0 & b & 0 \\ 0 & 0 & b \end{bmatrix},$$

$$\mathbf{B}_1 = \begin{bmatrix} a & 1 & 0 & 0 \\ 0 & a & 1 & 0 \\ 0 & 0 & a & 1 \\ 0 & 0 & 0 & a \end{bmatrix}, \quad \mathbf{B}_2 = \begin{bmatrix} a & 1 & 0 & 0 \\ 0 & a & 1 & 0 \\ 0 & 0 & a & 0 \\ 0 & 0 & 0 & a \end{bmatrix}, \quad \mathbf{B}_3 = \begin{bmatrix} a & 1 & 0 & 0 \\ 0 & a & 0 & 0 \\ 0 & 0 & a & 1 \\ 0 & 0 & 0 & a \end{bmatrix},$$

$$\mathbf{B}_4 = \begin{bmatrix} a & 1 & 0 & 0 \\ 0 & a & 0 & 0 \\ 0 & 0 & b & 1 \\ 0 & 0 & 0 & b \end{bmatrix}, \quad \mathbf{B}_5 = \begin{bmatrix} a & 1 & 0 & 0 \\ 0 & a & 0 & 0 \\ 0 & 0 & b & 0 \\ 0 & 0 & 0 & b \end{bmatrix}.$$

Exercice 16.— Déterminer le polynôme minimal des matrices suivantes de $\mathscr{M}_3(\mathbb{R})$:

$$\mathbf{A}_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{A}_2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{A}_3 = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{A}_4 = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix},$$

$$\mathbf{A}_5 = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{A}_6 = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{A}_7 = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}, \quad \mathbf{A}_8 = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}.$$

Exercice 17.— Soit $\mathbf{A}$ une matrice nilpotente de $\mathscr{M}_n(\mathbb{C})$.

1. Sans utiliser le polynôme minimal, montrer que le polynôme caractéristique de $\mathbf{A}$ est

$$p_{\mathbf{A}} = (-1)^n x^n.$$

2. Comment déterminer le polynôme caractéristique $p_{\mathbf{A}}$ en utilisant le polynôme minimal ?

3. Par récurrence, montrer que $\mathbf{A}$ est semblable à une matrice triangulaire supérieure avec des 0 sur la diagonale.

4. Inversement, montrer que toute matrice triangulaire de $\mathscr{M}_n(\mathbb{K})$ avec des 0 sur la diagonale est nilpotente d'indice de nilpotence $p \leqslant n$.

Exercice 18.— Trouver une condition nécessaire et suffisante sur les réels a,b,c,d,e,f pour que les matrices suivantes soient diagonalisables dans $\mathscr{M}_4(\mathbb{R})$.

$$\mathbf{A} = \begin{bmatrix} 1 & a & b & c \\ 0 & 1 & d & e \\ 0 & 0 & 1 & f \\ 0 & 0 & 0 & 1 \end{bmatrix}, \qquad \mathbf{B} = \begin{bmatrix} 0 & a & b & c \\ 0 & 1 & d & e \\ 0 & 0 & 1 & f \\ 0 & 0 & 0 & 1 \end{bmatrix}, \qquad \mathbf{C} = \begin{bmatrix} 0 & a & b & c \\ 0 & 0 & d & e \\ 0 & 0 & 1 & f \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Exercice 19.— Trouver une condition nécessaire et suffisante sur les réels a et b pour que la matrice suivante soit diagonalisables dans $\mathscr{M}_4(\mathbb{R})$.

$$\mathbf{A} = \begin{bmatrix} 1 & -a & -a & 1 \\ 1-b & a & a-1 & -b \\ b & -a & 1-a & 1+b \\ 0 & a & a & 0 \end{bmatrix}.$$

Exercice 20.— Soit $\mathbf{J}$ la matrice de $\mathscr{M}_n(\mathbb{C})$ définie par

$$\begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & & & \ddots & 1 \\ 1 & 0 & & \dots & 0 \end{bmatrix}$$

1. Calculer $\mathbf{J}^p$ pour tout entier $p \in [\![1,n]\!]$.

2. En déduire que $\mathbf{J}$ est diagonalisable.
3. Montrer que les matrices $\mathbf{1}_n, \mathbf{J}, \ldots, \mathbf{J}^{n-1}$ sont linéairement indépendantes.
4. Déterminer le polynôme minimal de $\mathbf{J}$.
5. Calculer les valeurs propres de $\mathbf{J}$.
6. Diagonaliser $\mathbf{J}$ en exhibant une matrice de passage.

Exercice 21.— Soit $\mathbf{A}$ la *matrice circulante* complexe suivante :

$$\begin{bmatrix} a_1 & a_2 & \ldots & a_n \\ a_n & a_1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_2 \\ a_2 & \ldots & a_n & a_1 \end{bmatrix}$$

1. Exprimer $\mathbf{A}$ comme un polynôme en la matrice $\mathbf{J}$, définie dans l'exercice 20
2. Montrer que pour tout polynôme g de $\mathbb{C}[x]$, $g(\mathbf{J})$ est diagonalisable et que

$$\mathrm{Sp}(\mathrm{g}(\mathbf{J})) = \{g(\lambda) \mid \lambda \in \mathrm{Sp}(\mathbf{J})\}.$$

3. En déduire que la matrice $\mathbf{A}$ est diagonalisable et calculer ses valeurs propres.
4. Calculer le déterminant de $\mathbf{A}$.

Exercice 22.— Soit $\mathbf{A}$ une matrice inversible de $\mathscr{M}_n(\mathbb{C})$.

1. Montrer que 0 ne peut pas être valeur propre de $\mathbf{A}$.
2. En déduire que $\mathbf{A}^{-1}$ est un polynôme en $\mathbf{A}$. [On pourra utiliser le fait que le polynôme x ne divise pas le polynôme caractéristique de $\mathbf{A}$.]

Exercice 23.— Soient $\mathbf{A}$ une matrice de $\mathscr{M}_n(\mathbb{K})$ et p un polynôme à coefficients dans $\mathbb{K}$.

1. Montrer que si le polynôme p est premier avec le polynôme minimal $m_{\mathbf{A}}$ de $\mathbf{A}$, alors la matrice $p(\mathbf{A})$ est inversible.
2. Inversement, montrer que si la matrice $p(\mathbf{A})$ est inversible, alors les polynômes p et $m_{\mathbf{A}}$ sont premiers entre eux.

Exercice 24.— Résoudre dans $\mathscr{M}_n(\mathbb{R})$ l'équation $\mathbf{X}^3 = \mathbf{X}$.

Exercice 25.— L'objectif est de résoudre dans $\mathscr{M}_3(\mathbb{R})$ l'équation

$$\mathbf{X}^3 + \mathbf{X} = \mathbf{0}.$$

Soit $\mathbf{A}$ une matrice non nulle solution de l'équation précédente.

1. Montrer que

$$\mathbb{R}^3 = \mathrm{Ker}\,\mathbf{A} \oplus \mathrm{Ker}\,(\mathbf{A}^2 + \mathbf{1}_3).$$

2. Déterminer le polynôme minimal de $\mathbf{A}$.
3. Montrer que si $\mathbf{x}$ n'appartient pas à $\mathrm{Ker}\,\mathbf{A}$, alors $(\mathbf{x}, \mathbf{A}\mathbf{x})$ est libre.

4. Montrer que Ker $\mathbf{A}$ est de dimension 1. En déduire que $\mathbf{A}$ est semblable à la matrice

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}.$$

Exercice 26.— Soient E un $\mathbb{K}$-espace vectoriel de dimension n et p le polynôme de $\mathbb{K}[x]$ défini par

$$P = x^n - a_{n-1}x^{n-1} - \ldots - a_1 x - a_0.$$

La matrice compagnon du polynôme p est la matrice

$$\mathbf{A} = \begin{bmatrix} 0 & \cdots & \cdots & 0 & a_0 \\ 1 & 0 & \cdots & 0 & a_1 \\ 0 & 1 & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & a_{n-2} \\ 0 & \cdots & 0 & 1 & a_{n-1} \end{bmatrix}.$$

On note u l'endomorphisme de E représenté par la matrice $\mathbf{A}$ dans une base $\mathscr{B} = (e_1, \ldots, e_n)$ de E fixée.

1. Montrer que le polynôme caractéristique de u est $p_u = (-1)^n p$.
2. Sans utiliser le théorème de Cayley-Hamilton, montrer que le polynôme p est annulateur de u.
3. En déduire que p est le polynôme minimal de u.

Exercice 27.— Soient v un endomorphisme de E et $\mathbf{x}$ un vecteur non nul de E. Soit p le plus grand entier tel que la famille $\mathscr{B}_{\mathbf{x}} = (\mathbf{x}, v(\mathbf{x}), \ldots, v^p(\mathbf{x}))$ soit libre.

1. Montrer que le sous-espace

$$E_{\mathbf{x}} = \mathrm{Vect}(\mathbf{x}, v(\mathbf{x}), \ldots, v^p(\mathbf{x})),$$

est stable par v.
2. Montrer que la matrice dans la base $\mathscr{B}_{\mathbf{x}}$ de la restriction de l'endomorphisme v au sous-espace $E_{\mathbf{x}}$ est une matrice compagnon.
3. Écrire le polynôme associé à cette matrice compagnon.
4. En déduire que le polynôme caractéristique de v vérifie $p_v(\mathbf{x}) = \mathbf{0}$.
5. En déduire le théorème de Cayley-Hamilton.

Exercice 28.— Soit E un $\mathbb{K}$-espace vectoriel de dimension finie n. Un endomorphisme u de E est dit cyclique s'il existe un vecteur $\mathbf{x}$ de E tel que la famille

$$\mathscr{B}_{\mathbf{x}} = (\mathbf{x}, u(\mathbf{x}), \ldots, u^{n-1}(\mathbf{x}))$$

soit une base de E.

1. Écrire la matrice de u dans la base $\mathscr{B}_{\mathbf{x}}$ est une matrice compagnon.
2. Montrer qu'un endomorphisme cyclique possède une unique matrice compagon.
3. Montrer qu'un endomorphisme cyclique de E est diagonalisable si et seulement s'il possède n valeurs propres distinctes.

Exercice 29.— Soient $\mathbb{K}$ un corps commutatif et E un $\mathbb{K}$-espace vectoriel de dimension finie n. On considère le $\mathbb{K}$-espace vectoriel $\mathscr{L}(E)$ formé des endomorphismes de E. Pour tout endomorphisme u de E, on définit l'application

$$\begin{aligned} D_u : \mathscr{L}(E) &\longrightarrow \mathscr{L}(E) \\ v &\longmapsto u \circ v. \end{aligned}$$

1. Montrer que D_u est un endomorphisme de $\mathscr{L}(E)$.
2. Montrer que, pour tout entier naturel n et tout endomorphisme $v \in \mathscr{L}(E)$, on a

$$(D_u)^n(v) = u^n \circ v.$$

En déduire que, pour tout polynôme f de $\mathbb{K}[x]$, on a

$$f(D_u) = D_{f(u)}.$$

3. Montrer que u est diagonalisable si, et seulement si, D_u est diagonalisable.
4. Soient $(\mathbf{e}_i)_{1\leqslant i\leqslant n}$ une base de E et u un endomorphisme de E dont la matrice dans la base $(\mathbf{e}_i)_{1\leqslant i\leqslant n}$ est notée M. Soit $(\mathbf{e}_{i,j})_{1\leqslant i,j\leqslant n}$ une base de $\mathscr{L}(E)$ définie par

$$\mathbf{e}_{i,j}(e_k) = \delta_{j,k}\mathbf{e}_i,$$

où $\delta_{j,k} = 1$ si $j = k$ et $\delta_{j,k} = 0$ si $j \neq k$. Montrer que la matrice de D_u dans la base

$$(\mathbf{e}_{1,1}, \ldots, \mathbf{e}_{n,1}, \mathbf{e}_{1,2}, \ldots, \mathbf{e}_{n,2}, \ldots, \mathbf{e}_{1,n}, \ldots, \mathbf{e}_{n,n}),$$

s'écrit

$$\begin{bmatrix} \mathbf{M} & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{0} & \mathbf{M} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{M} \end{bmatrix}.$$

Exercice 30.— Dans la suite, on suppose que E est un $\mathbb{R}$-espace vectoriel de dimension 2. On considère la base $\mathscr{B}$ de $\mathscr{L}(E)$, formée des endomorphismes représentés dans la base canonique par les matrices E_i suivantes :

$$\mathbf{E}_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad \mathbf{E}_2 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad \mathbf{E}_3 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad \mathbf{E}_4 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

Soit u l'endomorphisme de E représenté dans la base canonique par la matrice

$$\begin{bmatrix} 1 & 4 \\ 1 & 1 \end{bmatrix}.$$

1. Montrer que u est diagonalisable.
2. Écrire la matrice de D_u dans la base $\mathscr{B}$.
3. Diagonaliser D_u.

Exercice $^{(\mathsf{Sage})}$ **31.— 1.** Calculer le polynôme caractéristique et le polynôme minimal de la matrice suivante :

$$\mathbf{M} = \begin{bmatrix} -1 & 1 & 3 \\ -2 & 2 & 2 \\ -2 & 1 & 4 \end{bmatrix}$$

2. La matrice **M** est-elle diagonalisable ?
3. Déterminer le spectre de **M** ainsi qu'une base pour chaque sous-espace propre.

Exercice $^{(\mathsf{Sage})}$ **32.—** Avec le système Sage, dire si les matrices suivantes sont diagonalisables sur $\mathbb{Q}$, sur $\mathbb{R}$, sur $\mathbb{C}$?

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 2 \\ 1 & 1 & -1 & 1 \\ 1 & 2 & 0 & -1 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 1 & 1 & 2 & 3 \\ 1 & 2 & 3 & 2 \\ 2 & 3 & 3 & 4 \\ 3 & 2 & 4 & 4 \end{bmatrix},$$

$$\mathbf{C} = \begin{bmatrix} 4 & 2 & 4 \\ 0 & 0 & -1 \\ -3 & -2 & -3 \end{bmatrix}, \quad \mathbf{D} = \begin{bmatrix} 0 & -2 & 0 \\ 1 & 0 & -1 \\ 0 & 2 & 0 \end{bmatrix}.$$

Pour avoir la liste des valeurs propres dans $\mathbb{C}$ d'une matrice, on pourra calculer son polynôme caractéristique `p(x)`, puis les racines de ce polynôme. Attention aux erreurs d'approximation numérique. Pour afficher ces racines avec une approximation numérique avec 3 chiffres décimaux, on pourra utiliser le code suivant (cf. manuel en ligne de `solve`) :

```
valspres = solve([p(x) == 0], x, solution_dict = True)
for vp in valspres:
    print vp[x].n(digits=3)
```

Exercice $^{(\mathsf{Sage})}$ **33.—** On considère la matrice réelle suivante

$$\mathbf{A} = \begin{bmatrix} 1 & a^2 & a^2 & a \\ 1 & 1 & a & 1 \\ 1 & a & 1 & 1 \\ a & a^2 & a^2 & 1 \end{bmatrix}.$$

1. Calculer le polynôme caractéristique de **A**.
2. Pour quelles valeurs de a, la matrice **A** est-elle diagonalisable ?
3. Diagonaliser **A** en donnant une matrice de passage.

Exercice $^{(\mathsf{Sage})}$ **34 (Matrices compagnons).—** La procédure `CompanionMatrix` retourne la matrice compagnon du polynôme donné en argument :

```
> P := x^4 + x^3 + x^2 + x:
  CompanionMatrix(P,x);
```

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 \end{bmatrix}$$

Un des intérêts de la procédure `CompanionMatrix` est de construire des exemples de matrices dont on connaît par avance le polynôme caractéristique. On pourra aussi les procédures `CharacteristicMatrix`, `JordanBlockMatrix`.

1. Construire une matrice ayant pour polynôme caractéristique

$$p = (x^3-1)(x^2+x+1).$$

2. Cette matrice est-elle diagonalisable sur $\mathbb{R}$ ou sur $\mathbb{C}$?
3. Construire une matrice compagnon de $\mathscr{M}_3(\mathbb{R})$ diagonalisable et une matrice compagnon de $\mathscr{M}_3(\mathbb{R})$ non diagonalisable.

Exercice (Sage) **35 (Polynôme caractéristique et inverse).—** Soit $\mathbf{A}$ une matrice de $\mathscr{M}_n(\mathbb{K})$.

1. Montrer que si 0 n'est pas valeur propre de $\mathbf{A}$, la matrice $\mathbf{A}$ est inversible.
2. En utilisant le polynôme caractéristique de $\mathbf{A}$. Exprimer l'inverse de $\mathbf{A}$ comme un polynôme en $\mathbf{A}$.
3. Écrire un fonction qui prend en argument une matrice inversible et retourne son inverse en utilisant l'expression précédente. On pourra utiliser la méthode `shift(n)` qui multiplie un polynôme par x^n :

```
sage: x = PolynomialRing(QQ,'x').gen()
...   p = 3*x^3 + 2*x^2 + x
...   print p.shift(2)
...   print p.shift(-1)
3*x^5 + 2*x^4 + x^3
3*x^2 + 2*x + 1
```

4. Tester cette fonction. Calculer l'inverse de la matrice

$$\mathbf{A} = \begin{bmatrix} 5 & 2 & 2 & 2 & 2 \\ 2 & 5 & 2 & 2 & 2 \\ 2 & 2 & 5 & 2 & 2 \\ 2 & 2 & 2 & 5 & 2 \\ 2 & 2 & 2 & 2 & 5 \end{bmatrix}.$$

Exercice (Sage) **36.—** Écrire une procédure testant si une matrice est nilpotente.

Exercice (Sage) **37 (Calcul du polynôme minimal).—** On souhaite écrire une procédure permettant de calculer le polynôme minimal d'une matrice. Naturellement, on s'interdit l'utilisation de la procédure `MinimalPolynomial` qui réalise ce calcul.

Nous savons que toute matrice $\mathbf{A} \in \mathscr{M}_n(\mathbb{K})$ admet un polynôme minimal. En effet, la famille

$$(\mathbf{1}_n, \mathbf{A}, \mathbf{A}^2, \ldots, \mathbf{A}^{n^2}),$$

composée de n^2+1 vecteurs est liée. Il existe donc des scalaires non tous nuls, tels que

$$a_0\mathbf{1}_n + a_1\mathbf{A} + a_2\mathbf{A}^2 + \ldots + a_{n^2}\mathbf{A}^{n^2} = 0$$

et le polynôme non nul $g = a_0 + a_1x + a_2x^2 + \ldots + a_{n^2}x^{n^2}$ est ainsi annulateur de la matrice $\mathbf{A}$.

Afin d'écrire une procédure permettant de déterminer le polynôme minimal de $\mathbf{A}$, on observe que la première famille de la suite de familles

$$(\mathbf{A}^0, \mathbf{A}^1, \ldots, \mathbf{A}^k), \qquad k \in \{0, \ldots, n^2\},$$

est libre, car elle est réduite à la matrice identité, et que la dernière de ces familles est liée. Il existe donc un plus petit indice k_0, tel que la famille $(\mathbf{A}^i)_{0\leqslant i\leqslant k_0}$ soit liée. Par définition de k_0, la famille $(\mathbf{A}^0, \mathbf{A}^1, \ldots, \mathbf{A}^{k_0-1})$ est alors libre. Il existe donc une unique famille $(b_0, b_1, \ldots b_{k_0-1})$ de scalaires tels que

$$b_0\mathbf{1}_n + b_1\mathbf{A} + b_2\mathbf{A}^2 + \ldots + b_{k_0-1}\mathbf{A}^{k_0-1} + \mathbf{A}^{k_0} = 0.$$

Une méthode pour déterminer le polynôme minimal de la matrice $\mathbf{A}$, consiste donc à résoudre successivement le système de n^2 équations en les scalaires $b_0, b_1, \ldots b_{k-1}$:

$$b_0\mathbf{1}_n + b_1\mathbf{A} + b_2\mathbf{A}^2 + \ldots + b_{k-1}\mathbf{A}^{k-1} + \mathbf{A}^k = 0 \qquad\qquad (\mathscr{S}_k)$$

1. Montrer que le premier indice k, pour lequel le système $(\mathscr{S}_k)$ admet une solution est k_0 et que cette solution est unique.
2. Déduire de ce qui précède, une procédure qui prend en argument une matrice et retourne son polynôme minimal.
3. Comparer sur des exemples les résultats obtenus avec ceux obtenus avec la procédure `MinimalPolynomial`.

CHAPITRE
8

Décomposition spectrale des matrices

Sommaire

§ 1 Préliminaires

8.1.1. Diagonalisation par blocs.— Dans les chapitres précédents, nous avons vu qu'une matrice $\mathbf{A}$ de $\mathscr{M}_n(\mathbb{K})$ n'est pas toujours diagonalisable. Nous allons montrer dans ce chapitre que si le polynôme caractéristique de la matrice $\mathbf{A}$ est scindé, alors il est toujours possible de la « *diagonaliser par blocs* » autrement dit $\mathbf{A}$ est semblable à une matrice de la forme

$$\begin{bmatrix} \mathbf{B}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{B}_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{B}_p \end{bmatrix}$$

où les $\mathbf{B}_i$ sont des blocs carrés.

D'après le théorème de caractérisation des matrices diagonalisables, théorème 6.13, diagonaliser une matrice $\mathbf{A}$ de $\mathscr{M}_n(\mathbb{K})$ revient à trouver une décomposition de l'espace

$\mathbb{K}^n$ en une somme directe de sous-espaces propres :

$$\mathbb{K}^n = E_{\lambda_1} \oplus \ldots \oplus E_{\lambda_p}$$

associés à $\mathbf{A}$, i.e., pour tout $i \in [\![1,p]\!]$, $E_{\lambda_i} = \mathrm{Ker}\,(\mathbf{A} - \lambda_i \mathbf{1}_n)$. Diagonaliser une matrice $\mathbf{A}$ de $\mathscr{M}_n(\mathbb{K})$ par blocs consiste à trouver une décomposition de l'espace $\mathbb{K}^n$ en une somme directe

$$\mathbb{K}^n = N_1 \oplus \ldots \oplus N_p,$$

de sous-espaces vectoriels N_i stables par $\mathbf{A}$, i.e., pour tout $i \in [\![1,p]\!]$,

$$\mathbf{x} \in N_i \quad \Rightarrow \quad \mathbf{A}\mathbf{x} \in N_i.$$

On montre que ceci est équivalent à dire que la matrice $\mathbf{A}$ est semblable à une matrice diagonale par blocs, c'est-à-dire de la forme

$$\begin{bmatrix} \mathbf{B}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{B}_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{B}_p \end{bmatrix}$$

où, pour tout $i \in [\![1,p]\!]$, $\mathbf{B}_i$ est une matrice de $\mathscr{M}_{n_i}(\mathbb{K})$, avec $n_i = \dim N_{\lambda_i}$.

8.1.2. Formulation géométrique.— On peut reformuler ce problème de la façon suivante. Diagonaliser par blocs un endomorphisme de $\mathbb{K}^n$ revient à trouver une décomposition de $\mathbb{K}^n$ en une somme directe

$$\mathbb{K}^n = N_1 \oplus \ldots \oplus N_p,$$

de sous-espaces vectoriels stables par u, i.e., l'image par u du sous-espace N_i est un sous-espace de N_i. Considérant une base $\mathscr{B}_i$ pour chaque sous-espace N_i, la réunion de ces bases

$$\mathscr{B} = \mathscr{B}_1 \cup \ldots \cup \mathscr{B}_p$$

forme une base de $\mathbb{K}^n$. Exprimée dans cette base, la matrice de u est de la forme suivante :

$$[u]_{\mathscr{B}} = \begin{bmatrix} [u_{|N_1}]_{\mathscr{B}_1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & [u_{|N_2}]_{\mathscr{B}_2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & [u_{|N_p}]_{\mathscr{B}_p} \end{bmatrix}$$

Le polynôme caractéristique de l'endomorphisme u étant scindé, le polynôme caractéristique de chaque restriction $u_{|N_i}$ est scindé. D'après le théorème 6.1, il existe pour chaque espace N_i une base $\mathscr{B}_i$ de trigonalisation de $u_{|N_i}$. Ainsi, la matrice qui exprime u dans la base $\mathscr{B}$ est diagonale par blocs, et de plus ces blocs sont triangulaires.

§ 2 Matrices nilpotentes

8.2.1. Matrices strictement triangulaire.— On appelle matrice *strictement triangulaire* supérieure (resp. inférieure) une matrice triangulaire dont les coefficients diagonaux sont tous nuls.

8.1 Proposition.— Toute matrice strictement triangulaire est nilpotente.

Preuve. Soit $\mathbf{T}$ une matrice strictement triangulaire de $\mathscr{M}_n(\mathbb{K})$. Sa diagonale étant nulle, le spectre de $\mathbf{T}$ est réduit à $\{0\}$. Ainsi, le polynôme caractéristique de $\mathbf{T}$ est

$$p_{\mathbf{T}} = (-1)^n x^n.$$

Le polynôme minimal de $\mathbf{T}$ est donc de la forme x^k, avec $1 \leqslant k \leqslant n$. Autrement dit, $\mathbf{T}$ est nilpotente. □

Nous avons de nombreuses caractérisations des matrices nilpotentes :

8.2 Proposition.— Soit $\mathbf{A}$ une matrice de $\mathscr{M}_n(\mathbb{K})$ non nulle. Les assertions suivantes sont équivalentes :

- **i)** la matrice $\mathbf{A}$ est nilpotente,
- **ii)** le polynôme minimal de $\mathbf{A}$ est de la forme x^r, avec $r > 0$,
- **iii)** le polynôme caractéristique de $\mathbf{A}$ est $(-1)^n x^n$,
- **iv)** la puissance n-ième de $\mathbf{A}$ est nulle,
- **v)** le spectre de $\mathbf{A}$ est réduit à $\{0\}$,
- **vi)** la matrice $\mathbf{A}$ est semblable à une matrice strictement triangulaire,
- **vii)** $\operatorname{trace} \mathbf{A}^k = 0$, pour tout $k \in [\![1, n]\!]$.

Preuve. Si $\mathbf{A}$ est nilpotente, il existe un entier q tel que le polynôme x^q soit annulateur de $\mathbf{A}$, par suite, le polynôme minimal de $\mathbf{A}$ est de la forme x^r, où r est l'indice de nilpotence de la matrice $\mathbf{A}$. Ainsi **i)** implique **ii)**.

Par ailleurs, **ii)** implique **iii)**, **iii)** implique **iv)** et **iv)** implique **v)** sont immédiates.

Supposons que le spectre de $\mathbf{A}$ soit réduit à $\{0\}$, i.e., 0 est la seule valeur propre de $\mathbf{A}$. La matrice $\mathbf{A}$ est trigonalisable et semblable à une matrice strictement triangulaire supérieure $\mathbf{T}$: $\mathbf{A} = \mathbf{P}^{-1}\mathbf{T}\mathbf{P}$ avec $\mathbf{P}$ inversible. D'où **v)** implique **vi)**.

L'implication **vi)** implique **i)** est donnée par la proposition 8.1.

L'équivalence de **vii)** avec les autres propriétés est une conséquence immédiate de la proposition 6.3. □

§ 3 Les espaces spectraux

8.3.1. Premier exemple : le cas d'une matrice diagonalisable.— Soit $\mathbf{A}$ la matrice de $\mathscr{M}_3(\mathbb{R})$ suivante :

$$\mathbf{A} = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}.$$

Le polynôme caractéristique de $\mathbf{A}$ est $p_{\mathbf{A}} = -(x-4)(x-1)^2$ et son polynôme minimal est $m_{\mathbf{A}} = (x-4)(x-1)$. La matrice $\mathbf{A}$ est alors diagonalisable. Il existe donc une décomposition de l'espace $\mathbb{R}^3$ en somme directe de sous-espaces propres :

$$\mathbb{R}^3 = E_4 \oplus E_1.$$

La valeur propre 4 étant de multiplicité algébrique 1, on a $\dim E_4 = 1$. On en déduit que $\dim E_1 = 2$.

Déterminons les projections de l'espace $\mathbb{R}^3$ sur les sous-espaces propres E_4 et E_1. Nous noterons

$$\pi_4 : \mathbb{R}^3 \longrightarrow E_4$$

la projection de $\mathbb{R}^3$ sur le sous-espace E_4 parallèlement au sous-espace E_1 et

$$\pi_1 : \mathbb{R}^3 \longrightarrow E_1$$

la projection de $\mathbb{R}^3$ sur E_1 parallèlement à E_4. Ces deux projections sont entièrement caractérisées par les relations suivantes :

$$\begin{gathered} \pi_4^2 = \pi_4,\ \ \pi_1^2 = \pi_1,\ \ \pi_4\pi_1 = \pi_1\pi_4 = 0,\ \text{ et } \ \mathrm{id}_{\mathbb{R}^3} = \pi_4 + \pi_1, \\ \operatorname{Im}\pi_4 = E_4,\ \ \operatorname{Ker}\pi_4 = E_1,\ \ \operatorname{Im}\pi_1 = E_1,\ \ \operatorname{Ker}\pi_1 = E_4 \end{gathered} \tag{8.1}$$

On notera Π_1 et Π_4 les matrices des projections π_1 et π_4 exprimées dans la base canonique de $\mathbb{R}^3$. Nous allons voir que l'on peut exprimer ces deux matrices en fonction de la matrice $\mathbf{A}$.

Les polynômes $x-4$ et $x-1$ sont premiers entre eux. D'après l'identité de Bézout, théorème 0.4, il existe deux polynôme h_1 et h_2 de $\mathbb{R}[x]$ tels que

$$(x-4)h_1 + (x-1)h_2 = 1$$

Un calcul élémentaire nous permet de déterminer les polynômes h_1 et h_2, on a :

$$(x-4)(-\frac{1}{3}) + (x-1)(\frac{1}{3}) = 1. \tag{8.2}$$

Montrons que

$$\Pi_4 = \frac{1}{3}(\mathbf{A} - \mathbf{1}_3) \quad \text{et} \quad \Pi_1 = -\frac{1}{3}(\mathbf{A} - 4\mathbf{1}_3).$$

Supposons Π_4 et Π_1 ainsi définis et montrons qu'ils satisfont les relations 8.1. De l'équa-

tion 8.2, on déduit que
$$\Pi_4 + \Pi_1 = \mathbf{1}_3. \tag{8.3}$$
Montrons que
$$\operatorname{Ker}\Pi_4 = E_1 \quad \text{et} \quad \operatorname{Im}\Pi_4 = E_4.$$
Comme $\Pi_4 = \dfrac{1}{3}(\mathbf{A} - \mathbf{1}_3)$, la première égalité est immédiate. Pour la seconde égalité, supposons $\mathbf{y} \in \operatorname{Im}\Pi_4$, il existe alors un vecteur $\mathbf{x}$ de $\mathbb{R}^3$ tel que
$$\mathbf{y} = \Pi_4(\mathbf{x}) = \pi_4(\mathbf{x}) = \frac{1}{3}(\mathbf{A} - \mathbf{1}_3)(\mathbf{x}).$$
Par ailleurs,
$$(\mathbf{A} - 4\mathbf{1}_3)\frac{1}{3}(\mathbf{A} - \mathbf{1}_3)(\mathbf{x}) = \frac{1}{3}m_{\mathbf{A}}(\mathbf{A})(\mathbf{x}) = \mathbf{0}.$$
Par suite, $(\mathbf{A} - 4\mathbf{1}_3)(\mathbf{y}) = \mathbf{0}$, d'où $\mathbf{y} \in E_4$.

Inversement, si $\mathbf{y} \in E_4$, d'après (8.3) on a $\mathbf{y} = \Pi_4(\mathbf{y}) + \Pi_1(\mathbf{y})$. Or
$$\Pi_1(\mathbf{y}) = -\frac{1}{3}(\mathbf{A} - 4\mathbf{1}_3)(\mathbf{y}) = \mathbf{0}.$$
Donc $\mathbf{y} = \Pi_4(\mathbf{y})$, par suite $\mathbf{y} \in \operatorname{Im}\Pi_4$. Ainsi, $\operatorname{Im}\Pi_4 = E_4$.

De la même façon, on montre que
$$\operatorname{Ker}\Pi_1 = E_4 \quad \text{et} \quad \operatorname{Im}\Pi_1 = E_1.$$
Les relations $\Pi_4\Pi_1 = \Pi_1\Pi_4 = \mathbf{0}$ se déduisent immédiatement du fait que
$$\Pi_4\Pi_1 = \left(\frac{1}{3}(\mathbf{A} - \mathbf{1}_3)\right)\left(-\frac{1}{3}(\mathbf{A} - 4\mathbf{1}_3)\right) = -\frac{1}{9}m_{\mathbf{A}}(\mathbf{A}) = \mathbf{0}.$$
Les relations $\Pi_4^2 = \Pi_4$, $\Pi_1^2 = \Pi_1$ sont une conséquence immédiate des relations (8.3) et $\Pi_4\Pi_1 = \Pi_1\Pi_4 = \mathbf{0}$.

En multipliant les deux membres de la relation (8.3) à gauche par la matrice $\mathbf{A}$, on obtient la décomposition suivante de la matrice $\mathbf{A}$:
$$\mathbf{A} = \mathbf{A}\Pi_4 + \mathbf{A}\Pi_1.$$
Comme pour tout vecteur $\mathbf{x}$ de E_4, on a $\mathbf{A}\mathbf{x} = 4\mathbf{x}$ et $\Pi_4\mathbf{x} = \mathbf{x}$, on en déduit que
$$\mathbf{A}\Pi_4 = 4\Pi_4.$$
Par ailleurs, on a de façon similaire $\mathbf{A}\Pi_1 = \Pi_1$. Ainsi,
$$\mathbf{A} = 4\Pi_4 + \Pi_1. \tag{8.4}$$

La décomposition (8.4) de la matrice $\mathbf{A}$ est appelée la *décomposition spectrale* de $\mathbf{A}$. On montrera dans la suite que cette décomposition de la matrice $\mathbf{A}$ est très utile pour avoir des expressions simples de puissances de $\mathbf{A}$. En particulier, pour tout entier k, on montrera que

$$\mathbf{A}^k = 4^k\Pi_4 + \Pi_1.$$

Les matrices des projecteurs π_1 et π_4 dans la base canonique de $\mathbb{R}^3$ sont

$$\Pi_4 = \frac{1}{3}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}, \qquad \Pi_1 = -\frac{1}{3}\begin{bmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{bmatrix}.$$

On notera que $\operatorname{rang}\Pi_4 = 1 = \operatorname{trace}\Pi_4$ et que $\operatorname{rang}\Pi_1 = 2 = \operatorname{trace}\Pi_2$.

8.3.2. Deuxième exemple : le cas d'une matrice non diagonalisable.— Considérons la matrice suivante de $\mathscr{M}_4(\mathbb{R})$.

$$\mathbf{A} = \begin{bmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}.$$

Le polynôme caractéristique de $\mathbf{A}$ est $p_{\mathbf{A}} = (x-2)^3(x-3)$, son polynôme minimal est $m_{\mathbf{A}} = (x-2)^2(x-3)$. La matrice $\mathbf{A}$ n'est donc pas diagonalisable.

En appliquant le lemme des noyaux, théorème 7.6, au polynôme minimal, on obtient la décomposition :

$$\mathbb{R}^4 = \operatorname{Ker}\,(\mathbf{A} - 2\mathbf{1}_4)^2 \oplus E_3,$$

où E_3 est le sous-espace propre associé à la valeur propre 3. Notons

$$N_2 = \operatorname{Ker}\,(\mathbf{A} - 2\mathbf{1}_4)^2.$$

La dimension de E_3 est majorée par la multiplicité algébrique de la valeur propre 3. On a donc $\dim E_3 = 1$ et par suite $\dim N_2 = 3$.

Déterminons la matrice de la projection π_2 de $\mathbb{R}^4$ sur N_2 parallèlement à E_3 et celle de la projection π_3 de $\mathbb{R}^4$ sur E_3 parallèlement à N_2, exprimées dans la base canonique de $\mathbb{R}^4$.

Les polynômes $(x-2)^2$ et $x-3$ sont premiers entre eux ; on établit une relation de Bézout :

$$(x-2)^2 + (x-3)(-x+1) = 1.$$

Posons

$$\Pi_2 = (\mathbf{A} - 3\mathbf{1}_4)(-\mathbf{A} + \mathbf{1}_4) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{et} \quad \Pi_3 = (\mathbf{A} - 2\mathbf{1}_4)^2 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Ainsi définis, Π_2 et Π_3 satisfont :

$$\Pi_2^2 = \Pi_2, \quad \Pi_3^2 = \Pi_3, \quad \Pi_2\Pi_3 = \Pi_3\Pi_2 = 0, \quad \mathbf{1}_4 = \Pi_2 + \Pi_3.$$

De plus, on a

$$\mathrm{Im}\,\Pi_2 = N_2, \quad \mathrm{Ker}\,\Pi_3 = E_3, \quad \mathrm{Im}\,\Pi_3 = E_3, \quad \mathrm{Ker}\,\Pi_3 = N_2.$$

Ainsi, les matrices Π_2 et Π_3 sont bien celles des projections sur les sous-espaces N_2 et E_3 respectivement.

Posons

$$\mathbf{D} = 2\Pi_2 + 3\Pi_3 = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}, \quad \mathbf{N} = \mathbf{A} - \mathbf{D} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

On a $\mathbf{N}^2 = \mathbf{0}$ et $\mathbf{A} = \mathbf{D} + \mathbf{N}$. Ainsi, $\mathbf{A}$ s'écrit comme la somme d'une matrice diagonalisable et d'une matrice nilpotente. Nous montrons dans ce chapitre que cette décomposition est unique.

§ 4 Décomposition spectrale géométrique

8.4.1. Sous-espace spectral.— Soit $\mathbf{A}$ une matrice de $\mathscr{M}_n(\mathbb{K})$ dont le polynôme caractéristique est scindé :

$$p_{\mathbf{A}} = (-1)^n (x-\lambda_1)^{n_1} \ldots (x-\lambda_p)^{n_p},$$

avec $\lambda_i \neq \lambda_j$ si $i \neq j$ et $n_1 + \ldots + n_p = n$.

Le sous-espace vectoriel $\mathrm{Ker}\,(\mathbf{A} - \lambda_i \mathbf{1}_n)^{n_i}$ de $\mathbb{K}^n$ est appelé le *sous-espace spectral* de $\mathbf{A}$ associé à la valeur propre λ_i. On le notera N_{λ_i}. Certains textes utilisent aussi la terminologie de *sous-espace caractéristique*, ou encore de *sous-espace propre généralisé*.

De la même façon, si u est un endomorphisme de $\mathbb{K}^n$ dont le polynôme caractéristique est scindé :

$$p_u = (-1)^n (x-\lambda_1)^{n_1} \ldots (x-\lambda_p)^{n_p},$$

avec $\lambda_i \neq \lambda_j$ si $i \neq j$ et $n_1 + \ldots + n_p = n$, le sous-espace vectoriel

$$N_{\lambda_i} = \mathrm{Ker}\,(u - \lambda_i \mathrm{id}_{\mathbb{K}^n})^{n_i}$$

est appelé le sous-espace spectral de u associé à la valeur propre λ_i.

8.3 Proposition.— Soient $\mathbf{A}$ une matrice de $\mathscr{M}_n(\mathbb{K})$ et λ une valeur propre de $\mathbf{A}$. Le sous-espace propre E_λ est un sous-espace vectoriel du sous-espace spectral N_λ.

Preuve. Cela découle immédiatement du fait que pour toute matrice carrée $\mathbf{A}$, on a la suite d'inclusions de sous-espaces

$$\mathrm{Ker}\,\mathbf{A} \subset \mathrm{Ker}\,\mathbf{A}^2 \subset \ldots \subset \mathrm{Ker}\,\mathbf{A}^k \subset \ldots$$

□

8.4 Théorème (Théorème spectral géométrique).— Soit u un endomorphisme de $\mathbb{K}^n$ dont le polynôme caractéristique est scindé avec

$$p_u = (-1)^n (x-\lambda_1)^{n_1} \ldots (x-\lambda_p)^{n_p},$$

avec $\lambda_i \neq \lambda_j$ si $i \neq j$, $n_1 + \ldots + n_p = n$ et dont le polynôme minimal est

$$m_u = (x-\lambda_1)^{k_1} \ldots (x-\lambda_p)^{k_p}, \qquad \text{avec } 0 < k_i \leqslant n_i.$$

Alors, les sous-espaces spectraux de u satisfont, pour tout $i \in [\![1,p]\!]$, les assertions suivantes :

i) $\mathbb{K}^n = N_{\lambda_1} \oplus \ldots \oplus N_{\lambda_p}$,

ii) N_{λ_i} est stable par u,

iii) l'endomorphisme $u_{|N_{\lambda_i}} - \lambda_i \mathrm{id}_{N_{\lambda_i}}$, où $u_{|N_{\lambda_i}}$ est la restriction de u à N_{λ_i}, est nilpotent d'indice k_i,

iv) $\dim N_{\lambda_i} = n_i$.

Preuve. Montrons **i)**. D'après le théorème de Cayley-Hamilton 7.10, le polynôme caractéristique p_u est annulateur de u. La décomposition **i)** est une conséquence du lemme de décomposition en noyaux, proposition **??**, appliquée à la factorisation de p_u en facteurs irréductibles.

Montrons **ii)**. Soit $\mathbf{x} \in N_{\lambda_i}$. Alors $(u-\lambda_i \mathrm{id}_{\mathbb{K}^n})^{n_i}(\mathbf{x}) = \mathbf{0}$, donc $u \circ (u-\lambda_i \mathrm{id}_{\mathbb{K}^n})^{n_i}(\mathbf{x}) = \mathbf{0}$. Par suite, $(u-\lambda_i \mathrm{id}_{\mathbb{K}^n})^{n_i}(u(\mathbf{x})) = \mathbf{0}$, d'où $u(\mathbf{x}) \in N_{\lambda_i}$.

Montrons **iii)**. En appliquant le lemme de décomposition aux polynômes minimal et caractéristique, on a

$$\mathbb{K}^n = \bigoplus_{i=1}^{p} \mathrm{Ker}\,(u-\lambda_i \mathrm{id}_{\mathbb{K}^n})^{n_i} = \bigoplus_{i=1}^{p} \mathrm{Ker}\,(u-\lambda_i \mathrm{id}_{\mathbb{K}^n})^{k_i}.$$

Par ailleurs, pour tout $i \in [\![1,p]\!]$, on a $\mathrm{Ker}\,(u-\lambda_i \mathrm{id}_{\mathbb{K}^n})^{k_i} \subseteq \mathrm{Ker}\,(u-\lambda_i \mathrm{id}_{\mathbb{K}^n})^{n_i}$, on en déduit donc que :

$$\mathrm{Ker}\,(u-\lambda_i \mathrm{id}_{\mathbb{K}^n})^{k_i} = \mathrm{Ker}\,(u-\lambda_i \mathrm{id}_{\mathbb{K}^n})^{n_i} = N_{\lambda_i}.$$

Par suite, pour tout $\mathbf{x} \in N_{\lambda_i}$, on a $(u-\lambda_i \mathrm{id}_{\mathbb{K}^n})^{k_i}(\mathbf{x}) = \mathbf{0}$, soit

$$(u_{|N_{\lambda_i}} - \lambda_i \mathrm{id}_{N_{\lambda_i}})^{k_i} = 0.$$

Le polynôme $(x-\lambda_i)^{k_i}$ est donc annulateur de $u_{|N_{\lambda_i}}$, ainsi le polynôme minimal de $u_{|N_{\lambda_i}}$ est de la forme :

$$m_{u_{|N_{\lambda_i}}} = (x-\lambda_i)^{l_i}, \qquad \text{avec } l_i \leqslant k_i.$$

Le polynôme

$$(x-\lambda_1)^{l_1}\dots(x-\lambda_p)^{l_p}$$

est donc annulateur de u. Comme m_u est minimal, on en déduit que $l_i = k_i$ pour tout i. Donc $(u_{|N_{\lambda_i}} - \lambda_i \mathrm{id}_{N_{\lambda_i}})$ est nilpotent d'indice de nilpotence k_i.

Montrons **iv)**. On a

$$\begin{aligned} p_{u_{|N_{\lambda_i}}}(x+\lambda_i) &= \det\big(u_{|N_{\lambda_i}} - (x+\lambda_i)\mathrm{id}_{N_{\lambda_i}}\big), \\ &= p_{u_{|N_{\lambda_i}} - \lambda_i \mathrm{id}_{N_{\lambda_i}}}(x). \end{aligned}$$

Or, d'après **iii)**, l'endomorphisme $u_{|N_{\lambda_i}} - \lambda_i \mathrm{id}_{N_{\lambda_i}}$ de N_{λ_i} est nilpotent. Par suite, son polynôme caractéristique est

$$p_{u_{|N_{\lambda_i}} - \lambda_i \mathrm{id}_{N_{\lambda_i}}} = (-x)^{\dim N_{\lambda_i}}.$$

Donc le polynôme caractéristique de l'endomorphisme $u_{|N_{\lambda_i}}$ est

$$p_{u_{|N_{\lambda_i}}}(x) = (-1)^{\dim N_{\lambda_i}}(x-\lambda_i)^{\dim N_{\lambda_i}}.$$

Comme N_{λ_i} est stable par u, $p_{u_{|N_{\lambda_i}}}$ divise p_u, d'où $\dim N_{\lambda_i} \leqslant n_i$. Par ailleurs, $\mathbb{K}^n = N_{\lambda_1} \oplus \dots \oplus N_{\lambda_p}$, donc

$$n = \sum_{i=1}^{p} n_i = \sum_{i=1}^{p} \dim N_{\lambda_i}.$$

On en déduit que $\dim N_{\lambda_i} = n_i$, pour tout i. □

8.4.2. Remarque.— Avec les notations précédente, d'après **iv)** le polynôme caractéristique de la restriction $u_{|N_{\lambda_i}}$ de u au sous-espace spectral N_{λ_i} est

$$p_{u_{|N_{\lambda_i}}} = (-1)^{n_i}(x-\lambda_i)^{n_i}.$$

Donc l'endomorphisme $u_{|N_{\lambda_i}}$ est trigonalisable et admet une seule valeur propre λ_i d'ordre de multiplicité algébrique n_i. Il existe donc une base $\mathscr{B}_i$ de N_{λ_i} telle que dans $\mathscr{M}_{n_i}(\mathbb{K})$:

$$[u_{|N_{\lambda_i}}]_{\mathscr{B}_i} = \begin{bmatrix} \lambda_i & & * \\ & \ddots & \\ 0 & & \lambda_i \end{bmatrix}.$$

En considérant, pour tout $i \in [\![1,p]\!]$, une telle base de trigonalisation $\mathscr{B}_i$ pour chaque

N_{λ_i}, la réunion $\mathscr{B} = \mathscr{B}_1 \cup \ldots \cup \mathscr{B}_p$ forme une base de l'espace $\mathbb{K}^n$ et on a

$$[u]_{\mathscr{B}} = \begin{bmatrix} \begin{array}{c} n_1 \\ \begin{bmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_1 \end{bmatrix} n_1 \end{array} & & & \mathbf{0} \\ & \begin{array}{c} n_2 \\ \begin{bmatrix} \lambda_2 & & * \\ & \ddots & \\ 0 & & \lambda_2 \end{bmatrix} n_2 \end{array} & & \\ & & \ddots & \\ \mathbf{0} & & & \begin{array}{c} n_p \\ \begin{bmatrix} \lambda_p & & * \\ & \ddots & \\ 0 & & \lambda_p \end{bmatrix} n_p \end{array} \end{bmatrix}.$$

§ 5 Décomposition spectrale algébrique

Nous reformulons ici de façon algébrique le théorème spectral géométrique 8.4.

Rappelons que si $E_1, \ldots, E_p$ sont des sous-espaces vectoriels de $\mathbb{K}^n$, alors une condition nécessaire et suffisante pour que $\mathbb{K}^n = E_1 \oplus \ldots \oplus E_p$ est qu'il existe des applications linéaires

$$\pi_i : \mathbb{K}^n \longrightarrow E_i, \quad i \in [\![1, p]\!]$$

telles que

$$\mathrm{id}_{\mathbb{K}^n} = \pi_1 + \ldots + \pi_p, \qquad \pi_i \pi_j = 0, \text{ pour tout } i \neq j.$$

L'application linéaire π_i est appelée la projection sur E_i parallèlement au sous-espace $\underset{\substack{j=1\\ i\neq j}}{\overset{p}{\oplus}} E_j$.

Soit u un endomorphisme de $\mathbb{K}^n$ dont le polynôme caractéristique est scindé :

$$p_u = (-1)^n (x - \lambda_1)^{n_1} \ldots (x - \lambda_p)^{n_p},$$

avec $\lambda_i \neq \lambda_j$ si $i \neq j$, $n_1 + \ldots + n_p = n$. D'après le théorème 8.4, on a une décomposition

$$\mathbb{K}^n = N_{\lambda_1} \oplus \ldots \oplus N_{\lambda_p}.$$

8.5.1. Définition.— La projection sur le sous-espace spectral N_{λ_i} parallèlement au sous-espace $\underset{\substack{j=1\\ i\neq j}}{\overset{p}{\oplus}} N_{\lambda_j}$ est appelé *projecteur spectral associé à la valeur propre* λ_i. On le notera π_{λ_i}.

8.5 Proposition.— Les projecteur spectraux $\pi_{\lambda_1}, \ldots, \pi_{\lambda_p}$ sont des polynômes en u.

Preuve. Soit $i \in [\![1, p]\!]$. Notons

$$g_i = \prod_{\substack{j=1 \\ j \neq i}}^{p} (x - \lambda_j)^{n_j}.$$

Si $i \neq j$, le polynôme $g = g_i g_j = (-1)^n p_u$ est annulateur de u. Les polynômes $g_1, \ldots, g_p$ sont premiers entre eux dans leur ensemble. D'après le théorème de Bézout, il existe des polynômes h_1, …, h_p de $\mathbb{K}[x]$ tels que

$$g_1 h_1 + \ldots + g_p h_p = 1.$$

Par suite,

$$g_1(u) h_1(u) + \ldots + g_p(u) h_p(u) = \mathrm{id}_{\mathbb{K}^n}. \tag{8.5}$$

Pour tout $i \in [\![1, p]\!]$, posons

$$\pi_{\lambda_i} = g_i(u) h_i(u).$$

De la relation (8.5), on déduit que

$$\mathrm{id}_{\mathbb{K}} = \pi_{\lambda_1} + \ldots + \pi_{\lambda_p}. \tag{8.6}$$

D'autre part, si $j \neq i$, le polynôme $g_j g_i = g$ est annulateur de u, donc :

$$\begin{aligned}
\pi_{\lambda_i} \circ \pi_{\lambda_j} &= g_i(u) h_i(u) g_j(u) h_j(u) \\
&= g_i(u) g_j(u) h_i(u) h_j(u) \\
&= 0.
\end{aligned}$$

Il nous reste à montrer que

$$\mathrm{Im}\, \pi_{\lambda_i} = N_{\lambda_i}, \qquad \mathrm{Ker}\, \pi_{\lambda_i} = \bigoplus_{\substack{j=1 \\ j \neq i}}^{p} N_{\lambda_j}.$$

Montrons la première égalité. Soit $\mathbf{y} = \pi_{\lambda_i}(\mathbf{x}) \in \mathrm{Im}\, \pi_{\lambda_i}$, on a $\mathbf{y} = g_i(u) h_i(u)(\mathbf{x})$ et

$$\begin{aligned}
(u - \lambda_i \mathrm{id}_{\mathbb{K}})^{n_i}(\mathbf{y}) &= (x - \lambda_i)^{n_i}(u) g_i(u) h_i(u)(\mathbf{x}) \\
&= g_i(u) g(u)(\mathbf{x}) \\
&= \mathbf{0}.
\end{aligned}$$

Donc $\mathbf{y} \in N_{\lambda_i}$ et $\mathrm{Im}\, \pi_{\lambda_i} \subseteq N_{\lambda_i}$.

Inversement, soit $\mathbf{x} \in N_{\lambda_i}$. On a $\mathbf{x} = \pi_{\lambda_1}(\mathbf{x}) + \ldots + \pi_{\lambda_p}(\mathbf{x})$. Par ailleurs, si $i \neq j$ le polynôme $(x - \lambda_i)^{n_i}$ divise g_j. Le polynôme $(x - \lambda_i)^{n_i}$ étant annulateur de l'endomorphisme $u_{|N_{\lambda_i}}$, on a alors

$$\pi_{\lambda_j}(\mathbf{x}) = g_j(u)h_j(u)(\mathbf{x}) = \mathbf{0}, \qquad \text{si } i \neq j.$$

Donc $\mathbf{x} = \pi_{\lambda_i}(\mathbf{x}) \in \operatorname{Im} \pi_{\lambda_i}$ et $N_{\lambda_i} \subseteq \operatorname{Im} \pi_{\lambda_i}$.

Montrons la seconde égalité. Si $i \neq j$, le polynôme $(x-\lambda_j)^{n_j}$ divise g_i, donc pour tout $\mathbf{x} \in N_{\lambda_j}$, on a $\pi_i(\mathbf{x}) = g_i(u)h_i(u)(\mathbf{x}) = \mathbf{0}$. Donc $N_{\lambda_j} \subseteq \operatorname{Ker} \pi_{\lambda_i}$ si $i \neq j$. Par suite $\bigoplus\limits_{j\neq i} N_{\lambda_j} \subseteq \operatorname{Ker} \pi_{\lambda_i}$.

Inversement, soit $\mathbf{x} \in \operatorname{Ker} \pi_{\lambda_i}$. Comme $\mathbf{x} = \pi_{\lambda_1}(\mathbf{x}) + \ldots + \mathbf{0} + \ldots + \pi_{\lambda_p}(\mathbf{x})$, on a $\mathbf{x} \in \bigoplus\limits_{j\neq i} N_{\lambda_j}$. On montre ainsi la proposition. □

8.5.2. Lemme.— Soient u un endomorphisme diagonalisable de E et F un sous-espace vectoriel de E stable par u. La restriction de u au sous-espace F est un endomorphisme diagonalisable de F.

Preuve. Le sous-espace vectoriel F est stable par u, donc la restriction $u_{|F}$ est un endomorphisme de F. Montrons que le polynôme minimal de $u_{|F}$ divise le polynôme minimal de u. On a $m_u(u) = 0$, donc

$$m_u(u)(\mathbf{x}) = \mathbf{0}, \quad \text{pour tout } \mathbf{x} \in E.$$

En particulier, cette égalité reste vraie pour tout $\mathbf{x} \in F$, par suite m_u est un polynôme annulateur de $u_{|F}$. D'où $m_{u_{|F}}$ divise m_u.

Par ailleurs, u est diagonalisable, donc m_u est scindé et ne possède que des racines simples. Comme $m_{u_{|F}}$ divise m_u, il est donc scindé et ne possède que des racines simples. La restriction $u_{|F}$ est donc diagonalisable. □

8.5.3. Lemme.— Soient u et v deux endomorphismes de E qui commutent, i.e. $uv = vu$. Alors,

i) tout sous-espace propre de u est stable par v,

ii) u et v sont diagonalisables si, et seulement si, il existe une base commune de diagonalisation.

Preuve. Montrons **i)** Soit λ une valeur de u, montrons que le sous-espace propre E_λ de u associé à la valeur propre λ est stable par v. Soit $\mathbf{x} \in E_\lambda$, on a

$$vu(\mathbf{x}) = v(\lambda x) = \lambda v(\mathbf{x}).$$

Or $uv(\mathbf{x}) = vu(\mathbf{x})$, donc $u(v(\mathbf{x}))) = \lambda v(\mathbf{x})$. Par suite, $v(\mathbf{x}) \in E_\lambda$ et E_λ est stable par v.

Montrons **ii)** Supposons que u et v soient diagonalisable. Posons $\mathrm{Sp(u)} = \{\lambda_1, \ldots, \lambda_p\}$. L'endomorphisme v est diagonalisable et d'après **i)** E_{λ_i} est stable par v. D'après le lemme 8.5.2, la restriction $v_{|E_{\lambda_i}}$ de v au sous-espace E_{λ_i} est diagonalisable ; soit $\mathscr{B}_i$ une base de diagonalisation de $v_{|E_{\lambda_i}}$. Alors la réunion $\mathscr{B} = \mathscr{B}_1 \cup \ldots \cup \mathscr{B}_p$ forme une base de diagonalisation de v. C'est aussi une base de diagonalisation de u, car pour tout $\mathbf{e}_i \in \mathscr{B}_i$, on a $u(\mathbf{e}_i) = \lambda_i \mathbf{e}_i$.

La réciproque est immédiate.

□

Le théorème suivant est aussi connu sous le nom de « décomposition de Dunford ».

8.6 Théorème (Décomposition spectrale algébrique).— Soit u un endomorphisme de $\mathbb{K}^n$ dont le polynôme caractéristique p_u est scindé. Soient λ_1, …, λ_p les valeurs propres de u. Il existe un unique couple (d,n) d'endomorphismes de E tels que

i) $u = d + n$,

ii) d est diagonalisable et n est nilpotente,

iii) n et d sont des polynômes en u,

iv) $d \circ n = n \circ d$.

De plus, on a :

$$d = \lambda_1 \pi_{\lambda_1} + \ldots + \lambda_p \pi_{\lambda_p}$$

et

$$n = (u - \lambda_1 \mathrm{id}_E)\pi_{\lambda_1} + \ldots + (u - \lambda_p \mathrm{id}_E)\pi_{\lambda_p},$$

où $\pi_{\lambda_1}, \ldots, \pi_{\lambda_p}$ désignent les projecteurs spectraux de l'endomorphisme u.

Preuve. La preuve est une conséquence immédiate du théorème 8.4. Montrons l'existence de d et u satisfaisant **i)** et **ii)**. D'après 8.4 **i)**, on a

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

Donc $\mathrm{id}_E = \pi_{\lambda_1} + \ldots + \pi_{\lambda_p}$, par suite

$$u = u\pi_{\lambda_1} + \ldots + u\pi_{\lambda_p},$$

d'où

$$u = u_{|N_{\lambda_1}} \pi_{\lambda_1} + \ldots + u_{|N_{\lambda_p}} \pi_{\lambda_p}.$$

Par ailleurs, d'après 8.4 **iii)**, l'endomorphisme $n_i = u_{|N_{\lambda_i}} - \lambda_i \mathrm{id}_{N_{\lambda_i}}$ est nilpotent. On a donc

$$u_{|N_{\lambda_i}} = \lambda_i \mathrm{id}_{N_{\lambda_i}} + n_i.$$

Ainsi

$$u = (\lambda_1 \mathrm{id}_{N_{\lambda_1}} + n_1)\pi_{\lambda_1} + \ldots + (\lambda_p \mathrm{id}_{N_{\lambda_p}} + n_p)\pi_{\lambda_p}.$$

Posons

$$d = \lambda_1 \pi_{\lambda_1} + \ldots + \lambda_p \pi_{\lambda_p} \quad \text{et} \quad n = n_1 \pi_{\lambda_1} + \ldots + n_p \pi_{\lambda_p}.$$

Ainsi définis, les endomorphismes d et n varifient $u = d+n$. De plus, chaque n_i est nilpotent, donc n est nilpotent. Enfin, d est diagonalisable, car pour tout $\mathbf{x} \in N_{\lambda_i}$, $d(\mathbf{x}) = \lambda_i \mathbf{x}$ et toute base $\mathscr{B} = \mathscr{B}_1 \cup \ldots \cup \mathscr{B}_p$, où $\mathscr{B}_i$ est une base de N_{λ_i} est une base de diagonalisation de d. On montre ainsi l'existence de d diagonalisable et n nilpotente tels que $u = d+n$.

D'après la proposition 8.5 les projecteurs spectraux π_{λ_i}, $i \in [\![1,p]\!]$ sont des polynômes en u. Par suite, les endomorphismes d et n sont des polynômes en u. Ce qui montre **iii)**.

L'assertion **iv)** découle de **iii)** et du fait que, d'après la proposition **??**, deux polynômes en u commutent entre eux. Donc $d \circ n = n \circ d$.

Reste à montrer l'unicité de la décomposition. Supposons qu'il existe deux couples d'endomorphismes (d,n) et (d',n') vérifiant les propriétés **i**, **ii)**, **iii)** et **iv)**. On a

$$u = d+n = d'+n',$$

avec d' diagonalisable, n' nilpotente, $d \circ n = n \circ d$ et $d' \circ n' = n' \circ d'$. On a alors

$$d' \circ u = d' \circ (d'+n') = d'^2 + n' \circ d' = (d'+n') \circ d' = u \circ d'.$$

Donc d' commute avec u et par suite d' commute avec tous les polynômes en u. Ainsi, d' commute avec d. Par ailleurs, d et d' sont diagonalisables, d'après le lemme 8.5.3, ils sont donc diagonalisables dans une même base donc $d-d'$ est diagonalisable.

D'autre part, n et n' sont nilpotents et commutent, donc $n'-n$ est nilpotent. Ainsi $d-d' = n'-n$ est un endomorphisme diagonalisable et nilpotent, il est donc nul. Par suite $d = d'$ et $n = n'$. □

8.7 Proposition.— Un endomorphisme u de E est diagonalisable si, et seulement si, il existe une famille de projecteurs $\pi_1, \ldots, \pi_p$ et une famille de scalaires $\lambda_1, \ldots, \lambda_p$ distincts deux à deux vérifiant

- **i)** $\mathrm{id}_E = \pi_1 + \ldots + \pi_p$,
- **ii)** $\pi_i \pi_j = 0$, si $i \neq j$,
- **iii)** $u = \lambda_1 \pi_1 + \ldots + \lambda_p \pi_p$.

Preuve. Supposons que u soit diagonalisable. Soient $\lambda_1, \ldots, \lambda_p$ les valeurs propres de u. Le polynôme minimal de u est donc scindé à racines simples. On a donc pour tout $i \in [\![1,p]\!]$,

$$E_{\lambda_i} = \mathrm{Ker}\,(u - \lambda_i \mathrm{id}_E)^{n_i} = N_{\lambda_i}.$$

D'après le théorème 8.6, on a

$$u = \lambda_1 \pi_{\lambda_1} + \ldots + \lambda_p \pi_{\lambda_p},$$

où les π_{λ_i} sont les projections de E sur E_{λ_i} et vérifient donc **i)** et **ii)**.

Inversement, supposons qu'il existe des scalaires $\lambda_1, \ldots, \lambda_p$ et des projecteurs π_1, ..., π_p vérifiant les assertions **i)**, **ii)** et **iii)**. On a donc

$$E = \text{Im } \pi_1 \oplus \ldots \oplus \text{Im } \pi_p. \tag{8.7}$$

Montrons que, pour tout $i \in [\![1, p]\!]$, $\text{Im } \pi_i = \text{Ker } (u - \lambda_i \text{id}_E)$. Soient $i \in [\![1, p]\!]$ et $\mathbf{x} \in \text{Im } \pi_i$. Il existe $\mathbf{y}$, tel que $\mathbf{x} = \pi_i(\mathbf{y})$. Donc

$$\pi_i(\mathbf{x}) = \pi_i \pi_i(\mathbf{y}) = \pi_i(\mathbf{y}) = \mathbf{x}.$$

D'après **iii)**, on a

$$u(\mathbf{x}) = \lambda_1 \pi_1(\mathbf{x}) + \ldots + \lambda_p \pi_p(\mathbf{x}).$$

Par ailleurs, si $j \neq i$, on a $\pi_j(\mathbf{x}) = \pi_j(\pi_i(\mathbf{x})) = \mathbf{0}$. Donc

$$u(\mathbf{x}) = \lambda_i \pi_i(\mathbf{x}) = \lambda_i \mathbf{x}.$$

D'où $\mathbf{x} \in \text{Ker } (u - \lambda_i \text{id}_E)$. Inversement, si $\mathbf{x} \in \text{Ker } (u - \lambda_i \text{id}_E)$, on a $u(\mathbf{x}) = \lambda_i \mathbf{x}$. Or, d'après **i)** et **iii)**,

$$\mathbf{x} = \pi_1(\mathbf{x}) + \ldots + \pi_p(\mathbf{x}) \quad \text{et} \quad u(\mathbf{x}) = \lambda_1 \pi_1(\mathbf{x}) + \ldots + \lambda_p \pi_p(\mathbf{x}).$$

Donc

$$\lambda_i \mathbf{x} = \lambda_i(\pi_1(\mathbf{x}) + \ldots + \pi_p(\mathbf{x})) = u(\mathbf{x}) = \lambda_1 \pi_1(\mathbf{x}) + \ldots + \lambda_p \pi_p(\mathbf{x}). \tag{8.8}$$

D'après (8.7), on en déduit que

$$\lambda_i \pi_j(\mathbf{x}) = \lambda_j \pi_j(\mathbf{x}),$$

pour tout $j \in [\![1, p]\!]$. Soit $(\lambda_i - \lambda_j)\pi_j(\mathbf{x}) = \mathbf{0}$. Si $i \neq j$, on a $\lambda_i \neq \lambda_j$, d'où $\pi_j(\mathbf{x}) = \mathbf{0}$. Donc d'après (8.8), on a $\mathbf{x} = \pi_i(\mathbf{x})$, d'où $\mathbf{x} \in \text{Im } \pi_i$. □

§ 6 Calcul de la décomposition spectrale algébrique

Soit $\mathbf{A}$ une matrice de $\mathscr{M}_n(\mathbb{K})$ de spectre $\text{spec}(\mathbf{A}) = \{\lambda_1, \ldots, \lambda_p\}$. Étant donné un polynôme annulateur

$$g = (x - \lambda_1)^{n_1} \ldots (x - \lambda_p)^{n_p}$$

de $\mathbf{A}$, déterminons les projecteurs spectraux $\pi_{\lambda_1}, \ldots, \pi_{\lambda_p}$ de $\mathbf{A}$. En pratique, on prendra pour g le polynôme caractéristique $p_{\mathbf{A}}$ ou le polynôme minimal $m_{\mathbf{A}}$.

Étape 1 : On décompose la fraction rationnelle $\dfrac{1}{g}$ en éléments simples dans $\mathbb{K}(x)$:

$$\frac{1}{g} = \sum_{i=1}^{p} \sum_{j=1}^{n_i} \frac{a_{i,j}}{(x - \lambda_i)^j}.$$

Étape 2 : On pose

$$h_i = \sum_{j=1}^{n_i} a_{i,j}(x-\lambda_i)^{n_i-j} \quad \text{et} \quad g_i = \frac{1}{(x-\lambda_i)^{n_i}} g = \prod_{\substack{j=1 \\ j\neq i}}^{p} (x-\lambda_j)^{n_j}.$$

On a alors

$$\frac{1}{g} = \frac{h_1}{(x-\lambda_1)^{n_1}} + \ldots + \frac{h_p}{(x-\lambda_p)^{n_p}}.$$

Par suite,

$$1 = g_1 h_1 + \ldots + g_p h_p.$$

Étape 3 : On pose alors , pour tout $i \in [\![1,p]\!]$,

$$\Pi_{\lambda_i} = g_i(\mathbf{A}) h_i(\mathbf{A}).$$

On vérifie, qu'ainsi définis, pour tout i, Π_{λ_i} est la matrice du projecteur spectral π_{λ_i} de $\mathbb{K}^n$ sur N_{λ_i}, exprimée dans la base canonique de $\mathbb{K}^n$.

8.6.1. Exemple.— On considère la matrice $\mathbf{A}$ de $\mathscr{M}_3(\mathbb{R})$ définie par

$$\mathbf{A} = \begin{bmatrix} 3 & -1 & 1 \\ 2 & 0 & 1 \\ 1 & -1 & 2 \end{bmatrix}.$$

Son polynôme caractéristique est $p_{\mathbf{A}} = -(x-1)(x-2)^2$. Le polynôme $(x-1)(x-2)$ n'est pas annulateur de $\mathbf{A}$, donc le polynôme minimal de $\mathbf{A}$ est $m_{\mathbf{A}} = (x-1)(x-2)^2$ et $\mathbf{A}$ n'est pas diagonalisable.

Déterminons les projecteurs sur les espaces spectraux associés à la décomposition :

$$\mathbb{R}^3 = \text{Ker}\,(\mathbf{A} - \mathbf{1}_3) \oplus \text{Ker}\,(\mathbf{A} - 2\mathbf{1}_3)^2.$$

On considère le polynôme annulateur $g = (x-1)(x-2)^2$. On a :

$$\begin{aligned} \frac{1}{(x-1)(x-2)^2} &= \frac{1}{x-1} + \frac{-1}{x-2} + \frac{1}{(x-2)^2}, \\ &= \frac{1}{x-1} + \frac{-x+3}{(x-2)^2}. \end{aligned}$$

En posant

$$h_1 = 1,\ g_1 = (x-2)^2,\ h_2 = -x+3,\ g_2 = x-1,$$

on a

$$1 = g_1 h_1 + g_2 h_2.$$

La projection $\pi_1 : \mathbb{R}^3 \longrightarrow N_1$ sur l'espace spectral $N_1 = \text{Ker}\,(\mathbf{A} - \mathbf{1}_3)$ parallèlement à $N_2 = \text{Ker}\,(\mathbf{A} - 2\mathbf{1}_3)^2$ a pour matrice dans la base canonique de $\mathbb{K}^3$:

$$\Pi_1 = g_1(\mathbf{A}) h_1(\mathbf{A}) = (\mathbf{A} - 2\mathbf{1}_3)^2.$$

La projection $\pi_2 : \mathbb{R}^3 \longrightarrow N_2$ sur N_2 parallèlement à N_1 a pour matrice dans la base canonique de $\mathbb{K}^3$:

$$\Pi_2 = g_2(\mathbf{A})h_2(\mathbf{A}) = (-\mathbf{A}+3\mathbf{1}_3)(\mathbf{A}-\mathbf{1}_3).$$

Notons que l'on aurait pu aussi obtenir Π_2 par la relation :

$$\Pi_2 = \mathbf{1}_3 - \Pi_1,$$

soit

$$\Pi_2 = \mathbf{1}_3 - (\mathbf{A}-2\mathbf{1}_3)^2 = -\mathbf{A}^2 + 4\mathbf{A} - 3\mathbf{1}_3.$$

On a donc

$$\Pi_1 = \begin{bmatrix} 0 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & 1 & 0 \end{bmatrix}, \qquad \Pi_2 = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & -1 & 1 \end{bmatrix}.$$

On obtient

$$\mathbf{D} = \Pi_1 + 2\Pi_2 \quad \text{et} \quad \mathbf{N} = \mathbf{A} - \mathbf{D}$$

d'où

$$\mathbf{D} = \begin{bmatrix} 2 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & -1 & 2 \end{bmatrix} \quad \text{et} \quad \mathbf{N} = \begin{bmatrix} 1 & -1 & 1 \\ 1 & -1 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

La matrice $\mathbf{D}$ est diagonalisable (on peut vérifier que son polynôme minimal est $(x-2)(x-1)$) et la matrice $\mathbf{N}$ est nilpotente, on a $\mathbf{N}^2 = \mathbf{0}$.

8.6.2. Exemple.— Soit $\mathbf{A}$ la matrice de $\mathscr{M}_3(\mathbb{R})$ donnée par

$$\mathbf{A} = \begin{bmatrix} 1 & -4 & -4 \\ 8 & -11 & -8 \\ -8 & 8 & 5 \end{bmatrix}.$$

On calcule le polynôme caractéristique $p_{\mathbf{A}} = -(x-1)(x+3)^2$ et le polynôme minimal $m_{\mathbf{A}} = (x-1)(x+3)$. Ce dernier étant scindé et n'ayant que des racines simples, la matrice $\mathbf{A}$ est diagonalisable et il existe une décomposition :

$$\mathbb{R}^3 = E_1 \oplus E_{-3}.$$

Notons, Π_1 et Π_{-3} les matrices des projecteurs spectraux $\pi_1 : \mathbb{R}^3 \longrightarrow E_1$ et $\pi_{-3} : \mathbb{R}^3 \longrightarrow E_{-3}$ exprimées dans la base canonique de $\mathbb{R}^3$. Les matrices Π_1 et Π_{-3} vérifient les deux relations :

$$\left|\begin{array}{l} \mathbf{1}_3 = \Pi_1 + \Pi_{-3} \\ \mathbf{A} = \Pi_1 - 3\Pi_{-3} \end{array}\right.$$

qui s'écrivent sous la forme matricielle :

$$\begin{bmatrix} \mathbf{1}_3 \\ \mathbf{A} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -3 \end{bmatrix} \begin{bmatrix} \Pi_1 \\ \Pi_{-3} \end{bmatrix}$$

D'où

$$\begin{bmatrix} \Pi_1 \\ \Pi_{-3} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -3 \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{1}_3 \\ \mathbf{A} \end{bmatrix}$$

On en déduit :

$$\Pi_1 = \frac{1}{4}\mathbf{A} + \frac{3}{4}\mathbf{1}_3, \qquad \Pi_{-3} = -\frac{1}{4}\mathbf{A} + \frac{1}{4}\mathbf{1}_3. \tag{8.9}$$

—

§ 7 Exercices

Exercice 1.— Soit $\mathbf{A}$ la matrice de $\mathscr{M}_5(\mathbb{R})$ donnée par

$$\mathbf{A} = \begin{bmatrix} 2 & 1 & -1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & -1 & 1 \\ 0 & 0 & 0 & 2 & 0 \\ 1 & 0 & -1 & 0 & 2 \end{bmatrix}.$$

1. Déterminer les sous-espaces caractéristiques de $\mathbf{A}$.
2. Montrer que la matrice $\mathbf{A}$ est semblable à une matrice de la forme

$$\begin{bmatrix} \mathbf{T}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{T}_2 \end{bmatrix}$$

où $\mathbf{T}_1$ et $\mathbf{T}_2$ sont deux matrices triangulaires supérieures.
3. Exhiber la matrice de passage.

Exercice 2.— Soit $\mathbf{A}$ la matrice de $\mathscr{M}_3(\mathbb{R})$ définie par

$$\mathbf{A} = \begin{bmatrix} 5 & 1 & 3 \\ 4 & 3 & 4 \\ -1 & -1 & 1 \end{bmatrix}.$$

1. Déterminer en fonction de $\mathbf{A}$ les projecteurs spectraux de $\mathbf{A}$.
2. Exprimer, pour tout entier $k \geqslant 0$, la matrice $\mathbf{A}^k$ en fonction de $\mathbf{A}$.
3. Calculer la matrice de $\mathbf{A}^k$.
4. Répondre aux mêmes questions avec les matrices suivantes :

$$\mathbf{B} = \begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{bmatrix}, \qquad \mathbf{C} = \begin{bmatrix} -1 & 1 & 3 \\ -2 & 2 & 2 \\ -2 & 1 & 4 \end{bmatrix}.$$

Exercice 3.— On considère la matrice de $\mathscr{M}_4(\mathbb{R})$ suivante

$$\mathbf{A} = \begin{bmatrix} 1 & a^2 & a^2 & a \\ 1 & 1 & a & 1 \\ 1 & a & 1 & 1 \\ a & a^2 & a^2 & 1 \end{bmatrix}.$$

1. Déterminer le rang de la matrice $\mathbf{A} - (1-a)\mathbf{1}_4$. En déduire que $1-a$ est valeur propre de $\mathbf{A}$.
2. Si $a = 0$, la matrice $\mathbf{A}$ est-elle diagonalisable ?

   Dans la suite de l'exercice, on suppose que le réel a est non nul.
3. Déterminer toutes les valeurs propres de $\mathbf{A}$. La matrice $\mathbf{A}$ est-elle diagonalisable ?
4. Déterminer le polynôme caractéristique et le polynôme minimal de $\mathbf{A}$.
5. Exprimer en fonction de $\mathbf{A}$ les projecteurs spectraux de $\mathbf{A}$.

Exercice 4.— Soient E un $\mathbb{K}$-espace vectoriel de dimension finie $n \geqslant 1$ et u un endomorphisme de E.

1. Montrer que si λ est une racine d'ordre k du polynôme minimal m_u, i.e., $m_u = (x-\lambda)^k g$ avec $g(\lambda) \neq 0$, alors
$$\mathrm{Ker}\, g(u) = \mathrm{Im}\, (u - \lambda \mathrm{id}_E)^k.$$
2. Sous les mêmes hypothèses, montrer que
$$E = \mathrm{Ker}\, (u - \lambda \mathrm{id}_E)^k \oplus \mathrm{Im}\, (u - \lambda \mathrm{id}_E)^k.$$
3. Soit λ une valeur propre de u telle que
$$E = \mathrm{Ker}\, (u - \lambda \mathrm{id}_E) \oplus \mathrm{Im}\, (u - \lambda \mathrm{id}_E).$$
   Montrer que le polynôme $P = (x-\lambda)m'$, où m' est le polynôme minimal de la restriction de u au sous-espace $\mathrm{Im}\, (u - \lambda \mathrm{id}_E)$, est annulateur de u. En déduire que λ est racine simple de m_u.
4. Montrer que les deux égalités suivantes sont équivalents
$$E = \mathrm{Ker}\, (u - \lambda \mathrm{id}_E) \oplus \mathrm{Im}\, (u - \lambda \mathrm{id}_E), \qquad \mathrm{Ker}\, (u - \lambda \mathrm{id}_E) = \mathrm{Ker}\, (u - \lambda \mathrm{id}_E)^2.$$
5. Montrer que si u est diagonalisable, alors pour toute valeur propre λ on a :
$$\mathrm{Ker}\, (u - \lambda \mathrm{id}_E) = \mathrm{Ker}\, (u - \lambda \mathrm{id}_E)^2.$$

Exercice 5.— Soit u un endomorphisme d'un $\mathbb{K}$-espace vectoriel E de dimension finie, dont le polynôme caractéristique est scindé. Montrer que u est diagonalisable si et seulement si, pour tout scalaire $\lambda \in \mathbb{K}$,

$$\mathrm{rang}(u - \lambda \mathrm{id}_E) = \mathrm{rang}(u - \lambda \mathrm{id}_E)^2.$$

8.7.1. Calcul des projecteurs spectraux avec Sage.— Soit $\mathbf{A}$ une matrice de $\mathscr{M}_n(\mathbb{K})$. D'après le thèorème de décomposition en noyaux, si P_1 et P_2 sont deux polynômes premiers entre eux et tels que leur produit P_1P_2 soit annulateur de $\mathbf{A}$, alors on a une décomposition de $\mathbb{K}^n$ en somme directe :

$$\mathbb{K}^n = \operatorname{Ker} P_1(\mathbf{A}) \oplus \operatorname{Ker} P_2(\mathbf{A}).$$

L'objectif des exercices suivants est de calculer les projections de $\mathbb{K}^n$ sur les sous-espaces $\operatorname{Ker} P_1(\mathbf{A})$ et $\operatorname{Ker} P_2(\mathbf{A})$. On va pour cela exprimer les matrices de ces projections en fonction de la matrice $\mathbf{A}$.

D'après le théorème de Bézout, si f_1 et f_2 sont premiers entre eux, il existe des polynômes h_1 et h_2 tels que

$$f_1h_1 + f_2h_2 = 1.$$

Pour construire des polynômes à coefficients rationnels en une indéterminée :

```
sage: R = PolynomialRing(QQ,'x')
...   R
Univariate Polynomial Ring in x over Rational Field
```

Par exemple :

```
sage: f = x^3 + 2*x^2 + x + 1
...   print f in R
...   A = Matrix([[1,1,0],[0,1,0],[0,0,2]])
...   print A.charpoly() in R
True
True
```

La méthode `gcd()` retourne le pgcd de deux polynômes :

```
sage : x = PolynomialRing(QQ, 'x').gen()
...    f = (x-1)*(x+2)
...    g = x+1
...    f.gcd(g)
1
```

La méthode `xgcd` permet de calculer les coefficients de Bézout pour un couple de polynômes :

```
sage: d,u,v = xgcd(f,g)
...   print d,u,v
...   print f*u+g*v == d
-2 1 -x
True
```

Exercice (Sage) 6.— On considère la matrice suivante

$$\mathbf{B} = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}.$$

1. Calculer le polynôme caractéristique et le polynôme minimal de $\mathbf{B}$.
2. Notons λ_1 et λ_2 les deux valeurs propres de $\mathbf{B}$. Déterminer les sous-espaces propres E_{λ_1} et E_{λ_2}.
3. Calculer les projecteurs spectraux de la matrice $\mathbf{B}$.

4. Calculer les projecteurs spectraux des matrices suivantes

$$\mathbf{A} = \begin{bmatrix} 5 & 1 & 3 \\ 4 & 3 & 4 \\ -1 & -1 & 1 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} -1 & 1 & 3 \\ -2 & 2 & 2 \\ -2 & 1 & 4 \end{bmatrix}.$$

Exercice (Sage) **7.—** L'objectif est de construire une procédure `projecteursSpectraux(A)` qui prend en argument une matrice complexe **A** et retourne une liste

$$[[\lambda_1, n_1, \Pi_1], \ldots, [\lambda_p, n_p, \Pi_p]],$$

où λ_1, ..., λ_p sont les valeurs propres complexes de la matrice **A**, n_i est l'ordre de multiplicité algébrique de la valeur propre λ_i et Π_i est la matrice de la projection sur le sous-espace caractéristique $N_{\lambda_i} = \mathrm{Ker}\,(\mathbf{A} - \lambda_i \mathbf{1}_n)^{n_i}$, parallèlement au sous-espace $\bigoplus_{\substack{j=1 \\ i \neq j}}^{p} N_{\lambda_j}$.

On utilisera l'algorithme de la section 6. Considérons un polynôme annulateur g de $\mathbf{A} \in \mathscr{M}_n(\mathbb{C})$, par exemple le polynôme minimal de **A**. Le polynôme g est scindé sur $\mathbb{C}$:

$$g = (x - \lambda_1)^{l_1} \ldots (x - \lambda_p)^{l_p}.$$

En posant, pour tout $i \in [\![1, p]\!]$,

$$g_i = \prod_{\substack{j=1 \\ j \neq i}}^{p} (x - \lambda_j)^{l_j}, \tag{8.10}$$

les polynômes $g_1, \ldots, g_p$ sont premiers entre eux dans leur ensemble. On calcule les polynômes h_i de la décomposition de Bézout, i.e., vérifiant

$$g_1 h_1 + \ldots + g_p h_p = 1.$$

Nous avons montré qu'alors, la matrice Π_i de la projection sur N_{λ_i} est donnée par $\Pi_i = g_i(\mathbf{A}) h_i(\mathbf{A})$.

1. En utilisant la procédure `gcdex`, écrire une procédure `decompositionBezout(liste,x)` qui prend en premier argument une liste de polynômes $[g_1, ..., g_k]$ en l'indéterminée spécifiée en le deuxième argument et retourne la liste $[g, [h_1, ..., h_k]]$, où g est le pgcd des polynômes g_1, ..., g_k et les h_i vérifient

$$g_1 h_1 + \ldots + g_k h_k = g.$$

On pourra utiliser les procédures `collect`, `normal`, `seq`, `op`, `nops`.

2. Écrire une procédure `decomposePolyMin(A,x)` qui prend en argument une matrice complexe **A** et retourne une liste $[[\lambda_1, k_1, g_1], \ldots, [\lambda_p, k_p, g_p]]$, où les λ_i sont les valeurs propres de **A**, les k_i leur ordre de multiplicité dans le polynôme minimal et les g_i sont les polynômes en l'indéterminée `x`, comme définis dans (8.10) lorsque g est le polynôme minimal de **A**. On pourra utiliser les procédures `roots`, `factor`.

3. En utilisant les deux procédures ci-dessus, écrire la procédure `projecteursSpectraux`. On pourra utiliser la procédure `MatrixFunction`.

4. Tester la procédure avec les matrices suivantes, déjà rencontrées à plusieurs reprises,

$$\mathbf{A} = \begin{bmatrix} 5 & 1 & 3 \\ 4 & 3 & 4 \\ -1 & -1 & 1 \end{bmatrix}, \qquad \mathbf{B} = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}, \qquad \mathbf{C} = \begin{bmatrix} -1 & 1 & 3 \\ -2 & 2 & 2 \\ -2 & 1 & 4 \end{bmatrix}.$$

Troisième partie .

Applications de la réduction des matrices

CHAPITRE
9

Fonctions de matrices

Sommaire

Dans le chapitre 8, nous avons montré que toute matrice $\mathbf{A}$ de $\mathscr{M}_n(\mathbb{K})$ dont le polynôme caractéristique est scindé se décompose en une somme

$$\mathbf{A} = \mathbf{D} + \mathbf{N},$$

formée d'une matrice $\mathbf{D}$ diagonalisable et d'une matrice $\mathbf{N}$ nilpotente qui commutent.

Nous appliquons dans ce chapitre cette décomposition au calcul de puissances et d'-exponentielles de matrices, avec pour objectif l'étude des systèmes d'évolution discrets et les systèmes d'équations différentielles linéaire à coefficients constants.

§ 1 Calcul des puissances d'une matrice

9.1 Proposition.— Soit $\mathbf{A}$ une matrice de $\mathscr{M}_n(\mathbb{K})$ dont le polynôme caractéristique est scindé :

$$p_{\mathbf{A}} = (-1)^n (x - \lambda_1)^{n_1} \dots (x - \lambda_p)^{n_p}.$$

Soient $\Pi_{\lambda_1}, \dots, \Pi_{\lambda_p}$ les projecteurs spectraux de la matrice $\mathbf{A}$.

i) Si $\mathbf{A}$ est diagonalisable, alors, pour tout entier naturel $k > 0$, on a

$$\mathbf{A}^k = \lambda_1^k \Pi_{\lambda_1} + \dots + \lambda_p^k \Pi_{\lambda_p}. \tag{9.1}$$

ii) Si $\mathbf{A}$ n'est pas diagonalisable, alors, pour tout entier naturel $k > 0$, on a

$$\mathbf{A}^k = \sum_{i=1}^{p} \sum_{j=0}^{k_i-1} \binom{k}{j} \lambda_i^{k-j} (\mathbf{A} - \lambda_i \mathbf{1}_n)^j \Pi_{\lambda_i}, \tag{9.2}$$

où k_i est l'ordre de multiplicité de la valeur propre λ_i dans le polynôme minimal de $\mathbf{A}$.

Preuve. Montrons l'assertion **i)**. D'après la proposition 8.7, on a

$$\mathbf{A} = \lambda_1 \Pi_{\lambda_1} + \ldots + \lambda_p \Pi_{\lambda_p},$$

avec $\Pi_{\lambda_i}\Pi_{\lambda_j} = 0$, si $i \neq j$, et $\Pi_{\lambda_i}^2 = \Pi_{\lambda_i}$, pour tout i. On en déduit la relation (9.1).

Montrons l'assertion **ii)**. D'après le théorème de décomposition spectrale algébrique, théorème 8.6, il existe une décomposition

$$\mathbf{A} = \mathbf{D} + \mathbf{N},$$

avec

$$\mathbf{D} = \lambda_1 \Pi_{\lambda_1} + \ldots + \lambda_p \Pi_{\lambda_p} \qquad \text{et} \qquad \mathbf{N} = (\mathbf{A} - \lambda_1 \mathbf{1}_n)\Pi_{\lambda_1} + \ldots + (\mathbf{A} - \lambda_p \mathbf{1}_n)\Pi_{\lambda_p}.$$

Dans la décomposition de la matrice $\mathbf{N}$, la matrice $(\mathbf{A} - \lambda_i \mathbf{1}_n)\Pi_{\lambda_i}$ est nilpotent d'indice k_i, l'ordre de multiplicité de la racine λ_i dans le polynôme minimal de $\mathbf{A}$.

Les projecteurs spectraux satisfont $\Pi_{\lambda_i}\Pi_{\lambda_j} = 0$, si $i \neq j$, et $\Pi_{\lambda_i}^2 = \Pi_{\lambda_i}$, pour tout i. Par suite, pour tout entier k, on a

$$\mathbf{D}^k = \sum_{i=1}^{p} \lambda_i^k \Pi_{\lambda_i} \qquad \text{et} \qquad \mathbf{N}^k = \sum_{i=1}^{p} (\mathbf{A} - \lambda_i \mathbf{1}_n)^k \Pi_{\lambda_i}.$$

Les matrices $\mathbf{D}$ et $\mathbf{N}$ commutent, i.e., $\mathbf{DN} = \mathbf{ND}$, donc d'après la formule du binôme, on a :

$$\mathbf{A}^k = (\mathbf{D} + \mathbf{N})^k = \sum_{j=0}^{k} \binom{k}{j} \mathbf{D}^{k-j} \mathbf{N}^j. \tag{9.3}$$

Par suite,

$$\begin{aligned}
\mathbf{A}^k &= \sum_{j=0}^{k} \sum_{i=1}^{p} \binom{k}{j} \lambda_i^{k-j} (\mathbf{A} - \lambda_i \mathbf{1}_n)^j \Pi_{\lambda_i}, \\
&= \sum_{i=1}^{p} \sum_{j=0}^{k_i-1} \binom{k}{j} \lambda_i^{k-j} (\mathbf{A} - \lambda_i \mathbf{1}_n)^j \Pi_{\lambda_i}.
\end{aligned}$$

D'où la relation 9.2. □

9.1.1. Exemple.— Soit $\mathbf{A}$ la matrice de $\mathscr{M}_3(\mathbb{R})$ donnée par

$$\mathbf{A} = \begin{bmatrix} 1 & -4 & -4 \\ 8 & -11 & -8 \\ -8 & 8 & 5 \end{bmatrix}.$$

Nous avons vu en 8.6.2, que la matrice $\mathbf{A}$ est diagonalisable et de spectre $\mathrm{Sp}(\mathbf{A}) = \{-3, 1\}$. On a la décomposition spectrale :

$$\mathbf{A} = \Pi_1 - 3\Pi_{-3},$$

d'où, pour entier naturel k,

$$\mathbf{A}^k = \Pi_1 + (-3)^k \Pi_{-3}.$$

D'après l'expression (8.9) des projecteurs en fonction de $\mathbf{A}$, on déduit que :

$$\mathbf{A}^k = \frac{1}{4}(1-(-3)^k)\mathbf{A} + \frac{1}{4}(3+(-3)^k)\mathbf{1}_3.$$

9.1.2. Exemple.— Considérons la matrice $\mathbf{A}$ de $\mathscr{M}_3(\mathbb{R})$ définie par

$$\mathbf{A} = \begin{bmatrix} 3 & -1 & 1 \\ 2 & 0 & 1 \\ 1 & -1 & 2 \end{bmatrix}.$$

On a vu en 8.6.1 que la matrice $\mathbf{A}$, de spectre $\mathrm{Sp}(\mathbf{A}) = \{1, 2\}$, n'est pas diagonalisable. Il existe une décomposition

$$\mathbf{A} = \mathbf{D} + \mathbf{N}$$

en la somme d'une matrice diagonalisable et d'une matrice nilpotent, donnés par

$$\mathbf{D} = \Pi_1 + 2\Pi_2, \qquad \mathbf{N} = \mathbf{A} - \mathbf{D}.$$

On a $\mathbf{D}^k = \Pi_1 + 2^k\Pi_2$ et $\mathbf{N}^2 = \mathbf{0}$. Comme les matrices $\mathbf{D}$ et $\mathbf{N}$ commutent, et que $\mathbf{N}^2 = \mathbf{0}$, d'après la formule du binôme (9.3), on a :

$$\mathbf{A}^k = \binom{k}{0}\mathbf{D}^k + \binom{k}{1}\mathbf{D}^{k-1}\mathbf{N}.$$

D'où

$$\begin{aligned}\mathbf{A}^k &= \Pi_1 + 2^k\Pi_2 + k(\Pi_1 + 2^{k-1}\Pi_2)(\mathbf{A} - \Pi_1 - 2\Pi_2), \\ &= (1-k)\Pi_1 + k\mathbf{A}\Pi_1 + 2^k\Pi_2 + k2^{k-1}\mathbf{A}\Pi_2 - k2^k\Pi_2.\end{aligned}$$

soit

$$\mathbf{A}^k = ((1-k)\mathbf{1}_3 + k\mathbf{A})\Pi_1 + (2^k(1-k)\mathbf{1}_3 + k2^{k-1}\mathbf{A})\Pi_2.$$

§ 2 La fonction exponentielle

9.2.1. De l'exponentielle scalaire à l'exponentielle matricielle.— L'exponentielle réelle ou complexe peut se définir en utilisant le développement en série, pour tout réel ou complexe x, on a :

$$e^x = \sum_{k=0}^{+\infty} \frac{1}{k!} x^k.$$

Dans l'expression précédent, si l'on remplace formellement le scalaire x par une matrice $\mathbf{A}$ de $\mathscr{M}_n(\mathbb{K})$, on obtient une série de matrices :

$$\sum_{k=0}^{+\infty} \frac{1}{k!}\mathbf{A}^k = \mathbf{1}_n + \mathbf{A} + \frac{1}{2!}\mathbf{A}^2 + \frac{1}{3!}\mathbf{A}^3 + \ldots$$

On montre que cette série est convergente, sa valeur est appelée l'exponentielle de la matrice $\mathbf{A}$ et est noté $e^{\mathbf{A}}$. Dans ce cours, la preuve de la convergence de cette série de matrices sera admise.

On notera cependant que la convergence de la série $\displaystyle\sum_{k=0}^{+\infty} \frac{1}{k!}\mathbf{A}^k$ est une conséquence de la version scalaire de cette série. En effet, supposons que la matrice $\mathbf{A}$ soit diagonalisable, il existe une décomposition

$$\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^{-1}, \quad \text{avec} \quad \mathbf{D} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_p \end{bmatrix}$$

et où $\mathbf{P}$ est une matrice inversible. On a $\mathbf{A}^k = \mathbf{P}\mathbf{D}^k\mathbf{P}^{-1}$, d'où

$$\begin{aligned} \sum_{k=0}^{+\infty} \frac{1}{k!}\mathbf{A}^k &= \sum_{k=0}^{+\infty} \frac{1}{k!}\mathbf{P}\mathbf{D}^k\mathbf{P}^{-1} \\ &= \mathbf{P}\left(\sum_{k=0}^{+\infty} \frac{1}{k!}\mathbf{D}^k\right)\mathbf{P}^{-1} \\ &= \mathbf{P}\begin{bmatrix} e^{\lambda_1} & 0 & \cdots & 0 \\ 0 & e^{\lambda_2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & e^{\lambda_p} \end{bmatrix}\mathbf{P}^{-1} \end{aligned}$$

Ceci constitue une autre façon de définir l'exponentielle d'une matrice diagonalisable

A ; on pose

$$e^{\mathbf{A}} = \mathbf{P}\begin{bmatrix} e^{\lambda_1} & 0 & \cdots & 0 \\ 0 & e^{\lambda_2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & e^{\lambda_p} \end{bmatrix}\mathbf{P}^{-1}. \tag{9.4}$$

Notons que cette définition est indépendante du choix des vecteurs de la base de diagonalisation, i.e., indépendante de la matrice **P**. En effet, on a la décomposition spectrale, théorème 8.6,

$$\mathbf{A} = \lambda_1\Pi_{\lambda_1} + \ldots \lambda_p\Pi_{\lambda_p}.$$

En utilisant le même raisonnement, on montre que

$$e^{\mathbf{A}} = e^{\lambda_1}\Pi_{\lambda_1} + \ldots e^{\lambda_p}\Pi_{\lambda_p}.$$

La définition (9.4) est donc indépendante du choix des vecteurs propres, i.e., de la matrice **P** choisie.

9.2.2. Exponentielle d'une matrice.— Les preuves des deux résultats suivants sont admises.

9.2 Proposition.— Pour toute matrice **A** de $\mathscr{M}_n(\mathbb{K})$, la série

$$\sum_{k=0}^{+\infty} \frac{1}{k!}\mathbf{A}^k = \mathbf{1}_n + \mathbf{A} + \frac{1}{2!}\mathbf{A}^2 + \frac{1}{3!}\mathbf{A}^3 + \ldots,$$

est normalement convergente dans $\mathscr{L}(E)$.

9.3 Proposition.— La fonction exponentielle vérifie les propriétés suivantes :

i) pour toute matrice **A** de $\mathscr{M}_n(\mathbb{K})$ et toute matrice inversible **P** de $\mathscr{M}_n(\mathbb{K})$, on a

$$e^{\mathbf{P}^{-1}\mathbf{A}\mathbf{P}} = \mathbf{P}^{-1}e^{\mathbf{A}}\mathbf{P}.$$

ii) pour toute matrice **A** de $\mathscr{M}_n(\mathbb{K})$, la matrice $e^{\mathbf{A}}$ est inversible et

$$(e^{\mathbf{A}})^{-1} = e^{-\mathbf{A}}.$$

Preuve. Montrons **i)**. On a

$$e^{\mathbf{P}^{-1}\mathbf{A}\mathbf{P}} = \sum_{k=0}^{+\infty} \frac{1}{k!}(\mathbf{P}^{-1}\mathbf{A}\mathbf{P})^k = \sum_{k=0}^{+\infty} \frac{1}{k!}\mathbf{P}^{-1}\mathbf{A}^k\mathbf{P} = \mathbf{P}^{-1}e^{\mathbf{A}}\mathbf{P}.$$

L'assertion **ii)** découle de la proposition précédente. En effet, on a :

$$e^{\mathbf{A}}e^{-\mathbf{A}} = e^{-\mathbf{A}}e^{\mathbf{A}} = e^{\mathbf{A}-\mathbf{A}} = e^{\mathbf{0}} = \mathbf{1}_n.$$

□

9.2.3. Décomposition spectrale et exponentielle d'une matrice.— La décomposition spectrale des matrices nous permet d'exprimer l'exponentielle d'une matrice en terme de ses projecteurs spectraux.

9.4 Proposition.— Soit $\mathbf{A}$ une matrice de $\mathscr{M}_n(\mathbb{K})$ dont le polynôme caractéristique est scindé :

$$p_{\mathbf{A}} = (-1)^n(x-\lambda_1)^{n_1}\dots(x-\lambda_p)^{n_p}.$$

i) Si $\mathbf{A}$ est diagonalisable, alors

$$e^{\mathbf{A}} = e^{\lambda_1}\Pi_{\lambda_1} + \dots + e^{\lambda_p}\Pi_{\lambda_p}, \tag{9.5}$$

où les Π_{λ_i} désignent les projecteurs spectraux de $\mathbf{A}$.

ii) Si $\mathbf{A}$ n'est pas diagonalisable, alors

$$e^{\mathbf{A}} = \sum_{i=1}^{p}\sum_{k=0}^{k_i-1}\frac{e^{\lambda_i}}{k!}(\mathbf{A}-\lambda_i\mathbf{1}_n)^k\Pi_{\lambda_i}, \tag{9.6}$$

où les Π_{λ_i} désignent les projecteurs spectraux de $\mathbf{A}$.

Preuve. Montrons l'assertion **i)**. Supposons que $\mathbf{A}$ soit diagonalisable, d'après le théorème 8.6, on a

$$\mathbf{A} = \lambda_1\Pi_{\lambda_1} + \dots + \lambda_p\Pi_{\lambda_p}.$$

Pour tout $i \in [\![1,p]\!]$, on a $\Pi_{\lambda_i}^2 = \Pi_{\lambda_i}$ et pour tous $i, j \in [\![1,p]\!]$ tels que $i \neq j$, on a $\Pi_{\lambda_i}\Pi_{\lambda_j} = 0$. Par suite,

$$e^{\mathbf{A}} = \sum_{k=0}^{+\infty}\frac{1}{k!}\mathbf{A}^k = \sum_{k=0}^{+\infty}\sum_{i=1}^{p}\frac{\lambda_i^k}{k!}\Pi_{\lambda_i} = \sum_{i=1}^{p}\sum_{k=0}^{+\infty}\frac{\lambda_i^k}{k!}\Pi_{\lambda_i},$$

d'où la relation (9.5).

Montrons l'assertion **ii)**. Supposons que $\mathbf{A}$ soit non diagonalisable, d'après le théorème 8.6, il existe une décomposition

$$\mathbf{A} = \mathbf{D} + \mathbf{N},$$

avec

$$\mathbf{D} = \lambda_1\Pi_{\lambda_1} + \dots + \lambda_p\Pi_{\lambda_p}, \qquad \mathbf{N} = (\mathbf{A}-\lambda_1\mathbf{1}_n)\Pi_{\lambda_1} + \dots + (\mathbf{A}-\lambda_p\mathbf{1}_n)\Pi_{\lambda_p},$$

et où, pour tout $i \in [\![1,p]\!]$, la matrice $(\mathbf{A}-\lambda_i\mathbf{1}_n)\Pi_{\lambda_i}$ est nilpotente, d'indice de nilpotence l'ordre de multiplicité k_i de la racine λ_i dans le polynôme minimal de $\mathbf{A}$.

D'après **i)**, on a

$$e^{\mathbf{D}} = \sum_{i=1}^{p} e^{\lambda_i}\Pi_{\lambda_i}.$$

D'autre part, on a :

$$\begin{aligned} e^{\mathbf{N}} = \sum_{k=0}^{+\infty}\frac{1}{k!}\mathbf{N}^k &= \sum_{k=0}^{+\infty}\sum_{i=1}^{p}\frac{1}{k!}(\mathbf{A}-\lambda_i\mathbf{1}_n)^k\Pi_{\lambda_i},\\ &= \sum_{i=1}^{p}\sum_{k=0}^{k_i-1}\frac{1}{k!}(\mathbf{A}-\lambda_i\mathbf{1}_n)^k\Pi_{\lambda_i}. \end{aligned}$$

Enfin, comme les matrices **D** et **N** commutent, on en déduit la relation (9.6) :

$$e^{\mathbf{A}} = e^{\mathbf{D}}e^{\mathbf{N}} = \sum_{i=1}^{p}\sum_{k=0}^{k_i-1}\frac{e^{\lambda_i}}{k!}(\mathbf{A}-\lambda_i\mathbf{1}_n)^k\Pi_{\lambda_i}.$$

□

9.2.4. Fonction d'une matrice.— Plus généralement, on peut s'inspirer des sections précédentes pour définir la fonction $f(\mathbf{A})$ d'une matrice **A**, où f est une fonction définie sur les valeurs propres de **A**.

Soit **A** une matrice de $\mathscr{M}_n(\mathbb{K})$ dont le polynôme caractéristique est scindé :

$$p_{\mathbf{A}} = (-1)^n(x-\lambda_1)^{n_1}\ldots(x-\lambda_p)^{n_p}.$$

Soit $f:\mathbb{K}\longrightarrow\mathbb{K}$ une fonction définie sur les valeurs propres de **A**. On définit $f(\mathbf{A})$ comme la matrice

$$f(\mathbf{A}) = f(\lambda_1)\Pi_{\lambda_1}+\ldots+f(\lambda_p)\Pi_{\lambda_p},$$

où les Π_{λ_i} désignent les projecteurs spectraux de **A**.

—

§ 3 Exercices

Exercice 1.— Soit **A** une matrice de $\mathscr{M}_n(\mathbb{C})$.

1. Montrer que

$$\det e^{\mathbf{A}} = e^{\operatorname{trace}(\mathbf{A})}.$$

2. Montrer que si **A** est nilpotent, alors

$$\operatorname{Ker}(e^{\mathbf{A}}-\mathbf{1}_n) = \operatorname{Ker}\mathbf{A}.$$

3. Calculer l'exponentielle des matrices suivantes de $\mathcal{M}_2(\mathbb{C})$:

$$\begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}, \quad \begin{bmatrix} \lambda & \theta \\ 0 & \lambda \end{bmatrix}, \quad \begin{bmatrix} \lambda & 0 \\ \theta & \lambda \end{bmatrix}, \quad \begin{bmatrix} \lambda & \theta \\ \theta & \lambda \end{bmatrix}, \quad \begin{bmatrix} \lambda & -\theta \\ \theta & \lambda \end{bmatrix}.$$

Exercice 2.— Soit $\mathbf{A}$ une matrice de $\mathcal{M}_n(\mathbb{K})$ vérifiant

$$(\mathbf{A}-\mathbf{1}_n)^2(\mathbf{A}-2\mathbf{1}_n)=\mathbf{0}.$$

1. Montrer que

$$\mathbb{K}^n = \mathrm{Ker}\,(\mathbf{A}-\mathbf{1}_n)^2 \oplus \mathrm{Ker}\,(\mathbf{A}-2\mathbf{1}_n).$$

Soient Π_2 la matrice de la projection de $\mathbb{K}^n$ sur Ker $(\mathbf{A}-2\mathbf{1}_n)$ parallèlement à Ker $(\mathbf{A}-\mathbf{1}_n)^2$ et Π_1 la matrice de la projection sur Ker $(\mathbf{A}-\mathbf{1}_n)^2$ parallèlement à Ker $(\mathbf{A}-2\mathbf{1}_n)$ exprimées dans la base canonique de $\mathbb{K}^n$.

2. Établir les relations suivantes

$$e^{\mathbf{A}}\Pi_1 = e^2\Pi_1, \qquad e^{\mathbf{A}}\Pi_2 = e\mathbf{A}\Pi_2.$$

3. Exprimer en fonction de $\mathbf{A}$ les matrices Π_1 et Π_2.

4. En déduire une expression de $e^{\mathbf{A}}$ en fonction de $\mathbf{A}$.

Exercice 3.— Soit $\mathbf{A}$ la matrice de $\mathcal{M}_3(\mathbb{R})$ définie par

$$\mathbf{A} = \begin{bmatrix} 5 & 1 & 3 \\ 4 & 3 & 4 \\ -1 & -1 & 1 \end{bmatrix}.$$

1. Exprimer, pour tout entier $k \geqslant 0$, la matrice $\mathbf{A}^k$ en fonction de $\mathbf{A}$.

2. Exprimer, pour tout réel t, la matrice $e^{t\mathbf{A}}$ en fonction de $\mathbf{A}$.

3. Répondre aux mêmes questions avec les matrices suivantes

$$\mathbf{B} = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}, \qquad \mathbf{C} = \begin{bmatrix} -1 & 1 & 3 \\ -2 & 2 & 2 \\ -2 & 1 & 4 \end{bmatrix}.$$

Exercice 4.— On considère la matrice de $\mathcal{M}_{n+1}(\mathbb{R})$ suivante

$$\mathbf{A} = \begin{bmatrix} 1 & \cdots & 1 & -n \\ \vdots & & \vdots & \vdots \\ 1 & \cdots & 1 & -n \end{bmatrix}$$

dont les n premières colonnes sont égales.

1. Calculer le rang de $\mathbf{A}$.

2. Montrer que toutes les valeurs propres de $\mathbf{A}$ sont égales.

3. La matrice $\mathbf{A}$ est-elle diagonalisable ?

4. Calculer $e^{t\mathbf{A}}$, pour tout réel t.

5. Déterminer trois fonctions réelles $\alpha(t)$, $\beta(t)$ et $\gamma(t)$ telles que

$$\alpha'(t) = \beta'(t) = \gamma'(t) = \alpha(t) + \beta(t) - 2\gamma(t),$$

et $\alpha(0) = 0$, $\beta(0) = 1$ et $\gamma(0) = 2$.

6. Montrer que si $\mathbf{N} \in \mathscr{M}_n(\mathbb{R})$ est nilpotente, les solutions du système linéaire

$$\mathbf{x}(t)' = \mathbf{N}\mathbf{x}(t)$$

sont des polynômes de degré inférieur ou égal à n.

Exercice 5.— On considère dans $\mathscr{M}_n(\mathbb{R})$, $n \geqslant 2$, la matrice

$$\mathbf{A} = \begin{bmatrix} a & b & \dots & b \\ b & a & \ddots & \vdots \\ \vdots & \ddots & \ddots & b \\ b & \dots & b & a \end{bmatrix}.$$

1. Déterminer le polynôme minimal de $\mathbf{A}$.
2. Montrer que $\mathbf{A}$ est diagonalisable.
3. Donner une condition nécessaire et suffisante sur a et b pour que $\mathbf{A}$ soit inversible. Dans le cas où $\mathbf{A}$ est inversible, calculer l'inverse de $\mathbf{A}$.
4. Calculer $\mathbf{A}^k$, pour tout entier k.
5. Calculer $e^{t\mathbf{A}}$, pour tout réel t.

Exercice 6.— On considère la matrice suivante de $\mathscr{M}_4(\mathbb{R})$

$$\mathbf{A} = \begin{bmatrix} 1 & a^2 & a^2 & a \\ 1 & 1 & a & 1 \\ 1 & a & 1 & 1 \\ a & a^2 & a^2 & 1 \end{bmatrix}$$

étudiée dans l'exercice 3. On suppose que a est non nul.

Exprimer les matrices $\mathbf{A}^k$, pour tout entier naturel k, et $e^{\mathbf{A}}$ en fonction de la matrice $\mathbf{A}$.

Exercice $^{(\mathsf{Sage})}$ **7.—** On considère la matrice

$$\mathbf{A} = \begin{bmatrix} 5 & 2 & 2 & 2 & 2 \\ 2 & 5 & 2 & 2 & 2 \\ 2 & 2 & 5 & 2 & 2 \\ 2 & 2 & 2 & 5 & 2 \\ 2 & 2 & 2 & 2 & 5 \end{bmatrix}.$$

1. Sans utiliser la commande `**` sur les objets de type `Matrix`, calculer les puissances $\mathbf{A}^n$, $n \in \mathbb{N}$, de trois façons différentes :

- en utilisant les projecteurs spectraux,

- en exploitant le fait que le polynôme minimal de la matrice $\mathbf{A}$ est de degré 2,

- en exploitant le fait que la matrice $\mathbf{A}$ est diagonalisable. (La méthode `eigenmatrix_right()` retourne la matrice diagonale et une matrice de passage d'une matrice diagonalisable).

2. Écrire une fonction Sage qui prend une matrice `A` en argument et un entier `n` et retourne la puissance `A**n` en utilisant la troisième méthode.

Exercice $^{(\mathsf{Sage})}$ **8 (Combien de chemins mènent à Rome ?).—** Voici un problème exrait du très jolie livre de Pierre Gabriel, [Gab01],

$$\begin{tikzcd} Z \arrow[rr, shift right] \arrow[dr] & & W \arrow[ll, shift right] \\ & R \arrow[ur] & \end{tikzcd}$$

« Un pilote vole tous les jours vers une des trois villes R, W, Z. La figure ci-dessus indique les trajets journaliers autorisés : un trajet de R vers W, un trajet de W vers Z, un de Z vers R et un de Z vers W. À côté de ces trajets journaliers (de « longueur 1 »), nous considérons les itinéraires (composés) de longueur n, qui décrivent les chemins possibles parcourus pendant n jours consécutifs. Ainsi deux itinéraires de longueur 4 *mènent de Z à W, alors qu'il n'y en a qu'un de Z à R :*

$$Z \longrightarrow R \longrightarrow W \longrightarrow Z \longrightarrow W$$
$$Z \longrightarrow W \longrightarrow Z \longrightarrow R \longrightarrow W$$
$$Z \longrightarrow R \longrightarrow W \longrightarrow Z \longrightarrow R$$

Cherchons le nombre d'itinéraires de longueur 100 *de Z vers R. Pour cela, nous numérotons respectivement* $1,2,3$ *les villes* R,W,Z *et désignons par* $\mathbf{M}^i_j$ *le nombres de trajets (de longueur* 1*) de* j *à* i*.*

1. Écrire la matrice $\mathbf{M}$ formée des coefficients $\mathbf{M}^i_j$.
2. Montrer que le nombre d'itinéraires de longueur n de j à i est égal au coefficient $(\mathbf{M}^n)^i_j$ de la matrice $\mathbf{M}^n$.
3. Calculer le nombre d'itinéraires de longueur 100 de Z vers R.
4. Déterminer les valeurs propres de la matrice $\mathbf{M}$.
5. La matrice $\mathbf{M}$ est-elle diagonalisable ?

CHAPITRE

10

Systèmes dynamiques discrets

Sommaire

§ 1 Les suites récurrentes

10.1.1. Suites récurrentes couplées.— Le calcul des puissances d'une matrice peut s'appliquer à la résolution du problèmes de suites récurrentes couplées. Soient deux suites $(u_k)_k$ et $(v_k)_k$ récurrentes couplées de la façon suivante :

$$\left| \begin{array}{l} u_{k+1} = au_k + bv_k \\ v_{k+1} = cu_k + dv_k. \end{array} \right.$$

avec les conditions initiales (u_0, v_0). Naturellement se problème se pose en les mêmes termes pour un nombre arbitraire de suites.

En posant $\mathbf{x}_k = \begin{bmatrix} u_k \\ v_k \end{bmatrix}$, le système s'écrit

$$\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k, \quad \text{avec} \quad \mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

On montre par récurrence qu'alors

$$\mathbf{x}_k = \mathbf{A}^k \mathbf{x}_0, \quad \text{avec} \quad \mathbf{x}_0 = \begin{bmatrix} u_0 \\ v_0 \end{bmatrix}.$$

10.1.2. Suite récurrente d'ordre supérieur.— Les suites récurrentes géométriques d'ordre 1 sont les suites géométriques $(u_k)_k$ définies par la relation de récurrence :

$$u_{k+1} = qu_k.$$

Avec la condition initiale u_0, le terme général de la suite est $u_k = u_0 q^k$.

Plus généralement, une *suite récurrente géométrique d'ordre p*, où p est un entier supérieur à 1, est une suite $(u_k)_k$ à valeurs dans $\mathbb{K}$, définie pour tout entier k, par la relation de récurrence suivante :

$$u_{k+p} = a_0 u_k + a_1 u_{k+1} + \ldots + a_{p-1} u_{k+p-1}, \tag{10.1}$$

où les coefficients a_i sont dans $\mathbb{K}$. La suite $(u_k)_k$ peut s'écrire sous forme matricielle en considérant la transposée de la matrice compagnon du polynôme $a_0 + a_1 x + \ldots + a_{p-1} x^{p-1}$ définie par

$$\mathbf{C} = \begin{bmatrix} 0 & 1 & 0 & & \ldots & 0 \\ 0 & 0 & 1 & & & \vdots \\ \vdots & & & \ddots & & 0 \\ & & & & 1 & \\ 0 & 0 & & & 0 & 1 \\ a_0 & a_1 & a_2 & \ldots & a_{p-2} & a_{p-1} \end{bmatrix} \in \mathscr{M}_p(\mathbb{K}).$$

En posant

$$\mathbf{x}_k = \begin{bmatrix} u_k \\ u_{k+1} \\ \vdots \\ u_{k+p-1} \end{bmatrix},$$

la relation (10.1) s'écrit sous la forme

$$\mathbf{x}_{k+1} = \mathbf{C}\mathbf{x}_k.$$

Il s'agit d'une suite géométrique à valeurs dans $\mathbb{K}^p$ de raison $\mathbf{C}$. Étant donné une condition initiale $u_0, u_1, \ldots, u_p$, on obtient le terme général de la suite par récurrence en calculant les puissances de $\mathbf{C}$:

$$\mathbf{x}_k = \mathbf{C}^k \mathbf{x}_0.$$

§ 2 La suite de Fibonacci (1202)

10.2.1. Le problème d'évolution de Fibonacci.— Le problème de Fibonacci est un problème d'évolution d'une population à ages structurés posé en les termes suivants

> *« Un homme possède un couple de lapins dans un lieu clos et souhaite savoir combien il aura de couples au bout d'un an si par nature chaque couple de lapins donne naissance à partir de deux mois de vie à un nouveau couple de lapins tous les mois. »*

Notons x_k le nombre de couples de lapins le k-ième mois. Le nombre de couples de lapins au $(k+1)$-ième mois est égal à la somme du nombre de couples de lapins existants le k-ième mois et du nombre de couples de lapins nouveau-nés le $(k+1)$-ième mois, i.e., x_{k-1}. On suppose que les lapins ne meurent jamais.

Le premier mois, il existe un unique couple de lapin, on suppose qu'il s'agit de lapereaux, on a $x_1 = 1$. Les lapereaux ne deviennent adulte à partir de deux mois de vie, on a $x_2 = 1$; ils donnent alors naissance à couple de lapereaux le troisième mois : $x_3 = 2$. Le nombre de couples de lapin satisfait la relation de récurrence suivante :

$$x_{k+1} = x_k + x_{k-1}.$$

La suite $(x_k)_k$ ainsi définie est appelée la *suite de Fibonacci*.

10.2.2. La suite de Fibonacci en Sage.— Des fonctions Sage permettent de calculer les termes de la suites de Fibonacci. Par exemple, calcul des 20 premiers termes de cette suite :

```
sage: [fibonacci(i) for i in range(20)]
        [0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, 233,
             377, 610, 987, 1597, 2584, 4181]
```

Les termes de la suite entre le quinzième et le vingtième rang :

```
sage: [i for i in fibonacci_sequence(15, 20)]
        [610, 987, 1597, 2584, 4181]
```

Exercice $^{(\textsf{Sage})}$ **1.—** Écrire des fonctions Sage qui ont le même comportement que les fonctions `fibonacci(k)` et `fibonacci_sequence(k,l)`.

10.2.3. Description matricielle.— La suite de Fibonacci est une suite récurrence d'ordre deux, cf. 10.1.2. Ce système s'écrit matriciellement sous la forme

$$\begin{bmatrix} x_k \\ x_{k+1} \end{bmatrix} = \mathbf{C} \begin{bmatrix} x_{k-1} \\ x_k \end{bmatrix}, \quad \text{avec} \quad \mathbf{C} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}. \tag{10.2}$$

Le polynôme caractéristique de $\mathbf{C}$ est $p_{\mathbf{C}} = x^2 - x - 1$. Il admet deux racines réelles

$$\lambda_1 = \frac{1-\sqrt{5}}{2}, \qquad \lambda_2 = \frac{1+\sqrt{5}}{2}.$$

Les projecteurs spectraux Π_{λ_1} et Π_{λ_2} de $\mathbf{C}$ vérifient le système d'équations suivant :

$$\left|\begin{array}{l} \mathbf{1}_2 = \Pi_{\lambda_1} + \Pi_{\lambda_2} \\ \mathbf{C} = \lambda_1\Pi_{\lambda_1} + \lambda_2\Pi_{\lambda_2} \end{array}\right.$$

On en déduit que

$$\Pi_{\lambda_1} = \frac{1}{\lambda_1 - \lambda_2}(\mathbf{C} - \lambda_2\mathbf{1}_2) \qquad \text{et} \qquad \Pi_{\lambda_2} = \frac{1}{\lambda_2 - \lambda_1}(\mathbf{C} - \lambda_1\mathbf{1}_2).$$

La solution de l'équation (10.2) est alors

$$\begin{bmatrix} x_k \\ x_{k+1} \end{bmatrix} = \mathbf{C}^k \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \tag{10.3}$$

Or $\mathbf{C}^k = \lambda_1^k\Pi_{\lambda_1} + \lambda_2^k\Pi_{\lambda_2}$, on déduit de (10.3) l'expression de x_{k+1} :

$$x_{k+1} = \frac{\lambda_1^k}{\lambda_1 - \lambda_2}(1 - \lambda_2) + \frac{\lambda_2^k}{\lambda_2 - \lambda_1}(1 - \lambda_1).$$

Ainsi, le nombre de couples de lapins le k-ième mois de l'évolution de la population est

$$x_k = \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^k - \frac{1}{\sqrt{5}}\left(\frac{1-\sqrt{5}}{2}\right)^k.$$

FIGURE 10.1.: Leonardo Fibonacci

§ 3 Dynamique de populations

La dynamique des populations cherche à expliquer l'évolution dans le temps en nombre et en composition de populations biologiques. On peut s'intéresser à des populations humaines, animales, végétales, ou encore microbiennes.

Notre premier exemple est un problème de migration de populations entre deux zones géographiques.

10.3.1. Migration de populations.— On considère le problème de migration de population entre les zones urbaines et les zones rurales. Chaque année, un mouvement de population entre ces deux zones s'opère de la façon suivante :

- la moitié des habitants en zone urbaine partent habiter en zone rurale, alors que l'autre moitié reste résidante en zone hurbaine,
- un quart des habitants en zone rurale rejoignent une zone hurbaine, les trois quarts restant en zone rurale.

Le mouvement de population est indiqué par la figure 10.2

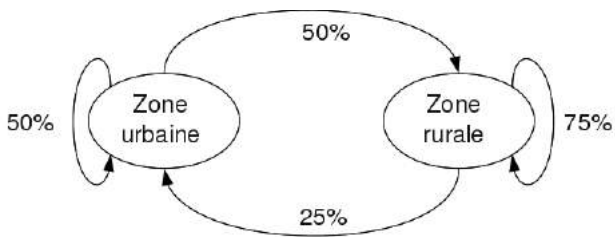


FIGURE 10.2.: Mouvement de population

Étant donné une répartition initiale, l'année $k = 0$, de la population entre les deux zones, quelle sera la répartition de population la k-ième année ? Est-ce que toute la population va au bout d'un certain temps toute se retrouver en zone rurale ?

Notons r_k la proportion de la population totale qui habite en zone rurale au terme de la k-ième année et u_k la proportion de population qui habite en zone urbaine au terme de cette même année. S'agissant de proportion de population, on a, pour toute année k,

$$r_k + u_k = 1.$$

L'évolution des deux populations est décrite par un système d'équations couplées :

$$\left| \begin{array}{l} u_{k+1} = 1/2\, u_k + 1/4\, r_k \\ r_{k+1} = 1/2\, u_k + 3/4\, r_k \end{array} \right.$$

qui s'écrit matriciellement sous la forme :

$$[u_{k+1} \;\; r_{k+1}] = [u_k \;\; r_k] \begin{bmatrix} 1/2 & 1/2 \\ 1/4 & 3/4 \end{bmatrix}.$$

La matrice

$$\mathbf{A} = \begin{bmatrix} 1/2 & 1/2 \\ 1/4 & 3/4 \end{bmatrix}$$

est appelée la *matrice de transition* du système. Si la répartition de population initiale, l'année 0, est $[u_0 \;\; r_0]$, on montre par récurrence que, pour tout k,

$$[u_k \;\; r_k] = [u_0 \;\; r_0]\mathbf{A}^k. \tag{10.4}$$

La relation (10.4) exprime la répartition de la population la k-ième année.

Déterminons la répartition de la population à terme. Il s'agit de calculer $\lim\limits_{k\to+\infty}[u_k \;\; r_k]$. Or

$$\lim_{k\to+\infty}[u_k \;\; r_k] = [u_0 \;\; r_0]\lim_{k\to+\infty}\mathbf{A}^k.$$

Calculons $\lim\limits_{k\to+\infty}\mathbf{A}^k$. Le polynôme caractéristique de $\mathbf{A}$ est $p_{\mathbf{A}} = (x-1)(x-1/4)$. Donc $\mathbf{A}$ est diagonalisable. Notons Π_1 et $\Pi_{\frac{1}{4}}$ les projecteurs spectraux de $\mathbf{A}$. On a

$$\mathbf{A} = \Pi_1 + \frac{1}{4}\Pi_{\frac{1}{4}}.$$

Donc $\mathbf{A}^k = \Pi_1 + \left(\frac{1}{4}\right)^k\Pi_{\frac{1}{4}}$. Par suite

$$\lim_{k\to+\infty}\mathbf{A}^k = \Pi_1.$$

Pour déterminer Π_1, on considère le système

$$\left|\begin{array}{l} \mathbf{1}_2 = \Pi_1 + \Pi_{\frac{1}{4}}, \\ \mathbf{A} = \Pi_1 + \frac{1}{4}\Pi_{\frac{1}{4}}. \end{array}\right.$$

D'on on déduit $\Pi_{\frac{1}{4}} = \mathbf{1}_2 - \Pi_1$ et $\Pi_1 = \dfrac{4}{3}(\mathbf{A} - \dfrac{1}{4}\mathbf{1}_2)$. D'où

$$\Pi_1 = \frac{1}{3}\begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix}.$$

On a donc

$$\begin{aligned} \lim_{k\to+\infty}[u_k \;\; r_k] &= \frac{1}{3}[u_0 \;\; r_0]\begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix}, \\ &= \frac{1}{3}[u_0 + r_0 \;\; 2u_0 + 2r_0], \\ &= [\frac{1}{3} \;\; \frac{2}{3}]. \end{aligned}$$

A terme, il y aura donc un tiers de la population totale en zone urbaine et deux tiers en zone rurale. Notons que cette proportion est indépendante de la répartition initiale des populations entre les deux zones.

Exercice 2.— Dans un pays, on étudie la migration de population entre les zones rurales et les zones urbaines. Chaque année, un tiers de la population des campagnes migre vers les villes, pendant qu'un quart de la population des villes va habiter dans des zones rurales. On notera u_k et r_k la proportion de population urbaine et rurale respectivement l'année k.

1. Étant donnée une répartition initiale u_0 et r_0, quelle est la répartition de population l'année k entre ces deux zones géographiques ?
2. Quelle est à terme la répartition des populations entre ces deux zones géographiques ?
3. Les zones urbaines seront-elles complètement désertées ?

—

§ 4 Exercices

Exercice 3.— On considère deux espèces A et B qui coexistent dans un même environnement naturel. On étudie deux situations d'évolution de ces espèces.

Dans une première situation, on suppose que les deux espèces sont en *compétition* : le nombre d'individus d'une espéce augmente proportionnellement au nombre d'individus de cette même espèce et décroit proportionnellement au nombre d'individus de l'autre espèce.

1. Si la population de chaque espèce augmente de deux fois le nombre d'individus de l'espèce et décroit d'une fois le nombre d'individus de l'autre espèce, déterminer à chaque instant le nombre d'individus de chaque espèce lorsque initiallement il y a 100 individus de l'espèce A et 200 individus de l'espèce B.
2. Est-ce qu'une des espèces est en voie d'instinction ? Si oui, au bout de combien de temps.

Dans une deuxième situation, on suppose que les deux espèces vivent en *symbiose* de la façon suivante. Le nombre d'individus de chaque espèce augmente d'une fois le nombre d'individus de l'autre espèce et décroit d'une fois le nombre d'individus de la même espèce.

3. Si initiallement les espèces A et B sont respectivement composées de 200 et 400 individus, déterminer à chaque instant la population de chaque espèce.
4. Que se passe-t-il à long terme ?

Exercice 4.— La suite de Fibonacci est définie par

$$u_k = u_{k-1} + u_{k-2}, \text{ si } k \geqslant 2, \qquad u_0 = u_1 = 1.$$

1. On considère la matrice suivante de $\mathscr{M}_2(\mathbb{R})$

$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}.$$

Montrer que

$$\mathbf{A}^k = \begin{bmatrix} u_k & u_{k-1} \\ u_{k-1} & u_{k-2} \end{bmatrix},$$

pour tout entier $k \geqslant 2$.

2. Diagonaliser **A** et en déduire une formule non-récurrente pour u_k.

Exercice 5.— Résoudre le système linéaire récurrent suivant :

$$\left|\begin{array}{ll} x_k & = x_{k-1} + z_{k-1} \\ y_k & = y_{k-1} + z_{k-1} \\ z_k & = 2z_{k-1} \end{array}\right.$$

avec $x_0, y_0, z_0 \in \mathbb{R}$.

CHAPITRE
11

Systèmes dynamiques continus

Sommaire

Dans ce chapitre, nous abordons le problème de la résolution de systèmes formés de plusieurs équations différentielles linéaires.

Dans l'exemple 4.1.2, nous avons exprimé l'évolution de la concentration $x(t)$ d'un polluant dans une cuve en fonction du temps par l'équation différentielle :

$$\frac{d}{dt}\mathbf{x}(t) = -\frac{r}{V}\mathbf{x}(t),$$

avec la condition initiale $\mathbf{x}(0) = p_0$. La solution de cette équation est

$$\mathbf{x}(t) = p_0 e^{-\frac{r}{V}t}.$$

Dans l'exemple 4.1.5, nous avons considéré le même problème mais avec trois cuves. L'évolution des concentrations $x_1(t)$, $x_2(t)$ et $x_3(t)$ dans chaque cuve est définie par le système différentiel suivant

$$\left|\begin{array}{ll}
\dfrac{dx_1(t)}{dt} & = -\dfrac{r}{V}x_1(t) \\
\dfrac{dx_2(t)}{dt} & = \dfrac{r}{V}x_1(t) - \dfrac{r}{V}x_2(t) \\
\dfrac{dx_3(t)}{dt} & = \dfrac{r}{V}x_2(t) - \dfrac{r}{V}x_3(t)
\end{array}\right.$$

Ce système s'écrit matriciellement sous la forme

$$\frac{d\mathbf{x}(t)}{dt} = \mathbf{A}\mathbf{x}(t), \tag{11.1}$$

avec

$$\mathbf{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix}, \quad \frac{d\mathbf{x}(t)}{dt} = \begin{bmatrix} \frac{dx_1}{dt} \\ \frac{dx_2}{dt} \\ \frac{dx_3}{dt} \end{bmatrix} \quad \text{et} \quad \mathbf{A} = \frac{r}{V}\begin{bmatrix} -1 & 0 & 0 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix}.$$

Nous montrons dans ce chapitreque la solution de l'équation différentielle (11.1) est

$$\mathbf{x}(t) = e^{t\mathbf{A}}p_0,$$

où $e^{t\mathbf{A}}$ est l'exponentielle de la matrice $t\mathbf{A}$. Plus généralement, nous donnons une méthode de résolution des systèmes différentiels de la forme :

$$\left|\begin{array}{ll} \dfrac{dx_1(t)}{dt} & = a_1^1 x_1(t) + \cdots + a_1^n x_n(t) \\ \quad\vdots & \qquad\qquad \vdots \\ \dfrac{dx_n(t)}{dt} & = a_n^1 x_1(t) + \cdots + a_n^n x_n(t) \end{array}\right.$$

où $a_i^j \in \mathbb{R}$ et les $x_i : \mathbb{R} \to \mathbb{R}$ sont des fonctions dérivables de variable t, qui s'écrit matriciellement sous la forme

$$\frac{d\mathbf{x}(t)}{dt} = \mathbf{A}\mathbf{x}(t),$$

avec

$$\mathbf{A} = \begin{bmatrix} a_1^1 & \ldots & a_1^n \\ \vdots & & \vdots \\ a_n^1 & \ldots & a_n^n \end{bmatrix}, \quad \text{et} \quad \mathbf{x}(t) = \begin{bmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{bmatrix},$$

où $\dfrac{d\mathbf{x}(t)}{dt}$ désigne la dérivé du vecteur $\mathbf{x}(t)$.

§ 1 Systèmes différentiels linéaires à coefficients constants

11.1.1. Matrices de variables réelle.— Une *matrice de variable réelle* t de $\mathscr{M}_{n,m}(\mathbb{K})$ est une application

$$\begin{aligned} \mathbf{A}(-) : \mathbb{R} &\longrightarrow \mathscr{M}_{n,m}(\mathbb{K}) \\ t &\longmapsto \mathbf{A}(t) \end{aligned}$$

Les coefficients de $\mathbf{A}$ sont alors des fonctions $a_i^j : \mathbb{R} \longrightarrow \mathbb{K}$. Une matrice est dite *constante* si tous ses coefficients sont constants.

La matrice $\mathbf{A}(t) = [a_i^j(t)]$ de variable réelle t est dite *dérivable*, si tous ses coefficients sont des fonctions dérivables. Sa dérivée est alors la matrice

$$\frac{d\mathbf{A}(t)}{dt} = \left(\frac{da_i^j(t)}{dt}\right).$$

L'application

$$\frac{d}{dt} : \mathscr{M}_{n,m}(\mathbb{K}) \longrightarrow \mathscr{M}_{n,m}(\mathbb{K})$$

est linéaire, i.e, pour toutes matrices $\mathbf{A}(t)$ et $\mathbf{B}(t)$ de $\mathscr{M}_{n,m}(\mathbb{K})$ de variable réelle t, et pour tous scalaires $\lambda, \mu \in \mathbb{K}$, on a

$$\frac{d}{dt}(\lambda\mathbf{A}(t) + \mu\mathbf{B}(t)) = \lambda\frac{d\mathbf{A}(t)}{dt} + \mu\frac{d\mathbf{B}(t)}{dt}.$$

11.1 Proposition.— L'application $\frac{d}{dt}$ satisfait les propriétés suivantes.

i) Pour toutes matrices compatibles $\mathbf{A}(t)$ et $\mathbf{B}(t)$, on a

$$\frac{d}{dt}(\mathbf{A}(t)\mathbf{B}(t)) = \frac{d\mathbf{A}(t)}{dt}\mathbf{B}(t) + \mathbf{A}(t)\frac{d\mathbf{B}(t)}{dt}.$$

ii) Si $\mathbf{P}$ est une matrice constante et inversible, pour toute matrice $\mathbf{A}(t)$, on a :

$$\frac{d}{dt}(\mathbf{P}^{-1}\mathbf{A}(t)\mathbf{P}) = \mathbf{P}^{-1}\frac{d\mathbf{A}(t)}{dt}\mathbf{P}.$$

Preuve. Montrons **i)**. Posons $\mathbf{A}(t) = [a_i^j(t)]$ et $\mathbf{B}(t) = [b_i^j(t)]$. On a

$$(\mathbf{AB})_i^j = \sum_k a_i^k b_k^j$$

Donc

$$\frac{d}{dt}(\mathbf{AB})_i^j = \frac{d}{dt}\left(\sum_k a_i^k b_k^j\right) = \sum_k \frac{d}{dt}(a_i^k b_k^j).$$

D'où

$$
\begin{aligned}
\frac{d}{dt}(\mathbf{AB})_i^j &= \sum_k \left(\frac{da_i^k}{dt} b_k^j + a_i^k \frac{db_k^j}{dt} \right) \\
&= \sum_k \frac{da_i^k}{dt} b_k^j + \sum_k a_i^k \frac{db_k^j}{dt} \\
&= \left(\frac{d\mathbf{A}}{dt}\mathbf{B} \right)_i^j + \left(\mathbf{A}\frac{d\mathbf{B}}{dt} \right)_i^j \\
&= \left(\frac{d\mathbf{A}}{dt}\mathbf{B} + \mathbf{A}\frac{d\mathbf{B}}{dt} \right)_i^j .
\end{aligned}
$$

L'assertion **ii)** est une conséquence immédiate de **i)**. Si **P** est une matrice constante, on a $\dfrac{d\mathbf{P}}{dt} = 0$. □

11.2 Proposition.— Pour toute matrice constante **A**, on a

$$
\frac{d}{dt} e^{t\mathbf{A}} = \mathbf{A} e^{t\mathbf{A}} = e^{t\mathbf{A}}\mathbf{A}. \tag{11.2}
$$

Preuve. Soit **A** une matrice diagonalisable de spectre $\mathrm{Sp}(\mathbf{A}) = \{\lambda_1, \ldots, \lambda_p\}$. D'après le théorème de décomposition spectrale, on a $\mathbf{A} = \lambda_1 \Pi_{\lambda_1} + \ldots + \lambda_p \Pi_{\lambda_p}$, d'où

$$
e^{t\mathbf{A}} = e^{\lambda_1 t}\Pi_{\lambda_1} + \ldots + e^{\lambda_p t}\Pi_{\lambda_p},
$$

où les Π_{λ_i} sont les projecteurs spectraux de **A**. On en déduit que

$$
\begin{aligned}
\frac{d}{dt} e^{t\mathbf{A}} &= \lambda_1 e^{\lambda_1 t}\Pi_{\lambda_1} + \ldots + \lambda_p e^{\lambda_p t}\Pi_{\lambda_p}, \\
&= \left(\sum_{i=1}^{p} \lambda_i \Pi_{\lambda_i} \right) \left(\sum_{i=1}^{p} e^{\lambda_i t}\Pi_{\lambda_i} \right) \\
&= \mathbf{A} e^{t\mathbf{A}}.
\end{aligned}
$$

Comme $e^{t\mathbf{A}}$ est un polynôme en **A**, on a $\mathbf{A}e^{t\mathbf{A}} = e^{t\mathbf{A}}\mathbf{A}$.

Si la matrice **A** n'est pas diagonalisable, la méthode est la même. On utilise la relation suivante qui découle de la décomposition spectrale algébrique :

$$
e^{t\mathbf{A}} = \sum_{i=1}^{p} \sum_{k=0}^{k_i - 1} t^k e^{\lambda_i t} \frac{(\mathbf{A} - \lambda_i \mathbf{1})^k}{k!} \Pi_{\lambda_i}.
$$

□

11.1.2. Systèmes différentiels linéaires.— Un *système différentiel linéaire* du premier ordre à coefficients constants réels (resp. complexes) est la donnée d'une équation

$$\frac{d\mathbf{x}(t)}{dt} = \mathbf{A}\mathbf{x}(t) + \mathbf{v}(t) \qquad (\mathscr{E})$$

où t parcourt un intervalle réel I, $\mathbf{A} \in \mathscr{M}_n(\mathbb{R})$ (resp. $\mathbf{A} \in \mathscr{M}_n(\mathbb{C})$) et $\mathbf{v} : I \longrightarrow \mathbb{R}^n$ (resp. $\mathbf{v} : I \longrightarrow \mathbb{C}^n$) est une application.

L'application $\mathbf{v}$ est appelée *second membre* de l'équation $(\mathscr{E})$. L'*équation homogène* associée à $(\mathscr{E})$ est l'équation différentielle sans second membre :

$$\frac{d\mathbf{x}(t)}{dt} = \mathbf{A}\mathbf{x}(t) \qquad (\mathscr{E}_0)$$

On appelle solution de $(\mathscr{E})$ toute application $\mathbf{x} : I \longrightarrow \mathbb{R}^n$ (resp. $\mathbf{x} : I \longrightarrow \mathbb{C}^n$) de classe C^1 et satisfaisant $(\mathscr{E})$, pour tout réel t.

11.3 Proposition.—

i) L'ensemble des solutions de l'équation homogène $(\mathscr{E}_0)$ forme un sous-espace vectoriel de l'espace vectoriel $\mathscr{C}^\infty(I, \mathbb{R}^n)$ (resp. $\mathscr{C}^\infty(I, \mathbb{C}^n)$).

ii) La solution générale de l'équation $(\mathscr{E})$ est la somme d'une solution particulière de $(\mathscr{E})$ et d'une solution générale de l'équation homogène $(\mathscr{E}_0)$.

Preuve. Montrons **i)**. De la relation $\dfrac{d\mathbf{x}(t)}{dt} = \mathbf{A}\mathbf{x}(t)$, on déduit que les solutions $\mathbf{x}(t)$ de l'équation homogène $(\mathscr{E}_0)$ sont de classe $\mathscr{C}^\infty$. Il est immédiat que toute combinaison linéaire de solutions de $(\mathscr{E}_0)$ est solution de $(\mathscr{E}_0)$.

Montrons **ii)**. Supposons que $\mathbf{y}(t)$ soit une solution particulière de l'équation $(\mathscr{E})$ et $\mathbf{x}(t)$ une solution générale de l'équation homogène $(\mathscr{E}_0)$. On a

$$\frac{d}{dt}(\mathbf{x}(t) - \mathbf{y}(t)) = \mathbf{A}(\mathbf{x}(t) - \mathbf{y}(t)).$$

Par suite $\mathbf{x}(t) - \mathbf{y}(t)$ est solution de l'équation homogène $(\mathscr{E}_0)$. $\square$

11.4 Théorème.— La solution de l'équation différentielle $(\mathscr{E})$ avec la condition initiale $\mathbf{x}(0) = \mathbf{x}_0$ existe et est unique. Elle est donnée par

$$\mathbf{x}(t) = e^{t\mathbf{A}}\mathbf{x}_0 + \int_0^t e^{(t-s)\mathbf{A}}\mathbf{v}(s)ds.$$

Preuve. Nous admettrons l'unicité.

L'existence peut se montrer en utilisant la méthode de la variation de la constante. D'après la relation (11.2), pour tout vecteur constant $\mathbf{y}$,

$$\mathbf{x}(t) = e^{t\mathbf{A}}\mathbf{y},$$

est solution de l'équation homogène $(\mathscr{E}_0)$. On cherche une solution de $(\mathscr{E})$ sous la forme

$$\mathbf{x}(t) = e^{t\mathbf{A}}\mathbf{y}(t),$$

où $\mathbf{y}(t)$ est un vecteur dépendant de t. On a :

$$\begin{aligned}\frac{d\mathbf{x}(t)}{dt} &= \mathbf{A}e^{t\mathbf{A}}\mathbf{y}(t) + e^{t\mathbf{A}}\frac{d\mathbf{y}(t)}{dt},\\ &= \mathbf{A}\mathbf{x}(t) + e^{t\mathbf{A}}\frac{d\mathbf{y}(t)}{dt}.\end{aligned}$$

Comme $\mathbf{x}(t)$ est solution de l'équation $(\mathscr{E})$, on obtient

$$e^{t\mathbf{A}}\frac{d\mathbf{y}(t)}{dt} = \mathbf{v}(t).$$

La matrice $e^{t\mathbf{A}}$ est inversible. On a donc

$$\frac{d\mathbf{y}(t)}{dt} = e^{-t\mathbf{A}}\mathbf{v}(t).$$

Soit

$$\mathbf{y}(t) = \int_0^t e^{-s\mathbf{A}}\mathbf{v}(s)ds + \mathbf{y}_0.$$

Donc

$$\begin{aligned}\mathbf{x}(t) &= e^{t\mathbf{A}}\mathbf{y}(t),\\ &= e^{t\mathbf{A}}\int_0^t e^{-s\mathbf{A}}\mathbf{v}(s)ds + e^{t\mathbf{A}}\mathbf{y}_0,\\ &= \int_0^t e^{(t-s)\mathbf{A}}\mathbf{v}(s)ds + e^{t\mathbf{A}}\mathbf{y}_0.\end{aligned}$$

Or $\mathbf{x}(0) = \mathbf{x}_0$ donc $\mathbf{y}_0 = \mathbf{x}_0$. On a ainsi

$$\mathbf{x}(t) = e^{t\mathbf{A}}\mathbf{x}_0 + \int_0^t e^{(t-s)\mathbf{A}}\mathbf{v}(s)ds.$$

□

11.5 Proposition.— La solution de l'équation homogène $(\mathscr{E}_0)$ avec la condition initiale $\mathbf{x}_0$ est

$$\mathbf{x}(t) = e^{t\mathbf{A}}\mathbf{x}_0.$$

De plus, l'espace des solutions de l'équation homogène ($\mathscr{E}_0$) est de dimension n. Une base de cet espace est donnée par les vecteurs colonnes de la matrice $e^{t\mathbf{A}}$.

Preuve. La première assertion est une conséquence immédiate du théorème précédent. Notons $\mathbf{c}_1, \ldots, \mathbf{c}_n$ les colonnes de la matrice $e^{t\mathbf{A}}$:

$$e^{t\mathbf{A}} = \big[\, \mathbf{c}_1 \,|\, \mathbf{c}_2 \,|\, \ldots \,|\, \mathbf{c}_n \,\big].$$

Si le vecteur $\mathbf{x}_0$ a pour composante x_0^i, on a

$$\mathbf{x}(t) = \big[\, \mathbf{c}_1 \,|\, \mathbf{c}_2 \,|\, \ldots \,|\, \mathbf{c}_n \,\big] \begin{bmatrix} x_0^1 \\ x_0^2 \\ \vdots \\ x_0^n \end{bmatrix}.$$

D'où $\mathbf{x}(t) = x_0^1\mathbf{c}_1 + \ldots + x_0^n\mathbf{c}_n$, ainsi la famille $(\mathbf{c}_1, \ldots, \mathbf{c}_n)$ forme une base de l'espace des solutions de l'équation homogène ($\mathscr{E}_0$). □

§ 2 Exemples

11.2.1. Exemple.— Le système différentiel

$$\frac{d\mathbf{x}(t)}{dt} = \mathbf{A}\mathbf{x}(t), \qquad \mathbf{x}(0) = \mathbf{x}_0, \qquad \mathbf{A} = \begin{bmatrix} 3 & 1 & -1 \\ 2 & 2 & -1 \\ 2 & 2 & 0 \end{bmatrix}, \tag{11.3}$$

a pour solution $\mathbf{x}(t) = e^{t\mathbf{A}}\mathbf{x}_0$. Il nous suffit donc de déterminer la matrice $e^{t\mathbf{A}}$.

Le polynôme caractéristique de la matrice $\mathbf{A}$ est $p_{\mathbf{A}} = -(x-1)(x-2)^2$, son polynôme minimal est $(x-1)(x-2)^2$. La matrice $\mathbf{A}$ n'est pas diagonalisable.

Déterminons une identité de Bézout entre les polynômes premiers entre eux $x-1$ et $(x-2)^2$. En effectuant un décomposition en éléments simples :

$$\frac{1}{(x-1)(x-2)^2} = \frac{1}{(x-1)} + \frac{-1}{(x-2)} + \frac{1}{(x-2)^2} = \frac{1}{(x-1)} + \frac{3-x}{(x-2)^2}.$$

D'où l'identité de Bézout :

$$1 = (x-2)^2 + (3-x)(x-1).$$

On en déduit les projecteurs spectraux de $\mathbf{A}$

$$\Pi_1 = (\mathbf{A} - 2\mathbf{1}_3)^2 = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 0 \\ 2 & -2 & 0 \end{bmatrix},$$
$$\Pi_2 = (-\mathbf{A} + 3\mathbf{1}_3)(\mathbf{A} - \mathbf{1}_3) = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 1 & 0 \\ -2 & 2 & 1 \end{bmatrix}.$$

Pour calculer ce dernier, on aurait pu aussi exploiter la relation $\Pi_2 = \mathbf{1}_3 - \Pi_1$, moins coûteuse en calculs.

De la relation (9.6), on déduit que

$$\begin{aligned} e^{t\mathbf{A}} &= t^0 \frac{e^t}{0!}(\mathbf{A} - \mathbf{1}_3)^0 \Pi_1 + t^0 \frac{e^{2t}}{0!}(\mathbf{A} - 2\mathbf{1}_3)^0 \Pi_2 + t\frac{e^{2t}}{1!}(\mathbf{A} - 2\mathbf{1}_3)\Pi_2, \\ &= e^t \mathbf{1}_3 \Pi_1 + e^{2t}(\mathbf{1}_3 + t(\mathbf{A} - 2\mathbf{1}_3))\Pi_2, \\ &= e^t \Pi_1 + e^{2t}\Pi_2 + te^{2t}(\mathbf{A} - 2\mathbf{1}_3)\Pi_2. \end{aligned}$$

D'où :

$$e^{t\mathbf{A}} = \begin{bmatrix} e^t + 2te^{2t} & -e^t + e^{2t} & -te^{2t} \\ 2te^{2t} & e^{2t} & -te^{2t} \\ 2e^t + 2(2t-1)e^{2t} & 2(e^{2t} - e^t) & (1-2t)e^{2t} \end{bmatrix}.$$

On en déduit la solution du système différentiel (11.3).

11.2.2. Exemple.— Le système différentiel

$$\frac{d\mathbf{x}(t)}{dt} = \mathbf{A}\mathbf{x}(t) + \mathbf{v}(t), \qquad \mathbf{x}(0) = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \qquad \mathbf{A} = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}, \quad \mathbf{v}(t) = \begin{bmatrix} 1 \\ t \\ 1 \end{bmatrix} \tag{11.4}$$

a pour solution

$$\mathbf{x}(t) = e^{t\mathbf{A}}\mathbf{x}(0) + \int_0^t e^{(t-s)\mathbf{A}}\mathbf{v}(s)ds.$$

D'après l'exemple 8.3.1, la matrice $\mathbf{A}$ est diagonalisable, de spectre $\mathrm{Sp}(\mathbf{A}) = \{1, 4\}$ et ses projecteurs spectraux sont

$$\Pi_1 = -\frac{1}{3}\begin{bmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{bmatrix}, \qquad \Pi_4 = \frac{1}{3}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}.$$

On a $\mathbf{A} = \Pi_1 + 4\Pi_4$, d'où, pour tous réel t et s, on a

$$e^{(t-s)\mathbf{A}} = e^{4(t-s)}\Pi_4 + e^{(t-s)}\Pi_1. \tag{11.5}$$

Par ailleurs,

$$\Pi_1 \mathbf{v}(s) = \frac{s-1}{3}\begin{bmatrix} -1 \\ 2 \\ -1 \end{bmatrix}, \qquad \Pi_4 \mathbf{v}(s) = \frac{2+s}{3}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}. \tag{11.6}$$

Des équations (11.5) et (11.6), on déduit que

$$\begin{aligned}\int_0^t e^{(t-s)\mathbf{A}}\mathbf{v}(s)ds &= \int_0^t e^{4(t-s)}\frac{2+s}{3}ds \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + \int_0^t e^{(t-s)}\frac{s-1}{3}ds \begin{bmatrix} -1 \\ 2 \\ -1 \end{bmatrix}, \\ &= \frac{e^{4t}}{3}\int_0^t e^{-4s}(2+s)ds \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + \frac{e^t}{3}\int_0^t e^{-s}(s-1)ds \begin{bmatrix} -1 \\ 2 \\ -1 \end{bmatrix}.\end{aligned}$$

Or, on a :

$$\int_0^t e^{-4s}(2+s)ds = -\frac{1}{4}\big((\frac{3}{4}+t)e^{-4t}-1\big),$$

et

$$\int_0^t e^{-s}(s-1)ds = (2-t)e^{-t}-2.$$

Par suite,

$$\int_0^t e^{(t-s)\mathbf{A}}\mathbf{v}(s)ds = -\frac{e^{4t}}{12}\big((\frac{3}{4}+t)e^{-4t}-1\big)\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + \frac{e^t}{3}\big((2-t)e^{-t}-2\big)\begin{bmatrix} -1 \\ 2 \\ -1 \end{bmatrix}.$$

Par ailleurs, on a

$$\begin{aligned} e^{t\mathbf{A}}\mathbf{x}(0) &= e^{4t}\Pi_4 \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + e^t \Pi_1 \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \\ &= \frac{e^{4t}}{3}(x_1+x_2+x_4)\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \frac{e^t}{3}\begin{bmatrix} -2x_1+x_2+x_3 \\ x_1-2x_2+x_3 \\ x_1+x_2-2x_3 \end{bmatrix}\end{aligned}$$

Le système différentiel admet donc pour solution

$$\mathbf{x}(t) = \frac{e^{4t}}{3}\big((x_1+x_2+x_4)-\frac{1}{4}\big((\frac{3}{4}+t)e^{-4t}-1\big)\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \frac{e^t}{3}\begin{bmatrix} -2x_1+x_2+x_3 \\ x_1-2x_2+x_3 \\ x_1+x_2-2x_3 \end{bmatrix} + \frac{e^t}{3}\big((2-t)e^{-t}-2\big)\Bigg[$$

Exercice 1.— Résoudre le système différentiel

$$\frac{d\mathbf{x}(t)}{dt} = \mathbf{A}\mathbf{x}(t) + \begin{bmatrix} 1 \\ t \\ 1 \end{bmatrix},$$

pour les matrices suivantes

$$\mathbf{A} = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}, \qquad \mathbf{A} = \begin{bmatrix} -1 & 1 & 3 \\ -2 & 2 & 2 \\ -2 & 1 & 4 \end{bmatrix}.$$

11.2.3. Exemple.— On considère un ensemble de trois cuves dans une usine chimique, reliées entre elles par un système de canalisations. Les cuves L_1, L_2 et L_3 ont chacune le même volume V. Le système de canalisation permet un échange du liquide contenu dans les trois cuves. Les canalisations et les débits de chacune d'elles sont représentés par la figure 11.1.

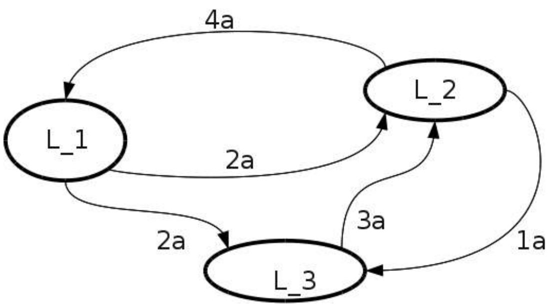


FIGURE 11.1.: Trois lacs reliés par des canaux

Par exemple, il circule de la cuve L_1 à la cuve L_2 un volume de $2a$ litres de liquide par seconde, où a est une constante fixée. On suppose que les échanges de liquide entre les cuves sont continus.

On suppose que les trois cuves L_1, L_2 et L_3 ne contiennent que de l'eau. À l'instant $t = 0$, on verse x grammes d'un produit chimique dans la cuve L_1. On souhaite étudier la diffusion de ce produit dans les autres cuves, en particulier déterminer la quantité de polluant dans chaque cuve à chaque instant, puis lorsque t tend vers $+\infty$.

Notons $x_i(t)$ la quantité du produit chimique, exprimée en gramme, dans la cuve L_i à l'instant t. Lorsque $t = 0$, on a $x_1(0) = x$, $x_2(0) = x_3(0) = 0$. La concentration de produit chimique dans la cuve L_i, à l'instant t est $\frac{1}{V}x_i(t)$ gramme par litre. LE taux de variation de la quantité du produit chimique dans la cuve L_i est défini par

$$\frac{dx_i(t)}{dt} = \left(\begin{array}{c}\text{quantité de produit chimique qui}\\ \text{entre dans la cuve } L_i \text{ par seconde}\end{array}\right) - \left(\begin{array}{c}\text{quantité de produit chimique qui}\\ \text{sort de la cuve } L_i \text{ par seconde}\end{array}\right)$$

D'où le système d'équations différentielles :

$$\left|\begin{array}{ll}\dfrac{dx_1(t)}{dt} & = 4\dfrac{a}{V}x_2(t) - 4\dfrac{a}{V}x_1(t)\\ \dfrac{dx_2(t)}{dt} & = 2\dfrac{a}{V}x_1(t) + 3\dfrac{a}{V}x_3(t) - 5\dfrac{a}{V}x_2(t)\\ \dfrac{dx_3(t)}{dt} & = 2\dfrac{a}{V}x_1(t) + \dfrac{a}{V}x_2(t) - 3\dfrac{a}{V}x_3(t)\end{array}\right.$$

Ce système s'écrit matriciellement sous la forme

$$\frac{d}{dt}\begin{bmatrix} x_1(t)\\ x_2(t)\\ x_3(t)\end{bmatrix} = \frac{a}{V}\begin{bmatrix} -4 & 4 & 0\\ 2 & -5 & 3\\ 2 & 1 & -3\end{bmatrix}\begin{bmatrix} x_1(t)\\ x_2(t)\\ x_3(t)\end{bmatrix}.$$

D'après le théorème 11.4, la solution de ce système est

$$\begin{bmatrix} x_1(t)\\ x_2(t)\\ x_3(t)\end{bmatrix} = e^{t\mathbf{A}}\begin{bmatrix} x\\ 0\\ 0\end{bmatrix},$$

avec

$$\mathbf{A} = \frac{a}{V}\begin{bmatrix} -4 & 4 & 0\\ 2 & -5 & 3\\ 2 & 1 & -3\end{bmatrix}.$$

On calcule le polynôme caractéristique de la matrice $\mathbf{A}$:

$$p_{\mathbf{A}} = -x(\frac{6a}{V} + x)^2.$$

Posons $\lambda_1 = 0$ et $\lambda_2 = -\frac{6a}{V}$. On détermine le sous-espace propre associé à la valeur propre λ_2 :

$$E_{\lambda_2} = \text{Vect}\begin{bmatrix} -2\\ 1\\ 1\end{bmatrix}.$$

Par suite, la matrice $\mathbf{A}$ n'est pas diagonalisable et son polynôme minimal est $m_{\mathbf{A}} = x(\dfrac{6a}{V} + x)^2$.

En effectuant un décomposition en éléments simples :

$$\frac{1}{x(x-\lambda_2)^2} = \frac{\frac{1}{\lambda_2^2}}{x} + \frac{ax+b}{(x-\lambda_2)^2}.$$

D'après le lemme de décomposition en noyaux, on a

$$\mathbb{R}^3 = \text{Ker }(\mathbf{A}) \oplus \text{Ker }(\mathbf{A} - \lambda_2\mathbf{1}_3)^2.$$

La matrice de la projection sur Ker $(\mathbf{A})$ parallèlement à Ker $(\mathbf{A}-2\mathbf{1}_3)^2$ est donc

$$\Pi_1 = (\mathbf{A}-\lambda_2\mathbf{1}_3)^2 = \frac{1}{3}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}.$$

Pour le calcul de Π_2, il est inutile de déterminer a et b, on exploite la relation $\Pi_2 = \mathbf{1}_3 - \Pi_1$.

On a donc

$$e^{t\mathbf{A}} = \Pi_1 + e^{t\lambda_2}\Pi_2 + e^{t\lambda_2}t(\mathbf{A}-\lambda_2\mathbf{1}_3)\Pi_2.$$

D'où

$$\mathbf{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix} = \Pi_1 \begin{bmatrix} x \\ 0 \\ 0 \end{bmatrix} + e^{t\lambda_2}\Pi_2 \begin{bmatrix} x \\ 0 \\ 0 \end{bmatrix} + e^{t\lambda_2}t(\mathbf{A}-\lambda_2\mathbf{1}_3)\Pi_2 \begin{bmatrix} x \\ 0 \\ 0 \end{bmatrix}.$$

Comme $\lambda_2 < 0$, on déduit que

$$\lim_{t\to+\infty} \mathbf{x}(t) = \Pi_1 \begin{bmatrix} x \\ 0 \\ 0 \end{bmatrix}.$$

Par suite, pour $i = 1,2,3$, on a

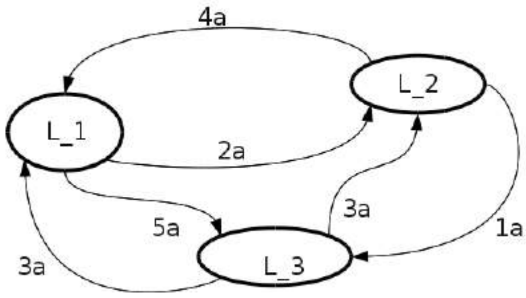
$$\lim_{t\to+\infty} x_i(t) = \frac{1}{3}x.$$

Autrement dit, à terme le produit chimique se sera diffusé dans les trois cuves de façon homogène, chaque cuve contenant un tiers de la quantité initiale du produit.

Exercice 2.— On considère la matrice réelle

$$\mathbf{A} = \begin{bmatrix} -7 & 4 & 3 \\ 2 & -5 & 3 \\ 5 & 1 & -6 \end{bmatrix}.$$

1. Montrer que le polynôme caractéristique de $\mathbf{A}$ est $p_{\mathbf{A}} = -x(x+9)^2$.
2. Déterminer la dimension du sous-espace propre associé à la valeur propre -9.
3. La matrice $\mathbf{A}$ est-elle diagonalisable ?
4. Déterminer le polynôme minimal de $\mathbf{A}$.
5. Exprimer, en fonction de la matrice $\mathbf{A}$, les projecteurs spectraux Π_0 et Π_{-9} de $\mathbf{A}$.
6. Exprimer, en fonction de $\mathbf{A}$, Π_0 et Π_{-9} la matrice $e^{t\mathbf{A}}$, où t est un réel.
7. On considère trois lacs L_1, L_2 et L_3, chacun de volume V, reliés entre eux par un système de canaux permettant de faire circuler l'eau entre les lacs. L'eau circule avec un taux indiqué par la figure suivante.



On suppose que les échanges sont continus. Les lacs L_1, L_2, L_3 contiennent, à l'instant $t = 0$, respectivement q_1, q_2, q_3 grammes de polluant.

8. Ècrire les équations décrivant l'évolution de la quantité de polluant dans chaque lac.

9. Déterminer la quantité de polluant dans chaque lac, lorsque t tend vers $+\infty$.

Exercice $^{(\mathsf{Sage})}$ **3.—** Calculer la solution des systèmes différentiels suivants :

$$\left|\begin{array}{l} \frac{dx(t)}{dt} = 2x(t)+y(t)+z(t)+t \\ \frac{dy(t)}{dt} = x(t)+2y(t)+z(t)+t^2 \\ \frac{dz(t)}{dt} = x(t)+y(t)+2z(t)+t^3 \end{array}\right. \qquad \left|\begin{array}{l} \frac{dx(t)}{dt} = -x(t)+y(t)+3z(t)+\cos(t) \\ \frac{dy(t)}{dt} = -2x(t)+2y(t)+2z(t)+\sin(t) \\ \frac{dz(t)}{dt} = -2x(t)+y(t)+4z(t)+\cos(t) \end{array}\right.$$

avec condition initiale $x(0) = 1$, $x(1) = 2$, $x(3) = 3$. On se ramenera à un système de la forme

$$\frac{d\mathbf{x}(t)}{dt} = \mathbf{A}\mathbf{x}(t) + \mathbf{v}(t)$$

et on définira une fonction Sage pour calculer la solution sous la forme

$$\mathbf{x}(t) = e^{t\mathbf{A}}\mathbf{x}(0) + \int_0^t e^{(t-s)\mathbf{A}}\mathbf{v}(s)ds.$$

11.2.4. Équations différentielles d'ordre *n*.—

11.2.5. Définition.— Une *équation différentielle d'ordre n* réelle (resp. complexe) est donnée de

$$\frac{d^n}{dt^n}z(t) + c_{n-1}\frac{d^{n-1}}{dt^{n-1}}z(t) + \ldots + c_1\frac{d}{dt}z(t) + c_0 z(t) = f(t),$$

où les c_i sont réels (resp. complexes) et f une fonction continue à valeurs réelles (resp. complexes).

Cette équation peut s'écrire sous la forme d'un système linéaire à coefficients constants. Pour cela on pose

$$x_1 = z, \quad x_2 = \frac{d}{dt}x_1 = \frac{d}{dt}z, \quad \ldots \quad x_n = \frac{d}{dt}x_{n-1} = \frac{d^{n-1}}{dt^{n-1}}z.$$

On a donc

$$\frac{d}{dt}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & & \dots & 0 \\ 0 & 0 & 1 & & & \vdots \\ \vdots & & & \ddots & & 0 \\ & & & & 1 & \\ 0 & 0 & & & 0 & 1 \\ -c0 & -c_1 & -c_2 & \dots & -c_{n-2} & -c_{n-1} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ f(t) \end{bmatrix}.$$

On remarquera que la matrice est une transposée de matrice *compagnon*.

Exercice $^{(\textsf{Sage})}$ **4.—** En se ramenant à un système différentiel de la forme

$$\frac{d}{dt}\mathbf{x}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{v}(t),$$

où $\mathbf{A}$ est une matrice compagnon à déterminer, calculer les fonctions $x : \mathbb{R} \longrightarrow \mathbb{C}$ de classe C^3 solutions de l'équation

$$\frac{d^3}{dt^3}x(t) + \frac{d^2}{dt^2}x(t) - \frac{d}{dt}x(t) - x(t) = \cos t.$$

—

§ 3 Exercices

Exercice 5.— 1. Montrer qu'une équation différentielle d'ordre n réelle (resp. complexe), donnée par

$$\frac{d^n}{dt^n}z(t) + c_{n-1}\frac{d^{n-1}}{dt^{n-1}}z(t) + \ldots + c_1\frac{d}{dt}z(t) + c_0 z(t) = f(t),$$

où les c_i sont réels (resp. complexes) et f une fonction continue à valeurs réelles (resp. complexes), peut s'écrire sous la forme d'un système linéaire à coefficients constants de la forme

$$\frac{d}{dt}\mathbf{x}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{v}(t).$$

2. Déterminer les fonctions x de $C^3(\mathbb{R},\mathbb{C})$ solutions de l'équation

$$\frac{d^3}{dt^3}x(t) + \frac{d^2}{dt^2}x(t) - \frac{d}{dt}x(t) - x(t) = \cos t, \quad t \in \mathbb{R}.$$

Exercice 6.— On considère la matrice de $\mathscr{M}_n(\mathbb{R})$ suivante

$$\mathbf{A} = \begin{bmatrix} a & b & \dots & b \\ b & a & \ddots & \vdots \\ \vdots & \ddots & \ddots & b \\ b & \dots & b & a \end{bmatrix}$$

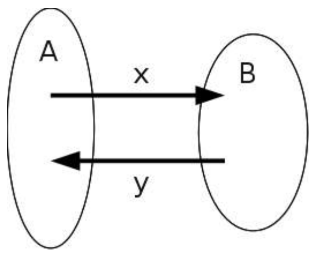
en supposant $n \geqslant 2$. Déterminer la solution du système différentiel

$$\frac{d\,\mathbf{x}(t)}{dt} = \mathbf{A}\mathbf{x}(t),$$

prenant en $t = 0$ la valeur $\mathbf{x}(0)$, contenue dans l'hyperplan de $\mathbb{R}^n$ d'équation $x_1 + \ldots + x_n = 0$.

Exercice 7.— Le problème de diffusion suivant apparaît dans de nombreuses situations biologiques. Considérons deux cellules contenant des composées chimiques. On souhaite étudier la concentration d'un composé injecté dans une cellule, alors que les deux cellules peuvent diffuser mutuellement ce composé.

À l'instant $t = 0$, on injecte une unité d'un composé chimique dans la cellule A. Le composé se diffuse selon les règles suivantes. À chaque instant t, la diffusion du composé de la cellule A vers la cellule B est de x fois la concentration du composé dans la cellule A et la diffusion du composé de la cellule B vers la cellule A est de y fois la concentration du composé dans la cellule B. On suppose $x, y > 0$.



a) Déterminer, à chaque instant t, la concentration du composé dans chacune des deux cellules.

b) Les concentrations du composé dans chaque cellule se stabilisent-elles au bout d'un certain temps ?

Exercice 8.— On considère le système différentiel suivant

$$\frac{d^2}{dt^2}\mathbf{x}(t) + \mathbf{B}\frac{d}{dt}\mathbf{x}(t) + \mathbf{A}\mathbf{x}(t) = \mathbf{v}(t), \qquad (\mathscr{S})$$

où $\mathbf{x}(t)$ est un vecteur de classe C^2 de l'espace $\mathbb{R}^n$, $\mathbf{A}$ et $\mathbf{B}$ sont des matrices de $\mathscr{M}_n(\mathbb{R})$ et $\mathbf{v}(t)$ est une fonction à valeur dans $\mathbb{R}^n$.

1. Par un changement de variable, montrer que le système différentiel $(\mathscr{S})$ se ramène à un système différentiel du premier ordre de la forme

$$\frac{d}{dt}\mathbf{y}(t) = \mathbf{C}\mathbf{y}(t) + \mathbf{w}(t),$$

où $\mathbf{y}(t)$ et $\mathbf{w}(t)$ sont des vecteur de $\mathbb{R}^{n^2}$ et $\mathbf{C}$ une matrice de $\mathscr{M}_{n^2}(\mathbb{R})$.

2. On considère le système différentiel

$$\frac{d^2}{dt^2}\mathbf{x}(t) + \mathbf{A}\mathbf{x}(t) = \mathbf{0}, \quad \mathbf{x}(0) = \mathbf{x}_0. \qquad (\mathscr{T})$$

Diagonaliser la matrice suivante définie par blocs

$$\begin{bmatrix} \mathbf{0} & \mathbf{1}_n \\ \mathbf{A} & \mathbf{0} \end{bmatrix}.$$

3. En déduire que le système $(\mathscr{T})$ admet pour solution

$$\mathbf{x}(t) = \cos(t\mathbf{A})\mathbf{x}_0 + \mathbf{A}^{-1}\sin(t\mathbf{A})\mathbf{y}_0,$$

où

$$\cos(t\mathbf{A}) = \frac{e^{it\mathbf{A}} + e^{-it\mathbf{A}}}{2}, \qquad \sin(t\mathbf{A}) = \frac{e^{it\mathbf{A}} - e^{-it\mathbf{A}}}{2i}.$$

Bibliographie

[Gab01] Pierre Gabriel. *Matrices, géométrie, algèbre linéaire*. Nouvelle bibliothèque mathématique. Cassini, 2001. Traduction de Gabrielle Arnaudiès. 10

[Gri11] Joseph Grifone. *Algèbre linéaire*. Cépaduès Éditions, Toulouse, 2011. 4e édition. 1

[LM03] François Liret and Dominique Martinais. *Algèbre* 1re *année*. Dunod, 2003. 2e édition. 1