Algèbre linéaire 2

Semestre d'été 2018

Université du Luxembourg

Gabor Wiese
gabor.wiese@uni.lu

Version du 11 juillet 2018

Table des matières

Préface

Ce cours est la suite du cours Algèbre Linéaire 1 enseigné au semestre d'hiver 2017/2018. Il est suivi par des étudiants de mathématiques et de physique.
Ces notes se sont developpées pendant des années. J'ai utilisé plusieurs sources, notamment le livre de Fischer *Lineare Algebra* (Vieweg-Verlag) et des notes de B. H. Matzat de l'université de Heidelberg.

Esch-sur-Alzette, 11 juillet 2018, Gabor Wiese

Littérature

Voici quelques références : ces livres sont disponibles dans la bibliothèque au Kirchberg.

- Lelong-Ferrand, Arnaudiès. *Cours de mathématiques, Tome 1, Algèbre*. Dunod. Ce livre est très complet et très détaillé. On peut l'utiliser comme ouvrage de référence.
- Siegfried Bosch : *Algebra* (en allemand), Springer-Verlag. Ce livre est très complet et bien lisible.
- Serge Lang : *Algebra* (en anglais), Springer-Verlag. C'est comme une encyclopédie de l'algèbre ; on y trouve beaucoup de sujets rassemblés, écrits de façon concise.
- Siegfried Bosch. *Lineare Algebra*, Springer-Verlag.
- Jens Carsten Jantzen, Joachim Schwermer. *Algebra*.
- Christian Karpfinger, Kurt Meyberg. *Algebra : Gruppen - Ringe - Körper*, Spektrum Akademischer Verlag.
- Gerd Fischer. *Lehrbuch der Algebra : Mit lebendigen Beispielen, ausführlichen Erläuterungen und zahlreichen Bildern*, Vieweg+Teubner Verlag.
- Gerd Fischer. *Lineare Algebra : Eine Einführung für Studienanfänger*, Vieweg+Teubner Verlag.
- Gerd Fischer, Florian Quiring. *Lernbuch Lineare Algebra und Analytische Geometrie : Das Wichtigste ausführlich für das Lehramts- und Bachelorstudium*, Springer Vieweg.
- Perrin. *Cours d'algèbre*, Ellipses.
- Guin, Hausberger. *Algèbre I. Groupes, corps et théorie de Galois*, EDP Sciences.
- Fresnel. *Algèbre des matrices*, Hermann.
- Tauvel. *Algèbre*.
- Combes. *Algèbre et géométrie*.

Prérequis

Le cours a un côté théorique et un côté pratique. Pour le côté pratique, (presque) tous les calculs peuvent se ramener à deux opérations fondamentales :

- résoudre des systèmes d'équations linéaires,
- calculer des déterminants.

Nous allons commencer le cours par deux sections de rappels : une sur les fondations des espaces vectoriels et une sur les déterminants.
L'algèbre linéaire peut se faire sur n'importe quel corps commutatif, pas seulement sur les nombres réels ou complexes. Les étudiants de mathématiques ont vu la définition de corps dans le cours « Structures mathématiques ». Nous la rappelons ici :

Définition 0.1. *Un* corps (commutatif) K *est un ensemble* K *qui contient deux éléments distincts notés* $0, 1$ *et qui admet deux applications*

$$\begin{aligned} + : K \times K \to K, &\quad (a,b) \mapsto a+b, && \text{« addition »} \\ \cdot : K \times K \to K, &\quad (a,b) \mapsto a \cdot b && \text{« multiplication »,} \end{aligned}$$

tel que pour tout $x, y, z \in K$*, les assertions suivantes sont satisfaites :*

- élément neutre pour l'addition : $x + 0 = x = 0 + x$ *;*
- associativité de l'addition : $(x + y) + z = x + (y + z)$ *;*
- existence d'inverse pour l'addition : *il existe un élément appelé* $-x$ *tel que* $x + (-x) = 0 = (-x) + x$ *;*
- commutativité de l'addition : $x + y = y + x$*.*
- élément neutre pour la multiplication : $x \cdot 1 = x = 1 \cdot x$ *;*
- associativité de la multiplication : $(x \cdot y) \cdot z = x \cdot (y \cdot z)$ *;*
- existence d'inverse pour la multiplication : *si* $x \neq 0$*, il existe un élément appelé* $x^{-1} = \frac{1}{x}$ *tel que* $x \cdot x^{-1} = 1 = x^{-1} \cdot x$ *;*
- commutativité de la multiplication : $x \cdot y = y \cdot x$*.*
- distributivité : $(x + y) \cdot z = x \cdot z + y \cdot z$*.*

Nous allons uniquement considérer des corps commutatifs ; on parlera donc simplement de « corps ».

Exemple 0.2.

- $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ *sont des corps.*
- *Si* p *est un nombre premier,* $\mathbb{Z}/p\mathbb{Z}$ *est un corps.*
- $\mathbb{Z}$ *et* $\mathbb{N}$ *ne sont pas des corps.*

Pour toute la suite, soit K un corps commutatif. Si cela vous aide pour la compréhension, vous pouvez prendre $K = \mathbb{R}$ ou $K = \mathbb{C}$.

1 Rappels : Espaces vectoriels, bases, dimension, homomorphismes

Objectifs :

- Maîtriser les notions d'espace vectoriel et de sous-espace ;
- maîtriser les notions de base et de dimension ;
- maîtriser les notions d'application linéaire ((homo)morphisme), de noyau, d'image ;
- connaître des exemples et savoir démontrer des propriétés simples.

On suppose connus et pratiqués la déscription par des matrices et la résolution par l'algorithme du pivot de Gauß des systèmes d'équations linéaires.

Définition d'espaces vectoriels

Définition 1.1. *Soit V un ensemble avec $0_V \in V$ un élément, et avec deux applications*

$$+_V : V \times V \to V, \quad (v_1, v_2) \mapsto v_1 + v_2 = v_1 + v_2$$

(appelée addition*) et*

$$\cdot_V : K \times V \to V, \quad (a, v) \mapsto a \cdot v = av$$

(appelée multiplication scalaire*).*
On appelle $(V, +_V, \cdot_V, 0_V)$ un K-espace vectoriel *si*

(A1) $\forall\, u, v, w \in V : (u +_V v) +_V w = u +_V (v +_V w)$*,*

(A2) $\forall\, v \in V : 0_V +_V v = v = v +_V 0_V$*,*

(A3) $\forall\, v \in V \exists\, w \in V : v +_V w = 0 = w +_V v$ *(on écrit $-v := w$),*

(A4) $\forall\, u, v \in V : u +_V v = v +_V u$*,*

(pour les mathématiciens : ces propriétés disent que $(V, +_V, 0_V)$ est un groupe abélien) et

(MS1) $\forall\, a \in K, \forall\, u, v \in V : a \cdot_V (u +_V v) = a \cdot_V u +_V a \cdot_V v$*,*

(MS2) $\forall\, a, b \in K, \forall v \in V : (a +_K b) \cdot_V v = a \cdot_V v +_V b \cdot_V v$*,*

(MS3) $\forall\, a, b \in K, \forall v \in V : (a \cdot_K b) \cdot_V v = a \cdot_V (b \cdot_V v)$*,*

(MS4) $\forall\, v \in V : 1 \cdot_V v = v$*.*

Pour plus de clarté, nous avons écrit $+_V$, $\cdot_V$ pour l'addition et la multiplication scalaire dans V, et $+_K$, $\cdot_K$ pour l'addition et la multiplication dans K. Dans la suite, on ne le fera plus.

Exemple 1.2. *Soit $n \in \mathbb{N}$. Le K-espace vectoriel canonique de dimension n est K^n, l'ensemble des vecteurs colonnes de taille n à coefficients dans K. Comme vous le savez, on peut additionner deux éléments de K^n de la façon suivante :*

$$\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} a_1+b_1 \\ a_2+b_2 \\ \vdots \\ a_n+b_n \end{pmatrix}.$$

Cette addition satisfait les propriétés suivantes :

(A1) $(\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}) + \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} + (\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} + \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}).$

(A2) $\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$

(A3) $\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} + \begin{pmatrix} -a_1 \\ -a_2 \\ \vdots \\ -a_n \end{pmatrix} = \begin{pmatrix} a_1-a_1 \\ a_2-a_2 \\ \vdots \\ a_n-a_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$

(A4) $\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} a_1+b_1 \\ a_2+b_2 \\ \vdots \\ a_n+b_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} + \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}.$

En plus, on dispose d'une multiplication scalaire : on multiplie un élément de K^n par un élément r de K ainsi :

$$r \cdot \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} ra_1 \\ ra_2 \\ \vdots \\ ra_n \end{pmatrix}.$$

L'addition et la multiplication sont compatibles de la manière suivante :

(MS1) $\forall\, r \in K, \forall \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}, \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} \in K^n : r \cdot (\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}) = r \cdot \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} + r \cdot \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix};$

(MS2) $\forall\, r, s \in K, \forall \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \in K^n : (r+s) \cdot \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = r \cdot \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} + s \cdot \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix};$

(MS3) $\forall\, r, s \in K, \forall \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \in K^n : r \cdot (s \cdot \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}) = (r \cdot s) \cdot \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix};$

(MS4) $\forall \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \in K^n : 1 \cdot \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}.$

Cela montre que K^n est bien un K-espace vectoriel.

La proposition suivante nous produit un grand nombre d'autres exemples.

Proposition 1.3. *Soit E un ensemble. On introduit la notation*

$$\mathcal{F}(E,K) := \{ f \mid f : E \to K \text{ application } \}$$

pour l'ensemble des applications de E dans K. On note l'application $E \to K$ telle que toutes ses valeurs sont 0 par $0_\mathcal{F}$ (concrètement : $0_\mathcal{F} : E \to K$ définie par la règle $0_\mathcal{F}(e) = 0$ pour tout $e \in E$). On définit l'addition

$$+_\mathcal{F} : \mathcal{F}(E,K) \times \mathcal{F}(E,K) \to \mathcal{F}(E,K), \quad (f,g) \mapsto f +_\mathcal{F} g \text{ où } \forall e \in E : (f +_\mathcal{F} g)(e) := f(e) + g(e)$$

et la mutliplication scalaire

$$\cdot_\mathcal{F} : K \times \mathcal{F}(E,K) \to \mathcal{F}(E,K), \quad (x,f) \mapsto x \cdot_\mathcal{F} f \text{ où } \forall e \in E : (x \cdot_\mathcal{F} f)(e) := x \cdot (f(e)).$$

Alors, $(\mathcal{F}(E,K), +_\mathcal{F}, \cdot_\mathcal{F}, 0_\mathcal{F})$ est un K-espace vectoriel.

Démonstration. Exercice. □

La plupart du temps, on n'écrira pas les indices, mais seulement $f + g$, $f \cdot g$, etc.

Exemple 1.4. *(a) $\{f \in \mathcal{F}(\mathbb{R},\mathbb{R}) \mid f(1) = 0\}$ est un K-espace vectoriel.*

(b) $\{f \in \mathcal{F}(\mathbb{R},\mathbb{R}) \mid f(0) = 1\}$ n'est pas un K-espace vectoriel.

Lemme 1.5. *Soit $(V, +_V, \cdot_V, 0_V)$ un K-espace vectoriel. Alors, les propriétés suivantes sont satisfaites pour tout $v \in V$ et tout $a \in K$:*

(a) $0 \cdot_V v = 0_V$;

(b) $a \cdot_V 0_V = 0_V$;

(c) $a \cdot_V v = 0_V \Rightarrow a = 0 \vee v = 0_V$;

(d) $(-1) \cdot_V v = -v$.

Démonstration. (a) $0 \cdot_V v = (0 + 0) \cdot_V v = 0 \cdot_V v + 0 \cdot_V v$, donc $0 \cdot_V v = 0_V$.
(b) $a \cdot_V 0_V = a \cdot_V (0_V + 0_V) = a \cdot_V 0_V + a \cdot_V 0_V$, donc $a \cdot_V 0_V = 0_V$.
(c) Supposons $a \cdot_V v = 0_V$. Si $a = 0$, l'assertion $a = 0 \vee v = 0_V$ est vraie. Supposons donc $a \neq 0$. Alors a^{-1} a un sens. En conséquence, $v = 1 \cdot_V v = (a^{-1} \cdot a) \cdot_V v = a^{-1} \cdot_V (a \cdot_V v) = a^{-1} \cdot_V 0_V = 0_V$ par (b).
(d) $v +_V (-1) \cdot_V v = 1 \cdot_V v +_V (-1) \cdot_V v = (1 + (-1)) \cdot_V v = 0 \cdot_V v = 0_V$ par (a). □

Au lieu de $(V, +_V, \cdot_V, 0_V)$ on écrira souvent seulement V.

Sous-espaces vectoriels

Définition 1.6. *Soit V un K-espace vectoriel. On dit qu'un sous-ensemble non-vide $W \subseteq V$ est un* sous-espace vectoriel de V *si*

$$\forall\, w_1, w_2 \in W, \forall a \in K : a \cdot w_1 + w_2 \in W.$$

Notation : $W \leq V$.

Exemple 1.7.

- *Soit V un K-espace vectoriel. L'ensemble $\{0\}$ est un sous-espace vectoriel de V, appelé* l'espace zéro, *noté 0 par simplicité (ne pas confondre avec l'élément 0).*
- *Soient $V = \mathbb{R}^2$ et $W = \{\left(\begin{smallmatrix} a \\ 0 \end{smallmatrix}\right) \mid a \in \mathbb{R}\} \subseteq V$. Alors, W est un sous-espace de V.*
- *Soient $V = \mathbb{R}^3$ et $W = \left\{\begin{pmatrix} a \\ b \\ 2b \end{pmatrix} \mid a, b \in \mathbb{R}\right\} \subseteq V$. Alors, W est un sous-espace de V.*
- *Soient $n, m \in \mathbb{N}_{\geq 1}$. On considère le système d'équations linéaires*

$$\begin{aligned} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,n}x_n &= b_1 \\ a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,n}x_n &= b_2 \\ &\vdots \\ a_{m,1}x_1 + a_{m,2}x_2 + \cdots + a_{m,n}x_n &= b_m \end{aligned}$$

avec $b_i, a_{i,j} \in K$ pour $1 \leq i \leq m$, $1 \leq j \leq n$.

(a) Soit S l'ensemble de toutes les solutions du système homogène avec $x_1, x_2, \dots, x_n \in K$, c'est-à-dire

$$S = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in K^n \mid \forall i \in \{1, 2, \dots, m\} : \sum_{j=1}^{n} a_{i,j}x_j = 0 \right\}.$$

Alors, S est un sous-espace vectoriel du K-espace vectoriel standard K^n.

(b) Soit $\begin{pmatrix} r_1 \\ r_2 \\ \vdots \\ r_n \end{pmatrix} \in K^n$ une solution du système d'équations linéaires, c'est-à-dire :

$$\forall i \in \{1, 2, \dots, m\} : \sum_{j=1}^{n} a_{i,j}r_j = b_i.$$

Soit S le sous-espace vectoriel de K^n défini en (a).
Alors, les solutions du système d'équations linéaires sont l'ensemble

$$\left\{ \begin{pmatrix} r_1 \\ r_2 \\ \vdots \\ r_n \end{pmatrix} + \begin{pmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{pmatrix} \mid \begin{pmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{pmatrix} \in S \right\}.$$

Voici une forme générale pour obtenir et écrire des sous-espaces :

Définition-Lemme 1.8. *Soient V un K-espace vectoriel et $E \subseteq V$ un sous-ensemble non-vide. On pose*

$$\langle E \rangle := \{\sum_{i=1}^{m} a_i e_i \mid m \in \mathbb{N}, e_1, \dots, e_m \in E, a_1, \dots, a_m \in K\}.$$

C'est un sous-espace vectoriel de V que l'on dit engendré par E.
Comme convention, en pose $\langle \emptyset \rangle = 0$, le sous-espace nul.

Démonstration. Comme $\langle E \rangle$ n'est pas vide (car E n'est pas vide), il suffit de vérifier la définition de sous-espace. Soient donc $w_1, w_2 \in \langle E \rangle$ et $a \in K$. Nous pouvons écrire

$$w_1 = \sum_{i=1}^{m} a_i e_i \text{ et } w_2 = \sum_{i=1}^{m} b_i e_i$$

pour $a_i, b_i \in K$ et $e_i \in E$ pour tout $i = 1, \dots, m$. Donc nous avons

$$a \cdot w_1 + w_2 = \sum_{i=1}^{m} (a a_i + b_i) e_i,$$

qui est bien un élément de $\langle E \rangle$. $\square$

Exemple 1.9. *L'ensemble* $\left\{ a \cdot \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} + b \cdot \begin{pmatrix} 0 \\ 0 \\ 7 \end{pmatrix} \mid a, b \in \mathbb{R} \right\}$ *est un sous-espace de $\mathbb{R}^3$.*

Parfois il est utile de caractériser le sous-espace engendré par un ensemble d'une façon plus théorique. Pour faire cela, nous avons besoin du lemme suivant.

Lemme 1.10. *Soit V un K-espace vectoriel et $W_i \leq V$ des sous-espaces pour $i \in I \neq \emptyset$. Alors, $W := \bigcap_{i \in I} W_i$ est un sous-espace vectoriel de V.*

Démonstration. Exercice. $\square$

Par contre, $\bigcup_{i \in I} W_i$ n'est pas un sous-espace en général (comme vous le voyez dans un exercice) !

Exemple 1.11. *Comment calculer l'intersection de deux sous-espaces ?*

(a) Le cas le plus simple est quand les deux sous-espaces sont donnés comme les ensembles des solutions de deux systèmes d'équations linéaires, par exemple

- *V est le sous-ensemble des* $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in K^n$ *tels que* $\sum_{i=1}^{n} a_{i,j} x_i = 0$ *pour* $j = 1, \dots, \ell$, *et*
- *W est le sous-ensemble des* $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in K^n$ *tels que* $\sum_{i=1}^{n} b_{i,k} x_i = 0$ *pour* $k = 1, \dots, m$.

Dans ce cas, le sous-espace $V \cap W$ est donné comme l'ensemble des solutions communes pour toutes les égalités.

(b) *Supposons maintenant que les sous-espaces sont donnés comme des sous-espaces de K^n engendrés par des ensembles finis de vecteurs : Soit $V = \langle E \rangle$ et $W = \langle F \rangle$ où*

$$E = \left\{ \begin{pmatrix} e_{1,1} \\ e_{2,1} \\ \vdots \\ e_{n,1} \end{pmatrix}, \dots, \begin{pmatrix} e_{1,m} \\ e_{2,m} \\ \vdots \\ e_{n,m} \end{pmatrix} \right\} \subseteq K^n \text{ et } F = \left\{ \begin{pmatrix} f_{1,1} \\ f_{2,1} \\ \vdots \\ f_{n,1} \end{pmatrix}, \dots, \begin{pmatrix} f_{1,p} \\ f_{2,p} \\ \vdots \\ f_{n,p} \end{pmatrix} \right\} \subseteq K^n.$$

Alors

$$V \cap W = \Big\{ \sum_{i=1}^{m} a_i \begin{pmatrix} e_{1,i} \\ e_{2,i} \\ \vdots \\ e_{n,i} \end{pmatrix} \mid$$
$$\exists\, b_1, \dots, b_p \in K : a_1 \begin{pmatrix} e_{1,1} \\ e_{2,1} \\ \vdots \\ e_{n,1} \end{pmatrix} + \dots + a_m \begin{pmatrix} e_{1,m} \\ e_{2,m} \\ \vdots \\ e_{n,m} \end{pmatrix} - b_1 \begin{pmatrix} f_{1,1} \\ f_{2,1} \\ \vdots \\ f_{n,1} \end{pmatrix} - \dots - b_p \begin{pmatrix} f_{1,p} \\ f_{2,p} \\ \vdots \\ f_{n,p} \end{pmatrix} = 0 \Big\}.$$

Voici un exemple concret : $E = \left\{ \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\} \subseteq K^n$ *et* $F = \left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} \right\} \subseteq K^n$*. On doit résoudre le système*

$$\begin{pmatrix} 1 & 0 & -1 & -2 \\ 1 & 1 & 0 & 0 \\ 2 & 0 & -1 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ y_1 \\ y_2 \end{pmatrix} = 0.$$

Avec des opérations sur des lignes, on obtient

$$\ker\left(\begin{pmatrix} 1 & 0 & -1 & -2 \\ 1 & 1 & 0 & 0 \\ 2 & 0 & -1 & -1 \end{pmatrix}\right) = \ker\left(\begin{pmatrix} 1 & 0 & -1 & -2 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 1 & 3 \end{pmatrix}\right) = \ker\left(\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 3 \end{pmatrix}\right),$$

donc on obtient comme sous-espace des solutions la droite engendrée par $\begin{pmatrix} -1 \\ 1 \\ -3 \\ 1 \end{pmatrix}$*, donc l'intersection est donné par la droite*

$$\langle -1 \cdot \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} + 1 \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \rangle = \langle -3 \cdot \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + 1 \cdot \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} \rangle = \langle \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} \rangle.$$

Voici la caractérisation alternative du sous-espace engendré par un ensemble :

Lemme 1.12. *Soient V un K-espace vectoriel et $E \subseteq V$ un sous-ensemble non-vide. Alors nous avons l'égalité*

$$\langle E \rangle = \bigcap_{W \leq V \text{ sous-espace t.q. } E \subseteq W} W$$

où le terme à droite de l'égalité est l'intersection de tous les sous-espaces W de V qui contiennent E.

Démonstration. Pour démontrer l'égalité de deux ensembles, il faut montrer les deux inclusions.
« $\subseteq$ » : Tout sous-espace W qui contient E, contient aussi toutes les combinaisons linéaires des éléments de E, donc W contient $\langle E \rangle$. En conséquence, $\langle E \rangle$ est contenu dans l'intersection à droite.
« $\supseteq$ » : Comme $\langle E \rangle$ fait partie des sous-espaces dans l'intersection à droite, il est clair que cette intersection est contenue dans $\langle E \rangle$. □

Définition 1.13. *Soient V un K-espace vectoriel et $E \subseteq V$ un sous-ensemble. On dit que V est* engendré par E (en tant que sous-espace vectoriel) *si $V = \langle E \rangle$.*
Cela revient à dire que tout élément de V s'écrit comme combinaison linéaire de vecteurs dans E.

Définition 1.14. *Soient V un K-espace vectoriel et $W_i \leq V$ des sous-espaces de V pour $i \in I \neq \emptyset$. On pose*

$$\sum_{i\in I} W_i := \langle \bigcup_{i\in I} W_i \rangle,$$

le sous-espace de V engendré par tous les éléments de tous les W_i. On l'appelle la somme des W_i, $i \in I$.
Si $I = \{1, 2, \ldots, n\}$, on peut décrire $\sum_{i=1}^n W_i$ explicitement comme

$$\sum_{i=1}^n W_i = \{\sum_{i=1}^n w_i \mid w_1 \in W_1, \ldots w_n \in W_n\}.$$

Pour un I général, ceci se généralise ainsi :

$$\sum_{i\in I} W_i = \{\sum_{i\in I} w_i \mid \big(\forall\, i \in I : w_i \in W_i\big) \text{ et } w_i \neq 0 \text{ pour seulement un nombre fini de } i \in I\}.$$

Nous utilisons la notation $\sum'_{i\in I} w_i$ pour indiquer $w_i \neq 0$ pour seulement un nombre fini de $i \in I$.

Exemple 1.15. *Comment calculer/obtenir la somme de deux sous-espaces ?*
La réponse est très facile si les deux sous-espaces sont donnés par des générateurs : Si $U = \langle E \rangle$ et $V = \langle F \rangle$ sont des sous-espaces d'un K-espace vectoriel V, alors $U + V = \langle E \cup F \rangle$.
(La question de donner une base pour la somme est une autre... voir plus loin.)

Quand est-ce que les $w_i \in W_i$ dans l'écriture $w = \sum'_{i\in I} w_i$ sont uniques ?

Définition 1.16. *Soient V un K-espace vectoriel et $W_i \leq V$ des sous-espaces de V pour $i \in I \neq \emptyset$. On dit que la somme $W = \sum_{i\in I} W_i$ est* directe *si pour tout $i \in I$ on a*

$$W_i \cap \sum_{j\in I\setminus\{i\}} W_j = 0.$$

Notation pour les sommes directes : $\bigoplus_{i\in I} W_i$.
Si $I = \{1, \ldots, n\}$, on note parfois les éléments d'une somme directe $\bigoplus_{i=1}^n W_i$ par $w_1 \oplus w_2 \oplus \cdots \oplus w_n$ (où, évidemment, $w_i \in W_i$ pour $i \in I$).

Exemple 1.17. *Dans l'exemple 1.11 (b), la somme $V + W$ n'est pas directe car l'intersection $V \cap W$ est une droite et donc non nulle.*

Proposition 1.18. *Soient V un K-espace vectoriel, $W_i \leq V$ des sous-espaces de V pour $i \in I \neq \emptyset$ et $W = \sum_{i\in I} W_i$. Alors les assertions suivantes sont équivalentes :*

(i) $W = \bigoplus_{i\in I} W_i$;

(ii) *pour tout $w \in W$ et tout $i \in I$ il existe un unique $w_i \in W_i$ tel que $w = \sum'_{i\in I} w_i$.*

Démonstration. « (i) ⇒ (ii) » : L'existence de tels $w_i \in W_i$ est claire. Démontrons donc l'unicité en prenant

$$w = \sum_{i\in I}{}' w_i = \sum_{i\in I}{}' w'_i$$

avec $w_i, w_i' \in W_i$ pour tout $i \in I$ (rappelons que la notation $\sum'$ indique que seul un nombre fini de w_i, w_i' est non nul). Cela implique pour $i \in I$:

$$w_i - w_i' = \sum_{j \in I \setminus \{i\}}{}' (w_j' - w_j) \in W_i \cap \sum_{j \in I \setminus \{i\}} W_j = 0.$$

Donc, $w_i - w_i' = 0$, alors $w_i = w_i'$ pour tout $i \in I$, montrant l'unicité.
« (ii) $\Rightarrow$ (i) » : Soient $i \in I$ et $w_i \in W_i \cap \sum_{j \in I \setminus \{i\}} W_j$. Donc, $w_i = \sum'_{j \in I \setminus \{i\}} w_j$ avec $w_j \in W_j$ pour tout $j \in I$. Nous pouvons maintenant écrire 0 de deux façons

$$0 = \sum_{i \in I}{}' 0 = -w_i + \sum_{j \in I \setminus \{i\}}{}' w_j.$$

Donc, l'unicité implique $-w_i = 0$. Alors, nous avons montré $W_i \cap \sum_{j \in I \setminus \{i\}} W_j = 0$. $\square$

Bases

Définition 1.19. *Soient V un K-espace vectoriel et $E \subseteq V$ un sous-ensemble.*
*On dit que E est K-*linéairement indépendant *si*

$$\forall\, n \in \mathbb{N}\, \forall\, a_1, \ldots, a_n \in K\, \forall\, e_1, \ldots, e_n \in E : \Big(\sum_{i=1}^{n} a_i e_i = 0 \in V \Rightarrow a_1 = a_2 = \cdots = a_n = 0 \Big)$$

*(c'est-à-dire, la seule combinaison K-linéaire d'éléments de E représentant $0 \in V$ est celle dans laquelle tous les coefficients sont 0). Dans le cas contraire, on dit que E est K-*linéairement dépendant*.*
*On appelle E une K-*base de V *si E engendre V et E est K-linéairement indépendant.*

Exemple 1.20. *Comment calculer si des vecteurs sont linéairement indépendants ? (Même réponse que presque toujours :) Résoudre un système d'équations linéaires.*
Soit le sous-ensemble

$$\left\{ \begin{pmatrix} e_{1,1} \\ e_{2,1} \\ \vdots \\ e_{n,1} \end{pmatrix}, \ldots, \begin{pmatrix} e_{1,m} \\ e_{2,m} \\ \vdots \\ e_{n,m} \end{pmatrix} \right\}$$

de K^n donné. Ces vecteurs sont linéairement indépendants si et seulement si la seule solution du système d'équations linéaires

$$\begin{pmatrix} e_{1,1} & e_{1,2} & \cdots & e_{1,m} \\ e_{2,1} & e_{2,2} & \cdots & e_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ e_{n,1} & e_{n,2} & \cdots & e_{n,m} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix} = 0$$

est nulle.

Exemple 1.21. *Soit $d \in \mathbb{N}_{>0}$. On pose* $e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ldots, e_d = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$ *et* $E = \{e_1, e_2, \ldots, e_d\}$*. Alors :*

- *E engendre K^d :*

  Tout vecteur $v = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_d \end{pmatrix}$ s'écrit comme K-combinaison linéaire : $v = \sum_{i=1}^d a_i e_i$.

- *E est K-linéairement indépendant :*

  Si l'on a une combinaison K-linéaire $0 = \sum_{i=1}^d a_i e_i$, alors clairement $a_1 = \cdots = a_d = 0$.

- *E est donc une K-base de K^d, car E engendre K^d et est K-linéairement indépendant. On l'appelle la* base canonique *de K^d.*

Le prochain théorème caractérise les bases.

Théorème 1.22. *Soient V un K-espace vectoriel et $E = \{e_1, e_2, \dots, e_n\} \subseteq V$ un sous-ensemble fini. Alors, les assertions suivantes sont équivalentes :*

(i) E est une K-base.

(ii) E est un ensemble minimal de générateurs de V, c'est-à-dire : E engendre V, mais pour tout $e \in E$, l'ensemble $E \setminus \{e\}$ n'engendre pas V.

(iii) E est un ensemble maximal K-linéairement indépendant, c'est-à-dire : E est K-linéairement indépendant, mais pour tout $e \in V \setminus E$, l'ensemble $E \cup \{e\}$ est K-linéairement dépendant.

(iv) Tout $v \in V$ s'écrit comme $v = \sum_{i=1}^n a_i e_i$ avec des uniques $a_1, \dots, a_n \in K$.

Corollaire 1.23. *Soient V un K-espace vectoriel et $E \subseteq V$ un ensemble fini qui engendre V. Alors, V possède une K-base contenue dans E.*

Dans l'appendice à cette section, nous allons démontrer à l'aide du lemme de Zorn que tout espace vectoriel possède une base.

Exemple 1.24. *(a) Soit $V = \left\{ \left(\begin{smallmatrix} a \\ a \\ b \end{smallmatrix}\right) \mid a, b \in \mathbb{R} \right\}$. Une base de V est $\left\{ \left(\begin{smallmatrix} 1 \\ 1 \\ 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} 0 \\ 0 \\ 1 \end{smallmatrix}\right) \right\}$.*

(b) Soit $V = \langle \left(\begin{smallmatrix} 1 \\ 2 \\ 3 \end{smallmatrix}\right), \left(\begin{smallmatrix} 2 \\ 3 \\ 4 \end{smallmatrix}\right), \left(\begin{smallmatrix} 3 \\ 5 \\ 7 \end{smallmatrix}\right) \rangle \subseteq \mathbb{Q}^3$.

L'ensemble $E = \left\{ \left(\begin{smallmatrix} 1 \\ 2 \\ 3 \end{smallmatrix}\right), \left(\begin{smallmatrix} 2 \\ 3 \\ 4 \end{smallmatrix}\right) \right\}$ est une $\mathbb{Q}$-base de V. Raison :

- *Le système d'équations linéaires*

$$a_1 \cdot \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + a_2 \cdot \begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix} + a_3 \cdot \begin{pmatrix} 3 \\ 5 \\ 7 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

  possède une solution non nulle (par exemple $a_1 = 1$, $a_2 = 1$, $a_3 = -1$). Cela implique que E engendre V car on peut exprimer le troisième générateur par les deux premiers.

- *Le système d'équations linéaires*

$$a_1 \cdot \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + a_2 \cdot \begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

  ne possède que $a_1 = a_2 = 0$ comme solution. Donc E est $\mathbb{Q}$-linéairement indépendant.

(c) Le $\mathbb{R}$-espace vectoriel

$$V = \{f : \mathbb{N} \to \mathbb{R} \mid \exists\, S \subseteq \mathbb{N} \text{ fini } \forall\, n \in \mathbb{N} \setminus S : f(n) = 0\}$$

possède $\{e_n \mid n \in \mathbb{N}\}$ avec $e_n(m) = \delta_{n,m}$ (Delta de Kronecker : $\delta_{n,m} = \begin{cases} 1 & \text{si } n = m, \\ 0 & \text{si } n \neq m. \end{cases}$) comme $\mathbb{R}$-base. Cela est donc une base avec une infinité d'éléments.

(d) Similaire à l'exemple précédent, le $\mathbb{R}$-espace vectoriel

$$V = \{f : \mathbb{R} \to \mathbb{R} \mid \exists\, S \subseteq \mathbb{R} \text{ fini } \forall\, x \in \mathbb{R} \setminus S : f(x) = 0\}$$

possède $\{e_x \mid x \in \mathbb{R}\}$ avec $e_x(y) = \delta_{x,y}$ comme $\mathbb{R}$-base. Cela est donc une base qui n'est pas dénombrable.

Exemple 1.25. *Comment calculer une base pour un espace vectoriel engendré par un ensemble fini de vecteurs ? (Même réponse que presque toujours :) Résoudre des systèmes d'équations linéaires. Soit V un K-espace vectoriel engendré par $\{e_1, e_2, \ldots, e_m\}$ (supposés tous non-nuls). On procède ainsi :*

- *Ajouter e_1 à la base.*
- *Si e_2 est linéairement indépendant de e_1 (c'est-à-dire que e_2 n'est pas un multiple scalaire de e_1), ajouter e_2 à la base et dans ce cas e_1, e_2 sont linéairement indépendants (sinon, ne rien faire).*
- *Si e_3 est linéairement indépendant des vecteurs déjà choisis pour la base, ajouter e_3 à la base et dans ce cas tous les éléments choisis pour la base sont linéairement indépendants (sinon, ne rien faire).*
- *Si e_4 est linéairement indépendant des vecteurs déjà choisis pour la base, ajouter e_4 à la base et dans ce cas tous les éléments choisis pour la base sont linéairement indépendants (sinon, ne rien faire).*
- *etc. jusqu'au dernier vecteur.*

Voici un exemple concret dans $\mathbb{R}^4$:

$$e_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 2 \end{pmatrix}, e_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, e_3 = \begin{pmatrix} 4 \\ 1 \\ 3 \\ 2 \end{pmatrix}, e_4 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}.$$

- *Ajouter e_1 à la base.*
- *Ajouter e_2 à la base car e_2 n'est clairement pas un multiple de e_1 (regarder, par exemple, le deuxième coefficient), donc e_1 et e_2 sont linéairement indépendants.*
- *Est-ce que e_1, e_2, e_3 sont linéairement indépendants ? On considère le système d'équations linéaires donné par la matrice*
$$\begin{pmatrix} 1 & 1 & 4 \\ 1 & 0 & 1 \\ 0 & 1 & 3 \\ 2 & 0 & 2 \end{pmatrix}.$$

Par des transformations de lignes, on obtient la matrice

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

On obtient la solution $\begin{pmatrix} 1 \\ 3 \\ -1 \end{pmatrix}$. *Donc on n'ajoute pas* e_3 *à la base car* e_3 *est linéairement dépendant de* e_1, e_2.

- *Est-ce que* e_1, e_2, e_4 *sont linéairement indépendants ? On considère le système d'équations linéaires donné par la matrice*

$$\begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix}.$$

Par des transformations de lignes, on obtient la matrice

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

Le système correspondant n'admet pas de solution non-nulle. Donc e_1, e_2, e_4 *sont linéairement indépendants. C'est la base cherchée.*

Dimension

Corollaire 1.26. *Soient* K *un corps et* V *un* K*-espace vectoriel qui possède une* K*-base finie. Alors, toutes les* K*-bases de* V *sont finies et ont la même cardinalité.*

Ce corollaire nous permet de faire une définition très importante, celle de la dimension d'un espace vectoriel. La dimension mesure la « taille » ou le « nombre de degrés de liberté » d'un espace vectoriel.

Définition 1.27. *Soient* K *un corps et* V *un* K*-espace vectoriel. Si* V *possède une* K*-base finie de cardinalité* n*, on dit que* V *est de* dimension n. *Si* V *ne possède pas de* K*-base finie, on dit que* V *est de* dimension infinie.
Notation : $\dim_K(V)$.

Exemple 1.28. *(a) La dimension du* K*-espace vectoriel standard* K^n *est égale à* n.

(b) Le K*-espace vectoriel nul* $(\{0\}, +, \cdot, 0)$ *est de dimension* 0 *(et c'est le seul).*

(c) Le $\mathbb{R}$*-espace vectoriel* $\mathcal{F}(\mathbb{N}, \mathbb{R})$ *est de dimension infinie.*

Lemme 1.29. *Soient* K *un corps,* V *un* K*-espace vectoriel de dimension* n *et* $W \leq V$ *un sous-espace.*

(a) $\dim_K(W) \leq \dim_K(V)$.

(b) Si $\dim_K(W) = \dim_K(V)$*, alors* $W = V$.

Le contenu de la proposition suivante est que tout ensemble K-linéairement indépendant peut être complété pour devenir une K-base.

Proposition 1.30 (Basisergänzungssatz)**.** *Soient V un K-espace vectoriel de dimension n, $E \subseteq V$ un ensemble fini tel que E engendre V et $\{e_1, \dots, e_r\} \subset V$ un sous-ensemble qui est K-linéairement indépendant.*
Alors $r \leq n$ et il existe $e_{r+1}, e_{r+2}, \dots, e_n \in E$ tels que $\{e_1, \dots, e_n\}$ est une K-base de V.

La proposition 1.30 se démontre de façon abstraite ou de façon constructive. Supposons que nous avons déjà des éléments $e_1, \dots, e_r$ qui sont K-linéairement indépendants. Si $r = n$, ces éléments sont une K-base par le lemme 1.29 (b) et il ne reste rien à faire. Supposons donc $r < n$. Nous parcourons maintenant les éléments de E jusqu'à trouver un $e \in E$ tel que $e_1, \dots, e_r, e$ sont K-linéairement indépendants. Un tel e doit exister car sinon l'ensemble E serait contenu dans le sous-espace engendré par $e_1, \dots, e_r$, il ne pourrait donc pas engendrer V. On nomme $e =: e_{r+1}$ et on a un ensemble K-linéairement indépendant de cardinalité $r + 1$. Il suffit maintenant de continuer ce processus jusqu'à arriver à un ensemble K-linéairement indépendant de n éléments, qui est automatiquement une K-base.

Corollaire 1.31. *Soit V un K-espace vectoriel de dimension finie n et soit $W \leq V$ un sous-espace vectoriel. Alors il existe un sous-espace vectoriel $U \leq V$ tel que $V = U \oplus V$. En plus, nous avons l'égalité $\dim(V) = \dim(W) + \dim(U)$.*
On appelle U un complément *de W dans V. Noter que ce complément n'est pas unique en général.*

Démonstration. On choisit une K-base $w_1, \dots, w_r$ de W et on utilise la proposition 1.30 pour obtenir des vecteurs $u_1, \dots, u_s \in V$ tels que $w_1, \dots, w_r, u_1, \dots, u_s$ forment une K-base de V. Mettons $U = \langle u_1, \dots, u_s \rangle$. Nous avons clairement $V = U + W$ et aussi $U \cap W = 0$, donc $V = U \oplus W$. L'assertion concernant les dimensions en suit. $\square$

Proposition 1.32. *Soit V un K-espace vectoriel de dimension finie n. Soit $B \subset V$ un sous-ensemble de cardinal n. Alors, les assertions suivantes sont équivalentes.*

(i) *B est une K-base.*

(ii) *B est K-linéairement indépendant.*

(iii) *B engendre V.*

Démonstration. Pour l'équivalence entre (i) et (ii) il suffit de remarquer qu'un ensemble K-linéairement indépendant de cardinal n est nécessairement maximal (donc une K-base par le théorème 1.22), car s'il n'était pas maximal, il y aurait un ensemble maximal K-linéairement indépendant de cardinal strictement supérieur à n, donc une K-base de cardinal différent de n ce qui n'est pas possible selon le corollaire 1.26.
Similairement, pour l'équivalence entre (i) et (iii) il suffit de remarquer qu'un ensemble de cardinal n qui engendre V est nécessairement minimal (donc une K-base par le théorème 1.22), car s'il n'était pas minimal, il y aurait un ensemble minimal de cardinal strictement inférieur à n qui engendre V, donc une K-base de cardinal différent de n. $\square$

Applications linéaires : les homomorphismes des espaces vectoriels

Nous commençons par l'idée principale :
Les (homo-)morphismes sont des applications qui respectent toutes les structures.

Définition 1.33. *Soient V, W des K-espaces vectoriels. Une application*

$$\varphi : V \to W$$

est appelée K-linéaire *ou* (homo-)morphisme de K-espaces vectoriels *si*

$$\forall\, v_1, v_2 \in V : \varphi(v_1 +_V v_2) = \varphi(v_1) +_W \varphi(v_2)$$

et

$$\forall\, v \in V, \forall a \in K : \varphi(a \cdot_V v) = a \cdot_W \varphi(v).$$

Un homomorphisme bijectif de K-espaces vectoriels s'appelle isomorphisme. *On note souvent les isomorphismes par un tilde : $\varphi : V \xrightarrow{\sim} W$. S'il existe un isomorphisme $V \to W$, on écrit souvent simplement $V \cong W$.*

Exemple 1.34. *(a) On commence par l'exemple le plus important. Soit $n \in \mathbb{N}$.*

Soit $M = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{pmatrix}$ une matrice à n colonnes, m lignes et à coefficients dans K (on note l'ensemble de ces matrices par $\mathrm{Mat}_{m\times n}(K)$; c'est aussi un K-espace vectoriel). Elle définit l'application K-linéaire

$$\varphi_M : K^n \to K^m, \quad v \mapsto Mv$$

où Mv est le produit habituel de matrices. Explicitement,

$$\varphi_M(v) = Mv = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^n a_{1,i} v_i \\ \sum_{i=1}^n a_{2,i} v_i \\ \vdots \\ \sum_{i=1}^n a_{m,i} v_i \end{pmatrix}.$$

La K-linéarité s'exprime comme

$$\forall\, a \in K\, \forall\, v, w \in V : M \circ (a \cdot v + w) = a \cdot (M \circ v) + M \circ w.$$

Cette égalité est très facile à vérifier (vous avez dû la voir dans votre cours d'algèbre linéaire 1).

(b) Soit $a \in \mathbb{R}$. Alors, $\varphi : \mathbb{R} \to \mathbb{R}$, $x \mapsto ax$ est $\mathbb{R}$-linéaire (c'est le cas spécial $n = m = 1$ de (a) si l'on regarde le scalaire a comme une matrice (a)). Par contre, si $0 \neq b \in \mathbb{R}$, alors $\mathbb{R} \to \mathbb{R}$, $x \mapsto ax + b$ n'est pas $\mathbb{R}$-linéaire !

(c) Soit $n \in \mathbb{N}$. Alors, l'application $\varphi : \mathcal{F}(\mathbb{N}, \mathbb{R}) \to \mathbb{R}$, $f \mapsto f(n)$ est K-linéaire.

Définition 1.35. *Soient V, W des K-espaces vectoriels et $\varphi : V \to W$ une application K-linéaire. Le* noyau de φ *est défini comme*

$$\ker(\varphi) = \{v \in V \mid \varphi(v) = 0\}.$$

Proposition 1.36. *Soient V, W des K-espaces vectoriels et $\varphi : V \to W$ une application K-linéaire.*

(a) $\mathrm{Im}(\varphi)$ *est un sous-espace vectoriel de W.*

(b) $\ker(\varphi)$ *est un sous-espace vectoriel de V.*

(c) φ *est surjectif si et seulement si* $\mathrm{Im}(\varphi) = W$.

(d) φ *est injectif si et seulement si* $\ker(\varphi) = 0$.

(e) Si φ est un isomorphisme, son inverse l'est aussi (en particulier, son inverse est aussi K-linéaire).

Définition 1.37. *Soit $M \in \mathrm{Mat}_{m\times n}(K)$ une matrice. On appelle* rang des colonnes de M *(anglais : rank) la dimension du sous-espace vectoriel de K^m engendré par les colonnes de M. On utilise la notation* $\mathrm{rk}(M)$.
Similairement, on définit le rang des lignes *de M la dimension du sous-espace vectoriel de K^n engendré par les lignes de M. Plus formellement, c'est le rang de M^{tr}, la matrice transposée.*

On verra vers la fin du cours que pour toute matrice le rang des colonnes est égal au rang des lignes. Cela explique pourquoi nous ne mentionnons pas le mot « colonnes » dans la notation pour le rang. Si $\varphi_M : K^n \to K^m$ est l'application K-linéaire associée à M, alors

$$\mathrm{rk}(M) = \dim(\mathrm{Im}(\varphi_M))$$

car l'image de φ_M est précisement l'espace vectoriel engendré par les colonnes de M.

Corollaire 1.38. *(a) Soit $\varphi : V \to X$ une application K-linéaire entre deux K-espaces vectoriels. On suppose V de dimension finie. Alors,*

$$\dim(V) = \dim(\ker(\varphi)) + \dim(\mathrm{Im}(\varphi)).$$

(b) Soit $M \in \mathrm{Mat}_{m\times n}(K)$ une matrice. Alors, on a

$$n = \dim(\ker(M)) + \mathrm{rk}(M).$$

Démonstration. (a) Soit $W = \ker(\varphi)$. On choisit un complément $U \leq V$ tel que $V = U \oplus W$ par le corollaire 1.31. Comme $U \cap W = 0$, l'application $\varphi|_U : U \to X$ est injective. En plus, $\varphi(V) = \varphi(U + W) = \varphi(U)$ montre que $\mathrm{Im}(\varphi)$ est égal à $\varphi(U)$. En conséquence, $\dim(\mathrm{Im}(\varphi)) = \dim(\varphi(U)) = \dim(U)$, d'où l'égalité voulue.
(b) suit directement de (a) par les considérations ci-dessus. $\square$

La partie (b) est très utile pour le calcul du noyau d'une matrice : si on connaît le rang de M, on en déduit la dimension du noyau par la formule

$$\dim(\ker(M)) = n - \mathrm{rk}(M).$$

L'algorithme de Gauß en termes de matrices

Nous considérons trois types de matrices :

Définition 1.39. *Pour $0 \neq \lambda \in K$ et $1 \leq i, j \leq n$, $i \neq j$, nous définissons les matrices suivantes dans $\mathrm{Mat}_{n \times n}(K)$, appelées* matrices élémentaires *:*

- *$P_{i,j}$ est égal à l'identité id_n sauf que la i-ième et la j-ième lignes sont échangées (ou, ce qui équivaut, la i-ième et la j-ième colonnes sont échangées) :* $P_{i,j} = \begin{pmatrix} 1 &&&&&&&& \\ & \ddots &&&&&&& \\ && 1 &&&&&& \\ &&& 0 &&& 1 && \\ &&&& 1 &&&& \\ &&&&& 1 &&& \\ &&& 1 &&& 0 && \\ &&&&&&& 1 & \\ &&&&&&&& \ddots & \\ &&&&&&&&& 1 \end{pmatrix}$.
- *$S_i(\lambda)$ est égal à l'identité id_n sauf que le coefficient (i,i) sur la diagonale est λ (au lieu de 1) :* $S_i(\lambda) = \begin{pmatrix} 1 &&&&& \\ & \ddots &&&& \\ && 1 &&& \\ &&& \lambda && \\ &&&& 1 & \\ &&&&& \ddots & \\ &&&&&& 1 \end{pmatrix}$.
- *$Q_{i,j}(\lambda)$ est égal à l'identité id_n sauf que le coefficient (i,j) est λ (au lieu de 0) :* $Q_{i,j}(\lambda) = \begin{pmatrix} 1 &&& \\ & \ddots && \\ && \ddots & \lambda \\ &&& \ddots & \\ &&&& 1 \end{pmatrix}$.

Les matrices élémentaires ont une signification pour les opérations de matrices.

Lemme 1.40. *Soient $\lambda \in K$, $i, j, n, m \in \mathbb{N}_{>0}$, $i \neq j$ et $M \in \mathrm{Mat}_{n \times m}(K)$.*

(a) $P_{i,j}M$ est la matrice obtenue à partir de M en échangeant la i-ième et la j-ième ligne.
$MP_{i,j}$ est la matrice obtenue à partir de M en échangeant la i-ième et la j-ième colonne.

(b) $S_i(\lambda)M$ est la matrice obtenue à partir de M en multipliant la i-ième ligne par λ.
$MS_i(\lambda)$ est la matrice obtenue à partir de M en multipliant la i-ième colonne par λ.

(c) $Q_{i,j}(\lambda)M$ est la matrice obtenue de M en additionnant λ fois la j-ième ligne sur la i-ième ligne.
$MQ_{i,j}(\lambda)$ est la matrice obtenue de M en additionnant λ fois la i-ième colonne sur la j-ième colonne.

Démonstration. Calculs simples. □

Proposition 1.41. *Soit $M \in \mathrm{Mat}_{n \times m}(K)$ une matrice et soit $N \in \mathrm{Mat}_{n \times m}(K)$ la matrice obtenue de M en faisant des opérations sur les lignes (comme dans l'algorithme de Gauß).*

(a) Alors il existe des matrices $C_1, \ldots, C_r$ (pour un $r \in \mathbb{N}$) choisies parmi les matrices de la définition 1.39 telles que $(C_1 \cdots C_r) \cdot M = N$.

(b) $\ker(M) = \ker(N)$ *et donc l'algorithme du pivot de Gauß peut être utilisé pour calculer le noyau d'une matrice.*

Démonstration. (a) Par le lemme 1.40 toute opération sur les lignes peut être éffectuée par la multiplication à gauche par une des matrices de la définition 1.39.
(b) Toutes les matrices de la définition 1.39 sont inversibles, donc ne changent pas le noyau. $\square$

Similairement à (b), toute opération sur les colonnes correspond à la multiplication à droite par une des matrices de la définition 1.39. Donc, si N est une matrice obtenue d'une matrice M en éffectuant des opérations sur les colonnes, il existe des matrices $C_1, \dots, C_r$ (pour un $r \in \mathbb{N}$) choisies parmi les matrices de la définition 1.39 telles que $M \cdot (C_1 \cdots C_r) = N$. Comme les C_i sont inversibles, on a aussi

$$\mathrm{im}(M) = \mathrm{im}(N),$$

et on particulier le rang de M est égal au rang de N.

Souvent on s'intéresse à connaître une matrice C telle que $CM = N$ où N est obtenu de M par des opérations sur les lignes. Pour l'obtenir il suffit de se rendre compte $C \cdot \mathrm{id} = C$, donc qu'appliquer C revient à faire les opérations sur les lignes correspondantes sur la matrice id. Dans l'exemple suivant, on voit comment le faire en pratique.

Exemple 1.42. *Soit* $M = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$. *On écrit la matrice augmentée et on fait les opérations sur les lignes comme toujours, mais sur toute la matrice.*

$$\begin{pmatrix} 1 & 0 & 0 & 1 & 2 & 3 \\ 0 & 1 & 0 & 4 & 5 & 6 \\ 0 & 0 & 1 & 7 & 8 & 9 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 & 0 & 1 & 2 & 3 \\ -4 & 1 & 0 & 0 & -3 & -6 \\ -7 & 0 & 1 & 0 & -6 & -12 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 & 0 & 1 & 2 & 3 \\ -4 & 1 & 0 & 0 & -3 & -6 \\ 1 & -2 & 1 & 0 & 0 & 0 \end{pmatrix}$$
$$\mapsto \begin{pmatrix} 1 & 0 & 0 & 1 & 2 & 3 \\ 4/3 & -1/3 & 0 & 0 & 1 & 2 \\ 1 & -2 & 1 & 0 & 0 & 0 \end{pmatrix} \mapsto \begin{pmatrix} -5/3 & 2/3 & 0 & 1 & 0 & -1 \\ 4/3 & -1/3 & 0 & 0 & 1 & 2 \\ 1 & -2 & 1 & 0 & 0 & 0 \end{pmatrix}$$

La moitié gauche de la matrice finale est la matrice C *cherchée :* $C = \begin{pmatrix} -5/3 & 2/3 & 0 \\ 4/3 & -1/3 & 0 \\ 1 & -2 & 1 \end{pmatrix}$. *La moitié droite est la matrice obtenue par les opérations sur les lignes.*
On sait que l'on a l'égalité suivante (pour s'en convaincre, on peut la vérfier par un petit calcul) :

$$CM = \begin{pmatrix} -5/3 & 2/3 & 0 \\ 4/3 & -1/3 & 0 \\ 1 & -2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix}.$$

Comme application de l'écriture de l'algorithme de Gauß par les matrices, on obtient que toute matrice carrée M inversible s'écrit comme produit des matrices de la définition 1.39. En effet, qu'on peut transformer M en l'identité par des opérations sur les lignes.

Matrices et représentation des applications linéaires

Dans l'exemple 1.34 (a) nous avons vu que les matrices donnent lieu à des applications K-linéaires. Il est très important et parfois appelé *théorème principal de l'algèbre linéaire* que l'assertion inverse est aussi vraie : **après choix de bases** toute application K-linéaire est donnée par une matrice.

Notation 1.43. *Soient V un K-espace vectoriel et $S = \{v_1, \dots, v_n\}$ une K-base de V. Nous rappelons que l'on a $v = \sum_{i=1}^n b_i v_i$ avec des uniques $b_1, \dots, b_n \in K$; ce sont les coordonnées de v pour la base S. Nous utilisons la notation suivante :*

$$v_S = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} \in K^n.$$

Exemple 1.44. *(a) Soient $n \in \mathbb{N}$ et $e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, e_n = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$.*

Donc $E = \{e_1, e_2, \dots, e_n\}$ est la K-base canonique de K^n. Alors, pour tout $v = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_n \end{pmatrix} \in K^n$ on a $v_E = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_n \end{pmatrix}$.

(b) Soit $V = \mathbb{R}^2$ et $S = \{\left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right), \left(\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right)\}$. C'est une $\mathbb{R}$-base de V (car la dimension est 2 et les deux vecteurs sont $\mathbb{R}$-linéairement indépendants). Soit $v = \left(\begin{smallmatrix} 4 \\ 2 \end{smallmatrix}\right) \in V$. Alors, $v = 3 \cdot \left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right) + \left(\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right)$, donc $v_S = \left(\begin{smallmatrix} 3 \\ 1 \end{smallmatrix}\right)$.

La proposition suivante dit que tout K-espace vectoriel de dimension n est isomorphe à K^n.

Proposition 1.45. *Soit V un K-espace vectoriel de dimension finie n avec K-base $S = \{v_1, \dots, v_n\}$. Alors, l'application $\varphi = ()_S : V \to K^n$ donnée par $v \mapsto v_S$ est un K-isomorphisme.*

Démonstration. Soient $v, w \in V$ et $a \in K$. On écrit v et w en coordonnées pour la base S : $v = \sum_{i=1}^n b_i v_i$ et $w = \sum_{i=1}^n c_i v_i$. Donc, nous avons $av + w = \sum_{i=1}^n (ab_i + c_i) v_i$. Ecrit comme vecteurs on trouve alors :

$$v_S = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}, \quad w_S = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} \text{ et } (av + w)_S = \begin{pmatrix} ab_1 + c_1 \\ ab_2 + c_2 \\ \vdots \\ ab_n + c_n \end{pmatrix},$$

donc l'égalité $(a \cdot v + w)_S = a \cdot v_S + w_S$. Cela montre que l'application φ est K-linéaire. On démontre qu'elle est bijective.

Injectivité : Soit $v \in V$ tel que $v_S = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$, i.e. $v \in \ker(\varphi)$. Cela veut dire que $v = \sum_{i=1}^n 0 \cdot v_i = 0$.

Le noyau de φ ne contient donc que 0, alors, φ est injective.

Surjectivité : Soit $\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \in K^n$. On pose $v := \sum_{i=1}^n a_i \cdot v_i$. Nous avons $\varphi(v) = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$ et la surjectivité est démontrée.

□

Théorème 1.46. *Soient V, W deux K-espaces vectoriels de dimensions finies n et m et $\varphi : V \to W$ une application K-linéaire. Soient $S = \{v_1, \dots, v_n\}$ une K-base de V et $T = \{w_1, \dots, w_m\}$ une K-base de W. Pour tout $1 \leq i \leq n$, le vecteur $\varphi(v_i)$ appartient à W. On peut donc l'exprimer en tant que combinaison K-linéaire des vecteurs dans la base T ainsi :*

$$\varphi(v_i) = \sum_{j=1}^{m} a_{j,i} w_j.$$

Nous « rassemblons » les coefficients $a_{j,i}$ dans une matrice :

$$M_{T,S}(\varphi) := \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{pmatrix} \in \mathrm{Mat}_{m \times n}(K).$$

Alors, pour tout $v \in V$ on a

$$(\varphi(v))_T = M_{T,S}(\varphi) \circ v_S.$$

C'est-à-dire que le produit matriciel $M_{T,S}(\varphi) \circ v_S$ donne les coordonnées dans la base T de l'image $\varphi(v)$. Alors, la matrice $M_{T,S}(\varphi)$ décrit l'application K-linéaire φ en coordonnées.

Remarquons qu'il est facile d'écrire la matrice $M_{T,S}(\varphi)$: la i-ème colonne de $M_{T,S}(\varphi)$ est $(\varphi(v_i))_T$.

Démonstration. Nous faisons un calcul matriciel très facile :

$$M_{T,S}(\varphi) \circ (v_i)_S = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{pmatrix} \circ \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} a_{1,i} \\ a_{2,i} \\ \vdots \\ a_{m,i} \end{pmatrix} = (\varphi(v_i))_T,$$

où le 1 est dans la i-ième ligne du vecteur. Nous avons donc obtenu le résultat pour les vecteurs v_i dans la base S.
L'assertion générale suit par linéarité : Soit $v = \sum_{i=1}^n b_i v_i$. Alors nous obtenons

$$\begin{aligned} M_{T,S}(\varphi) \circ (\sum_{i=1}^{n} b_i v_i)_S &= \sum_{i=1}^{n} b_i \cdot \big(M_{T,S}(\varphi) \circ (v_i)_S\big) \\ &= \sum_{i=1}^{n} b_i \cdot (\varphi(v_i))_T = (\sum_{i=1}^{n} b_i \cdot \varphi(v_i))_T = (\varphi(\sum_{i=1}^{n} b_i \cdot v_i))_T = (\varphi(v))_T. \end{aligned}$$

Cela montre le théorème. □

Exemple 1.47. *$\mathbb{C}$ possède la $\mathbb{R}$-base $B = \{1, i\}$. Soit $z = x + iy \in \mathbb{C}$ avec $x, y \in \mathbb{R}$, donc $z_B = \left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right)$. Soit $a = r + is$ avec $r, s \in \mathbb{R}$. L'application*

$$\varphi : \mathbb{C} \to \mathbb{C}, \quad z \mapsto a \cdot z$$

est $\mathbb{R}$-linéaire. Nous décrivons $M_{B,B}(\varphi)$. La première colonne est $(a \cdot 1)_B = (r + is)_B = \left(\begin{smallmatrix} r \\ s \end{smallmatrix}\right)$, et la deuxième colonne est $(a \cdot i)_B = (-s + ir)_B = \left(\begin{smallmatrix} -s \\ r \end{smallmatrix}\right)$, alors $M_{B,B}(\varphi) = \left(\begin{smallmatrix} r & -s \\ s & r \end{smallmatrix}\right)$.

Définition 1.48. *Notons $\mathrm{Hom}_K(V, W)$ l'ensemble de toutes les applications $\varphi : V \to W$ qui sont K-linéaires.*
Dans le cas spécial $W = V$, une application K-linéaire $\varphi : V \to V$ est aussi appelée endomorphisme *de V et nous écrivons*

$$\mathrm{End}_K(V) := \mathrm{Hom}_K(V, V).$$

Corollaire 1.49. *Soient K un corps, V, W deux K-espaces vectoriels de dimensions finies n et m. Soient $S = \{v_1, \dots, v_n\}$ une K-base de V et $T = \{w_1, \dots, w_m\}$ une K-base de W.*
Alors, l'application

$$\mathrm{Hom}_K(V, W) \to \mathrm{Mat}_{m \times n}(K), \quad \varphi \mapsto M_{T,S}(\varphi)$$

est une bijection.

Il est important de souligner que les bases dans le corollaire sont fixées ! La même matrice peut exprimer des applications linéaires différentes si on change les bases.

Démonstration. **Injectivité :** Supposons $M_{T,S}(\varphi) = M_{T,S}(\psi)$ pour $\varphi, \psi \in \mathrm{Hom}_K(V, W)$. Alors pour tout $v \in V$, on a $(\varphi(v))_T = M_{T,S}(\varphi) \circ v_S = M_{T,S}(\psi) \circ v_S = (\psi(v))_T$. Comme l'écriture en coordonnées est unique, nous trouvons $\varphi(v) = \psi(v)$ pour tout $v \in V$, donc $\varphi = \psi$.

Surjectivité : Soit $M \in \mathrm{Mat}_{m \times n}(K)$ une matrice. On définit $\varphi \in \mathrm{Hom}_K(V, W)$ par

$$(\varphi(v))_T = M \circ v_S$$

pour $v \in V$. Il est clair que φ est K-linéaire. En plus, nous avons

$$M_{T,S}(\varphi) \circ v_S = (\varphi(v))_T = M \circ v_S$$

pour tout $v \in V$. Prenant $v = v_i$ de façon que $(v_i)_S$ est le vecteur dont la i-ème coordonnée est 1 et le reste est 0, on obtient que les i-èmes colonnes de $M_{T,S}(\varphi)$ et de M sont les mêmes. Cela montre $M = M_{T,S}(\varphi)$.

□

Définition-Lemme 1.50. *Soit V un K-espace vectoriel de dimension finie n. Soient S_1, S_2 deux K-bases de V. On pose*

$$C_{S_2,S_1} := M_{S_2,S_1}(\mathrm{id}_V)$$

et on l'appelle matrice de changement de bases.

(a) C_{S_2,S_1} est une matrice à n colonnes et n lignes.

(b) Pour tout $v \in V$:

$$v_{S_2} = C_{S_2,S_1} \circ v_{S_1}.$$

En mots : la multiplication de la matrice de changement de bases par le vecteur v exprimé en coordonnées pour la base S_1, donne le vecteur v exprimé en coordonnées pour la base S_2.

(c) C_{S_2,S_1} est inversible d'inverse C_{S_1,S_2}.

Il est facile d'écrire la matrice C_{S_2,S_1} : sa j-ième colonne est formée des coordonnées dans la base S_2 du j-ième vecteur de la base S_1.

Démonstration. (a) C'est clair.
(b) $C_{S_2,S_1} \circ v_{S_1} = M_{S_2,S_1}(\mathrm{id}_V) \circ v_{S_1} = (\mathrm{id}_V(v))_{S_2} = v_{S_2}$.
(c) $C_{S_1,S_2} \circ C_{S_2,S_1} \circ v_{S_1} = C_{S_1,S_2} \circ v_{S_2} = v_{S_1}$ pour tout $v \in V$. Cela montre que $C_{S_1,S_2} \circ C_{S_2,S_1}$ est l'identité. Le même raisonnement marche avec les rôles de S_1 et S_2 inversés. □

Proposition 1.51. *Soient V, W des K-espaces vectoriels de dimension finie, soient S_1, S_2 deux K-bases de V, soient T_1, T_2 deux K-bases de W, et soit $\varphi \in \mathrm{Hom}_K(V, W)$. Alors,*

$$M_{T_2,S_2}(\varphi) = C_{T_2,T_1} \circ M_{T_1,S_1}(\varphi) \circ C_{S_1,S_2}.$$

Démonstration. $C_{T_2,T_1} \circ M_{T_1,S_1}(\varphi) \circ C_{S_1,S_2} \circ v_{S_2} = C_{T_2,T_1} \circ M_{T_1,S_1}(\varphi) v_{S_1} = C_{T_2,T_1} \circ (\varphi(v))_{T_1} = (\varphi(v))_{T_2}$. □

Proposition 1.52. *Soient V, W, Z des K-espaces vectoriels de dimension finie, soient S une K-base de V, T une K-base de W et U une K-base de Z. Soient $\varphi \in \mathrm{Hom}_K(V, W)$ et $\psi \in \mathrm{Hom}_K(W, Z)$. Alors,*

$$M_{U,T}(\psi) \circ M_{T,S}(\varphi) = M_{U,S}(\psi \circ \varphi).$$

En mots : le produit matriciel correspond à la composition d'applications.

Démonstration. $M_{U,T}(\psi) \circ M_{T,S}(\varphi) \circ v_S = M_{U,T}(\psi) \circ (\varphi(v))_T = (\psi(\phi(v)))_U = M_{U,T}(\psi \circ \varphi) \circ v_S$. □

Appendice : existence de bases

Pour manque de temps, cette section ne sera ni enseignée ni examinée.
Dans le cours « Structures mathématiques » nous avons introduit les ensembles d'un point de vue intuitif et non rigoureux. Un traitement strict ne peut se faire que dans un cours de logique à un moment plus avancé (un tel cours n'est pas offert à l'UL en ce moment – vous pouvez regarder des livres pour plus de détails). Dans la théorie des ensembles il y a un axiome important : « l'axiome du choix ».[1] Dans la théorie des ensembles on montre le « lemme de Zorn » qui dit que l'axiome du choix est équivalent à l'assertion suivante.

[1] L'axiome du choix : Soit X un ensemble dont les éléments sont des ensembles non vides. Alors il existe une fonction f définie sur X qui à chaque $M \in X$ associe un élément de M. Une telle fonction est appelée « fonction du choix ».

Axiome 1.53 (Lemme de Zorn)**.** *Soit S un ensemble non-vide et $\leq$ une relation d'ordre sur S.*[2] *On fait l'hypothèse suivante : Tout sous-ensemble $T \subseteq S$ qui est totalement ordonné*[3] *possède un majorant.*[4] *Alors, S contient un élément maximal.*[5]

Pour montrer comment appliquer le lemme de Zorn, nous démontrons que tout espace vectoriel possède une base. Si vous avez vu cette assertion dans votre cours d'algèbre linéaire 1, alors c'était pour des espaces vectoriels de dimension finie car le cas général est en fait équivalent à l'axiome du choix (et donc au lemme de Zorn).

Proposition 1.54. *Soit K un corps et $V \neq \{0\}$ un K-espace vectoriel. Alors, V possède une K-base.*

Démonstration. On rappelle quelques notions d'algèbre linéaire. Un sous-ensemble fini $G \subseteq V$ est appelé *K-linéairement indépendant* si la seule combinaison linéaire $0 = \sum_{g \in G} a_g g$ avec $a_g \in K$ est celle où $a_g = 0$ pour tout $g \in G$. Plus généralement, un sous-ensemble non nécessairement fini $G \subseteq V$ est appelé *K-linéairement indépendant* si tout sous-ensemble fini $H \subseteq G$ est K-linéairement indépendant. Un sous-ensemble $G \subseteq V$ est appelé une *K-base* s'il est K-linéairement indépendant et engendre V.[6]
On veut utiliser le lemme de Zorn 1.53. Soit

$$S := \{G \subseteq V \text{ sous-ensemble} \mid G \text{ est } K\text{-linéairement indépendant } \}.$$

L'ensemble S est non-vide car $G = \{v\}$ est K-linéairement indépendant pour tout $0 \neq v \in V$. L'inclusion d'ensembles « $\subseteq$ » définit une relation d'ordre sur S (c'est évident – voir Algèbre 1).
On vérifie que l'hypothèse du lemme de Zorn est satisfaite : Soit $T \subseteq S$ un sous-ensemble totalement ordonné. On doit produire un majorant $E \in S$ pour T. On pose $E := \bigcup_{G \in T} G$. Il est clair que $G \subseteq E$ pour tout $G \in T$. Il faut montrer que $E \in S$, donc que E est K-linéairement indépendant. Soit $H \subseteq E$ un sous-ensemble de cardinal n. On montre par récurrence en n qu'il existe $G \in T$ tel que $H \subseteq G$. L'assertion est claire pour $n = 1$. Supposons-la démontrée pour $n - 1$ et écrivons $H = H' \sqcup \{h\}$. Il existe $G', G \in T$ tels que $H' \subseteq G'$ (par l'hypothèse de récurrence car le cardinal de H' est $n - 1$) et $h \in G$ (par le cas $n = 1$). Par le fait que T est totalement ordonné, on a $G \subseteq G'$ ou $G' \subseteq G$. Dans les deux cas on obtient que H est un sous-ensemble de G ou de G'. Comme H est un sous-ensemble fini d'un ensemble qui est K-linéairement indépendant, H l'est aussi. Donc, E est K-linéairement indépendant.
Le lemme de Zorn nous donne un élément maximal $B \in S$. On montre que B est une K-base de V. En tant qu'élément de S, B est K-linéairement indépendant. Il faut démontrer que B engendre V. Supposons que cela ne soit pas le cas et prenons $v \in V$ qui ne s'écrit pas comme combinaison K-linéaire

[2] On rappelle que par définition les trois points suivants sont satisfaits :

- $s \leq s$ pour tout $s \in S$.
- Si $s \leq t$ et $t \leq s$ pour $s, t \in S$, alors $s = t$.
- Si $s \leq t$ et $t \leq u$ pour $s, t, u \in S$, alors $s \leq u$.

[3] T est totalement ordonné si T est ordonné et pour tout couple $s, t \in T$ on a $s \leq t$ ou $t \leq s$.
[4] $g \in S$ est un majorant pour T si $t \leq g$ pour tout $t \in T$.
[5] $m \in S$ est maximal si pour tout $s \in S$ tel que $m \leq s$ on a $m = s$.
[6] C'est-à-dire : tout élément $v \in V$ s'écrit comme $v = \sum_{i=1}^{n} a_i g_i$ avec $n \in \mathbb{N}$, $a_1, \ldots, a_n \in K$ et $g_1, \ldots, g_n \in G$.

des éléments dans B. Alors l'ensemble $G := B \cup \{v\}$ est aussi K-linéairement indépendant, car toute combinaison K-linéaire $0 = av + \sum_{i=1}^{n} a_i b_i$ avec $n \in \mathbb{N}$, $a, a_1, \ldots, a_n \in K$ et $b_1, \ldots, b_n \in B$ avec $a \neq 0$ donnerait la contradiction $v = \sum_{i=1}^{n} \frac{-a_i}{a} b_i$ (noter que $a = 0$ correspond à une combinaison K-linéaire dans B qui est K-linéairement indépendant). Mais, $B \subsetneq G \in S$ contredit la maximalité de B. $\square$

2 Rappel : Déterminants

Objectifs :

- Maîtriser la définition et les propriétés fondamentales des déterminants ;
- savoir calculer des déterminants ;
- connaître des exemples et savoir démontrer des propriétés simples.

Définition et premières propriétés

Les déterminants ont été introduits au semestre précédent. Ici on les rappelle d'un autre point de vue : on part des règles de calculs. En fait, notre première proposition peut être utilisée comme une définition ; c'est l'axiomatique de Weierstraß (voir le livre de Fischer).

Dans cette section nous permettons que K soit un anneau commutatif (mais vous pouvez encore prendre $K = \mathbb{R}$ ou $K = \mathbb{C}$ sans perdre d'information).

Si $M = \begin{pmatrix} m_{1,1} & m_{1,2} & \cdots & m_{1,n} \\ m_{2,1} & m_{2,2} & \cdots & m_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ m_{n,1} & m_{n,2} & \cdots & m_{n,n} \end{pmatrix}$ est une matrice, nous notons $m_i = \begin{pmatrix} m_{i,1} & m_{i,2} & \cdots & m_{i,n} \end{pmatrix}$ sa i-ième ligne, donc $M = \begin{pmatrix} m_1 \\ m_2 \\ \vdots \\ m_n \end{pmatrix}$.

Proposition 2.1. *Soit $n \in \mathbb{N}_{>0}$. Le* déterminant *est une application*

$$\det : \mathrm{Mat}_{n \times n}(K) \to K, \quad M \mapsto \det(M)$$

telle que

D1 $\det$ *est* K-linéaire *en toute ligne, c'est-à-dire pour tout* $1 \leq i \leq n$*, si* $m_i = r + \lambda s$ *avec* $\lambda \in K$*,* $r = \begin{pmatrix} r_1 & r_2 & \cdots & r_n \end{pmatrix}$ *et* $s = \begin{pmatrix} s_1 & s_2 & \cdots & s_n \end{pmatrix}$*, alors*

$$\det \begin{pmatrix} m_1 \\ \vdots \\ m_{i-1} \\ m_i \\ m_{i+1} \\ \vdots \\ m_n \end{pmatrix} = \det \begin{pmatrix} m_1 \\ \vdots \\ m_{i-1} \\ r+\lambda s \\ m_{i+1} \\ \vdots \\ m_n \end{pmatrix} = \det \begin{pmatrix} m_1 \\ \vdots \\ m_{i-1} \\ r \\ m_{i+1} \\ \vdots \\ m_n \end{pmatrix} + \lambda \cdot \det \begin{pmatrix} m_1 \\ \vdots \\ m_{i-1} \\ s \\ m_{i+1} \\ \vdots \\ m_n \end{pmatrix}.$$

D2 $\det$ *est* alternant*, c'est-à-dire, si deux des lignes de M sont égales, alors* $\det(M) = 0$.

D3 $\det$ *est* normalisé*, c'est-à-dire,* $\det(\mathrm{id}_n) = 1$ *où* id_n *est l'identité.*

Démonstration. Cela a été démontré dans le cours d'algèbre linéaire au semestre précédent. □

Souvent nous utilisons la notation

$$\begin{vmatrix} m_{1,1} & m_{1,2} & \cdots & m_{1,n} \\ m_{2,1} & m_{2,2} & \cdots & m_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ m_{n,1} & m_{n,2} & \cdots & m_{n,n} \end{vmatrix} := \det \begin{pmatrix} m_{1,1} & m_{1,2} & \cdots & m_{1,n} \\ m_{2,1} & m_{2,2} & \cdots & m_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ m_{n,1} & m_{n,2} & \cdots & m_{n,n} \end{pmatrix}.$$

Proposition 2.2. *Les propriétés suivantes sont satisfaites.*

D4 *Pour tout $\lambda \in K$, on a $\det(\lambda \cdot M) = \lambda^n \det(M)$.*

D5 *Si une ligne est égale à 0, alors $\det(M) = 0$.*

D6 *Si $\tilde{M}$ est obtenu de M en échangeant deux lignes, alors $\det(M) = -\det(\tilde{M})$.*

D7 *Soient $\lambda \in A$ et $i \neq j$. Si $\tilde{M}$ est obtenu de M en additionnant λ fois la j-ième ligne à la i-ième ligne, alors $\det(M) = \det(\tilde{M})$.*

Démonstration. **D4** Cela résulte de la linéarité (D1).
D5 Cela résulte de la linéarité (D1).

D6 Disons que la i-ième et la j-ième ligne sont échangées. Donc $M = \begin{pmatrix} m_1 \\ \vdots \\ m_i \\ \vdots \\ m_j \\ \vdots \\ m_n \end{pmatrix}$ et $\tilde{M} = \begin{pmatrix} m_1 \\ \vdots \\ m_j \\ \vdots \\ m_i \\ \vdots \\ m_n \end{pmatrix}$.

$$\begin{aligned} \det(M) + \det(\tilde{M}) &= \det \begin{pmatrix} m_1 \\ \vdots \\ m_i \\ \vdots \\ m_j \\ \vdots \\ m_n \end{pmatrix} + \det \begin{pmatrix} m_1 \\ \vdots \\ m_j \\ \vdots \\ m_i \\ \vdots \\ m_n \end{pmatrix} \\ &\overset{\text{D2}}{=} \det \begin{pmatrix} m_1 \\ \vdots \\ m_j \\ \vdots \\ m_j \\ \vdots \\ m_n \end{pmatrix} + \det \begin{pmatrix} m_1 \\ \vdots \\ m_i \\ \vdots \\ m_j \\ \vdots \\ m_n \end{pmatrix} + \det \begin{pmatrix} m_1 \\ \vdots \\ m_j \\ \vdots \\ m_i \\ \vdots \\ m_n \end{pmatrix} + \det \begin{pmatrix} m_1 \\ \vdots \\ m_i \\ \vdots \\ m_i \\ \vdots \\ m_n \end{pmatrix} \\ &\overset{\text{D1}}{=} \det \begin{pmatrix} m_1 \\ \vdots \\ m_i + m_j \\ \vdots \\ m_j \\ \vdots \\ m_n \end{pmatrix} + \det \begin{pmatrix} m_1 \\ \vdots \\ m_i + m_j \\ \vdots \\ m_i \\ \vdots \\ m_n \end{pmatrix} = \det \begin{pmatrix} m_1 \\ \vdots \\ m_i + m_j \\ \vdots \\ m_i + m_j \\ \vdots \\ m_n \end{pmatrix} \overset{\text{D2}}{=} 0. \end{aligned}$$

D7 Nous avons

$$\det(\tilde{M}) = \det\begin{pmatrix} m_1 \\ \vdots \\ m_i+\lambda m_j \\ \vdots \\ m_j \\ \vdots \\ m_n \end{pmatrix} \stackrel{\text{D1}}{=} \det\begin{pmatrix} m_1 \\ \vdots \\ m_i \\ \vdots \\ m_j \\ \vdots \\ m_n \end{pmatrix} + \lambda\cdot\det\begin{pmatrix} m_1 \\ \vdots \\ m_j \\ \vdots \\ m_j \\ \vdots \\ m_n \end{pmatrix} \stackrel{\text{D2}}{=} \det(M) + \lambda\cdot 0 = \det(M).$$

□

Proposition 2.3. *Les propriétés suivantes sont aussi satisfaites.*

D8 *Si M est de forme triangulaire (supérieure)*

$$\begin{pmatrix} \lambda_1 & m_{1,2} & m_{1,3} & \cdots & m_{1,n} \\ 0 & \lambda_2 & m_{2,3} & \cdots & m_{2,n} \\ 0 & 0 & \lambda_3 & \cdots & m_{3,n} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{pmatrix},$$

alors $\det(M) = \prod_{i=1}^n \lambda_i$.

D9 *Si M est une matrice en blocs $\left(\begin{smallmatrix} A & B \\ 0 & C \end{smallmatrix}\right)$ avec des matrices carrées A et C, alors* $\det(M) = \det(A)\cdot\det(C)$.

Démonstration. Laissée au lecteur. □

Formule de Leibniz

Lemme 2.4. *Pour $1 \le i \le n$, soit $e_i := \left(\begin{smallmatrix} 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \end{smallmatrix}\right)$ où le 1 est à la i-ème position. Soit $\sigma : \{1,\dots,n\} \to \{1,\dots,n\}$ une application. Soit $M = \begin{pmatrix} e_{\sigma(1)} \\ e_{\sigma(2)} \\ \vdots \\ e_{\sigma(n)} \end{pmatrix}$. Alors*

$$\det(M) = \begin{cases} 0 & \text{si } \sigma \text{ n'est pas bijectif,} \\ \operatorname{sgn}(\sigma) & \text{si } \sigma \text{ est bijectif } (\sigma \in S_n). \end{cases}$$

Démonstration. Si σ n'est pas bijectif, alors la matrice possède deux fois la même ligne, donc le déterminant est 0. Si σ est bijectif, alors σ est un produit de transpositions $\sigma = \tau_r \circ \cdots \circ \tau_1$ (voir Algèbre 1). Donc $\operatorname{sgn}(\sigma) = (-1)^r$. Commençons par $\sigma = \mathrm{id}$. Dans ce cas le déterminant est 1 et donc égal à $\operatorname{sgn}(\sigma)$. On continue par récurrence et on suppose donc (hypothèse de récurrence) que le résultat est vrai pour $r-1$ transpositions (avec $r \ge 1$). Soit M' la matrice qui correspond à $\sigma' = \tau_{r-1}\circ\cdots\circ\tau_1$; son déterminant est $(-1)^{r-1} = \operatorname{sgn}(\sigma')$ par l'hypothèse de récurrence. La matrice M est obtenue par M' en échangeant deux lignes, donc $\det(M) = -\det(M') = -(-1)^{r-1} = (-1)^r$. □

Proposition 2.5 (Formule de Leibniz). *Soit $M = \begin{pmatrix} m_{1,1} & m_{1,2} & \cdots & m_{1,n} \\ m_{2,1} & m_{2,2} & \cdots & m_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ m_{n,1} & m_{n,2} & \cdots & m_{n,n} \end{pmatrix} \in \mathrm{Mat}_{n\times n}(K)$. Alors,*

$$\det(M) = \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma)\cdot m_{1,\sigma(1)}\cdot m_{2,\sigma(2)}\cdot\ldots\cdot m_{n,\sigma(n)}.$$

Démonstration. La linéarité des lignes (D1) nous donne

$$\begin{aligned}
\det(M) &= \sum_{i_1=1}^{n} m_{1,i_1} \begin{vmatrix} e_{i_1} \\ m_2 \\ m_3 \\ \vdots \\ m_n \end{vmatrix} = \sum_{i_1=1}^{n}\sum_{i_2=1}^{n} m_{1,i_1} m_{2,i_2} \begin{vmatrix} e_{i_1} \\ e_{i_2} \\ m_3 \\ \vdots \\ m_n \end{vmatrix} \\
&= \cdots = \sum_{i_1=1}^{n}\sum_{i_2=1}^{n}\cdots\sum_{i_n=1}^{n} m_{1,i_1} m_{2,i_2}\cdots m_{n,i_n} \begin{vmatrix} e_{i_1} \\ e_{i_2} \\ e_{i_3} \\ \vdots \\ e_{i_n} \end{vmatrix} \\
&= \sum_{\sigma\in S_n} m_{1,\sigma(1)} m_{2,\sigma(2)}\cdots m_{n,\sigma(n)} \cdot \operatorname{sgn}(\sigma),
\end{aligned}$$

où la dernière égalité résulte du lemme 2.4. Noter que le déterminant de la matrice $\begin{pmatrix} e_{i_1} \\ e_{i_2} \\ e_{i_3} \\ \vdots \\ e_{i_n} \end{pmatrix}$ est non-nul seulement si les i_j sont tous différents ; cela nous permet de l'identifier avec la permutation $\sigma(j) = i_j$. Que le déterminant est unique est clair parce qu'il est une fonction des coefficients de la matrice. □

Corollaire 2.6. *Soit* $M \in \mathrm{Mat}_{n\times n}(K)$. *On note* M^{tr} *la matrice transposée. Alors,* $\det(M) = \det(M^{\mathrm{tr}})$.

Démonstration. Nous utilisons la formule de Leibniz 2.5. Noter d'abord que $\operatorname{sgn}(\sigma) = \operatorname{sgn}(\sigma^{-1})$ pour tout σ dans S_n car sgn est un homomorphisme de groupes, $1^{-1} = 1$ et $(-1)^{-1} = -1$. Noter maintenant

$$\begin{aligned}
m_{1,\sigma(1)} m_{2,\sigma(2)}\cdots m_{n,\sigma(n)} &= m_{\sigma^{-1}(\sigma(1)),\sigma(1)} m_{\sigma^{-1}(\sigma(2)),\sigma(2)}\cdots m_{\sigma^{-1}(\sigma(n)),\sigma(n)} \\
&= m_{\sigma^{-1}(1),1} m_{\sigma^{-1}(2),2}\cdots m_{\sigma^{-1}(n),n},
\end{aligned}$$

où pour la dernière égalité nous avons seulement écrit le produit dans un autre ordre car les valeurs $\sigma(1), \sigma(2), \ldots, \sigma(n)$ parcourent $1, 2, \ldots, n$ (juste dans un autre ordre).
Nous avons donc

$$\begin{aligned}
\det(M) &= \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) m_{1,\sigma(1)} m_{2,\sigma(2)}\cdots m_{n,\sigma(n)} \\
&= \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma^{-1}) m_{\sigma^{-1}(1),1} m_{\sigma^{-1}(2),2}\cdots m_{\sigma^{-1}(n),n} \\
&= \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma^{-1}) m^{\mathrm{tr}}_{1,\sigma^{-1}(1)} m^{\mathrm{tr}}_{2,\sigma^{-1}(2)}\cdots m^{\mathrm{tr}}_{n,\sigma^{-1}(n)} \\
&= \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) m^{\mathrm{tr}}_{1,\sigma(1)} m^{\mathrm{tr}}_{2,\sigma(2)}\cdots m^{\mathrm{tr}}_{n,\sigma(n)} \\
&= \det(M^{\mathrm{tr}}),
\end{aligned}$$

où nous avons utilisé la bijection $S_n \to S_n$ donnée par $\sigma \mapsto \sigma^{-1}$; il est donc indifférent si la somme parcourt $\sigma \in S_n$ ou les inverses. □

Corollaire 2.7. *Les règles* **D1** *à* **D9** *sont aussi vraies pour les colonnes au lieu des lignes.*

Démonstration. En prenant la transposée d'une matrice, les lignes deviennent les colonnes, mais par le corollaire 2.6, le déterminant ne change pas. $\square$

Développement de Laplace

Définition 2.8. *Soient* $n \in \mathbb{N}_{>0}$ *et* $M = \begin{pmatrix} m_{1,1} & m_{1,2} & \cdots & m_{1,n} \\ m_{2,1} & m_{2,2} & \cdots & m_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ m_{n,1} & m_{n,2} & \cdots & m_{n,n} \end{pmatrix} \in \mathrm{Mat}_{n \times n}(K)$*. Pour* $1 \leq i, j \leq n$ *nous définissons les matrices*

$$M_{i,j} = \begin{pmatrix} m_{1,1} & \cdots & m_{1,j-1} & 0 & m_{1,j+1} & \cdots & m_{1,n} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ m_{i-1,1} & \cdots & m_{i-1,j-1} & 0 & m_{i-1,j+1} & \cdots & m_{i-1,n} \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ m_{i+1,1} & \cdots & m_{i+1,j-1} & 0 & m_{i+1,j+1} & \cdots & m_{i+1,n} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ m_{n,1} & \cdots & m_{n,j-1} & 0 & m_{n,j+1} & \cdots & m_{n,n} \end{pmatrix} \in \mathrm{Mat}_{n \times n}(A)$$

et

$$M'_{i,j} = \begin{pmatrix} m_{1,1} & \cdots & m_{1,j-1} & m_{1,j+1} & \cdots & m_{1,n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ m_{i-1,1} & \cdots & m_{i-1,j-1} & m_{i-1,j+1} & \cdots & m_{i-1,n} \\ m_{i+1,1} & \cdots & m_{i+1,j-1} & m_{i+1,j+1} & \cdots & m_{i+1,n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ m_{n,1} & \cdots & m_{n,j-1} & m_{n,j+1} & \cdots & m_{n,n} \end{pmatrix} \in \mathrm{Mat}_{n-1 \times n-1}(A).$$

En plus, soit $\tilde{M}_{i,j}$ *la matrice obtenue à partir de* M *en remplaçant la* j*-ième colonne par* $\begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$*, où le* 1 *est à la* i*-ième position.*
Les déterminants $\det(M'_{i,j})$ *sont appelés les* mineurs *de* M.

Lemme 2.9. *Soient* $n \in \mathbb{N}_{>0}$ *et* $M \in \mathrm{Mat}_{n \times n}(K)$*. Pour tout* $1 \leq i, j \leq n$*, nous avons*

(a) $\det(M_{i,j}) = (-1)^{i+j} \cdot \det(M'_{i,j})$,

(b) $\det(\tilde{M}_{i,j}) = \det(M_{i,j})$.

Démonstration. (a) En échangeant i lignes, la ligne avec les zéros est la première. En échangeant j colonnes, nous obtenons la matrice

$$\begin{pmatrix} 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & m_{1,1} & \cdots & m_{1,j-1} & m_{1,j+1} & \cdots & m_{1,n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & m_{i-1,1} & \cdots & m_{i-1,j-1} & m_{i-1,j+1} & \cdots & m_{i-1,n} \\ 0 & m_{i+1,1} & \cdots & m_{i+1,j-1} & m_{i+1,j+1} & \cdots & m_{i+1,n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & m_{n,1} & \cdots & m_{n,j-1} & m_{n,j+1} & \cdots & m_{n,n} \end{pmatrix} \in \mathrm{Mat}_{n \times n}(A)$$

dont le déterminant est $\det(M'_{i,j})$ (à cause de **D9**), ce qui montre le résultat.
(b) Additionner $-m_{i,k}$ fois la j-ième colonne sur la k-ième colonne de $\tilde{M}_{i,j}$ rend le coefficient (i,k) égal à 0 pour $k \neq i$ sans changer le déterminant (corollaire 2.7). $\square$

Proposition 2.10 (Développement de Laplace pour les lignes)**.** *Soit* $n \in \mathbb{N}_{>0}$*. Pour tout* $1 \leq i \leq n$*, nous avons l'égalité*

$$\det(M) = \sum_{j=1}^{n} (-1)^{i+j} m_{i,j} \det(M'_{i,j})$$

Démonstration. Par l'axiome **D2** (linéarité dans les lignes), nous avons

$$\det(M) = \sum_{j=1}^{n} m_{i,j} \begin{vmatrix} m_1 \\ \vdots \\ m_{i-1} \\ e_i \\ m_{i+1} \\ \vdots \\ m_n \end{vmatrix} = \sum_{j=1}^{n} m_{i,j} \det(M_{i,j}) = \sum_{j=1}^{n} (-1)^{i+j} m_{i,j} \det(M'_{i,j}).$$

□

Corollaire 2.11 (Développement de Laplace pour les colonnes)**.** *Pour tout* $n \in \mathbb{N}_{>0}$ *et tout* $1 \leq j \leq n$*, nous avons la formule*

$$\det(M) := \sum_{i=1}^{n} (-1)^{i+j} m_{i,j} \det(M'_{i,j}).$$

Démonstration. Il suffit d'appliquer la proposition 2.10 à la matrice transposée et de se rappeler (corollaire 2.6) que le déterminant de la matrice transposée est le même. □

Noter que les formules de Laplace peuvent être écrites comme

$$\det(M) = \sum_{j=1}^{n} m_{i,j} \det(M_{i,j}) = \sum_{i=1}^{n} m_{i,j} \det(M_{i,j}).$$

Matrices adjointes

Définition 2.12. *La* matrice adjointe $\operatorname{adj}(M) = M^{\#} = (m^{\#}_{i,j})$ *de la matrice* $M \in \operatorname{Mat}_{n\times n}(A)$ *est définie par* $m^{\#}_{i,j} := \det(M_{j,i}) = (-1)^{i+j} \det(M'_{j,i})$.

Proposition 2.13. *Pour toute matrice* $M \in \operatorname{Mat}_{n\times n}(K)$*, nous avons l'égalité*

$$M^{\#} \cdot M = M \cdot M^{\#} = \det(M) \cdot \mathrm{id}_n.$$

Démonstration. Soit $N = (n_{i,j}) := M \cdot M^{\#}$. Nous calculons $n_{i,j}$:

$$n_{i,j} = \sum_{k=1}^{n} m^{\#}_{i,k} m_{k,j} = \sum_{k=1}^{n} \det(M_{k,i}) m_{k,j}.$$

Si $i = j$, nous trouvons $n_{i,i} = \det(M)$ par la formule de Laplace. Mais nous n'avons pas besoin d'utiliser cette formule et nous continuons en généralité en utilisant $\det(M_{k,i}) = \det(\tilde{M}_{k,i})$ par le lemme 2.9 (b). La linéarité en la i-ième colonne montre que $\sum_{k=1}^{n} \det(\tilde{M}_{k,i}) m_{k,j}$ est le déterminant de la matrice dont la i-ième colonne est remplacée par la j-ième colonne. Si $i = j$, cette matrice est M, alors $m_{i,i} = \det(M)$. Si $i \neq j$, ce déterminant (et donc $n_{i,j}$) est 0 car deux des colonnes sont égales.

La démonstration pour $M^{\#} \cdot M$ est similaire. □

Corollaire 2.14. *Soit* $M \in \mathrm{Mat}_{n\times n}(K)$.

(a) Si $\det(M)$ *est inversible dans* K *(pour* K *un corps cela veut die* $\det(M) \neq 0$*), alors* M *est inversible et la matrice inverse* M^{-1} *est égale à* $\frac{1}{\det(M)} M^{\#}$.

(b) Si M *est inversible, alors* $M^{-1}\det(M) = M^{\#}$.

Démonstration. Proposition 2.13. □

Nous finissons ce rappel par le résultat fondamental suivant.

Proposition 2.15. *Soient* $M, N \in \mathrm{Mat}_{n\times n}(K)$.

(a) $\det(M \cdot N) = \det(M) \cdot \det(N)$.

(b) Les assertions suivantes sont équivalentes :

- *(i)* M *est inversible ;*
- *(ii)* $\det(M)$ *est inversible.*

Dans ce cas $\det(M^{-1}) = \frac{1}{\det(M)}$.

Démonstration. (a) a été démontré au cours d'algèbre linéaire.
(b) est évident à cause de la Proposition 2.13. □

3 Valeurs propres

Objectifs :

- Maîtriser la définition et les propriétés fondamentales des valeurs et vecteurs propres ;
- savoir calculer des espaces propres ;
- connaître des exemples et savoir démontrer des propriétés simples.

Exemple 3.1. *(a) Considérons* $M = \left(\begin{smallmatrix} 3 & 0 \\ 0 & 2 \end{smallmatrix}\right) \in \mathrm{Mat}_{2\times 2}(\mathbb{R})$. *Nous avons :*

- $\left(\begin{smallmatrix} 3 & 0 \\ 0 & 2 \end{smallmatrix}\right)\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right) = 3 \cdot \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)$ *et*
- $\left(\begin{smallmatrix} 3 & 0 \\ 0 & 2 \end{smallmatrix}\right)\left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right) = 2 \cdot \left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right)$.

(b) Considérons $M = \left(\begin{smallmatrix} 3 & 1 \\ 0 & 2 \end{smallmatrix}\right) \in \mathrm{Mat}_{2\times 2}(\mathbb{R})$. *Nous avons :*

- $\left(\begin{smallmatrix} 3 & 1 \\ 0 & 2 \end{smallmatrix}\right)\left(\begin{smallmatrix} a \\ 0 \end{smallmatrix}\right) = 3 \cdot \left(\begin{smallmatrix} a \\ 0 \end{smallmatrix}\right)$ *pour tout* $a \in \mathbb{R}$.
- $\left(\begin{smallmatrix} 3 & 1 \\ 0 & 2 \end{smallmatrix}\right)\left(\begin{smallmatrix} a \\ b \end{smallmatrix}\right) = \left(\begin{smallmatrix} 3a+b \\ 2b \end{smallmatrix}\right) = 2 \cdot \left(\begin{smallmatrix} a \\ b \end{smallmatrix}\right) \Leftrightarrow a = -b$. *Donc pour tout* $a \in \mathbb{R}$, *nous avons* $\left(\begin{smallmatrix} 3 & 1 \\ 0 & 2 \end{smallmatrix}\right)\left(\begin{smallmatrix} a \\ -a \end{smallmatrix}\right) = 2 \cdot \left(\begin{smallmatrix} a \\ -a \end{smallmatrix}\right)$.

(c) Considérons $M = \left(\begin{smallmatrix} 5 & 1 \\ -4 & 10 \end{smallmatrix}\right) \in \mathrm{Mat}_{2\times 2}(\mathbb{R})$. *Nous avons :*

- $\left(\begin{smallmatrix} 5 & 1 \\ -4 & 10 \end{smallmatrix}\right)\left(\begin{smallmatrix} 1 \\ 4 \end{smallmatrix}\right) = 9 \left(\begin{smallmatrix} 1 \\ 4 \end{smallmatrix}\right)$ *et*

- $\begin{pmatrix} 5 & 1 \\ -4 & 10 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 6 \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

(d) Considérons $M = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix} \in \mathrm{Mat}_{2\times 2}(\mathbb{R})$*. Nous avons :*

- $\begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} a \\ 0 \end{pmatrix} = 2 \cdot \begin{pmatrix} a \\ 0 \end{pmatrix}$ *pour tout* $a \in \mathbb{R}$.
- *Soit* $\lambda \in \mathbb{R}$*. Nous regardons* $\begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 2a+b \\ 2b \end{pmatrix} = \lambda \cdot \begin{pmatrix} a \\ b \end{pmatrix} \Leftrightarrow (2a + b = \lambda a \wedge 2b = \lambda b) \Leftrightarrow (b = 0 \wedge (\lambda = 2 \vee a = 0)) \vee (\lambda = 2 \wedge b = 0) \Leftrightarrow b = 0 \wedge (\lambda = 2 \vee a = 0)$.

  Donc, les seules solutions de $M \begin{pmatrix} a \\ b \end{pmatrix} = \lambda \cdot \begin{pmatrix} a \\ b \end{pmatrix}$ *avec un vecteur* $\begin{pmatrix} a \\ b \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ *sont de la forme* $M \begin{pmatrix} a \\ 0 \end{pmatrix} = 2 \cdot \begin{pmatrix} a \\ 0 \end{pmatrix}$ *avec* $a \in \mathbb{R}$.
- *Considérons* $M = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in \mathrm{Mat}_{2\times 2}(\mathbb{R})$*. Nous regardons* $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} b \\ -a \end{pmatrix}$*. Ce vecteur est égal à* $\lambda \cdot \begin{pmatrix} a \\ b \end{pmatrix}$ *si et seulement si* $b = \lambda \cdot a$ *et* $a = -\lambda \cdot b$*. Cela donne* $a = -\lambda^2 \cdot a$*. Donc il n'existe aucun* $\lambda \in \mathbb{R}$ *avec cette propriété si* $\begin{pmatrix} a \\ b \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}$.

Nous allons étudier ces phénomènes en général. Soit K un corps (commutatif, comme toujours) et V un K-espace vectoriel. Nous rappelons qu'une application K-linéaire $\varphi : V \to V$ est aussi appelée *endomorphisme* et que nous écrivons $\mathrm{End}_K(V) := \mathrm{Hom}_K(V, V)$.

Définition 3.2. *Soient* V *un* K*-espace vectoriel et* $\varphi \in \mathrm{End}_K(V)$.

- $\lambda \in K$ *est appelé* valeur propre *de* φ *s'il existe* $0 \neq v \in V$ *t.q.* $\varphi(v) = \lambda v$ *(ou de manière équivalente :* $\ker(\varphi - \lambda \cdot \mathrm{id}_V) \neq 0$*).*
- *On pose* $E_\varphi(\lambda) := \ker(\varphi - \lambda \cdot \mathrm{id}_V)$*. En tant que noyau d'une application* K*-linéaire,* $E_\varphi(\lambda)$ *est un* K*-sous-espace de* V*. Si* λ *est une valeur propre de* φ*, on appelle* $E_\varphi(\lambda)$ l'espace propre pour λ.
- *Tout* $0 \neq v \in E_\varphi(\lambda)$ *est appelé* vecteur propre pour la valeur propre λ.
- *On pose* $\mathrm{Spec}(\varphi) = \{\lambda \in K \mid \lambda \text{ est valeur propre de } \varphi\}$.
- *Soit* $M \in \mathrm{Mat}_{n\times n}(K)$*. Nous savons que l'application*

$$\varphi_M : K^n \to K^n, \quad \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \mapsto M \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$$

  est K*-linéaire, donc* $\varphi_M \in \mathrm{End}_K(K^n)$*. Dans ce cas, nous parlons souvent de valeurs/vecteurs propres de* M *(au lieu de* φ_M*).*

Proposition 3.3. *Les espaces propres* $E_\varphi(\lambda)$ *et* $E_M(\lambda)$ *sont des sous-espaces vectoriels.*

Démonstration. C'est clair car les espaces propres sont définis comme des noyaux d'une matrice/d'un endomorphisme linéaire, et nous savons que les noyaux sont des sous-espaces vectoriels. □

Nous reprenons l'exemple précédent.

Exemple 3.4.

(a) Soit $M = \begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix} \in \mathrm{Mat}_{2\times 2}(\mathbb{R})$.

- $\mathrm{Spec}(M) = \{2, 3\}$;
- $E_M(2) = \langle \left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right) \rangle$;
- $E_M(3) = \langle \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right) \rangle$.
- *La matrice M est diagonale et la base standard $\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right)$ est formée de vecteurs propres de M.*

(b) Soit $M = \left(\begin{smallmatrix} 3 & 1 \\ 0 & 2 \end{smallmatrix}\right) \in \mathrm{Mat}_{2\times 2}(\mathbb{R})$.

- $\mathrm{Spec}(M) = \{2, 3\}$;
- $E_M(2) = \langle \left(\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right) \rangle$;
- $E_M(3) = \langle \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right) \rangle$.
- *La matrice M n'est pas diagonale, mais K^2 possède la base $\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right)$ dont les éléments sont des vecteurs propres de M.*
- *Définissons la matrice dont les colonnes sont la base précédente $C := \left(\begin{smallmatrix} 1 & 1 \\ 0 & -1 \end{smallmatrix}\right)$. Cette matrice est inversible (car les colonnes forment une base) et nous avons*

  $$C^{-1}MC = \left(\begin{smallmatrix} 3 & 0 \\ 0 & 2 \end{smallmatrix}\right),$$

  une matrice diagonale avec les valeurs propres sur la diagonale ! Noter que nous n'avons pas besoin de faire le calcul de matrices, le produit matriciel n'est qu'une reformulation des assertions que nous avons vues.

(c) Soit $M = \left(\begin{smallmatrix} 5 & 1 \\ -4 & 10 \end{smallmatrix}\right) \in \mathrm{Mat}_{2\times 2}(\mathbb{R})$.

- $\mathrm{Spec}(M) = \{6, 9\}$;
- $E_M(6) = \langle \left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right) \rangle$;
- $E_M(9) = \langle \left(\begin{smallmatrix} 1 \\ 4 \end{smallmatrix}\right) \rangle$;
- *Les vecteurs propres $\left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right), \left(\begin{smallmatrix} 1 \\ 4 \end{smallmatrix}\right)$ forment une base de K^2 et donc la matrice $C := \left(\begin{smallmatrix} 1 & 1 \\ 1 & 4 \end{smallmatrix}\right)$ dont les colonnes sont les vecteurs de la base est inversible et*

  $$C^{-1}MC = \left(\begin{smallmatrix} 6 & 0 \\ 0 & 9 \end{smallmatrix}\right),$$

  encore une matrice diagonale avec les valeurs propres sur la diagonale !

(d) Soit $M = \left(\begin{smallmatrix} 2 & 1 \\ 0 & 2 \end{smallmatrix}\right) \in \mathrm{Mat}_{2\times 2}(\mathbb{R})$.

- $\mathrm{Spec}(M) = \{2\}$;
- $E_M(2) = \langle \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right) \rangle$;
- *K^2 ne possède pas de base formée de vecteurs propres de M, donc nous ne pouvons pas adapter la procédure des exemples précédents à ce cas.*

(e) Soit $M = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right) \in \mathrm{Mat}_{2\times 2}(\mathbb{R})$.

- $\mathrm{Spec}(M) = \emptyset$;

- *La matrice M ne possède pas de vecteur propre dans $\mathbb{R}$.*

Exemple 3.5. *Soient $K = \mathbb{R}$ et $V = \mathcal{C}^\infty(\mathbb{R})$ le $\mathbb{R}$-espace vectoriel des fonctions $f : \mathbb{R} \to \mathbb{R}$ infiniment continuement dérivables. Soit $D : V \to V$ la dérivation $f \mapsto Df = \frac{df}{dx} = f'$. C'est une application $\mathbb{R}$-linéaire, donc $D \in \mathrm{End}_\mathbb{R}(V)$.*
Considérons $f_r(x) = \exp(rx)$ pour $r \in \mathbb{R}$. Il est connu du cours d'analyse que $D(f_r) = r \cdot \exp(rx) = r \cdot f_r$. Donc $f_r \in V$ est un vecteur propre pour la valeur propre r.
Nous trouvons donc $\mathrm{Spec}(D) = \mathbb{R}$.

Dans plusieurs exemples nous avons rencontré des matrices M telles qu'il existe une matrice inversible C avec la propriété que $C^{-1}MC$ est une matrice diagonale. Mais nous avons aussi vu des exemples où nous n'avons pas trouvé une telle matrice C.

Définition 3.6. *(a) Une matrice M est dite* diagonalisable *s'il existe une matrice inversible C telle que $C^{-1}MC$ est diagonal.*

(b) Soient V un K-espace vectoriel et $\varphi \in \mathrm{End}_K(V)$. On dit que φ est diagonalisable *si V possède une K-base dont les éléments sont des vecteurs propres de φ.*

Cette définition exprime précisement l'idée de la diagonalisation suggérée avant, comme le prochain lemme l'affirme, dont la preuve nous renseigne sur comment trouver la matrice C (qui n'est pas unique, en général).

Lemme 3.7. *Soit V un K-espace vectoriel de dimension finie. Soit $\varphi \in \mathrm{End}_K(V)$ et $\mathrm{Spec}(\varphi) = \{\lambda_1, \ldots, \lambda_r\}$. Les assertions suivantes sont équivalentes :*

(i) φ est diagonalisable.

(ii) Il existe une base S de V t.q.

$$M_{S,S}(\varphi) = \begin{pmatrix} \lambda_1 & & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \ddots & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \lambda_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \lambda_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \ddots & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \lambda_2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \ddots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ddots & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda_r & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda_r \end{pmatrix}.$$

Démonstration. « (i) $\Rightarrow$ (ii) » : Par définition il existe une K-base de V consistant en vecteurs propres. Nous les trions selon les valeurs propres :

$$S = \{v_{1,1}, \ldots, v_{1,e_1}, v_{2,1}, \ldots, v_{2,e_2}, \ldots, \ldots, \ldots, v_{r,1}, \ldots, v_{r,e_r}\}$$

où pour tout $1 \leq i \leq r$ les vecteurs $v_{i,1}, \ldots, v_{i,e_i}$ sont des vecteurs propres pour la valeur propre λ_i. La forme de la matrice $M_{S,S}(\varphi)$ est claire.
« (ii) $\Rightarrow$ (i) » : La base S consiste en vecteurs propres, donc par définition φ est diagonalisable. $\square$

Proposition 3.8. *Soient $M \in \mathrm{Mat}_{n\times n}(K)$ et φ_M l'application K-linéaire $K^n \to K^n$ donnée par $\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \mapsto M \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$. Les assertions suivantes sont équivalentes.*

(i) *φ_M est diagonalisable.*

(ii) *Il existe $C \in \mathrm{Mat}_{n\times n}(K)$ inversible telle que $C^{-1}MC$ est une matrice diagonale ; donc M est diagonalisable.*

Démonstration. « (i) $\Rightarrow$ (ii) » : Soit S la K-base de K^n qui existe par la définition de la diagonalisabilité de φ_M. Il suffit de prendre C comme la matrice dont les colonnes sont les éléments de la base S.
« (ii) $\Rightarrow$ (i) » : Soit e_i le i-ème vecteur standard. C'est un vecteur propre pour la matrice $C^{-1}MC$, disons avec valeur propre λ_i. L'égalité $C^{-1}MCe_i = \lambda_i \cdot e_i$ donne $MCe_i = \lambda_i \cdot Ce_i$, i.e. Ce_i est un vecteur propre pour la matrice M de même valeur propre. Mais, Ce_i n'est rien d'autre que la i-ème colonne de C. Donc, les colonnes de C forment une base de K^n consistant en vecteurs propres. $\square$

La question qui nous intéresse bien évidemment maintenant, est : comment décider si φ (ou M) est diagonalisable et, si cela est le cas, comment calculer la matrice C ? En fait, il est utile de considérer deux « sous-questions » séparément :

- Comment calculer $\mathrm{Spec}(\varphi)$?
- Pour $\lambda \in \mathrm{Spec}(\varphi)$, comment calculer l'espace propre $E_\varphi(\lambda)$?

Nous allons répondre à la première question dans la prochaine section. Pour l'instant nous regardons la deuxième question. Commençons par $E_M(\lambda)$. Il s'agit donc de calculer $E_M(\lambda) = \ker(M - \lambda \cdot \mathrm{id}_n)$. Ce calcul se fait en utilisant la réduction de Gauß.

Exemple 3.9. *(a)* *Pour la matrice $M = \left(\begin{smallmatrix} 5 & 1 \\ -4 & 10 \end{smallmatrix}\right) \in \mathrm{Mat}_{2\times 2}(\mathbb{R})$ et la valeur propre 9 nous devons calculer le noyau de $\left(\begin{smallmatrix} 5 & 1 \\ -4 & 10 \end{smallmatrix}\right) - 9 \cdot \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} -4 & 1 \\ -4 & 1 \end{smallmatrix}\right)$. Se rappeler que pour calculer le noyau d'une matrice, on doit seulement utiliser des opérations sur les lignes (et jamais des opérations sur les colonnes car elles mixent les variables). Nous avons donc*

$$\ker(\left(\begin{smallmatrix} -4 & 1 \\ -4 & 1 \end{smallmatrix}\right)) = \ker(\left(\begin{smallmatrix} -4 & 1 \\ 0 & 0 \end{smallmatrix}\right)) = \langle \left(\begin{smallmatrix} 1 \\ 4 \end{smallmatrix}\right) \rangle.$$

Pour la valeur propre 6 nous faisons un calcul similaire :

$$\ker(\left(\begin{smallmatrix} 5 & 1 \\ -4 & 10 \end{smallmatrix}\right) - 6 \cdot \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)) = \ker(\left(\begin{smallmatrix} -1 & 1 \\ -4 & 4 \end{smallmatrix}\right)) = \ker(\left(\begin{smallmatrix} -1 & 1 \\ 0 & 0 \end{smallmatrix}\right)) = \langle \left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right) \rangle.$$

(b) *La matrice $M = \begin{pmatrix} 2 & 1 & 1 \\ 3 & 2 & 3 \\ -3 & -1 & -2 \end{pmatrix} \in \mathrm{Mat}_{3\times 3}(\mathbb{R})$ possède les valeurs propres $-1, 1, 2$.*
Pour la valeur propre 1, nous calculons le noyau

$$\begin{aligned} \ker\Big(\begin{pmatrix} 2 & 1 & 1 \\ 3 & 2 & 3 \\ -3 & -1 & -2 \end{pmatrix} - 1 \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \Big) &= \ker\Big(\begin{pmatrix} 1 & 1 & 1 \\ 3 & 1 & 3 \\ -3 & -1 & -3 \end{pmatrix} \Big) \\ &= \ker\Big(\begin{pmatrix} 1 & 1 & 1 \\ 0 & -2 & 0 \\ 0 & 0 & 0 \end{pmatrix} \Big) = \ker\Big(\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \Big) = \langle \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \rangle \end{aligned}$$

Pour la valeur propre -1, nous calculons le noyau

$$\ker\left(\begin{pmatrix} 2 & 1 & 1 \\ 3 & 2 & 3 \\ -3 & -1 & -2 \end{pmatrix} + 1 \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\right) = \ker\left(\begin{pmatrix} 3 & 1 & 1 \\ 3 & 3 & 3 \\ -3 & -1 & -1 \end{pmatrix}\right)$$
$$= \ker\left(\begin{pmatrix} 3 & 1 & 1 \\ 0 & 2 & 2 \\ 0 & 0 & 0 \end{pmatrix}\right) = \ker\left(\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}\right) = \langle\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}\rangle$$

Pour la valeur propre 2, nous calculons le noyau

$$\ker\left(\begin{pmatrix} 2 & 1 & 1 \\ 3 & 2 & 3 \\ -3 & -1 & -2 \end{pmatrix} - 2 \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\right) = \ker\left(\begin{pmatrix} 0 & 1 & 1 \\ 3 & 0 & 3 \\ -3 & -1 & -4 \end{pmatrix}\right) = \ker\left(\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}\right) = \langle\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}\rangle$$

On écrit ces vecteur dans la matrice $C = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ -1 & -1 & -1 \end{pmatrix}$ de façon à ce que l'on a

$$C^{-1} \cdot M \cdot C = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

Cela explique comment trouver les vecteurs propres dans des exemples. Si l'on veut plus abstraitement calculer l'espace propre $E_\varphi(\lambda) = \ker(\varphi - \lambda \cdot \mathrm{id}_V)$, on doit choisir une K-base S de V et représenter φ par la matrice $M = M_{S,S}(\varphi)$. Dans la base S, $E_\varphi(\lambda)$ s'exprime comme le noyau $\ker(M - \lambda \cdot \mathrm{id}_n)$, et nous avons déjà vu comment calculer celui-ci.

Donnons finalement une réformulation de la diagonalisabilité plus abstraite, mais très utile. D'abord une préparation.

Lemme 3.10. *Soit V un K-espace vectoriel. Soient $\varphi \in \mathrm{End}_K(V)$ et $\lambda_1, \dots, \lambda_r \in K$ deux-à-deux distincts. Alors, $\sum_{i=1}^r E_\varphi(\lambda_i) = \bigoplus_{i=1}^r E_\varphi(\lambda_i)$.*

Démonstration. Nous faisons une récurrence pour $r \geq 1$. Le cas $r = 1$ est trivial. Nous supposons le résultat démontré pour $r - 1 \geq 1$ et nous le montrons pour r. Il faut prouver que pour tout $1 \leq i \leq r$ on a

$$0 = E_\varphi(\lambda_i) \cap \sum_{j=1, j\neq i}^{r} E_\varphi(\lambda_j) = E_\varphi(\lambda_i) \cap \bigoplus_{j=1, j\neq i}^{r} E_\varphi(\lambda_j),$$

où la deuxième égalité provient de l'hypothèse de récurrence (la somme a $r - 1$ sommands). Soit $v \in E_\varphi(\lambda_i) \cap \bigoplus_{j=1,j\neq i}^r E_\varphi(\lambda_j)$. Alors, $v = \sum_{j=1,j\neq i}^r v_j$ avec $v_j \in E_\varphi(\lambda_j)$. Nous avons

$$\varphi(v) = \lambda_i \cdot v = \sum_{j=1, j\neq i}^{r} \lambda_i \cdot v_j = \varphi(\sum_{j=1, j\neq i}^{r} v_j) = \sum_{j=1, j\neq i}^{r} \varphi(v_j) = \sum_{j=1, j\neq i}^{r} \lambda_j \cdot v_j,$$

donc

$$0 = \sum_{j=1, j\neq i}^{r} (\lambda_j - \lambda_i) \cdot v_j.$$

Comme la somme est directe et $\lambda_j - \lambda_i \neq 0$ pour tout $i \neq j$, on conclut que $v_j = 0$ pour tout $1 \leq j \leq r$, $j \neq i$, donc $v = 0$. $\square$

Proposition 3.11. *Soit $\varphi \in \mathrm{End}_K(V)$. Les assertions suivantes sont équivalentes :*

(i) φ est diagonalisable.

(ii) $V = \bigoplus_{\lambda \in \mathrm{Spec}(\varphi)} E_\varphi(\lambda)$.

Démonstration. « (i) $\Rightarrow$ (ii) » : Nous avons l'inclusion $\sum_{\lambda \in \mathrm{Spec}(\varphi)} E_\varphi(\lambda) \subseteq V$. Par le lemme 3.10, la somme est directe, donc nous avons l'inclusion $\bigoplus_{\lambda \in \mathrm{Spec}(\varphi)} E_\varphi(\lambda) \subseteq V$. Comme φ est diagonalisable, il existe une K-base de V consistant en vecteurs propres pour φ. Donc, tout élément de cette base appartient déjà à $\bigoplus_{\lambda \in \mathrm{Spec}(\varphi)} E_\varphi(\lambda)$, d'où l'égalité $\bigoplus_{\lambda \in \mathrm{Spec}(\varphi)} E_\varphi(\lambda) = V$.
« (ii) $\Rightarrow$ (i) » : Pour tout $\lambda \in \mathrm{Spec}(\varphi)$ soit S_λ une K-base de l'espace propre $E_\varphi(\lambda)$. Donc $S = \bigcup_{\lambda \in \mathrm{Spec}(\varphi)} S_\lambda$ est une K-base de V consistant en vecteurs propres, montrant la diagonalisabilité de φ. $\square$

4 Excursion : division euclidienne et pgcd de polynômes

Objectifs :

- Maîtriser la division euclidienne et l'algorithme d'Euclide ;
- savoir calculer la division euclidienne, le pgcd et une relation de Bézout à l'aide de l'algorithme d'Euclide.

Nous supposons connu du lycée et/ou des autres cours la notion de polynôme. Comme notation pour l'ensemble de tous les polynômes à coefficients dans K, on utilise $K[X]$ où X désigne la variable. Un polynôme s'écrit alors comme une somme finie $\sum_{i=0}^{d} a_i X^i$ avec $a_0, \dots, a_d \in K$. On peut bien évidemment choisir n'importe quel autre symbole pour la variable, par exemple x, T, $\square$; dans ce cas, nous écrivons$\sum_{i=0}^{d} a_i x^i$, $\sum_{i=0}^{d} a_i T^i$, $\sum_{i=0}^{d} a_i \square^i$, $K[x]$, $K[T]$, $K[\square]$, etc.
Le degré d'un polynôme f sera noté $\deg(f)$ avec la convention $\deg(0) = -\infty$. Rappelons que pour tout pair $f, g \in K[X]$ on a $\deg(fg) = \deg(f) + \deg(g)$ et $\deg(f+g) \le \max\{\deg(f), \deg(g)\}$.

Définition 4.1. *Un polynôme $f = \sum_{i=0}^{d} a_i X^i$ de degré d est appelé* unitaire *si $a_d = 1$.*
Un polynôme $f \in K[X]$ de degré ≥ 1 est appelé irréductible *s'il ne s'écrit pas comme produit $f = gh$ avec $g, h \in K[X]$ de degré ≥ 1.*

C'est un fait que les seuls polynômes irréductible dans $\mathbb{C}[X]$ sont les polynômes de degré 1. (On dit que $\mathbb{C}$ est *algébriquement clos*.) Tout polynôme irréductible dans $\mathbb{R}[X]$ est soit de degré 1 (et trivialement, tous les polynômes de degré 1 sont irréductibles), soit de degré 2 (il existe des polynômes de degré 2 irréductibles comme $X^2 + 1$, mais aussi des polynômes réductibles comme $X^2 - 1 = (X-1)(X+1)$; plus précisement, un polynôme de degré 2 est irréductible si et seulement si son discriminant est négatif).

Définition 4.2. *Un polynôme $f \in K[X]$ est appelé* diviseur *d'un polynôme $g \in K[X]$ s'il existe $q \in K[X]$ tel que $g = qf$. Nous utilisons la notation $f \mid g$.*

Si f divise g, nous avons bien évidemment $\deg(f) \le \deg(g)$.
Pour tout ce que nous allons faire dans le cours avec des polynômes, la division euclidienne sera centrale. Nous démontrons ici son existence.

Théorème 4.3 (Division euclidienne)**.** *Soit* $g = \sum_{i=0}^{d} b_i X^i \in K[X]$ *un polynôme de degré* $d \geq 0$. *Alors, pour tout polynôme* $f \in K[X]$ *il existe des uniques polynômes* $q, r \in K[X]$ *tels que*

$$f = qg + r \quad \text{et} \quad \deg(r) < d.$$

On appelle r *le* reste *de la division.*

Démonstration. Soit $f(X) = \sum_{i=0}^{n} a_i X^i \in K[X]$ de degré n.
Existence : Nous montrons l'existence par récurrence sur n. Si $n < d$, on pose $q = 0$ et $r = f$ et on a terminé. Supposons donc $n \geq d$ et que l'existence est déjà connue pour tous les polynômes de degré strictement plus petit que n. On pose

$$f_1(X) := f(X) - a_n \cdot b_d^{-1} X^{n-d} g(X).$$

C'est un polynôme de degré au plus $n - 1$ parce que nous avons annulé le coefficient devant X^n. Alors, par l'hypothèse de récurrence il existe $q_1, r_1 \in K[X]$ tels que $f_1 = q_1 g + r_1$ et $\deg(r_1) < d$. Donc

$$f(X) = f_1(X) + a_n b_d^{-1} g(X) X^{n-d} = q(X) g(X) + r_1(X)$$

où $q(X) := q_1(X) + a_n b_d^{-1} X^{n-d}$ et nous avons démontré l'existence.
Unicité : Supposons $f = qg + r = q_1 g + r_1$ avec $q, q_1, r, r_1 \in K[X]$ et $\deg(r), \deg(r_1) < d$. Alors $g(q - q_1) = r_1 - r$. Si $q = q_1$, alors $r = r_1$ et on a terminé. Si $q \neq q_1$, alors $\deg(q - q_1) \geq 0$ et on trouve $\deg(r_1 - r) = \deg(g(q - q_1)) \geq \deg(g) = d$. Cela est une contradiction, donc $q \neq q_1$ ne peut pas apparaître. □

Dans les exercices, vous allez vous entraîner pour calculer des divisions euclidiennes.

Corollaire 4.4. *Soit* $f \in K[X]$ *un polynôme de degré* $\deg(f) \geq 1$ *et soit* $a \in K$. *Alors, les assertions suivantes sont équivalentes :*

(i) $f(a) = 0$

(ii) $(X - a) \mid f$

Démonstration. (i) ⇒ (ii) : Supposons $f(a) = 0$ et faisons la division euclidienne $f(X)$ par $X - a$:

$$f(X) = q(X)(X - a) + r$$

pour $r \in K$ (un polynôme de degré < 1). Evaluons cette égalité en a, nous avons $0 = f(a) = q(a)(a - a) + r = r$, donc le reste est zéro.
(ii) ⇒ (i) : Supposons que $X - a$ divise $f(X)$. Alors on a $f(X) = q(X) \cdot (X - a)$ pour un polynôme $q \in K[X]$. Evaluons-le en a pour obtenir $f(a) = q(a) \cdot (a - a) = 0$. □

Proposition 4.5. *Soient* $f, g \in K[X]$ *deux polynômes tels que* $f \neq 0$. *Alors il existe un unique polynôme unitaire* $d \in K[X]$, *appelé* plus grand diviseur commun $\mathrm{pgcd}(f, g)$, *tel que*

- $d \mid f$ *et* $d \mid g$ *(*diviseur commun*) et*

- *pour tout* $e \in K[X]$ *on a* $((e \mid f$ *et* $e \mid g) \Rightarrow e \mid d)$ *(*plus grand *dans le sens que tout autre diviseur commun divise* d*).*

En plus, il existe des polynômes $a, b \in K[X]$ *tels qu'on a une* relation de Bézout

$$d = af + bg.$$

Démonstration. On montre que l'algorithme d'Euclide donne le résultat.

- Préparation : On pose

$$\begin{cases} f_0 = f, \ f_1 = g & \text{si } \deg(f) \geq \deg(g), \\ f_0 = g, \ f_1 = f & \text{sinon.} \end{cases}$$

  On pose aussi $B_0 = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$.

- Si $f_1 = 0$, on **arrête** et on pose $d := f_0$.
  Si $f_1 \neq 0$, on fait la division euclidienne

$$f_0 = f_1 q_1 + f_2 \quad \text{où } q_1, f_2 \in A \text{ tels que } (f_2 = 0 \text{ ou } \deg(f_2) < \deg(f_1)).$$

  On pose $A_1 := \left(\begin{smallmatrix} -q_1 & 1 \\ 1 & 0 \end{smallmatrix}\right)$, $B_1 := A_1 B_0$.
  On a $\left(\begin{smallmatrix} f_2 \\ f_1 \end{smallmatrix}\right) = A_1 \left(\begin{smallmatrix} f_1 \\ f_0 \end{smallmatrix}\right) = B_1 \left(\begin{smallmatrix} f_1 \\ f_0 \end{smallmatrix}\right)$.

- Si $f_2 = 0$, on **arrête** et on pose $d := f_1$.
  Si $f_2 \neq 0$, on fait la division euclidienne

$$f_1 = f_2 q_2 + f_3 \quad \text{où } q_2, f_3 \in A \text{ tels que } (f_3 = 0 \text{ ou } \deg(f_3) < \deg(f_2)).$$

  On pose $A_2 := \left(\begin{smallmatrix} -q_2 & 1 \\ 1 & 0 \end{smallmatrix}\right)$, $B_2 := A_2 B_1$.
  On a $\left(\begin{smallmatrix} f_3 \\ f_2 \end{smallmatrix}\right) = A_2 \left(\begin{smallmatrix} f_2 \\ f_1 \end{smallmatrix}\right) = B_2 \left(\begin{smallmatrix} f_1 \\ f_0 \end{smallmatrix}\right)$.

- Si $f_3 = 0$, on **arrête** et on pose $d := f_2$.
  Si $f_3 \neq 0$, on fait la division euclidienne

$$f_2 = f_3 q_3 + f_4 \quad \text{où } q_3, f_4 \in A \text{ tels que } (f_4 = 0 \text{ ou } \deg(f_4) < \deg(f_3)).$$

  On pose $A_3 := \left(\begin{smallmatrix} -q_3 & 1 \\ 1 & 0 \end{smallmatrix}\right)$, $B_3 := A_3 B_2$.
  On a $\left(\begin{smallmatrix} f_4 \\ f_3 \end{smallmatrix}\right) = A_3 \left(\begin{smallmatrix} f_3 \\ f_2 \end{smallmatrix}\right) = B_3 \left(\begin{smallmatrix} f_1 \\ f_0 \end{smallmatrix}\right)$.

- $\cdots$

- Si $f_n = 0$, on **arrête** et on pose $d := f_{n-1}$.
  Si $f_n \neq 0$, on fait la division euclidienne

$$f_{n-1} = f_n q_n + f_{n+1} \quad \text{où } q_n, f_{n+1} \in A \text{ tels que } (f_{n+1} = 0 \text{ ou } \deg(f_{n+1}) < \deg(f_n)).$$

  On pose $A_n := \left(\begin{smallmatrix} -q_n & 1 \\ 1 & 0 \end{smallmatrix}\right)$, $B_n := A_n B_{n-1}$.
  On a $\left(\begin{smallmatrix} f_{n+1} \\ f_n \end{smallmatrix}\right) = A_n \left(\begin{smallmatrix} f_n \\ f_{n-1} \end{smallmatrix}\right) = B_n \left(\begin{smallmatrix} f_1 \\ f_0 \end{smallmatrix}\right)$.

- $\cdots$

Il est clair que l'algorithme ci-dessus (c'est l'algorithme d'Euclide !) s'arrête car

$$\deg(f_n) < \deg(f_{n-1}) < \cdots < \deg(f_2) < \deg(f_1)$$

sont des nombres naturels ou $-\infty$.
Supposons que l'algorithme se termine avec $f_n = 0$. Donc, $d = f_{n-1}$. Nous avons par construction :

$$\begin{pmatrix} f_n \\ f_{n-1} \end{pmatrix} = \begin{pmatrix} 0 \\ d \end{pmatrix} = B_{n-1} \begin{pmatrix} f_1 \\ f_0 \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ r & s \end{pmatrix} \begin{pmatrix} f_1 \\ f_0 \end{pmatrix} = \begin{pmatrix} \alpha f_1 + \beta f_0 \\ r f_1 + s f_0 \end{pmatrix},$$

montrant

$$d = r f_1 + s f_0. \tag{4.1}$$

Noter que le déterminant de A_i est -1 pour tout i, donc $\det(B_{n-1}) = (-1)^{n-1}$. Alors $C := (-1)^{n-1} \begin{pmatrix} s & -\beta \\ -r & \alpha \end{pmatrix}$ est l'inverse de B_{n-1}. Donc

$$\begin{pmatrix} f_1 \\ f_0 \end{pmatrix} = C B_{n-1} \begin{pmatrix} f_1 \\ f_0 \end{pmatrix} = C \begin{pmatrix} 0 \\ d \end{pmatrix} = \begin{pmatrix} d(-1)^n \beta \\ d(-1)^{n-1} \alpha \end{pmatrix},$$

montrant $d \mid f_1$ et $d \mid f_0$. Ceci montre que d est un diviseur commun de f_0 et f_1. Si e est n'importe quel diviseur commun de f_0 et f_1, alors par l'équation (4.1) e divise d. Finalement on divise d, r, s par le coefficient dominant de d pour rendre d unitaire.
Si nous avons d_1, d_2 des pgcd unitaires, alors d_1 divise d_2 et d_2 divise d_1. Comme les deux sont unitaires, il en suit $d_1 = d_2$, donc l'unicité. $\square$

Dans les exercices, vous allez vous entraîner à calculer le pgcd de deux polynômes. On n'exigera pas l'utilisation de matrices pour obtenir la relation de Bézout ; il suffira simplement de « remonter » les égalités pour l'obtenir.

5 Polynôme caractéristique

Objectifs :

- Maîtriser la définition du polynôme caractéristique ;
- connaître sa signification pour le calcul des valeurs propres ;
- savoir calculer des polynômes caractéristiques ;
- connaître des exemples et savoir démontrer des propriétés simples.

Dans la section 3 nous avons vu comment calculer l'espace propre pour une valeur propre donnée. Ici nous allons répondre à la question : **comment trouver les valeurs propres ?**
Commençons par l'idée principale. Soit $\lambda \in K$ et M une matrice carrée. Rappelons

$$E_M(\lambda) = \ker(\lambda \cdot \mathrm{id} - M).$$

Nous avons l'équivalence des assertions suivantes :

(i) λ est une valeur propre pour M.

(ii) $E_M(\lambda) \neq 0$.

(iii) La matrice $\lambda \cdot \mathrm{id} - M$ n'est pas inversible.

(iv) $\det(\lambda \cdot \mathrm{id} - M) = 0$.

L'idée principale est de considérer λ comme une variable X. Ainsi le déterminant de la matrice $X \cdot \mathrm{id} - M$ devient un polynôme dans $K[X]$. C'est le *polynôme caractéristique*. Par les assertions équivalentes ci-dessus, ses zéros sont précisement les valeurs propres de M.

Définition 5.1. • *Soit $M \in \mathrm{Mat}_{n\times n}(K)$ une matrice. Le* polynôme caractéristique de M *est défini comme*

$$\mathrm{car}_M(X) := \det\big(X \cdot \mathrm{id}_n - M\big) \in K[X].$$

• *Soient V un K-espace vectoriel de dimension finie et $\varphi \in \mathrm{End}_K(V)$ et S une K-base de V. Le* polynôme caractéristique de φ *est défini comme*

$$\mathrm{car}_\phi(X) := \mathrm{car}_{M_{S,S}(\varphi)}(X).$$

Remarque 5.2. *Informations pour les 'spécialistes' : Noter que la définition du polynôme caractéristique utilise les déterminants dans l'anneau $K[X]$. C'est la raison pour laquelle nous avons présenté les déterminants d'une façon plus générale dans le rappel. Alternativement, on peut aussi travailler dans le corps des fonctions rationnelles sur K, c'est-à-dire, le corps dont les éléments sont des fractions de polynômes à coefficient dans K.*

Lemme 5.3. *Soit $M \in \mathrm{Mat}_{n\times n}(K)$.*

(a) $\mathrm{car}_M(X)$ *est un polynôme unitaire de degré n.*

(b) $\mathrm{car}_M(X)$ *est invariant par conjugaison, c'est-à-dire, pour tout $N \in \mathrm{GL}_n(K)$ on a l'égalité* $\mathrm{car}_M(X) = \mathrm{car}_{N^{-1}MN}(X)$.

Démonstration. (a) Cela se démontre par récurrence en n. Le cas $n = 1$ est clair car la matrice est $(X - m_{1,1})$, donc son déterminant est $X - m_{1,1}$.
Pour l'hérédité, rappelons la notation $M'_{i,j}$ pour la matrice obtenue de M en supprimant la i-ème ligne et le j-ème colonne. Supposons maintenant le résultat démontré pour $n - 1$. Par le développement de Laplace, nous avons

$$\mathrm{car}_M(X) = (X - m_{1,1})\,\mathrm{car}_{M'_{1,1}}(X) - \sum_{i=2}^{n} (-1)^i m_{i,1} \cdot \det\big(X \cdot \mathrm{id} - M\big)'_{i,1}.$$

Par l'hypothèse de récurrence, $\mathrm{car}_{M_{1,1}}(X)$ est un polynôme unitaire de degré $n - 1$, donc $(X - m_{1,1})\,\mathrm{car}_{M_{1,1}}(X)$ est unitaire de degré n. Dans la matrice $\big(X \cdot \mathrm{id} - M\big)'_{i,1}$ avec $i \neq 1$, la variable X n'apparaît que $n - 2$ fois. Donc dans le polynôme caractéristique, elle ne peut apparaître qu'à la $n - 2$-ème puissance au maximum. En conséquence, $\mathrm{car}_M(X)$ est unitaire.

(b) Nous utilisons la multiplicativité du déterminant pour l'anneau $K[X]$ (proposition 2.15).

$$\begin{aligned}\mathrm{car}_{N^{-1}MN}(X) &= \det(X\cdot \mathrm{id}_n - N^{-1}MN) = \det(N^{-1}(X\cdot \mathrm{id}_n - M)N)\\ &= \det(N)^{-1}\det(X\cdot \mathrm{id}_n - M)\det(N) = \det(X\cdot \mathrm{id}_n - M) = \mathrm{car}_M(X).\end{aligned}$$

□

Corollaire 5.4. *Soit V un K-espace vectoriel de dimension finie n.*

(a) $\mathrm{car}_\varphi(X)$ *est un polynôme unitaire de degré n.*

(b) $\mathrm{car}_\varphi(X)$ *ne dépend pas du choix de la base de V qui entre dans sa définition.*

Démonstration. (a) Lemme 5.3 (a).
(b) Soient S et T deux bases de V. L'assertion suit du lemme 5.3 (b) et l'égalité $M_{T,T}(\varphi) = C_{S,T}^{-1} \circ M_{S,S}(\varphi) \circ C_{S,T}$. □

Nous reprenons les exemples de la section 3.

Exemple 5.5. *(a) Soit* $M = \begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix} \in \mathrm{Mat}_{2\times 2}(\mathbb{R})$. *Nous trouvons*

$$\mathrm{car}_M(X) = (X-3)(X-2).$$

(Il est important de connaître la factorisation en polynômes irréductibles du polynôme caractéristique. Donc il est inutile de l'écrire comme $X^2 - 5X + 6$.)

(b) Soit $M = \begin{pmatrix} 3 & 1 \\ 0 & 2 \end{pmatrix} \in \mathrm{Mat}_{2\times 2}(\mathbb{R})$. *Nous trouvons encore*

$$\mathrm{car}_M(X) = (X-3)(X-2).$$

(c) Soit $M = \begin{pmatrix} 5 & 1 \\ -4 & 10 \end{pmatrix} \in \mathrm{Mat}_{2\times 2}(\mathbb{R})$. *Nous trouvons encore*

$$\mathrm{car}_M(X) = (X-5)(X-10) + 4 = (X-6)(X-9).$$

Noter que pour simplifier le calcul, le lemme 5.3 (b) nous permet d'utiliser la matrice conjuguée $\begin{pmatrix} 1 & 1 \\ 1 & 4 \end{pmatrix}^{-1} \begin{pmatrix} 5 & 1 \\ -4 & 10 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 4 \end{pmatrix} = \begin{pmatrix} 6 & 0 \\ 0 & 9 \end{pmatrix}$ *pour le calcul du polynôme caractéristique, donc on peut directement écrire la factorisation en facteurs linéaires (en général, cela ne sera pas possible).*

(d) Soit $M = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix} \in \mathrm{Mat}_{2\times 2}(\mathbb{R})$. *Nous trouvons*

$$\mathrm{car}_M(X) = (X-2)^2,$$

un polynôme avec une double racine.

(e) *Soit* $M = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in \mathrm{Mat}_{2\times 2}(\mathbb{R})$. *Nous trouvons*

$$\mathrm{car}_M(X) = X^2 + 1,$$

un polynôme qui ne se factorise pas en produit de polynômes linéaires dans $\mathbb{R}[X]$.

(f) *Soit* $M = \left(\begin{smallmatrix} 2 & 1 & 1 \\ 3 & 2 & 3 \\ -3 & -1 & -2 \end{smallmatrix}\right) \in \mathrm{Mat}_{3\times 3}(\mathbb{R})$. *Pour le polynôme caractéristique, nous calculons le déterminant*

$$\begin{aligned} \left|\begin{smallmatrix} X-2 & -1 & -1 \\ -3 & X-2 & -3 \\ 3 & 1 & X+2 \end{smallmatrix}\right| &= (X-2)\cdot \left|\begin{smallmatrix} X-2 & -3 \\ 1 & X+2 \end{smallmatrix}\right| + 3\cdot \left|\begin{smallmatrix} -1 & -1 \\ 1 & X+2 \end{smallmatrix}\right| + 3\cdot \left|\begin{smallmatrix} -1 & -1 \\ X-2 & -3 \end{smallmatrix}\right| \\ &= (X-2)\big((X-2)(X+2)+3\big) + 3\cdot\big(-(X+2)+1\big) + 3\cdot\big(3+(X-2)\big) \\ &= (X-2)(X^2-1) = (X-2)(X-1)(X+1) \end{aligned}$$

Proposition 5.6. *(a)* *Pour* $M \in \mathrm{Mat}_{n\times n}(K)$ *nous avons*

$$\mathrm{Spec}(M) = \{a \in K \mid \mathrm{car}_M(a) = 0\} = \{a \in K \mid \ (X-a) \mid \mathrm{car}_M(X)\}.$$

(b) *Pour* $\varphi \in \mathrm{End}_K(V)$ *avec un* K*-espace vectoriel* V *de dimension finie nous avons*

$$\mathrm{Spec}(\varphi) = \{a \in K \mid \mathrm{car}_\varphi(a) = 0\} = \{a \in K \mid \ (X-a) \mid \mathrm{car}_\varphi(X)\}.$$

Démonstration. Il suffit de montrer (a). La première égalité suit comme ceci (avec $a \in K$) :

$$a \in \mathrm{Spec}(M) \Leftrightarrow \ker(a\cdot \mathrm{id}_n - M) \neq 0 \Leftrightarrow \underbrace{\det(a\cdot \mathrm{id}_n - M)}_{=\mathrm{car}_M(a)} = 0 \Leftrightarrow \mathrm{car}_M(a) = 0.$$

La deuxième égalité est seulement le fait que $a \in K$ est un zéro d'un polynôme f si et seulement si $(X-a)|f$ (corollaire 4.4). $\square$

Nous avons donc identifié les valeurs propres comme les zéros du polynôme caractéristique. Cela répond à la question du début : **pour calculer les valeurs propres d'une matrice, calculer son polynôme caractéristique et trouver ses zéros**.
Mais le polynôme caractéristique possède encore une autre propriété importante qui a été découverte par Cayley et Hamilton. D'abord il faut introduire de la terminologie.

Définition 5.7. *(a)* *Soit* $M \in \mathrm{Mat}_{n\times n}(K)$ *une matrice. Si* $f(X) = \sum_{i=0}^d a_i X^i \in K[X]$ *est un polynôme, alors nous posons* $f(M) := \sum_{i=0}^d a_i M^i \in \mathrm{Mat}_{n\times n}(K)$. *Noter :* $M^0 = \mathrm{id}_n$.

(b) *Soit* $\varphi \in \mathrm{End}_K(V)$ *un endomorphisme d'un* K*-espace vectoriel* V. *Si* $f(X) = \sum_{i=0}^d a_i X^i \in K[X]$ *est un polynôme, alors nous posons* $f(\varphi) := \sum_{i=0}^d a_i \varphi^i$*, ce qui est encore un endomorphisme dans* $\mathrm{End}_K(V)$. *Attention :* $\varphi^i = \underbrace{\varphi \circ \varphi \circ \cdots \circ \varphi}_{i \text{ fois}}$ *et* $\varphi^0 = \mathrm{id}_V$.

Définition-Lemme 5.8 (Pour les mathématiciens uniquement)**.** *(a)* *L'application « évaluation »*

$$\mathrm{ev}_M : K[X] \to \mathrm{Mat}_{n\times n}(K), \quad f(X) \mapsto f(M)$$

est un homomorphisme d'anneaux (même de K*-algèbres).*

(b) L'application « évaluation »

$$\mathrm{ev}_\varphi : K[X] \to \mathrm{End}_K(V), \quad f(X) \mapsto f(\varphi)$$

est un homomorphisme d'anneaux (même de K-algèbres).

Démonstration. Calculs faciles. □

Théorème 5.9 (Cayley-Hamilton). *Soit $M \in \mathrm{Mat}_{n\times n}(K)$. Alors,*

$$\mathrm{car}_M(M) = 0_n \in \mathrm{Mat}_{n\times n}(K).$$

Démonstration. L'astuce est d'utiliser les matrices adjointes. Nous avons dans $\mathrm{Mat}_{n\times n}(K[X])$

$$(X \cdot \mathrm{id}_n - M)^\# \cdot (X \cdot \mathrm{id}_n - M) = \det(X \cdot \mathrm{id}_n - M) \cdot \mathrm{id}_n \overset{\text{déf}}{=} \mathrm{car}_M(X) \cdot \mathrm{id}_n. \tag{5.2}$$

L'idée de la preuve est très simple : si l'on remplace X par M dans (5.2), on obtient 0, car à gauche nous avons le facteur $(M \cdot \mathrm{id}_n - M) = M - M = 0$. Le problème est que dans $\mathrm{Mat}_{n\times n}(K[X])$ le X apparaît dans les coefficients des matrices, et nous n'avons certainement pas le droit de remplacer un coefficient d'une matrice par une matrice. Ce qu'on fait est d'écrire une matrice dont les coefficients sont des polynômes comme un polynôme dont les coefficients sont des matrices :

$$\begin{pmatrix} \sum_{k=0}^d a_{1,1,k}X^k & \cdots & \sum_{k=0}^d a_{1,n,k}X^k \\ \vdots & \ddots & \vdots \\ \sum_{k=0}^d a_{n,1,k}X^k & \cdots & \sum_{k=0}^d a_{n,n,k}X^k \end{pmatrix} = \sum_{k=0}^d \begin{pmatrix} a_{1,1,k} & \cdots & a_{1,n,k} \\ \vdots & \ddots & \vdots \\ a_{n,1,k} & \cdots & a_{n,n,k} \end{pmatrix} \cdot X^k \cdot \mathrm{id}_n.$$

Ayant fait cela, il faudrait montrer que l'évaluation de ce polynôme à coefficients matriciels en une matrice donne lieu à un homomorphisme d'anneaux. Malheureusement, l'anneau des matrices n'est pas commutatif, donc la théorie développée ne s'applique pas. La preuve que nous donnons contourne ce problème en faisant une comparaison de coefficients au lieu d'une évaluation, mais est basée sur l'idée exposée.

La définition de la matrice adjointe montre que la plus grande puissance de X qui peut apparaître dans un coefficient de la matrice $(X \cdot \mathrm{id}_n - M)^\#$ est $n-1$. Comme indiqué ci-dessus, nous pouvons alors écrire cette matrice en tant que polynôme de degré $n-1$ à coefficients dans $\mathrm{Mat}_{n\times n}(K)$:

$$(X \cdot \mathrm{id}_n - M)^\# = \sum_{i=0}^{n-1} B_i X^i \quad \text{avec} \quad B_i \in \mathrm{Mat}_{n\times n}(K).$$

Nous écrivons $\mathrm{car}_M(X) = \sum_{i=0}^n a_i X^i$ (où $a_n = 1$) et reprenons l'équation (5.2) dans $\mathrm{Mat}_{n\times n}(K)$:

$$\begin{aligned} \mathrm{car}_M(X) \cdot \mathrm{id}_n = \sum_{i=0}^n a_i \cdot \mathrm{id}_n \cdot X^i &= \big(\sum_{i=0}^{n-1} B_i X^i\big)(X \cdot \mathrm{id}_n - M) \\ &= \sum_{i=0}^{n-1} (B_i X^{i+1} - B_i M X^i) = -B_0 M + \sum_{i=1}^{n-1} (B_{i-1} - B_i M) X^i + B_{n-1} X^n. \end{aligned}$$

Nous comparons les coefficients (encore des matrices !) pour obtenir

$$a_0 \cdot \mathrm{id}_n = -B_0 M, \quad a_i \cdot \mathrm{id}_n = B_{i-1} - B_i M \text{ pour } 1 \leq i \leq n-1 \quad \text{et} \quad B_{n-1} = \mathrm{id}_n.$$

Cette comparaison de coefficients nous permet de continuer nos calculs dans $\mathrm{Mat}_{n\times n}(K)$ pour obtenir $\mathrm{car}_M(M) = 0_n$ ainsi :

$$\begin{aligned}
&\mathrm{car}_M(M)\cdot \mathrm{id}_n = \sum_{i=0}^{n} a_i \cdot M^i = -B_0 M + \sum_{i=1}^{n-1}(B_{i-1} - B_i M)M^i + B_{n-1}M^n \\
&= -B_0 M + B_0 M - B_1 M^2 + B_1 M^2 - B_2 M^3 + B_2 M^3 - \cdots - B_{n-1}M^n + B_{n-1}M^n = 0_n.
\end{aligned}$$

□

Le théorème de Cayley-Hamilton reste évidemment vrai si l'on remplace la matrice M par un endomorphisme $\varphi \in \mathrm{End}_K(V)$.

Théorème 5.10 (Cayley-Hamilton pour endomorphismes)**.** *Soit V un K-espace vectoriel de dimension finie et $\varphi \in \mathrm{End}_K(\varphi)$. Alors, $\mathrm{car}_\varphi(\varphi) = 0 \in \mathrm{End}_K(V)$.*

Démonstration. Nous avons par définition, $\mathrm{car}_\varphi(X) = \mathrm{car}_{M_{S,S}(\varphi)}(X)$ et par le théorème 5.9

$$0 = \mathrm{car}_{M_{S,S}(\varphi)}(M_{S,S}(\varphi)) = M_{S,S}(\mathrm{car}_{M_{S,S}(\varphi)}(\varphi)) = M_{S,S}(\mathrm{car}_\varphi(\varphi)),$$

donc $\mathrm{car}_\varphi(\varphi) = 0$. Ce calcul repose sur $M_{S,S}(\varphi^i) = \big(M_{S,S}(\varphi)\big)^i$ (voir exercices) □

6 Polynôme minimal

Objectifs :

- Maîtriser la définition du polynôme minimal ;
- connaître sa signification pour le calcul des valeurs propres ;
- savoir calculer des polynômes minimaux ;
- connaître des exemples et savoir démontrer des propriétés simples.

A part le polynôme caractéristique, nous allons également introduire le *polynôme minimal*.

Définition-Lemme 6.1. *Soit $M \in \mathrm{Mat}_{n\times n}(K)$ une matrice.*

(a) Il existe un unique polynôme unitaire $\mathrm{mipo}_M(X) \in K[X]$ de degré minimal avec la propriété $\mathrm{mipo}_M(M) = 0_n$. Ce polynôme est appelé le polynôme minimal de M.

(b) Tout polynôme $f \in K[X]$ avec la propriété $f(M) = 0_n$ est un multiple de $\mathrm{mipo}_M(X)$.

(c) Pour toute matrice inversible $N \in \mathrm{Mat}_{n\times n}(K)$, nous avons $\mathrm{mipo}_{N^{-1}MN}(X) = \mathrm{mipo}_M(X)$.

(d) *Soit $\varphi \in \mathrm{End}_K(V)$ pour un K-espace vectoriel V de dimension finie avec K-base S. Nous posons*

$$\mathrm{mipo}_\varphi(X) := \mathrm{mipo}_{M_{S,S}(\varphi)}(X)$$

et l'appellons polynôme minimal de φ. *Ce polynôme ne dépend pas du choix de la base S.*

Démonstration. (a,b) Par le théorème de Cayley-Hamilton 5.9 il existe un polynôme $0 \neq f \in K[X]$ qui annule M. Considérons maintenant l'ensemble de tels polynômes

$$E = \{f \in K[X] \mid f \neq 0 \text{ et } f(M) = 0\}.$$

On choisit un $g \in E$ unitaire de degré minimal parmi les éléments de E.
Nous allons utiliser la division euclidienne pour montrer l'unicité et (b). Soit $f \in E$. On a donc $q, r \in K[X]$ tels que $r = 0$ ou $\deg(r) < \deg(g)$ et

$$f = qg + r,$$

ce qui implique

$$0 = f(M) = q(M)g(M) + r(M) = q(M) \cdot 0 + r(M) = r(M).$$

En conséquence, soit $r = 0$, soit $r \in E$. La dernière possibilité est exclue car le degré de r est strictement plus petit que celui de g qui est minimal. Le fait $r = 0$ signifie $f = qg$, donc tout autre polynôme dans E est un multiple de g. Cela implique aussi l'unicité : si f est de même degré que g et aussi unitaire, alors $f = g$.
(c) Il suffit de noter $(N^{-1}MN)^i = N^{-1}M^iN$, donc pour tout $f \in K[X]$

$$f(N^{-1}MN) = N^{-1}f(M)N = 0_n \Leftrightarrow f(M) = 0_n.$$

(d) L'indépendance du choix de la base est une conséquence de (c) et l'égalité $M_{T,T}(\varphi) = C_{S,T}^{-1} \circ M_{S,S}(\varphi) \circ C_{S,T}$ pour toute autre base T. $\square$

Proposition 6.2. *Soient V un K-espace vectoriel de dimension finie et $\varphi \in \mathrm{End}_K(V)$. Alors,* $\mathrm{Spec}(\varphi) = \{a \in K \mid \ (X-a) \mid \mathrm{mipo}_\varphi(X)\} = \{a \in K \mid \mathrm{mipo}_\varphi(a) = 0\}$.

Evidemment, la même assertion est vraie pour des matrices $M \in \mathrm{Mat}_{n\times n}(K)$. Comparer cette proposition avec la proposition 5.6.

Démonstration. La deuxième égalité est claire (même argument que dans la démonstration de la proposition 5.6). Pour voir la première égalité supposons d'abord $(X-a) \nmid \mathrm{mipo}_\varphi(X)$. De cela nous déduisons que le pgcd de $(X-a)$ et $\mathrm{mipo}_\varphi(X)$ est égal à 1, ce qui nous permet (par l'algorithme d'Euclide/Bézout) de trouver $b, c \in K[X]$ tels que $1 = b(X)(X-a) + c(X)\,\mathrm{mipo}_\varphi(X)$. Soit maintenant $v \in V$ t.q. $\varphi(v) = av$. Nous avons

$$v = \mathrm{id}_V v = b(\varphi)(\varphi(v) - av) + c(\varphi)\,\mathrm{mipo}_\varphi(\varphi)v = 0 + 0 = 0,$$

alors $a \notin \mathrm{Spec}(\varphi)$.

Supposons maintenant $(X-a) \mid \mathrm{mipo}_\varphi(X)$ ce qui nous permet d'écrire $\mathrm{mipo}_\varphi(X) = (X-a)g(X)$ pour un $g \in K[X]$. Puisque le degré de g est strictement plus petit que celui de $\mathrm{mipo}_\varphi(X)$, il doit y avoir un $v \in V$ tel que $w := g(\varphi)v \neq 0$ (sinon, le polynôme minimal $\mathrm{mipo}_\varphi(X)$ serait un diviseur de $g(X)$ ce qui est absurde). Nous avons alors

$$(\varphi - a)w = \mathrm{mipo}_\varphi(\varphi)v = 0,$$

alors $a \in \mathrm{Spec}(\varphi)$. □

Il est utile de remarquer que les propositions 5.6 et 6.2 disent que $\mathrm{car}_\varphi(X)$ et $\mathrm{mipo}_\varphi(X)$ ont les mêmes facteurs de degré 1. En plus, le polynôme caractéristique $\mathrm{car}_\varphi(X)$ est toujours un multiple du polynôme minimal $\mathrm{mipo}_\varphi(X)$, par le théorème de Cayley-Hamilton comme nous le voyons maintenant.

Corollaire 6.3. *Soit $M \in \mathrm{Mat}_{n\times n}(K)$. Alors, le polynôme minimal $\mathrm{mipo}_M(X)$ est un diviseur du polynôme caractéristique $\mathrm{car}_M(X)$. Nous avons aussi la même assertion pour $\varphi \in \mathrm{End}_K(V)$.*

Démonstration. Par le théorème de Cayley-Hamilton 5.9 $\mathrm{car}_M(M) = 0_n$, donc $\mathrm{mipo}_M(X)$ divise $\mathrm{car}_M(X)$ par le lemme 6.1. □

Exemple 6.4. *Voici des exemples clés pour comprendre la différence entre polynôme minimal et polynôme caractéristique :*

- *Les trois matrices suivantes ont le même polynôme caractéristique, $(X-1)^2$:*

  $$M_1 := \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad M_2 := \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad M_3 := \begin{pmatrix} 1 & 691 \\ 0 & 1 \end{pmatrix}.$$

  Le polynôme minimal de M_1 est $X-1$. Puisque $M_2 - 1 \cdot \mathrm{id}_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \neq 0_2$ et $M_3 - 1 \cdot \mathrm{id}_2 = \begin{pmatrix} 0 & 691 \\ 0 & 0 \end{pmatrix} \neq 0_2$, le polynôme minimal est $(X-1)^2$ dans ces deux cas. Noter que nous avons utilisé le fait que les seuls diviseurs normalisés non-constants de $(X-1)^2$ sont $X-1$ et $(X-1)^2$, alors le polynôme minimal doit être un parmi les deux.

- *Les mêmes arguments donnent les polynômes minimaux des matrices suivantes (mais, noter qu'il y a une possibilité de plus) :*

  $$M_4 := \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, M_5 := \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, M_6 := \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Exemple 6.5. *Nous faisons encore un exemple un peu plus complexe. Soit*

$$M = \begin{pmatrix} 4 & 3 & -3 & 7 \\ 7 & 0 & -3 & 7 \\ 6 & -1 & -2 & 6 \\ -1 & -4 & 4 & -4 \end{pmatrix}.$$

Il y a (au moins) deux façons de procéder :

(I) Calculer le polynôme caractéristique et en déduire le polynôme minimal.

Un calcul montre :

$$\mathrm{car}_M(X) = X^4 + 2X^3 - 11X^2 - 12X + 36 = (X+3)^2 \cdot (X-2)^2.$$

Nous savons que les facteurs linéaires dans le polynôme minimal sont les mêmes que dans le polynôme caractéristique. On sait donc

$$\mathrm{mipo}_M(X) = (X+3)^a \cdot (X-2)^b$$

pour $1 \le a, b \le 2$.

On calcule le polynôme minimal en essayant les possibilités.

- *On part de la possiblité du plus petit degré :*

  $$M_{-3} := M + 3 \cdot \mathrm{id} = \begin{pmatrix} 7 & 3 & -3 & 7 \\ 7 & 3 & -3 & 7 \\ 6 & -1 & 1 & 6 \\ -1 & -4 & 4 & -1 \end{pmatrix}, \quad M_2 := M - 2 \cdot \mathrm{id} = \begin{pmatrix} 2 & 3 & -3 & 7 \\ 7 & -2 & -3 & 7 \\ 6 & -1 & -4 & 6 \\ -1 & -4 & 4 & -6 \end{pmatrix}$$

  et on calcule

  $$M_{-3} \cdot M_2 = \begin{pmatrix} 10 & -10 & 10 & 10 \\ 10 & -10 & 10 & 10 \\ 5 & -5 & 5 & 5 \\ -5 & 5 & -5 & -5 \end{pmatrix} \neq 0.$$

  Donc $(X-3)(X+2)$ *n'est pas le polynôme minimal.*

- *On augmente les exposants, l'un après l'autre.*

  On calcule

  $$M_{-3}^2 \cdot M_2 = \begin{pmatrix} 50 & -50 & 50 & 50 \\ 50 & -50 & 50 & 50 \\ 25 & -25 & 25 & 25 \\ -25 & 25 & -25 & -25 \end{pmatrix} \neq 0.$$

  Donc le polynôme minimal n'est pas $(X-3)^2(X+2)$.

  On doit continuer et on calcule

  $$M_{-3} \cdot M_2^2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

  Nous avons donc terminé et trouvé

  $$\mathrm{mipo}_M(X) = (X+3) \cdot (X-2)^2 = X^3 - X^2 - 8X + 12.$$

(II) *Si l'on ne connaît pas le polynôme caractéristique et si l'on ne veut pas le calculer, on peut procéder autrement. Cela nous amenera à la réponse standard :* Pour calculer le polynôme minimal, il faut résoudre des systèmes d'équations linéaires.

Nous procédons par récurrence sur le degré (potentiel) d *du polynôme minimal.*

$d = 1$ *Si le degré était* 1, *la matrice serait scalaire. Cela n'est visiblement pas le cas.*

$d = 2$ *Nous calculons*

$$M^2 = \begin{pmatrix} 12 & -13 & 13 & 3 \\ 3 & -4 & 13 & 3 \\ -1 & -4 & 13 & -1 \\ -4 & 9 & -9 & 5 \end{pmatrix}.$$

Maintenant, il faut considérer le système d'équations linéaires :

$$0 = a_2 M^2 + a_1 M + a_0 = \\ a_2 \cdot \begin{pmatrix} 12 & -13 & 13 & 3 \\ 3 & -4 & 13 & 3 \\ -1 & -4 & 13 & -1 \\ -4 & 9 & -9 & 5 \end{pmatrix} + a_1 \cdot \begin{pmatrix} 4 & 3 & -3 & 7 \\ 7 & 0 & -3 & 7 \\ 6 & -1 & -2 & 6 \\ -1 & -4 & 4 & -4 \end{pmatrix} + a_0 \cdot \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Ce sont donc 16 équations linéaires. En pratique, on peut écrire les coefficients dans une grande matrice. La première ligne contient les coefficients $(1,1)$ *des trois matrices, la deuxième ligne les coefficients* $(1,2)$*, etc., jusqu'à la 16ème ligne qui contient les coefficients* $(4,4)$ *:*

$$\begin{pmatrix} 12 & 4 & 1 \\ -13 & 3 & 0 \\ 13 & -3 & 0 \\ 3 & 7 & 0 \\ 3 & 7 & 0 \\ -4 & 0 & 1 \\ 13 & -3 & 0 \\ 3 & 7 & 0 \\ -1 & 6 & 0 \\ -4 & -1 & 0 \\ 13 & -2 & 1 \\ -1 & 6 & 0 \\ -4 & -1 & 0 \\ 9 & -4 & 0 \\ -9 & 4 & 0 \\ 5 & -4 & 1 \end{pmatrix}.$$

On trouve que ce système n'admet pas de solution non-zéro car le rang de la matice est 3*.*

$d = 3$ *Nous calculons*

$$M^3 = \begin{pmatrix} 32 & 11 & -11 & 59 \\ 59 & -16 & -11 & 59 \\ 47 & -12 & -15 & 47 \\ -12 & -23 & 23 & -39 \end{pmatrix}.$$

Maintenant, il faut considérer le système d'équations linéaires :

$$0 = a_3 M^3 + a_2 M^2 + a_1 M + a_0 =$$
$$a_3 \cdot \begin{pmatrix} 32 & 11 & -11 & 59 \\ 59 & -16 & -11 & 59 \\ 47 & -12 & -15 & 47 \\ -12 & -23 & 23 & -39 \end{pmatrix} + a_2 \cdot \begin{pmatrix} 12 & -13 & 13 & 3 \\ 3 & -4 & 13 & 3 \\ -1 & -4 & 13 & -1 \\ -4 & 9 & -9 & 5 \end{pmatrix} + a_1 \cdot \begin{pmatrix} 4 & 3 & -3 & 7 \\ 7 & 0 & -3 & 7 \\ 6 & -1 & -2 & 6 \\ -1 & -4 & 4 & -4 \end{pmatrix} + a_0 \cdot \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Ce sont donc encore 16 équations linéaires. Nous écrivons encore la matrice des coefficients (noter qu'il suffit d'ajouter la première colonne). Nous donnons également un générateur du noyau (obtenu par l'algorithme de Gauß (en général)) :

$$\begin{pmatrix} 32 & 12 & 4 & 1 \\ 11 & -13 & 3 & 0 \\ -11 & 13 & -3 & 0 \\ 59 & 3 & 7 & 0 \\ 59 & 3 & 7 & 0 \\ -16 & -4 & 0 & 1 \\ -11 & 13 & -3 & 0 \\ 59 & 3 & 7 & 0 \\ 47 & -1 & 6 & 0 \\ -12 & -4 & -1 & 0 \\ -15 & 13 & -2 & 1 \\ 47 & -1 & 6 & 0 \\ -12 & -4 & -1 & 0 \\ -23 & 9 & -4 & 0 \\ 23 & -9 & 4 & 0 \\ -39 & 5 & -4 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ -1 \\ -8 \\ 12 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

On voit que le résultat est le polynôme $X^3 - X^2 - 8X + 12$*, le même que dans (I).*

7 Diagonalisation et décomposition spectrale

Objectifs :

- Connaître et maîtriser la décomposition spectrale ;
- savoir décider si une matrice/un endomorphisme est diagonalisable ; si oui, savoir calculer la forme diagonale et une matrice de changement de bases ;

- savoir calculer la décomposition spectrale d'une matrice/d'un endomorphisme ;
- connaître des exemples et savoir démontrer des propriétés simples.

Une forme diagonale est certainement la forme la plus simple qu'on puisse espérer pour une matrice. Mais nous avons déjà vu qu'en général les matrices n'admettent pas une telle forme. La *décomposition spectrale* et la *forme de Jordan* sont des formes simples qu'on peut toujours obtenir. Dans les cas les plus avantageux, ces formes sont diagonales.
Soient V un K-espace vectoriel (de dimension n) et $\varphi \in \mathrm{End}_K(V)$ un endomorphisme. Nous faisons d'abord une observation fondamentale, mais simple, concernant les matrices en blocs.

Lemme 7.1. *(a) Soit $W \leq V$ un sous-espace tel que $\varphi(W) \subseteq (W)$. Soit S_1 une base de W que nous étendons à une base S de V. Alors,*

$$M_{S,S}(\varphi) = \begin{pmatrix} \boxed{M_1} & \boxed{???} \\ \boxed{0} & \boxed{???} \end{pmatrix}$$

avec $M_1 = M_{S_1,S_1}(\varphi|_W)$.

(b) Soit $V = W_1 \oplus W_2$ tel que $\varphi(W_i) \subseteq W_i$ pour $i = 1, 2$. Soit S_i une K-base de W_i pour $i = 1, 2$; donc, $S = S_1 \cup S_2$ est une K-base de V. Alors,

$$M_{S,S}(\varphi) = \begin{pmatrix} \boxed{M_1} & \boxed{0} \\ \boxed{0} & \boxed{M_2} \end{pmatrix}$$

avec $M_1 = M_{S_1,S_1}(\varphi|_{W_1})$ et $M_2 = M_{S_2,S_2}(\varphi|_{W_2})$.

Démonstration. Il suffit d'appliquer les règles pour écrire la matrice $M_{S,S}(\varphi)$. □

Nous poursuivrons par un lemme vers ces formes spéciales.

Lemme 7.2. *Soit $\varphi \in \mathrm{End}_K(V)$.*

(a) Soit $f \in K[X]$ et $W := \ker(f(\varphi))$. Alors, W est un sous-espace de V stable par φ, c'est à dire : pour tout $w \in W$ on a $\varphi(w) \in W$. Ceci nous permet de restreindre φ à W ; on notera l'application restreinte par $\varphi|_W : W \to W$.

(b) Soient $f, g \in K[X]$ deux polynômes premiers entre eux, c'est à dire : $\mathrm{pgcd}(f(X), g(X)) = 1$. Alors,

$$\underbrace{\ker(f(\varphi) \cdot g(\varphi))}_{=:W} = \underbrace{\ker(f(\varphi))}_{=:W_1} \oplus \underbrace{\ker(g(\varphi))}_{=:W_2}.$$

Avant la preuve, un petit mot sur la notation : $f(\varphi)$ est une application K-linéaire $V \to V$, alors on peut l'appliquer à un vecteur $v \in V$. Notre notation pour ceci est : $f(\varphi)(v)$ ou bien $f(\varphi)v$. Noter les rôles distincts des deux paires de paranthèses dans la première expression. On pourrait aussi l'écrire $(f(\varphi))(v)$, mais je trouve cette écriture un peu lourde.

Démonstration. (a) Nous savons que le noyau de toute application K-linéaire est un sous-espace. Ecrivons $f(X) = \sum_{i=0}^{d} a_i X^i$. Soit alors $w \in W$, i.e. $f(\varphi)w = \sum_{i=0}^{d} a_i \varphi^i(w) = 0$. Nous calculons

$$f(\varphi)(\varphi(w)) = \sum_{i=0}^{d} a_i \varphi^i(\varphi(w)) = \sum_{i=0}^{d} a_i \varphi^{i+1}(w) = \varphi\big(\sum_{i=0}^{d} a_i \varphi^i(w)\big) = \varphi(0) = 0.$$

(b) Il est clair que $W_1 \subseteq W$ et $W_2 \subseteq W$, alors $W_1 + W_2 \subseteq W$. Nous devons démontrer

- $W_1 \cap W_2 = 0$ (le K-espace vectoriel zéro) et
- $W_1 + W_2 = W$.

Puisque $K[X]$ est un anneau euclidien, nous pouvons utiliser l'algorithme d'Euclide (de Bézout) pour obtenir deux autres polynômes $a, b \in K[X]$ tels que $1 = a(X)f(X) + b(X)g(X)$. Soit d'abord $w \in W_1 \cap W_2$. Alors

$$w = \mathrm{id}_V(w) = a(\varphi)f(\varphi)w + b(\varphi)g(\varphi)w = 0 + 0 = 0,$$

ce qui montre le premier point. Pour le deuxième soit $w \in W$. L'équation qu'on vient d'utiliser s'écrit comme

$$w = w_2 + w_1 \text{ avec } w_2 := a(\varphi)f(\varphi)w \text{ et } w_1 := b(\varphi)g(\varphi)w.$$

Mais, on a

$$f(\varphi)(w_1) = b(\varphi)f(\varphi)g(\varphi)w = b(\varphi)0 = 0 \Rightarrow w_1 \in W_1$$

et

$$g(\varphi)(w_2) = a(\varphi)f(\varphi)g(\varphi)w = a(\varphi)0 = 0 \Rightarrow w_2 \in W_2,$$

achevant la démonstration. □

Théorème 7.3 (Décomposition spectrale)**.** *Soit $\varphi \in \mathrm{End}_K(V)$ un endomorphisme avec polynôme minimal $\mathrm{mipo}_\varphi(X) = f_1^{e_1}(X) \cdot f_2^{e_2}(X) \cdot \ldots \cdot f_r^{e_r}(X)$ où les polynômes $f_i(X)$ sont irréductibles (ce sont alors des éléments premiers dans l'anneau principal $K[X]$) et premiers entre eux, c'est à dire $\mathrm{pgcd}(f_i, f_j) = 1$ pour tout $1 \le i < j \le n$ (si l'on choisit les f_i unitaires, alors la condition ne revient qu'à dire que les polynômes sont distincts). Posons $W_i := \ker(f_i^{e_i}(\varphi))$. Alors, les assertions suivantes sont vraies.*

(a) $V = \bigoplus_{i=1}^{r} W_i$.

(b) Si l'on choisit une base S_i du sous-espace W_i pour $1 \le i \le r$, alors $S = S_1 \cup S_2 \cup \cdots \cup S_r$ est une base de W pour laquelle on a :

$$M_{S,S}(\varphi) = \begin{pmatrix} \boxed{M_1} & \boxed{0} & \boxed{0} & \ldots & \boxed{0} \\ \boxed{0} & \boxed{M_2} & \boxed{0} & \ldots & \boxed{0} \\ \vdots & & \ddots & \ddots & \vdots \\ \boxed{0} & \ldots & \boxed{0} & \boxed{M_{r-1}} & \boxed{0} \\ \boxed{0} & \ldots & \boxed{0} & \boxed{0} & \boxed{M_r} \end{pmatrix}$$

avec $M_i := M_{S_i,S_i}(\varphi|_{W_i})$ pour $1 \le i \le r$.

Démonstration. (a) suit du lemme 7.2 (b) par récurrence.
(b) est clair : Ecrire la matrice selon les règles pour obtenir cette forme. Noter que les blocs hors la diagonale sont zéros parce que $\varphi(W_i) \subseteq W_i$. $\square$

Le cas le plus important pour nous est celui où $f_i(X) = X - a_i$ avec $a_i \neq a_j$ pour $i \neq j$ (ce qui implique que les f_i sont irréductibles et distincts). La décomposition spectrale n'est en fait qu'un pas (décisif !) vers la réduction de Jordan. Nous voyons dans la prochaine proposition aussi son utilité pour la diagonalisation. Pour l'instant nous illustrons l'effet de la décomposition spectrale à l'aide d'un exemple. Avant cela, il peut être utile de se rappeller comment appliquer les résultats pour les applications linéaires φ aux matrices.

Remarque 7.4. *Soit $M \in \mathrm{Mat}_{n\times n}(K)$. On peut appliquer la décomposition spectrale à la matrice M comme suit. Pour la base canonique $B := (\begin{pmatrix} 1\\0\\0\\ \vdots\\0\\0\end{pmatrix}, \begin{pmatrix} 0\\1\\0\\ \vdots\\0\\0\end{pmatrix}, \dots, \begin{pmatrix} 0\\0\\0\\ \vdots\\0\\1\end{pmatrix})$ la matrice M décrit une application K-linéaire $\varphi = \varphi_M$ et l'on a $M = M_{B,B}(\varphi)$.*
La décomposition spectrale nous donne une base S. Soit $C := M_{B,S}(\mathrm{id})$ la matrice de changement de bases entre la base S et la base canonique. Alors, nous avons

$$M_{S,S}(\varphi) = C^{-1}MC.$$

Pour être encore un peu plus concret, rappelons comment écrire la matrice C. Si $S = (v_1, \dots, v_n)$ et les vecteurs v_i sont donnés en coordonnées pour la base standard, alors la i-ième colonne de C est juste le vecteur v_i.
Alors, la décomposition spectrale peut être utilisée pour calculer une matrice semblable (par définition, deux matrices A, B sont semblables *si l'une est une conjugée de l'autre : il existe une matrice inversible C telle que $B = C^{-1}AC$) à M ayant la jolie forme du théorème.*

Exemple 7.5. *(a) Soit $M := \begin{pmatrix} 1 & 2 & 3\\ 0 & 1 & 4\\ 0 & 0 & 5\end{pmatrix}$ à coefficients dans $\mathbb{R}$. Le polynôme caractéristique est $(X-1)^2(X-5)$. Il est clair que $\ker(M - 5\cdot \mathrm{id}_3)$ est de dimension 1 ; c'est à dire que 5 est une valeur propre de multiplicité 1 (par définition : son espace propre est de dimension 1). Sans calcul il est clair que $\dim\ker((M-\mathrm{id}_3)^2) = 3-1 = 2$.*

Le théorème 7.3 implique l'existence d'une matrice C telle que

$$C^{-1}\cdot M\cdot C = \begin{pmatrix} 1 & x & 0\\ 0 & 1 & 0\\ 0 & 0 & 5\end{pmatrix}$$

pour un $x \in \mathbb{R}$ qui reste à être déterminé.

En fait, on voit facilement que $x \neq 0$, car dans ce cas le polynôme minimal serait $(X-1)(X-5)$ ce qui est faux (voir aussi la Proposition 7.7).

Calculons une telle matrice C. Pour cela, il nous faut calculer une base du noyau de la matrice

$$(M-\mathrm{id}_3)^2 = \begin{pmatrix} 0 & 2 & 3 \\ 0 & 0 & 4 \\ 0 & 0 & 4 \end{pmatrix} \begin{pmatrix} 0 & 2 & 3 \\ 0 & 0 & 4 \\ 0 & 0 & 4 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 20 \\ 0 & 0 & 16 \\ 0 & 0 & 16 \end{pmatrix}.$$

On peut donc juste prendre $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$.

Il nous faut aussi calculer une base du noyau de la matrice

$$M - 5\cdot \mathrm{id}_3 = \begin{pmatrix} -4 & 2 & 3 \\ 0 & -4 & 4 \\ 0 & 0 & 0 \end{pmatrix}.$$

Pour calculer ce noyau, nous additionons $\frac{1}{2}$ fois la deuxième ligne sur la première pour obtenir $\begin{pmatrix} -4 & 0 & 5 \\ 0 & -4 & 4 \\ 0 & 0 & 0 \end{pmatrix}$. *Le noyau est donc engendré par le vecteur* $\begin{pmatrix} 5 \\ 4 \\ 4 \end{pmatrix}$.

La matrice C recherchée est donc $\begin{pmatrix} 1 & 0 & 5 \\ 0 & 1 & 4 \\ 0 & 0 & 4 \end{pmatrix}$. *Pour nous convaincre de l'exactitude du calcul, nous le vérifions*

$$C^{-1}MC = \begin{pmatrix} 1 & 0 & -5/4 \\ 0 & 1 & -1 \\ 0 & 0 & 1/4 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 4 \\ 0 & 0 & 5 \end{pmatrix} \begin{pmatrix} 1 & 0 & 5 \\ 0 & 1 & 4 \\ 0 & 0 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 5 \end{pmatrix}.$$

Le théorème sur la réduction de Jordan nous dira (plus tard) que nous pouvons choisir une autre matrice C telle que même le 2 dans la matrice est remplacé par 1.

(b) Soit $M := \begin{pmatrix} 2 & -1 & 3 \\ -2 & 1 & -4 \\ 1 & 1 & 0 \end{pmatrix}$ *à coefficients dans $\mathbb{R}$. D'abord nous calculons son polynôme caractéristique :*

$$\mathrm{car}_M(X) = \det(\begin{pmatrix} X-2 & 1 & -3 \\ 2 & X-1 & 4 \\ -1 & -1 & X \end{pmatrix})$$
$$= (X-2)(X-1)X-4+6-3(X-1)+4(X-2)-2X = X^3-3X^2+X-3 = (X-3)(X^2+1).$$

Pour ce calcul nous avons utilisé la règle de Sarrus. Pour obtenir la factorisation, nous pouvons essayer des petits entiers pour trouver un zéro (ici 3). Puis l'autre facteur X^2+1 provient de la division de X^3-3X^2+X-3 par $(X-3)$. Noter que X^2+1 est irréductible dans $\mathbb{R}[X]$ (mais pas dans $\mathbb{C}[X]$).

Commençons par le calcul de

$$E_M(3) = \ker(M - 3 \cdot \mathrm{id}_n) = \ker(\begin{pmatrix} -1 & -1 & 3 \\ -2 & -2 & -4 \\ 1 & 1 & -3 \end{pmatrix}).$$

Maintenant on devrait faire des opérations sur les lignes pour obtenir la forme échelonnée de la matrice pour en déduire le noyau. Mais nous avons de la chance, on peut juste 'voir' un vecteur dans le noyau, notamment $\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$. *Ce vecteur génère donc* $E_M(3)$ *(la dimension ne peut être* 2 *car dans ce cas* $(X-3)^2$ *serait un diviseur du polynôme caractéristique).*

Calculons mainteant

$$\ker(M^2 + M^0) = \ker(\begin{pmatrix} 10 & 0 & 10 \\ -10 & -0 & -10 \\ 0 & 0 & 0 \end{pmatrix}).$$

Ce noyau est clairement de dimension 2 *engendré par* $\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$.

Donc nous pouvons écrire la matrice recherchée : $C = \begin{pmatrix} 1 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}$.

Nous vérifions que tout est allé correctement avec le calcul :

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

Avant de donner une autre caractérisation de la diagonalisabilité, nous rappelons dans un lemme de propriétés faciles des matrices diagonales.

Lemme 7.6. *Soit* $D \in \mathrm{Mat}_{n\times n}(K)$ *une matrice diagonale avec coefficients* $\lambda_1, \lambda_2, \dots, \lambda_n$ *sur la diagonale.*

(a) $\mathrm{Spec}(D) = \{\lambda_i \mid i = 1, \dots, n\}$.

Noter que $\#\,\mathrm{Spec}(D) < n$ *si et seulement s'il existent* $1 \leq i < j \leq n$ *tel que* $\lambda_i = \lambda_j$.

(b) $\mathrm{mipo}_D(X) = \prod_{\lambda \in \mathrm{Spec}(D)} (X - \lambda)$.

Démonstration. Ces assertions sont claires. □

La forme du polynôme minimal dans le lemme, nous permet de donner encore une autre caractérisation de la diagonalisabilité :

Proposition 7.7. *Soient V un K-espace vectoriel de dimension finie et $\varphi \in \mathrm{End}_K(V)$. Les assertions suivantes sont équivalentes :*

(i) *φ est diagonalisable.*

(ii) $\mathrm{mipo}_\varphi(X) = \prod_{a \in \mathrm{Spec}(\varphi)}(X - a)$.

Evidemment, les mêmes assertions sont vraies pour des matrices $M \in \mathrm{Mat}_{n\times n}(K)$.

Démonstration. On écrit $\mathrm{Spec}(\varphi) = \{a_1, \dots, a_r\}$.
« (i) $\Rightarrow$ (ii) » : On choisit une base S telle que $M := M_{S,S}(\varphi)$ est diagonale (voir la proposition 3.11). Un calcul très simple montre que $\prod_{i=1}^r (M - a_i) = 0_n$. Alors, $\mathrm{mipo}_\varphi(X)$ est un diviseur de $\prod_{i=1}^r (X - a_i)$. Mais la proposition 6.2 montre que pour tout i on a $(X - a_i) \mid \mathrm{mipo}_\varphi(X)$. Donc, $\mathrm{mipo}_\varphi(X) = \prod_{i=1}^r (X - a_i)$ (les deux polynômes sont unitaires).
« (ii) $\Rightarrow$ (i) » : On applique la décomposition spectrale 7.3 et il suffit de noter que les matrices M_i sont diagonales car $W_i = E_\varphi(a_i)$ est l'espace propre pour la valeur propre a_i. $\square$

Exemple 7.8. *Considérons la matrice* $M := \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 4 \end{pmatrix}$ *à coefficients dans $\mathbb{R}$. Son polynôme minimal est $(X-1)(X-4)$, alors, elle est diagonalisable.*
(Pour obtenir le polynôme minimal il suffit de voir que l'espace propre pour la valeur propre 1 est de dimension 2.)

8 Réduction de Jordan

Objectifs :

- Connaître et maîtriser la réduction de Jordan ;
- savoir décider les différentes possibilités pour la réduction de Jordan si on connaît le polynôme minimal et le polynôme caractéristique ;
- savoir calculer la réduction de Jordan d'une matrice/d'un endomorphisme ainsi que d'une matrice de changement de bases si le polynôme caractéristique se factorise en facteurs linéaires ;
- connaître des exemples et savoir démontrer des propriétés simples.

Nous avons vu dans la proposition 3.11 que les matrices diagonalisables sont semblables à des matrices diagonales. L'utilité d'une matrice diagonale pour des calculs est évidente. Malheureusement, les matrices ne sont pas toutes diagonalisables. Notre but maintenant est de choisir une base S de V de façon à ce que $M_{S,S}(\varphi)$ ait une forme « simple, jolie et élégante » et soit le plus proche possible de la forme diagonale.
Nous avons aussi vu que la décomposition spectrale 7.3 nous donne une forme diagonale « en blocs ». Notre but pour la réduction de Jordan sera de rendre les matrices dans les blocs le plus simple possible.

Nous présentons la *réduction de Jordan* (la *forme normale de Jordan*) d'un point de vue algorithmique. Les preuves peuvent être abrégées un peu si on travaille sans coordonnées, mais dans ce cas, le calcul de la réduction n'est pas clair.
Pour la suite, soient V un K-espace vectoriel de dimension n et $\varphi \in \mathrm{End}_K(V)$ un endomorphisme.

Définition 8.1. *Soit $v \in V$. Nous posons*

$$\langle v \rangle_\varphi := \langle \varphi^i(v) \mid i \in \mathbb{N} \rangle,$$

le sous-espace de V engendré par $v, \varphi(v), \varphi^2(v), \ldots$.

Remarque 8.2. *Les assertions suivantes sont claires et seront utilisées sans être mentionnées explicitement.*

(a) $\langle v \rangle_\varphi$ est stable par φ, c'est-à-dire, $\varphi(\langle v \rangle_\varphi) \subseteq \langle v \rangle_\varphi$.

(b) Si $W \subseteq V$ est un sous-espace vectoriel stable par φ et si $v \in W$, alors $\langle v \rangle_\varphi \subseteq W$.

Lemme 8.3. *Le polynôme minimal de la matrice dans $\mathrm{Mat}_{n \times n}(K)$*

$$\begin{pmatrix} a & 1 & 0 & 0 & \ldots & 0 \\ 0 & a & 1 & 0 & \ldots & 0 \\ 0 & 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 1 & 0 \\ 0 & 0 & \ldots & 0 & a & 1 \\ 0 & 0 & \ldots & 0 & 0 & a \end{pmatrix}$$

est égal à $(X - a)^n$.

Démonstration. Exercice. □

Cette matrice apparaît de façon très naturelle, comme nous le voyons maintenant.

Lemme 8.4. *Soient $a \in K$, $e \in \mathbb{N}_{>0}$ et $v \in V$ tels que*

$$(\varphi - a \cdot \mathrm{id})^e(v) = 0 \quad \text{et} \quad (\varphi - a \cdot \mathrm{id})^{e-1}(v) \neq 0.$$

Nous posons :

$$\begin{aligned} v_e &:= v, \\ v_{e-1} &:= (\varphi - a \cdot \mathrm{id})(v), \\ &\ldots \\ v_2 &:= (\varphi - a \cdot \mathrm{id})^{e-2}(v), \\ v_1 &:= (\varphi - a \cdot \mathrm{id})^{e-1}(v). \end{aligned}$$

(a) Nous avons :

$$\begin{aligned}\varphi(v_1) &= av_1,\\ \varphi(v_2) &= v_1 + av_2,\\ \varphi(v_3) &= v_2 + av_3,\\ &\dots,\\ \varphi(v_e) &= v_{e-1} + av_e.\end{aligned}$$

(b) $\langle v\rangle_\varphi = \langle v_1, \dots, v_e\rangle$*, le sous-espace de* V *engendré par les* $v_1, \dots, v_e$*.*

(c) Le polynôme minimal de φ *agissant sur* $\langle v\rangle_\varphi$ *est égal à* $(X-a)^e$*.*

(d) Les $v_1, \dots, v_e$ *sont* K*-linéairement indépendants et, en conséquence, forment une base* S *de* $\langle v\rangle_\varphi$*.*

(e) $M_{S,S}(\varphi|_{\langle v\rangle_\varphi}) = \begin{pmatrix} a & 1 & 0 & 0 & \dots & 0\\ 0 & a & 1 & 0 & \dots & 0\\ 0 & 0 & \ddots & \ddots & \ddots & \vdots\\ \vdots & \vdots & \ddots & \ddots & 1 & 0\\ 0 & 0 & \dots & 0 & a & 1\\ 0 & 0 & \dots & 0 & 0 & a \end{pmatrix}.$

Démonstration. (a) C'est un calcul très facile :

$$\begin{aligned}(\varphi - a\cdot \mathrm{id})v_1 &= (\varphi - a\cdot\mathrm{id})^e v = 0 && \Rightarrow \varphi(v_1) = av_1.\\ (\varphi - a\cdot\mathrm{id})v_2 &= v_1 && \Rightarrow \varphi(v_2) = v_1 + av_2.\\ &\dots\\ (\varphi - a\cdot\mathrm{id})v_e &= v_{e-1} && \Rightarrow \varphi(v_e) = v_{e-1} + av_e.\end{aligned}$$

(b) Les équations dans (a) montrent que $\langle v_1, \dots, v_e\rangle$ est stable par φ. Comme $v = v_e \in \langle v\rangle_\varphi$, on obtient l'inclusion $\langle v\rangle_\varphi \subseteq \langle v_1, \dots, v_e\rangle$. L'inclusion inverse se voit par définition :

$$v_{e-i} = (\varphi - a\cdot\mathrm{id})^i(v) = \sum_{k=0}^{i} \binom{i}{k} (-a)^{i-k}\varphi^k(v). \tag{8.3}$$

(c) Le polynôme $(X-a)^e$ annule v et donc $\langle v\rangle_\varphi$. Comme $(X-a)^{e-1}$ n'annule pas v, le polynôme minimal de $\varphi|_{\langle v\rangle_\varphi}$ est $(X-a)^e$.

(d) Supposons que nous avons une combinaison linéaire non-triviale de la forme

$$0 = \sum_{i=0}^{j} \alpha_i v_{e-i}$$

pour $\alpha_j \neq 0$ et $0 \leq j \leq e-1$. Par l'égalité (8.3), nous obtenons

$$0 = \sum_{i=0}^{j} \alpha_i \sum_{k=0}^{i} \binom{i}{k} (-a)^{i-k}\varphi^k(v) = \sum_{k=0}^{j-1} \Big(\sum_{i=k}^{j} \alpha_i \binom{i}{k} (-a)^{i-k}\Big)\varphi^k(v) + \alpha_j\varphi^j(v).$$

Nous avons donc un polynôme non-zéro de degré $j \leq e - 1$ qui annule v et donc $\langle v \rangle_\varphi$. C'est une contradiction avec (c).
(e) La partie (a) donne précisement les renseignements pour pouvoir écrire la matrice selon les règles. $\square$

Nous allons maintenant préciser ce que nous entendons par « la forme de Jordan ».

Définition 8.5. *Une matrice $M \in \mathrm{Mat}_{n\times n}(K)$ est dite ayant « la forme de Jordan » si M est diagonale en blocs et tout bloc a la forme du lemme 8.4(e).*
Plus précisement, M a la forme de Jordan si

$$M = \begin{pmatrix} \boxed{M_1} & \boxed{0} & \boxed{0} & \dots & \boxed{0} \\ \boxed{0} & \boxed{M_2} & \boxed{0} & \dots & \boxed{0} \\ \vdots & & \ddots & \ddots & \vdots \\ \boxed{0} & \dots & \boxed{0} & \boxed{M_{r-1}} & \boxed{0} \\ \boxed{0} & \dots & \boxed{0} & \boxed{0} & \boxed{M_r} \end{pmatrix}$$

(matrice diagonale en blocs), où, pour tout $1 \leq i \leq r$,

$$M_i = \begin{pmatrix} a_i & 1 & 0 & 0 & \dots & 0 \\ 0 & a_i & 1 & 0 & \dots & 0 \\ 0 & 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 1 & 0 \\ 0 & 0 & \dots & 0 & a_i & 1 \\ 0 & 0 & \dots & 0 & 0 & a_i \end{pmatrix}.$$

(On ne demande pas, ici, que les a_i soient deux-à-deux distincts. Mais on peut rassembler les blocs ayant le même a_i ; cela va être le cas dans le théorème principal 8.8.)
La procédure pour trouver une matrice inversible C telle que $C^{-1}MC$ a la forme de Jordan s'appelle réduction de Jordan*. On appelle également* réduction de Jordan *la procédure (à présenter) pour trouver une base S telle que $M_{S,S}(\varphi)$ possède la forme de Jordan (pour un endomorphisme φ). Il peut arriver que l'on appelle la matrice obtenue aussi* la réduction de Jordan de M ou de φ*.*

Exemple 8.6. *Nous reprenons les matrices de l'exemple 6.4.*

(a) Les matrices $M_1 := \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$, $M_2 := \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ ont la forme de Jordan, mais pas la matrice $M_3 := \left(\begin{smallmatrix} 1 & 691 \\ 0 & 1 \end{smallmatrix}\right)$ (sa réduction de Jordan est M_2).

(b) Les matrices

$$M_4 := \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, M_5 := \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, M_6 := \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

ont également déjà la forme de Jordan.

(c) La/une réduction de Jordan de la matrice $\begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 5 \end{pmatrix}$ *obtenue de la décomposition spectrale dans l'exemple 7.5(a) est* $\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 5 \end{pmatrix}$ *(expliqué plus loin).*

Attention : avec nos définitions, il existe des matrices qui ne possèdent pas de réduction de Jordan (sauf si l'on travaille sur $\mathbb{C}$, mais pas sur $\mathbb{R}$) ; on peut relaxer un peu les exigences pour avoir une réduction de Jordan pour toute matrice ; nous n'allons pas poursuivre ceci dans ce cours par manque de temps. Dans les exercices, vous voyez quelques pas vers le cas général.

Nous présentons maintenant l'algorithme de réduction de Jordan. Pour cela nous posons :

- $\varphi_a := \varphi - a \cdot \mathrm{id}$,
- $V_i = \ker(\varphi_a^i)$

Pour l'instant, nous faisons l'hypothèse

$$\mathrm{mipo}_\varphi(X) = (X - a)^e.$$

De cela nous obtenons

$$V = V_e \supset V_{e-1} \supset V_{e-2} \supset \cdots \supset V_1 = E_\varphi(a) \supset V_0 = 0.$$

Avant de donner l'algorithme en général, nous regardons les cas spéciaux $\dim(V) \leq 4$.

- $\underline{\dim(V) = 1}$: Dans ce cas nous avons $e = 1$. Soient $0 \neq v \in V$ n'importe quel vecteur non-zéro et $S = \{v\}$. Alors $M_{S,S}(\varphi) = \begin{pmatrix} a \end{pmatrix}$ est la réduction de Jordan cherchée.
- $\underline{\dim(V) = 2}$: On distingue deux cas :

  (I) $e = 1$. Dans ce cas, $M_{S,S}(\varphi) = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}$ est scalaire pour toute base S de V car V est égal à l'espace propre de φ pour la valeur propre a.

  (II) $e = 2$. Dans ce cas, on peut choisir n'importe quel vecteur $v \in V_2 \setminus V_1$. Si on prend $S = \{\varphi_a(v), v\}$, alors $M_{S,S}(\varphi) = \begin{pmatrix} a & 1 \\ 0 & a \end{pmatrix}$.

- $\underline{\dim(V) = 3}$: On distingue trois cas :

  (I) $e = 1$. Dans ce cas, $M_{S,S}(\varphi) = \begin{pmatrix} a & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & a \end{pmatrix}$ est scalaire pour toute base S de V car V est égal à l'espace propre de φ pour la valeur propre a.

(II) $e = 2$. Dans ce cas, il doit y avoir deux blocs de Jordan. On choisit n'importe quel vecteur $v \in V_2 \setminus V_1$. Puis, on prend n'importe quel vecteur $w \in V_1 \setminus \langle \varphi_a(v) \rangle = V \setminus \langle \varphi_a(v), v \rangle$. Finalement, on pose $S = \{\varphi_a(v), v, w\}$ pour obtenir $M_{S,S}(\varphi) = \begin{pmatrix} a & 1 & 0 \\ 0 & a & 0 \\ 0 & 0 & a \end{pmatrix}$.

(III) $e = 3$. Dans ce cas, on peut choisir n'importe quel vecteur $v \in V_3 \setminus V_2$. Si on prend $S = \{\varphi_a^2(v), \varphi_a(v), v\}$, alors $M_{S,S}(\varphi) = \begin{pmatrix} a & 1 & 0 \\ 0 & a & 1 \\ 0 & 0 & a \end{pmatrix}$.

- $\underline{\dim(V) = 4}$: On distingue quatre cas :

(I) $e = 1$. Dans ce cas, $M_{S,S}(\varphi) = \begin{pmatrix} a & 0 & 0 & 0 \\ 0 & a & 0 & 0 \\ 0 & 0 & a & 0 \\ 0 & 0 & 0 & a \end{pmatrix}$ est scalaire pour toute base S de V car V est égal à l'espace propre de φ pour la valeur propre a.

(II) $e = 2$. Dans ce cas, il doit y a encore deux possibilités :

(a) Il y a trois blocs de Jordan. Donc l'un d'eux est de taille 2, les deux autres de taille 1. On choisit n'importe quel vecteur $v \in V_2 \setminus V_1$. Puis, on prend n'importe quel vecteur $w_1 \in V_1 \setminus \langle \varphi_a(v) \rangle = V \setminus \langle \varphi_a(v), v \rangle$. Après cela, on choisit n'importe quel vecteur $w_2 \in V_1 \setminus \langle \varphi_a(v), w_1 \rangle = V \setminus \langle \varphi_a(v), v, w_1 \rangle$. Finalement, on pose $S = \{\varphi_a(v), v, w_1, w_2\}$ pour obtenir $M_{S,S}(\varphi) = \begin{pmatrix} a & 1 & 0 & 0 \\ 0 & a & 0 & 0 \\ 0 & 0 & a & 0 \\ 0 & 0 & 0 & a \end{pmatrix}$.

(b) Il y a deux blocs de Jordan. Tous les deux sont de taille 2. On choisit n'importe quel vecteur $v \in V_2 \setminus V_1$. Puis, on prend n'importe quel vecteur $w \in V_2 \setminus \langle v \rangle$. Finalement, on pose $S = \{\varphi_a(v), v, \varphi_a(w), w\}$ pour obtenir $M_{S,S}(\varphi) = \begin{pmatrix} a & 1 & 0 & 0 \\ 0 & a & 0 & 0 \\ 0 & 0 & a & 1 \\ 0 & 0 & 0 & a \end{pmatrix}$.

(III) $e = 3$. Il y a alors deux blocs de Jordan, un de taille 3, l'autre de taille 1. Dans ce cas, on peut choisir n'importe quel vecteur $v \in V_3 \setminus V_2$. Puis, on prend n'importe quel vecteur $w \in V_1 \setminus \langle \varphi_a^2(v) \rangle = V \setminus \langle \varphi_a^2(v), \varphi_a(v), v \rangle$. Si on prend $S = \{\varphi_a^2(v), \varphi_a(v), v, w\}$, alors $M_{S,S}(\varphi) = \begin{pmatrix} a & 1 & 0 & 0 \\ 0 & a & 1 & 0 \\ 0 & 0 & a & 0 \\ 0 & 0 & 0 & a \end{pmatrix}$.

(IV) $e = 4$. Il n'y a alors qu'un seul bloc de Jordan. Dans ce cas, on peut choisir n'importe quel vecteur $v \in V_4 \setminus V_3$. Si on prend $S = \{\varphi_a^3(v), \varphi_a^2(v), \varphi_a(v), v\}$, alors $M_{S,S}(\varphi) =$

$$\begin{pmatrix} a & 1 & 0 & 0 \\ 0 & a & 1 & 0 \\ 0 & 0 & a & 1 \\ 0 & 0 & 0 & a \end{pmatrix}.$$

Passons maintenant à l'algorithme en général, en gardant les notations introduites en haut. On peut s'imaginer l'espace vectoriel V à l'intérieur d'une boîte rectangulaire :

$$\begin{array}{r||cccccccc}
V_e \setminus V_{e-1} & \blacksquare & & & & & & & \blacksquare \\
V_{e-1} \setminus V_{e-2} & & \blacksquare & \blacksquare & \blacksquare & & & & \\
V_{e-2} \setminus V_{e-3} & \blacksquare & & \blacksquare & \blacksquare & \blacksquare & & & \\
\vdots & & & & & & & & \\
V_2 \setminus V_1 & & \blacksquare & \blacksquare & & \blacksquare & \blacksquare & & \blacksquare \\
V_1 & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare
\end{array}$$

Tout bloc noir représente un vecteur non-zéro, et l'ensemble des vecteurs dans le diagram est linéairement indépendant. Il s'agit dans l'algorithme de mettre de l'ordre dans la boîte. Pour l'instant, on a mis les blocs noirs de façon arbitraire pour indiquer que nous n'avons pas encore beaucoup de connaissances sur ses vecteurs (il n'y a pas de sens profond dans l'image). Le fait qu'il y a deux blocs dans la première ligne veut dire que $\dim V_e - \dim V_{e-1} = 2$, etc. On peut faire l'observation que le nombre de blocs par ligne ne diminue pas en allant d'en haut vers le bas.

(1.) Nous choisissons un vecteur $x_1 \in V_e \setminus V_{e-1}$. Alors, nous avons les vecteurs non-zéros $\varphi_a(x_1) \in V_{e-1}$, $\varphi_a^2(x_1) \in V_{e-2}$, et plus généralement, $\varphi_a^i(x_1) \in V_{e-i}$ pour $i = 0, \ldots, e-1$. Nous modifions notre image :

$$\begin{array}{r||c|ccccccc}
V_e \setminus V_{e-1} & x_1 & & & & & & & \blacksquare \\
V_{e-1} \setminus V_{e-2} & \varphi_a(x_1) & & \blacksquare & \blacksquare & & & & \\
V_{e-2} \setminus V_{e-3} & \varphi_a^2(x_1) & & \blacksquare & \blacksquare & \blacksquare & & & \\
\vdots & & & & & & & & \\
V_2 \setminus V_1 & \varphi_a^{e-2}(x_1) & \blacksquare & & & \blacksquare & \blacksquare & & \blacksquare \\
V_1 & \varphi_a^{e-1}(x_1) & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare
\end{array}$$

La première colonne contient donc une base de $\langle x_1 \rangle_\varphi$.

Si $\langle x_1 \rangle_\varphi = V$ (si aucun bloc noir ne reste), nous avons terminé. Sinon, on continue.

(2.) Maintenant nous calculons l'entier k tel que $\langle x_1 \rangle_\varphi + V_k = V$, mais $\langle x_1 \rangle_\varphi + V_{k-1} \neq V$. Dans notre exemple c'est $k = e$.

Nous choisissons un vecteur x_2 dans $V_k \setminus (\langle x_1 \rangle_\varphi + V_{k-1}) = V_k \setminus (\langle x_1 \rangle + V_{k-1})$. Cela veut dire que nous prenons un vecteur x_2 dans $V_k \setminus V_{k-1}$ qui est linéairement indépendant du vecteur x_1. Nous avons donc les vecteurs non-zéros $\varphi_a^i(x_2) \in V_{k-i}$ pour $i = 0, \ldots, k-1$. Nous modifions notre image :

$$\begin{array}{r||cccccccc}
V_e \setminus V_{e-1} & x_1 & x_2 & & & & & & \\
V_{e-1} \setminus V_{e-2} & \varphi_a(x_1) & \varphi_a(x_2) & \blacksquare & & & & & \\
V_{e-2} \setminus V_{e-3} & \varphi_a^2(x_1) & \varphi_a^2(x_2) & \blacksquare & & \blacksquare & & & \\
\vdots & & & & & & & & \\
V_2 \setminus V_1 & \varphi_a^{e-2}(x_1) & \varphi_a^{e-2}(x_2) & & \blacksquare & \blacksquare & & \blacksquare & \\
V_1 & \varphi_a^{e-1}(x_1) & \varphi_a^{e-1}(x_2) & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare
\end{array}$$

La deuxième colonne contient donc une base de $\langle x_2 \rangle_\varphi$. Le lemme 8.7 nous dit que la somme $\langle x_1 \rangle_\varphi + \langle x_2 \rangle_\varphi$ est directe.

Si $\langle x_1 \rangle_\varphi \oplus \langle x_2 \rangle_\varphi = V$ (si aucun bloc noir ne reste), nous avons terminé. Sinon, on continue.

(3.) Maintenant nous calculons l'entier k tel que $\langle x_1 \rangle_\varphi \oplus \langle x_2 \rangle_\varphi + V_k = V$, mais $\langle x_1 \rangle_\varphi \oplus \langle x_2 \rangle_\varphi + V_{k-1} \neq V$. Dans notre exemple c'est $k = e - 1$.

Nous choisissons un vecteur x_3 dans $V_k \setminus (\langle x_1 \rangle_\varphi \oplus \langle x_2 \rangle_\varphi + V_{k-1})$. Dans l'exemple il faut choisir x_3 dans $V_k \setminus V_{k-1}$ linéairement indépendant de $\varphi_a(x_1), \varphi_a(x_2)$. En fait, par bloc il faut « éviter » précisement un vecteur, celui qui appartient à $V_k \setminus V_{k-1}$. Nous avons donc les vecteurs non-zéros $\varphi_a^i(x_3) \in V_{k-i}$ pour $i = 0, \dots, k-1$. Nous modifions notre image :

$$\begin{array}{r||cccccccc}
V_e \setminus V_{e-1} & x_1 & x_2 & & & & & & \\
V_{e-1} \setminus V_{e-2} & \varphi_a(x_1) & \varphi_a(x_2) & x_3 & & & & & \\
V_{e-2} \setminus V_{e-3} & \varphi_a^2(x_1) & \varphi_a^2(x_2) & \varphi_a(x_3) & & \blacksquare & & & \\
\vdots & & & & & & & & \\
V_2 \setminus V_1 & \varphi_a^{e-2}(x_1) & \varphi_a^{e-2}(x_2) & \varphi_a^{e-3}(x_3) & & \blacksquare & & \blacksquare & \\
V_1 & \varphi_a^{e-1}(x_1) & \varphi_a^{e-1}(x_2) & \varphi_a^{e-2}(x_3) & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare
\end{array}$$

La troisième colonne contient donc une base de $\langle x_3 \rangle_\varphi$. Le lemme 8.7 nous dit que la somme $\langle x_1 \rangle_\varphi \oplus \langle x_2 \rangle_\varphi + \langle x_3 \rangle_\varphi$ est directe.

Si $\langle x_1 \rangle_\varphi \oplus \langle x_2 \rangle_\varphi \oplus \langle x_3 \rangle_\varphi = V$ (si aucun bloc noir ne reste), nous avons terminé. Sinon, on continue.

(...) Ainsi on continue jusqu'à ce qu'aucun bloc noir ne reste. Dans l'exemple on obtiendra l'image :

$$\begin{array}{r||cccccccc}
V_e \setminus V_{e-1} & x_1 & x_2 & & & & & & \\
V_{e-1} \setminus V_{e-2} & \varphi_a(x_1) & \varphi_a(x_2) & x_3 & & & & & \\
V_{e-2} \setminus V_{e-3} & \varphi_a^2(x_1) & \varphi_a^2(x_2) & \varphi_a(x_3) & x_4 & & & & \\
\vdots & & & & & & & & \\
V_2 \setminus V_1 & \varphi_a^{e-2}(x_1) & \varphi_a^{e-2}(x_2) & \varphi_a^{e-3}(x_3) & \varphi_a^{e-4}(x_4) & x_5 & & & \\
V_1 & \varphi_a^{e-1}(x_1) & \varphi_a^{e-1}(x_2) & \varphi_a^{e-2}(x_3) & \varphi_a^{e-3}(x_4) & \varphi_a(x_5) & x_6 & x_7 & x_8
\end{array}$$

Chaque colonne contient une base de $\langle x_i \rangle_\varphi$ et correspond à un bloc. Plus précisement, nous mettons les vecteurs qui se trouvent dans la boîte dans une base S, en commençant à gauche en bas, puis on remonte la première colonne, puis on commence en bas de la deuxième colonne et on remonte, puis la

troisième colonne d'en bas vers le haut, etc. Alors, $M_{S,S}(\varphi)$ sera une matrice en blocs. Chaque bloc a a sur la diagonale principale et des 1 sur la diagonale au-dessus de la diagonale principale. Chaque colonne correspond à un bloc, et la taille du bloc est la hauteur de la colonne.
Dans notre exemple, on a donc 8 blocs, deux de taille e, un de taille $e-1$, un de taille $e-2$, un de taille 2 et trois de taille 1.

Dans l'algorithme, nous avons

$$V_k \setminus (\langle x_1\rangle_\varphi \oplus \cdots \oplus \langle x_r\rangle_\varphi + V_{k-1}) = V_k \setminus (\langle \varphi_a(x_1)^{d_1}, \ldots, \varphi_a(x_r)^{d_r}\rangle + V_{k-1})$$

où pour $i = 1, \ldots, r$, l'entier d_i est l'unique tel que $\varphi_a^{d_i}(x_i) \in V_k \setminus V_{k-1}$. Cela veut dire que par blocs, il faut (et suffit de) « éviter » un seul vecteur. L'égalité ci-dessus suit du fait que la somme $\langle x_1\rangle_\varphi \oplus \cdots \oplus \langle x_r\rangle_\varphi$ est directe ainsi du fait que les vecteurs $x, \varphi_a(x), \ldots, \varphi_a^d(x)$ sont linéairement indépendants si d est strictement inférieur au degré du polynôme minimal de a sur K.

Pour justifier l'algorithme, il faut encore démontrer le lemme suivant.

Lemme 8.7. *Soit $L = \langle x_1\rangle_\varphi \oplus \langle x_2\rangle_\varphi \oplus \cdots \oplus \langle x_i\rangle_\varphi$ construit dans l'algorithme précédent. Par l'algorithme, on a en particulier*

$$\dim_K \langle x_1\rangle_\varphi \geq \dim_K \langle x_2\rangle_\varphi \geq \cdots \geq \dim_K \langle x_i\rangle_\varphi.$$

(La dimension ici est égale à la hauteur de la colonne correspondante.)
Soit k l'entier tel que $L + V_k = V$ et $L + V_{k-1} \neq V$. On a $V_k \not\subseteq L + V_{k-1}$.
Par l'algorithme, on a aussi $k \leq \dim_K \langle x_i\rangle_\varphi$.
Si $y \in V_k \setminus (L + V_{k-1})$ est n'importe quel vecteur, alors la somme

$$L + \langle y\rangle_\varphi = \langle x_1\rangle_\varphi \oplus \langle x_2\rangle_\varphi \oplus \cdots \oplus \langle x_i\rangle_\varphi \oplus \langle y\rangle_\varphi$$

est directe.

Démonstration. Si $V_k \subseteq L + V_{k-1}$, alors $V_k + L = V_{k-1} + L$ (comme $V_{k-1} \subseteq V_k$). Cela implique la première assertion : $V_k \not\subseteq L + V_{k-1}$.
Démontrons maintenant que la somme $L + \langle y\rangle_\varphi$ est directe, c'est-à-dire, $L \cap \langle y\rangle_\varphi = 0$. Soit $w \in L \cap \langle y\rangle_\varphi$. On suppose $w \neq 0$. Soit j le maximum tel que $w \in V_{k-j}$. On a $0 \leq j \leq k-1$. En conséquence, on peut écrire $w = \sum_{q=0}^{k-1-j} c_q \varphi_a^{q+j}(y)$ pour $c_q \in K$ avec $c_0 \neq 0$. Donc

$$w = \varphi_a^j\big(c_0 y + \sum_{q=1}^{k-j-1} c_q \varphi_a^q(y)\big).$$

Par construction de L, on peut écrire

$$w = \varphi_a^j(\ell)$$

pour $\ell \in L$. C'est le cas car $L \cap V_{k-j}$ est engendré par $\varphi_a^{e_m}(x_m)$ pour $1 \leq m \leq i$ et $j \leq e_m = \dim_K \langle x_m\rangle_\varphi - (k-j)$.
Nous obtenons donc

$$0 = \varphi_a^j\big(c_0 y - \ell + \sum_{q=1}^{k-j-1} c_q \varphi_a^q(y)\big).$$

Cela implique

$$z := c_0 y - \ell + \sum_{q=1}^{k-j-1} c_q \varphi_a^q(y) \in V_j \subseteq V_{k-1}.$$

Utilisant que $\sum_{q=1}^{k-j-1} c_q \varphi_a^q(y) \in V_{k-1}$, nous obtenons finalement

$$y = \frac{1}{c_0}\ell + \frac{1}{c_0}z + \frac{1}{c_0}\sum_{q=1}^{k-j-1} c_q \varphi_a^q(y) \in L + V_{k-1},$$

une contradiction. Donc $w = 0$. □

En mettant ensemble la décomposition spectrale avec l'algorithme ci-dessus, nous obtenons finalement le théorème sur la réduction de Jordan.

Théorème 8.8 (Réduction de Jordan)**.** *Supposons que le polynôme minimal de φ est égal à*

$$\mathrm{mipo}_\varphi(X) = \prod_{i=1}^{r}(X - a_i)^{e_i}$$

avec différents $a_i \in K$ et $e_i > 0$ (ceci est toujours le cas lorsque K est « algébriquement clos » (voir Algèbre 3), par exemple $K = \mathbb{C}$).

Alors, φ possède une réduction de Jordan.

On peut décrire la réduction de Jordan précisement, comme suit. En calculant $V_i := \ker\big((\varphi - a_i \cdot \mathrm{id})^{e_i}\big)$, on obtient la décomposition spectrale *(voir le théorème 7.3), c'est à dire :*

$$V = \bigoplus_{i=1}^{r} V_i \quad \text{et} \quad \varphi(V_i) \subseteq V_i \text{ pour tout } 1 \leq i \leq r.$$

Pour tout $1 \leq i \leq r$, on applique l'algorithme ci-dessus pour construire des $x_{i,1}, \ldots, x_{i,s_i} \in V_i$ tels que

$$V_i = \langle x_{i,1}\rangle_\varphi \oplus \cdots \oplus \langle x_{i,s_i}\rangle_\varphi \quad \text{et} \quad \varphi(\langle x_{i,j}\rangle_\varphi) \subseteq \langle x_{i,j}\rangle_\varphi.$$

Soit $e_{i,j}$ l'entier positif minimal tel que $(\varphi - a_i \cdot \mathrm{id})^{e_{i,j}}(x_{i,j}) = 0$ pour tout $1 \leq i \leq r$ et $1 \leq j \leq s_i$. Pour tout espace $\langle x_{i,j}\rangle_\varphi$ on chosit la base $S_{i,j}$ comme dans le lemme 8.4. On pose

$$S := S_{1,1} \cup S_{1,2} \cup \cdots \cup S_{1,s_1} \cup S_{2,1} \cup S_{2,2} \cup \cdots \cup S_{2,s_2} \cup \ldots \ldots \cdots \cup S_{r,s_r}.$$

Alors, S est une K-base de V telle que

$$M_{S,S}(\varphi) = \begin{pmatrix} \boxed{M_1} & \boxed{0} & \boxed{0} & \ldots & \boxed{0} \\ \boxed{0} & \boxed{M_2} & \boxed{0} & \ldots & \boxed{0} \\ \vdots & & \ddots & \ddots & \vdots \\ \boxed{0} & \ldots & \boxed{0} & \boxed{M_{r-1}} & \boxed{0} \\ \boxed{0} & \ldots & \boxed{0} & \boxed{0} & \boxed{M_r} \end{pmatrix}$$

(matrice diagonale de blocs), où, pour tout $1 \leq i \leq r$*,*

$$M_i = \begin{pmatrix} \boxed{N_{i,1}} & \boxed{0} & \boxed{0} & \dots & \boxed{0} \\ \boxed{0} & \boxed{N_{i,2}} & \boxed{0} & \dots & \boxed{0} \\ \vdots & & \ddots & \ddots & \vdots \\ \boxed{0} & \dots & \boxed{0} & \boxed{N_{i,s_i-1}} & \boxed{0} \\ \boxed{0} & \dots & \boxed{0} & \boxed{0} & \boxed{N_{i,s_i}} \end{pmatrix}$$

(matrice diagonale de blocs), où, pour tout $1 \leq j \leq s_i$*,*

$$N_{i,j} = \begin{pmatrix} a_i & 1 & 0 & 0 & \dots & 0 \\ 0 & a_i & 1 & 0 & \dots & 0 \\ 0 & 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 1 & 0 \\ 0 & 0 & \dots & 0 & a_i & 1 \\ 0 & 0 & \dots & 0 & 0 & a_i \end{pmatrix},$$

qui est d'ordre $e_{i,j}$*. On appelle les* $N_{i,j}$ *les* blocs de Jordan *(pour la valeur propre* a_i*).*

Remarque 8.9. *Explicitement, la base* S *est la suivante :*

$$\begin{array}{ccccc} (\varphi - a_1 \cdot \mathrm{id})^{e_{1,1}-1}(x_{1,1}), & (\varphi - a_1 \cdot \mathrm{id})^{e_{1,1}-2}(x_{1,1}), & \dots & (\varphi - a_1 \cdot \mathrm{id})(x_{1,1}), & x_{1,1}, \\ (\varphi - a_1 \cdot \mathrm{id})^{e_{1,2}-1}(x_{1,2}), & (\varphi - a_1 \cdot \mathrm{id})^{e_{1,2}-2}(x_{1,2}), & \dots & (\varphi - a_1 \cdot \mathrm{id})(x_{1,2}), & x_{1,2}, \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ (\varphi - a_1 \cdot \mathrm{id})^{e_{1,s_1}-1}(x_{1,s_1}), & (\varphi - a_1 \cdot \mathrm{id})^{e_{1,s_1}-2}(x_{1,s_1}), & \dots & (\varphi - a_1 \cdot \mathrm{id})(x_{1,s_1}), & x_{1,s_1}, \\ (\varphi - a_2 \cdot \mathrm{id})^{e_{2,1}-1}(x_{2,1}), & (\varphi - a_2 \cdot \mathrm{id})^{e_{2,1}-2}(x_{2,1}), & \dots & (\varphi - a_2 \cdot \mathrm{id})(x_{2,1}), & x_{2,1}, \\ (\varphi - a_2 \cdot \mathrm{id})^{e_{2,2}-1}(x_{2,2}), & (\varphi - a_2 \cdot \mathrm{id})^{e_{2,2}-2}(x_{2,2}), & \dots & (\varphi - a_2 \cdot \mathrm{id})(x_{2,2}), & x_{2,2}, \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ (\varphi - a_2 \cdot \mathrm{id})^{e_{2,s_2}-1}(x_{2,s_2}), & (\varphi - a_2 \cdot \mathrm{id})^{e_{2,s_2}-2}(x_{2,s_2}), & \dots & (\varphi - a_2 \cdot \mathrm{id})(x_{2,s_2}), & x_{2,s_2}, \\ (\varphi - a_3 \cdot \mathrm{id})^{e_{3,1}-1}(x_{3,1}), & (\varphi - a_3 \cdot \mathrm{id})^{e_{3,1}-2}(x_{3,1}), & \dots & (\varphi - a_3 \cdot \mathrm{id})(x_{3,1}), & x_{3,1}, \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ (\varphi - a_r \cdot \mathrm{id})^{e_{r,s_r}-1}(x_{r,s_r}), & (\varphi - a_r \cdot \mathrm{id})^{e_{r,s_r}-2}(x_{r,s_r}), & \dots & (\varphi - a_r \cdot \mathrm{id})(x_{r,s_r}), & x_{r,s_r} \end{array}.$$

Noter que la réduction de Jordan n'est pas unique en général (on peut, par exemple, permuter les blocs). Alors, pour être précis, on devrait parler plutôt d'une réduction de Jordan, ce que nous allons faire parfois. Si S est une base telle que $M_{S,S}(\varphi)$ a la forme du théorème, on dira que $M_{S,S}(\varphi)$ est *la/une réduction de Jordan* ou bien qu'elle a *la/une forme de Jordan*.

Pour appliquer le théorème 8.8 aux matrices, regarder (encore une fois) la remarque 7.4.

Exemple 8.10. *(a) La/une réduction de Jordan de la matrice* $\begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 5 \end{pmatrix}$ *obtenue de la décomposition spectrale dans l'exemple 7.5(a) est* $\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 5 \end{pmatrix}$ *pour la raison suivante.*

La matrice satisfait aux hypothèses du théorème 8.8, donc elle possède une réduction de Jordan. Comme elle n'est pas diagonalisable, il ne peut y avoir qu'un seul bloc avec 1 *sur la diagonale, mais le polynôme caractéristique montre que* 1 *doit apparaître deux fois sur la diagonale. Donc, il n'y a pas d'autre possibilité.*

(b) Considérons la matrice $M := \begin{pmatrix} 1 & 1 & 0 \\ -1 & 3 & 0 \\ -1 & 1 & 2 \end{pmatrix}$ *à coefficients dans* $\mathbb{R}$.

Un calcul montre que $\mathrm{car}_M(X) = (X-2)^3$. *Alors,* $r = 1$ *dans les notation du théorème 8.8 et, alors, la réduction de Jordan doit être parmi les trois matrices suivantes :*

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

On calcule facilement que $\mathrm{mipo}_M(X) = (X-2)^2$. *De ce fait nous pouvons déjà déduire que la réduction de Jordan est* $\begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$.

La question devient plus désagréable si on nous demande de calculer une matrice C *telle que* $C^{-1}MC = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$. *Mais cela n'est pas aussi difficile que ça. Nous suivons l'algorithme de la réduction de Jordan :*

- *On a* $M - 2\mathrm{id}_3 = \begin{pmatrix} -1 & 1 & 0 \\ -1 & 1 & 0 \\ -1 & 1 & 0 \end{pmatrix}$.
- *Alors,* $\ker(M - 2\mathrm{id}_3) = \langle \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \rangle$.
- *Nous savons que* $\mathrm{mipo}_M(X) = (X-2)^2$ *(ce qu'on revérifie facilement :* $(M - 2 \cdot \mathrm{id}_3)^2 = 0_3$*). Selon l'algorithme, nous choisissons*

$$x_1 \in \ker((M - 2\mathrm{id}_3)^2) \setminus \ker(M - 2\mathrm{id}_3) = \mathbb{R}^3 \setminus \langle \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \rangle,$$

par exemple $x_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$.

- *Nous commençons à écrire notre base* S. *Le premier vecteur de la base est, selon l'algorithme,*

$$v_1 := (M - 2\mathrm{id}_3)x_1 = \begin{pmatrix} -1 & 1 & 0 \\ -1 & 1 & 0 \\ -1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \\ -1 \end{pmatrix}$$

  et le deuxième est juste $v_2 := x_1$.

- *Dans la deuxième étape, nous devons choisir un vecteur* $y \in \ker(M - 2\mathrm{id}_3) \setminus \langle v_1, v_2 \rangle = \langle \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \rangle \setminus \langle \begin{pmatrix} -1 \\ -1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \rangle$.

  On choisit $y = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ *et on pose directement* $v_3 = y$.

- *Il suffit d'écrire les vecteurs* v_1, v_2, v_3 *dans les colonnes d'une matrice :*

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

Le théorème 8.8 nous dit que

$$C^{-1}MC = \begin{pmatrix} 0 & -1 & 0 \\ 1 & -1 & 0 \\ 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ -1 & 3 & 0 \\ -1 & 1 & 2 \end{pmatrix} \begin{pmatrix} -1 & 1 & 0 \\ -1 & 0 & 0 \\ -1 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix},$$

ce qu'on peut vérifier pour se convaincre des calculs.

Remarque 8.11. *Dans des exemples et des exercices vous avez vu/voyez que la connaissance du polynôme minimal nous donne déjà beaucoup de renseignements sur la réduction de Jordan.*

Plus précisement, si a *est une valeur propre de* φ *et* $(X - a)^e$ *est la plus grande puissance de* $X - a$ *qui divise le polynôme minimal* $\mathrm{mipo}_\varphi(X)$, *alors la taille du plus grand bloc de Jordan avec* a *sur la diagonale est* e.

En général, on n'obtient pas toute la réduction de Jordan de cette manière ; si, par exemple, $(X - a)^{e+2}$ *est la plus grande puissance de* $X - a$ *qui divise* $\mathrm{car}_\varphi(X)$, *alors, on a deux possibilités : (1) il y a deux blocs de Jordan pour la valeur propre* a *de taille* e *et* 2 *; ou (2) il y a trois blocs de Jordan pour* a *de taille* e, 1 *et* 1.

Exemple 8.12. *Nous faisons encore un exemple. Soit*

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

Son polynôme caractéristique est

$$\mathrm{car}_M(X) = X^6 - 18X^5 + 135X^4 - 540X^3 + 1215X^2 - 1458X + 729 = (X-3)^6.$$

Calculons d'abord

$$M_3 := M + 3\mathrm{id} = \begin{pmatrix} -5 & -1 & -5 & -3 & 6 & -4 \\ -1 & -1 & -1 & -1 & 1 & 0 \\ 2 & 1 & 1 & 2 & -2 & 1 \\ 4 & 2 & 4 & 3 & -5 & 2 \\ 0 & 1 & -1 & 1 & 0 & -1 \\ 1 & -1 & 1 & 0 & -1 & 2 \end{pmatrix},$$

puis

$$M_3^2 = \begin{pmatrix} 0 & 5 & -1 & 3 & -2 & -5 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & -1 & 0 & 1 \\ 0 & -3 & 1 & -2 & 1 & 3 \\ 0 & 1 & 1 & 0 & -1 & -1 \\ 0 & -2 & 0 & -1 & 1 & 2 \end{pmatrix}, \quad M_3^3 = \begin{pmatrix} 0 & 3 & 3 & 0 & -3 & -3 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & -1 & 0 & 1 & 1 \\ 0 & -2 & -2 & 0 & 2 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & -1 & 0 & 1 & 1 \end{pmatrix}, \quad M_3^4 = 0.$$

Nous avons donc

$$\mathrm{mipo}_M(X) = (X-3)^4$$

et

$$V_4 = \ker(M_3^4) = \mathbb{R}^6 \supsetneq V_3 \supsetneq V_2 \supsetneq V_1 = E_M(3) \supsetneq 0.$$

Nous calculons d'abord

$$V_3 = \ker(M_3^3) = \ker \begin{pmatrix} 0 & 1 & 1 & 0 & -1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} = \langle \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \rangle.$$

En fait, il n'est pas nécessaire pour l'algorithme de donner toute une base pour V_3, il suffit de trouver un vecteur qui n'appartient pas au noyau. C'est très facile. On prendra $x_1 = \begin{pmatrix} 0\\1\\0\\0\\0\\0 \end{pmatrix}$ *et on calcule*

$$x_1 = \begin{pmatrix} 0\\1\\0\\0\\0\\0 \end{pmatrix}, M_3x_1 = \begin{pmatrix} -1\\-1\\1\\2\\1\\-1 \end{pmatrix}, M_3^2x_1 = \begin{pmatrix} 5\\0\\-1\\-3\\1\\-2 \end{pmatrix}, M_3^3x_1 = \begin{pmatrix} 3\\0\\-1\\-2\\0\\-1 \end{pmatrix}.$$

Nous avons alors déjà un bloc de Jordan de taille 4*. Il y a donc soit encore un bloc de taille* 2*, ou deux blocs de taille* 1*. Nous calculons maintenant*

$$V_2 = \ker(M_3^2) = \ker \begin{pmatrix} 0 & 1 & 1 & 0 & -1 & -1\\ 0 & 0 & -6 & 3 & 3 & 0\\ 0 & 0 & 2 & -1 & -1 & 0\\ 0 & 0 & 4 & -2 & -2 & 0\\ 0 & 0 & 2 & -1 & -1 & 0\\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} = \ker \begin{pmatrix} 0 & 1 & 1 & 0 & -1 & -1\\ 0 & 0 & 1 & -1/2 & -1/2 & 0\\ 0 & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$= \ker \begin{pmatrix} 0 & 1 & 0 & 1/2 & -1/2 & -1\\ 0 & 0 & 1 & -1/2 & -1/2 & 0\\ 0 & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} = \langle \begin{pmatrix} 1\\0\\0\\0\\0\\0 \end{pmatrix}, \begin{pmatrix} 0\\1\\0\\0\\0\\1 \end{pmatrix}, \begin{pmatrix} 0\\1/2\\1/2\\0\\1\\0 \end{pmatrix}, \begin{pmatrix} 0\\-1/2\\1/2\\1\\0\\0 \end{pmatrix} \rangle.$$

Finalement, nous calculons l'espace propre pour la valeur propre 3 *:*

$$V_1 = \ker(M_3) = \ker \begin{pmatrix} -5 & -1 & -5 & -3 & 6 & -4\\ -1 & -1 & -1 & -1 & 1 & 0\\ 2 & 1 & 1 & 2 & -2 & 1\\ 4 & 2 & 4 & 3 & -5 & 2\\ 0 & 1 & -1 & 1 & 0 & -1\\ 1 & -1 & 1 & 0 & -1 & 2 \end{pmatrix} = \ker \begin{pmatrix} 1 & -1 & 1 & 0 & -1 & 2\\ 0 & 1 & -1 & 1 & 0 & -1\\ 0 & -6 & 0 & -3 & 1 & 6\\ 0 & -2 & 0 & -1 & 0 & 2\\ 0 & 3 & -1 & 2 & 0 & -3\\ 0 & 6 & 0 & 3 & -1 & -6 \end{pmatrix}$$

$$= \ker \begin{pmatrix} 1 & 0 & 0 & 1 & -1 & 1\\ 0 & 1 & -1 & 1 & 0 & -1\\ 0 & 0 & -6 & 3 & 1 & 0\\ 0 & 0 & -2 & 1 & 0 & 0\\ 0 & 0 & 2 & -1 & 0 & 0\\ 0 & 0 & 6 & -3 & -1 & 0 \end{pmatrix} = \ker \begin{pmatrix} 1 & 0 & 0 & 1 & -1 & 1\\ 0 & 1 & -1 & 1 & 0 & -1\\ 0 & 0 & 1 & -1/2 & 0 & 0\\ 0 & 0 & -6 & 3 & 1 & 0\\ 0 & 0 & 2 & -1 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$= \ker \begin{pmatrix} 1 & 0 & 0 & 1 & -1 & 1\\ 0 & 1 & 0 & 1/2 & 0 & -1\\ 0 & 0 & 1 & -1/2 & 0 & 0\\ 0 & 0 & 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} = \ker \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 1\\ 0 & 1 & 0 & 1/2 & 0 & -1\\ 0 & 0 & 1 & -1/2 & 0 & 0\\ 0 & 0 & 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} = \langle \begin{pmatrix} -1\\1\\0\\0\\0\\1 \end{pmatrix}, \begin{pmatrix} -1\\-1/2\\1/2\\1\\0\\0 \end{pmatrix} \rangle.$$

Il y a donc 2 *vecteurs propres, donc deux blocs en total. Alors le deuxième bloc est de taille* 2*. Nous devons trouver un vecteur dans* V_2 *qui n'appartient pas à* $V_1 + \langle x_1, M_3x_1, M_3^2x_1, M_3^3x_1 \rangle$*, donc un* *élément dans*

$$\left\langle \begin{pmatrix}1\\0\\0\\0\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\\0\\0\\1\end{pmatrix}, \begin{pmatrix}0\\1/2\\1/2\\0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\-1/2\\1/2\\1\\0\\0\end{pmatrix} \right\rangle \setminus \left\langle \begin{pmatrix}-1\\1\\0\\0\\0\\1\end{pmatrix}, \begin{pmatrix}-1\\-1/2\\1/2\\1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\\0\\0\\0\end{pmatrix}, \begin{pmatrix}-1\\-1\\1\\2\\1\\-1\end{pmatrix}, \begin{pmatrix}5\\0\\-1\\-3\\1\\-2\end{pmatrix}, \begin{pmatrix}3\\0\\-1\\-2\\0\\-1\end{pmatrix} \right\rangle.$$

Pour trouver un tel élément, on teste l'un après l'autre si les vecteurs dans la base de V_2 sont linéairement indépendants de l'espace à droite. Nous avons de la chance que déjà $x_2 := \left(\begin{smallmatrix}1\\0\\0\\0\\0\\0\end{smallmatrix}\right)$ marche (comme on le voit en faisant un calcul standard). Nous calculons donc

$$x_2 = \begin{pmatrix}1\\0\\0\\0\\0\\0\end{pmatrix}, \quad M_3 x_2 = \begin{pmatrix}-5\\-1\\2\\4\\0\\1\end{pmatrix}.$$

Nous pouvons maintenant écire la matrice

$$C := \begin{pmatrix} 3 & 5 & -1 & 0 & -5 & 1 \\ 0 & 0 & -1 & 1 & -1 & 0 \\ -1 & -1 & 1 & 0 & 2 & 0 \\ -2 & -3 & 2 & 0 & 4 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ -1 & -2 & -1 & 0 & 1 & 0 \end{pmatrix}$$

et un calcul vérifie

$$C^{-1}MC = \begin{pmatrix} 3 & 1 & 0 & 0 & 0 & 0 \\ 0 & 3 & 1 & 0 & 0 & 0 \\ 0 & 0 & 3 & 1 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 0 & 3 \end{pmatrix}.$$

Remarque 8.13. *Voici quelques remarques faciles à montrer qui sont parfois utiles pour des calculs. Supposons* $\mathrm{mipo}_M(X) = (X-a)^e$.

(a) La taille du plus grand bloc de Jordan est e.

(b) Chaque bloc de Jordan contient un espace propre de dimension 1 *pour la valeur propre* a.

(c) Le nombre de blocs de Jordan est égal à la dimension de l'espace propre pour la valeur propre a.

9 Espaces hermitiens

Objectifs :

- Connaître la définition d'espaces euclidiens et hermitiens ;
- connaître les propriétés fondamentales des espaces euclidiens et hermitiens ;
- savoir calculer des bases orthonormales par la méthode de Gram-Schmidt ;
- connaître des exemples et savoir démontrer des propriétés simples.

Nous commençons par une motivation de quelques-uns des sujets qui vont suivre.
Soit $M \in \mathrm{Mat}_{n\times n}(K)$ une matrice. Considérons l'application :

$$\langle\ ,\ \rangle_M : K^n \times K^n \to K, \quad \langle \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}, \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} \rangle_M := (a_1 \ \ a_2 \ \ \cdots \ \ a_n) M \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}.$$

Nous avons donc l'égalité

$$\langle x, y\rangle_M = x^{\mathrm{tr}} M y.$$

Si M est l'identité, alors on n'écrit pas l'indice et

$$\langle \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}, \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} \rangle = (a_1 \ \ a_2 \ \ \cdots \ \ a_n) \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = \sum_{i=1}^n a_i b_i.$$

C'est le produit scalaire canonique bien connu. Cela donne en plus (si $K = \mathbb{R}$)

$$\langle \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}, \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \rangle = a_1^2 + a_2^2 + \cdots + a_n^2 > 0$$

pour tout $\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \neq 0$.
Faisons un autre exemple : $M = \left(\begin{smallmatrix} 1 & 2 \\ 3 & 4 \end{smallmatrix}\right)$. Alors

$$\langle \left(\begin{smallmatrix} a_1 \\ a_2 \end{smallmatrix}\right), \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \rangle_M = (a_1 \ \ a_2) \left(\begin{smallmatrix} 1 & 2 \\ 3 & 4 \end{smallmatrix}\right) \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = a_1 b_1 + 2 a_1 b_2 + 3 a_2 b_1 + 4 a_2 b_2.$$

En général, nous observons immédiatement les propriétés suivantes :

(a) *Linéarité dans la première variable :* Pour tout $y \in K^n$, l'application

$$\langle \cdot, y\rangle_M : K^n \to K, \quad x \mapsto \langle x, y\rangle_M$$

est K-linéaire, c'est-à-dire, pour tout $x_1, x_2 \in K^n$ et tout $a \in K$, on a

$$\langle x_1 + a x_2, y\rangle_M = \langle x_1, y\rangle_M + a\langle x_2, y\rangle_M.$$

(b) *Linéarité dans la deuxième variable :* Pour tout $x \in K^n$, l'application

$$\langle x, \cdot\rangle_M : K^n \to K, \quad y \mapsto \langle x, y\rangle_M$$

est K-linéaire, c'est-à-dire, pour tout $y_1, y_2 \in K^n$ et tout $a \in K$, on a

$$\langle x, y_1 + ay_2\rangle_M = \langle x, y_1\rangle_M + a\langle x, y_2\rangle_M.$$

Question : Quand avons-nous que $\langle\ ,\ \rangle_M$ est symétrique, c'est-à-dire, $\langle x, y\rangle_M = \langle y, x\rangle_M$ pour tout $x, y \in K^n$?
Pour voir la réponse à cette question, prenons $x = e_i$ comme le i-ème vecteur canonique et $y = e_j$. Alors

$$\langle e_i, e_j\rangle_M = e_i^{\mathrm{tr}} M e_j = e_i(j\text{-ème colonne de } M) = i\text{-ème coeff. de } (j\text{-ème colonne de } M) = m_{i,j}.$$

Donc, $\langle e_i, e_j\rangle_M = \langle e_j, e_i\rangle_M$ implique $m_{i,j} = m_{j,i}$ pour tout $1 \le i, j \le n$, en d'autres mots, la matrice M est symétrique $M = M^{\mathrm{tr}}$.
Inversement, partons d'une matrice symétrique $M = M^{\mathrm{tr}}$. Nous faisons un petit calcul formel et assez élégant :

$$\langle x, y\rangle_M = x^{\mathrm{tr}} M y = (x^{\mathrm{tr}} M y)^{\mathrm{tr}} = y^{\mathrm{tr}} M^{\mathrm{tr}} (x^{\mathrm{tr}})^{\mathrm{tr}} = y^{\mathrm{tr}} M x = \langle y, x\rangle_M,$$

où nous avons utilisé que $x^{\mathrm{tr}} M y$ est une matrice de taille 1, donc égalé à sa transposée, ainsi que le lemme suivant :

Lemme 9.1. *Soient $M \in \mathrm{Mat}_{n,m}(K)$ et $N \in \mathrm{Mat}_{m,\ell}(K)$ des matrices. Alors*

$$(M \cdot N)^{\mathrm{tr}} = N^{\mathrm{tr}} \cdot M^{\mathrm{tr}}.$$

Démonstration. Exercice. □

Nous avons donc obtenu l'équivalence :

$\langle\ ,\ \rangle_M$ est symétrique $\Longleftrightarrow$ M est symétrique : $M = M^{\mathrm{tr}}$.

Question : Pour $K = \mathbb{R}$, quand avons-nous $\langle x, x\rangle_M \ge 0$ pour tout $x \in \mathbb{R}^n$?
Nous avons vu que c'est le cas si M est l'identité et $\langle\ ,\ \rangle$ est donc le produit scalaire canonique. Nous allons revoir cette question plus tard.
Pour l'instant, passons à $K = \mathbb{C}$. Nous notons $\overline{z} = x - iy$ le conjugué complexe de $z = x + iy \in \mathbb{C}$ avec $x = \mathrm{Re}(z)$ et $y = \mathrm{Im}(z)$.
Pour des nombres complexes, il n'est pas vrai que $\sum_{i=1}^n z_i^2$ est plus grand ou égal à 0, en fait, cela n'est même pas une question admise car z_i^2 n'est pas réel en général, donc demander si c'est plus grand que zéro n'a pas de sense. Par contre, la valeur absolue $z_i\overline{z_i} = |z_i|^2$ est toujours réelle et non-négatif. Donc il est utile de modifier la définition dans le cas $K = \mathbb{C}$:

$$\langle\ ,\ \rangle_M : \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}, \quad \langle x, y\rangle_M := x^{\mathrm{tr}} M \overline{y}$$

où $\overline{y}$ est le vecteur obtenu en appliquant la conjugaison complexe à tout coefficient. Noter que cette définition revient à la définition en haut si $K = \mathbb{R}$ car la conjugaison complexe n'a pas d'effet sur les nombres réels.
Avec M l'identité, cela donne

$$\langle \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}, \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} \rangle = (a_1 \ \ a_2 \ \ \cdots \ \ a_n) \begin{pmatrix} \overline{b_1} \\ \overline{b_2} \\ \vdots \\ \overline{b_n} \end{pmatrix} = \sum_{i=1}^{n} a_i \overline{b_i}.$$

C'est encore le produit scalaire canonique bien connu. On obtient en plus

$$\langle \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}, \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \rangle = |a_1|^2 + |a_2|^2 + \cdots + |a_n|^2 > 0$$

pour tout $\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \neq 0$.
Regardons encore les propriétés suivantes :

(a) *Linéarité dans la première variable :* Inchangé !

(b) *Sesqui-linéarité dans la deuxième variable :* Pour tout $x \in \mathbb{C}^n$, l'application

$$\langle x, \cdot \rangle_M : K^n \to K, \quad y \mapsto \langle x, y \rangle_M$$

est sesqui-linéaire, c'est-à-dire, pour tout $y_1, y_2 \in K^n$ et tout $a \in K$, on a

$$\langle x, y_1 + a y_2 \rangle_M = \langle x, y_1 \rangle_M + \overline{a} \langle x, y_2 \rangle_M.$$

Par les même calculs que ci-dessus, nous obtenons l'équivalence :

$$\langle x, y \rangle_M = \overline{\langle y, x \rangle_M} \text{ pour tout } x, y \in \mathbb{C}^n \iff M = \overline{M^{\mathrm{tr}}}.$$

Une matrice M telle que $M = \overline{M^{\mathrm{tr}}}$ est appelée *hermitienne*.
Pour la suite, de cette section nous posons $K = \mathbb{R}$ ou $K = \mathbb{C}$.

Définition 9.2. *Soit V un K-espace vectoriel. Une application*

$$\langle \cdot, \cdot \rangle : V \times V \to K, \quad (v, w) \mapsto \langle v, w \rangle$$

est appelée forme hermitienne *si pour tout $v, v_1, v_2, w, w_1, w_2 \in V$ et pour tout $a, b \in K$ on a*

- $\langle a v_1 + v_2, w \rangle = a \langle v_1, w \rangle + \langle v_2, w \rangle$ *(linéarité dans la première variable),*
- $\langle v, b w_1 + w_2 \rangle = \overline{b} \langle v, w_1 \rangle + \langle v, w_2 \rangle$ (sesquilinéarité *dans la deuxième variable) et*
- $\langle v, w \rangle = \overline{\langle w, v \rangle}$.

Une forme hermitienne $\langle \cdot, \cdot \rangle$ est dite positive *si*

- $\forall\, v \in V : \langle v, v\rangle \geq 0$. *(Noter $\langle v, v\rangle = \overline{\langle v, v\rangle}$, donc $\langle v, v\rangle \in \mathbb{R}$.)*

Elle est dite définie positive *si*

- $\forall\, 0 \neq v \in V : \langle v, v\rangle > 0$.

Une forme hermitienne définie positive s'appelle aussi produit scalaire.
On appelle espace hermitien *tout pair* $(V, \langle\cdot,\cdot\rangle)$ *où* $\langle\cdot,\cdot\rangle$ *est une forme hermitienne définie positive.*

Noter que le deuxième point dans la définition est superflu car on a

$$\langle v, bw_1 + w_2\rangle = \overline{\langle bw_1 + w_2, v\rangle} = \overline{b\langle w_1, v\rangle + \langle w_2, v\rangle} = \overline{b}\ \overline{\langle w_1, v\rangle} + \overline{\langle w_2, v\rangle} = \overline{b}\langle v, w_1\rangle + \langle v, w_2\rangle.$$

Remarque 9.3. *Noter que pour $K = \mathbb{R}$ les deux dernières conditions de la définition d'une forme hermitienne deviennent*

- $\langle v, bw_1 + w_2\rangle = b\langle v, w_1\rangle + \langle v, w_2\rangle$ *(linéarité dans la deuxième variable) et*
- $\forall\, v \in V\ \forall\, w \in W : \langle v, w\rangle = \langle w, v\rangle$.

On parle d'une forme bilinéaire symétrique.
Si $K = \mathbb{R}$, dans la littérature on utlise plutôt le nom espace euclidien *au lieu d'espace hermitien (qui est souvent reservé pour $K = \mathbb{C}$). Ici, pour simplifier la terminologie, nous allons toujours parler d'espaces hermitiens, même si $K = \mathbb{R}$.*

Nous avons déjà vu les produits scalaires canonique ci-dessus pour $\mathbb{R}^n$ et $\mathbb{C}^n$. On peut faire des définitions similaires aussi dans des espaces de fonctions (de dimension infinie) :

Exemple 9.4. *(a) Le produit scalaire canonique $\langle\ ,\ \rangle_M$ pour M l'identité est en effet un produit scalaire si $K = \mathbb{R}$ ou $K = \mathbb{C}$.*

(b) Soit $\mathcal{C} = \{f : [0,1] \to \mathbb{R} \mid f$ est continu *$\}$ l'ensemble de toutes les fonctions continues de $[0,1]$ dans $\mathbb{R}$. C'est un $\mathbb{R}$-espace vectoriel pour $+$ et $\cdot$ définis point par point. L'application*

$$\langle\cdot,\cdot\rangle : \mathcal{C} \times \mathcal{C} \to \mathbb{R}, \quad \langle f, g\rangle = \int_0^1 f(x)g(x)dx$$

est une forme hermitienne définie positive.

(c) Soit $\mathcal{C} = \{f : [0,1] \to \mathbb{C} \mid f$ est continu *$\}$ l'ensemble de toutes les fonctions continues de $[0,1]$ dans $\mathbb{C}$. C'est un $\mathbb{C}$-espace vectoriel pour $+$ et $\cdot$ définis point par point. L'application*

$$\langle\cdot,\cdot\rangle : \mathcal{C} \times \mathcal{C} \to \mathbb{C}, \quad \langle f, g\rangle = \int_0^1 f(x)\overline{g(x)}dx$$

est une forme hermitienne définie positive.

Définition 9.5. *Soit $(V, \langle\cdot,\cdot\rangle)$ un K-espace hermitien.*
On dit que $v, w \in V$ sont orthogonaux $v \perp w$ *si $\langle v, w\rangle = 0$. Noter : $v \perp w \Leftrightarrow w \perp v$.*
Soit $W \leq V$ un sous-espace. On dit que $v \in V$ et W sont orthogonaux $v \perp W$ *si $v \perp w$ pour tout $w \in W$. Noter : $v \perp W \Leftrightarrow W \perp v$ (avec les définitions évidentes).*

Soit $U \leq V$ encore un sous-espace. On dit que U et W sont orthogonaux $U \perp W$ *si $U \perp w$ pour tout $w \in W$. Noter : $U \perp W \Leftrightarrow W \perp U$.*
Le complément orthogonal de W *est défini comme*

$$W^{\perp} = \{v \in V \mid v \perp W\}.$$

La norme (« longueur ») *de $v \in V$ est définie comme $|v| := \sqrt{\langle v, v\rangle}$ et $|v-w|$ est dit la* distance *entre v et w.*

Proposition 9.6. *Soit $(V, \langle\cdot,\cdot\rangle)$ un K-espace hermitien.*

(a) Pour tout $v \in V$ on a $|v| \geq 0$ et $|v| = 0 \Leftrightarrow v = 0$.

(b) Pour tout $v \in V$ et tout $a \in K$ on a : $\underbrace{|a \cdot v|}_{|\cdot| \text{ dans } V} = \underbrace{|a|}_{|\cdot| \text{ dans } K} \cdot \underbrace{|v|}_{|\cdot| \text{ dans } V}$.

*(c) Pour tout $v, w \in V$ on a $\underbrace{|\langle v, w\rangle|}_{|\cdot| \text{ dans } K} \leq \underbrace{|v|}_{|\cdot| \text{ dans } V} \cdot \underbrace{|w|}_{|\cdot| \text{ dans } V}$ (*inégalité de Cauchy-Schwarz*).*

*(d) Pour tout $v, w \in V$ on a $\underbrace{|v + w|}_{|\cdot| \text{ dans } V} \leq \underbrace{|v|}_{|\cdot| \text{ dans } V} + \underbrace{|w|}_{|\cdot| \text{ dans } V}$ (*inégalité triangulaire*).*

Démonstration. (a) Définition.
(b) $|a \cdot v|^2 = \langle a \cdot v, a \cdot v\rangle = a \cdot \overline{a} \cdot \langle v, v\rangle = |a|^2 \cdot |v|^2$.
(c) 1er cas : $w = 0$. Alors, $\langle v, w\rangle = \langle v, 0 \cdot w\rangle = 0\langle v, w\rangle = 0$, donc $|\langle v, w\rangle| = 0 = |v| \cdot |w|$.
2ème cas : $w \neq 0$. Soit $c := \frac{\langle v,w\rangle}{|w|^2}$. Alors

$$\begin{aligned} 0 &\leq |w|^2 \cdot \langle v - c \cdot w, v - c \cdot w\rangle \\ &= |w|^2 \cdot \langle v, v\rangle - |w|^2 \cdot c \cdot \langle w, v\rangle - |w|^2 \cdot \overline{c} \cdot \langle v, w\rangle + |w|^2 \cdot c \cdot \overline{c} \cdot \langle w, w\rangle \\ &= |w|^2 \cdot |v|^2 - \underbrace{\langle v, w\rangle \cdot \langle w, v\rangle}_{=|\langle v,w\rangle|} - \underbrace{\overline{\langle v, w\rangle} \cdot \langle v, w\rangle + \langle v, w\rangle \cdot \overline{\langle v, w\rangle}}_{=0}. \end{aligned}$$

(d)

$$\begin{aligned} |v + w|^2 &= \langle v + w, v + w\rangle \\ &= \langle v, v\rangle + \langle v, w\rangle + \langle w, v\rangle + \langle w, w\rangle \\ &= |v|^2 + |w|^2 + \langle v, w\rangle + \overline{\langle v, w\rangle} \\ &= |v|^2 + |w|^2 + 2 \cdot \mathrm{Re}(\langle v, w\rangle) \\ &\leq |v|^2 + |w|^2 + 2 \cdot |\langle v, w\rangle| \\ &\leq |v|^2 + |w|^2 + 2 \cdot |v| \cdot |w| \\ &= (|v| + |w|)^2. \end{aligned}$$

□

Proposition 9.7 (Pythagore). *Si $v \perp w$, alors $|v + w|^2 = |v|^2 + |w|^2$.*

Démonstration. $|v + w|^2 = \langle v + w, v + w\rangle = \langle v, v\rangle + \langle w, w\rangle = |v|^2 + |w|^2$. □

Noter que toute forme hermitienne définie positive est non-dégénérée : si $\langle v, w\rangle = 0$ pour tout $w \in W$, alors en particulier $\langle v, v\rangle = |v|^2 = 0$, donc $v = 0$. Le même argument montre aussi que $w = 0$ si $\langle v, w\rangle = 0$ pour tout $v \in V$.

Définition 9.8. *Soient* $(V, \langle\cdot,\cdot\rangle)$ *un* K*-espace hermitien et* $S = \{s_i \mid i \in I\}$ *(avec un ensemble* I*, par exemple,* $S = \{s_1, \dots, s_n\}$ *si* $I = \{1, 2, \dots, n\}$*).*
On dit que S *est un* système orthogonal *si*

- $\langle s_i, s_i\rangle > 0$ *pour tout* $i \in I$ *et*
- $\langle s_i, s_j\rangle = 0$ *pour tout* $i, j \in I$*,* $i \neq j$*.*

On dit qu'un système orthogonal S *est un* orthonormal *si* $\langle s_i, s_j\rangle = \delta_{i,j}$ *pour tout* $i, j \in I$*.*
Si S *est une base de* V *qui est un système orthogonal/orthonormal, on parle d'une* base orthogonale/orthonormale.

Exemple 9.9. *La base canonique de* $\mathbb{R}^n$ *(ou de* $\mathbb{C}^n$*) est une base orthonormale pour le produit scalaire canonique de l'exemple 9.4.*

Proposition 9.10 (Orthonormalisation de Gram-Schmidt)**.** *Soient* $(V, \langle\cdot,\cdot\rangle)$ *un* K*-espace hermitien et des vecteurs* K*-linéairement indépendants* $s_1, s_2, \dots, s_n \in V$*.*
La méthode de Gram-Schmidt *(voir la démonstration) calcule des vecteurs* $t_1, t_2, \dots, t_n \in V$ *tels que*

- $\langle t_i, t_j\rangle = \delta_{i,j}$ *pour tout* $1 \le i, j \le n$ *et*
- $\langle s_1, s_2, \dots, s_r\rangle = \langle t_1, t_2, \dots, t_r\rangle$ *pour tout* $1 \le r \le n$ *(les sous-espaces de* V *engendrés par* $s_1, s_2, \dots, s_r$ *et par* $t_1, t_2, \dots, t_r$ *sont égaux).*

Démonstration. Nous présentons la méthode de Gram-Schmidt.
C'est une récurrence pour $r = 1, 2, \dots, n$; donc il y a n étapes.
$\underline{r = 1.}$ $t_1 := \frac{s_1}{|s_1|}$.
$\underline{r \Rightarrow r + 1.}$ Par hypothèse de récurrence nous avons déjà $t_1, \dots, t_r$ tels que $\langle t_i, t_j\rangle = \delta_{i,j}$ pour tout $1 \le i, j \le r$ et $\langle s_1, s_2, \dots, s_r\rangle = \langle t_1, t_2, \dots, t_r\rangle$.
Nous devons trouver t_{r+1}. D'abord on définit

$$w_{r+1} := s_{r+1} - \sum_{i=1}^{r} \langle s_{r+1}, t_i\rangle t_i.$$

Ce vecteur satisfait pour tout $1 \le j \le r$

$$\begin{aligned}
\langle w_{r+1}, t_j\rangle &= \langle s_{r+1} - \sum_{i=1}^{r} \langle s_{r+1}, t_i\rangle t_i, t_j\rangle \\
&= \langle s_{r+1}, t_j\rangle - \sum_{i=1}^{r} \langle \langle s_{r+1}, t_i\rangle t_i, t_j\rangle \\
&= \langle s_{r+1}, t_j\rangle - \langle \langle s_{r+1}, t_j\rangle t_j, t_j\rangle \\
&= \langle s_{r+1}, t_j\rangle - \langle s_{r+1}, t_j\rangle \cdot \langle t_j, t_j\rangle \\
&= 0
\end{aligned}$$

Comme $\langle s_1, s_2, \ldots, s_r \rangle = \langle t_1, t_2, \ldots, t_r \rangle$, on a $w_{r+1} \notin \langle t_1, t_2, \ldots, t_r \rangle$, donc, en particulier, $w_{r+1} \neq 0$. Cela nous permet de définir

$$t_{r+1} := \frac{w_{r+1}}{|w_{r+1}|}.$$

Ce vecteur satisfait clairement $\langle t_{r+1}, t_i \rangle = \delta_{r+1,i}$ pour tout $1 \leq i \leq r+1$ et $\langle s_1, s_2, \ldots, s_r, s_{r+1} \rangle = \langle t_1, t_2, \ldots, t_r, t_{r+1} \rangle$. $\square$

Exemple 9.11. *Nous appliquons la méthode de Gram-Schmidt aux vecteurs suivants :*

$$s_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}, s_2 = \begin{pmatrix} 1 \\ 3 \\ -2 \\ 3 \\ -2 \\ 5 \end{pmatrix}, s_3 = \begin{pmatrix} -1 \\ 5 \\ 2 \\ -3 \\ -6 \\ 3 \end{pmatrix}$$

sur $\mathbb{R}^6$ avec le produit scalaire canonique.

(1) Calculons la longeur de s_1 :

$$|s_1| = \sqrt{4} = 2.$$

Donc

$$t_1 = \frac{1}{2} s_1 = \begin{pmatrix} 1/2 \\ 1/2 \\ 0 \\ 1/2 \\ 0 \\ 1/2 \end{pmatrix}.$$

(2) Calculons maintenant

$$\langle s_2, t_1 \rangle = \langle \begin{pmatrix} 1 \\ 3 \\ -2 \\ 3 \\ -2 \\ 5 \end{pmatrix}, \begin{pmatrix} 1/2 \\ 1/2 \\ 0 \\ 1/2 \\ 0 \\ 1/2 \end{pmatrix} \rangle = 6.$$

Donc

$$w_2 := s_2 - \langle s_2, t_1 \rangle t_1 = \begin{pmatrix} 1 \\ 3 \\ -2 \\ 3 \\ -2 \\ 5 \end{pmatrix} - 6 \begin{pmatrix} 1/2 \\ 1/2 \\ 0 \\ 1/2 \\ 0 \\ 1/2 \end{pmatrix} = \begin{pmatrix} -2 \\ 0 \\ -2 \\ 0 \\ -2 \\ 2 \end{pmatrix}.$$

La longeur de w_2 est

$$|w_2| = \sqrt{16} = 4.$$

Donc

$$t_2 = \frac{1}{4} w_2 = \begin{pmatrix} -1/2 \\ 0 \\ -1/2 \\ 0 \\ -1/2 \\ 1/2 \end{pmatrix}.$$

(3) Calculons maintenant

$$\langle s_3, t_1 \rangle = \langle \begin{pmatrix} -1 \\ 5 \\ 2 \\ -3 \\ -6 \\ 3 \end{pmatrix}, \begin{pmatrix} 1/2 \\ 1/2 \\ 0 \\ 1/2 \\ 0 \\ 1/2 \end{pmatrix} \rangle = 2$$

et

$$\langle s_3, t_2\rangle = \langle \begin{pmatrix} -1\\ 5\\ 2\\ -3\\ -6\\ 3\end{pmatrix}, \begin{pmatrix} -1/2\\ 0\\ -1/2\\ 0\\ -1/2\\ 1/2\end{pmatrix}\rangle = 4.$$

Donc

$$w_3 := s_3 - \langle s_3, t_1\rangle t_1 - \langle s_3, t_2\rangle t_2 = \begin{pmatrix} -1\\ 5\\ 2\\ -3\\ -6\\ 3\end{pmatrix} - 2\begin{pmatrix} 1/2\\ 1/2\\ 0\\ 1/2\\ 0\\ 1/2\end{pmatrix} - 4\begin{pmatrix} -1/2\\ 0\\ -1/2\\ 0\\ -1/2\\ 1/2\end{pmatrix} = \begin{pmatrix} 0\\ 4\\ 4\\ -4\\ -4\\ 0\end{pmatrix}.$$

La longeur de w_3 est

$$|w_3| = \sqrt{64} = 8.$$

Donc

$$t_3 = \frac{1}{8} w_3 = \begin{pmatrix} 0\\ 1/2\\ 1/2\\ -1/2\\ -1/2\\ 0\end{pmatrix}.$$

Corollaire 9.12. *Soit $(V, \langle\cdot,\cdot\rangle)$ un K-espace hermitien de dimension finie (ou même dénombrable). Alors, V possède une K-base orthonormale.*

Démonstration. Conséquence directe de Gram-Schmidt 9.10. □

Corollaire 9.13. *Soit $(V, \langle\cdot,\cdot\rangle)$ un K-espace hermitien et $W \leq V$ un sous-espace de dimension finie. Soit $s_1, \dots, s_n \in W$ une K-base orthonormale de W (qui existe à cause du corollaire 9.12). Nous définissons*

$$\pi_W : V \to W, \quad v \mapsto \sum_{i=1}^{n} \langle v, s_i\rangle s_i.$$

Cette application est appelée la projection orthogonale sur W.

(a) *π_W est K-linéaire et satisfait $\pi_W \circ \pi_W = \pi_W$.*

(b) *$V = W \oplus W^{\perp}$.*

(c) *Pour tout $v \in V$, on a*

$$|\pi_W(v)|^2 = \sum_{i=1}^{n} |\langle v, s_i\rangle|^2 \leq |v|^2.$$

C'est l'inégalité de Bessel.

(d) *Pour tout $v \in V$, $\pi_W(v)$ peut être caracterisé comme l'unique $w \in W$ tel que $|v-w|$ est minimal. L'application π_W est donc indépendante du choix de la base.*

Démonstration. (a) Calculs simples.

(b) Soit $v \in V$. Nous écrivons $v = \pi_W(v) + (v - \pi_W(v))$. Nous avons clairement $\pi_W(v) \in W$. Montrons que $v - \pi_W(v) \in W^\perp$; pour cela il suffit de prouver que $\langle v - \pi_W(v), s_j\rangle = 0$ pour tout $1 \le j \le n$:

$$\langle v - \pi_W(v), s_j\rangle = \langle v, s_j\rangle - \langle \sum_{i=1}^n \langle v, s_i\rangle s_i, s_j\rangle = \langle v, s_j\rangle - \sum_{i=1}^n \langle v, s_i\rangle \cdot \langle s_i, s_j\rangle = \langle v, s_j\rangle - \langle v, s_j\rangle = 0.$$

Cela nous donne $V = W + W^\perp$, donc il suffit de montrer que la somme est directe. Soit $w \in W \cap W^\perp$. En particulier, $w \perp w$, c'est-à-dire, $\langle w, w\rangle = |w|^2 = 0$, donc $w = 0$.
(c) Nous venons de voir $\pi_W(v) \perp (v - \pi_W(v))$, donc par Pythagore 9.7 nous avons

$$|v|^2 = |\pi_W(v)|^2 + |v - \pi_W(v)|^2,$$

donc $|\pi_W(v)|^2 \le |v|^2$. C'est déjà l'inégalité. Calculons maintenant l'égalité :

$$|\pi_W(v)|^2 = \langle \pi_W(v), \pi_W(v)\rangle = \sum_{j=1}^n \sum_{k=1}^n \langle v, s_j\rangle \overline{\langle v, s_k\rangle} \langle s_j, s_k\rangle = \sum_{j=1}^n |\langle v, s_j\rangle|^2.$$

(d) Nous utilisons encore une fois Pythagore 9.7 pour obtenir pour $w \in W$

$$|v - w|^2 = |\underbrace{(v - \pi_W(v))}_{\in W^\perp} + \underbrace{(\pi_W(v) - w)}_{\in W}|^2 = \underbrace{|v - \pi_W(v)|^2}_{\text{indépendant de } w} + |\pi_W(v) - w|^2.$$

Donc $|v - w|$ est minimal si et seulement si $|\pi_W(v) - w| = 0$, donc si et seulement si $w = \pi_W(v)$. $\square$

10 Opérateurs adjoints, autoadjoints, normaux et isométries

Objectifs :

- Maîtriser les concepts d'opérateur adjoint, normal et autoadjoint ;
- maîtriser la notion d'isométrie, et les notions de matrice unitaire et orthogonale ;
- connaître les propriétés fondamentales des opérateurs normaux et autoadjoints et des isométries ;
- savoir décider si ces notions sont satisfaites ;
- connaître des exemples et savoir démontrer des propriétés simples.

On continue avec $K \in \{\mathbb{R}, \mathbb{C}\}$. Dans cette section, on s'intéresse à la question quand dans un espace hermitien une application linéaire est « compatible » avec le produit scalaire ; plus précisément, on voudrait comparer

$$\langle Mv, w\rangle, \langle Mv, Mw\rangle, \langle v, Mw\rangle, \text{ et } \langle v, w\rangle$$

où M est une matrice et v, w sont des vecteurs.
Cela nous amenera aux matrices symétriques, hermitiennes, orthogonaux, unitaires et aux isométries.
On démontrera plus loin notamment que toute matrice symétrique à coefficients réels est diagonalisable, et des généralisations.
On fait/rappelle les définitions suivantes :

Définition 10.1. *(a) Soit $M \in \mathrm{Mat}_{n\times n}(\mathbb{C})$. La matrice $M^{\mathrm{ad}} := \overline{M}^{\mathrm{tr}} = \overline{M^{\mathrm{tr}}}$ est appelée* matrice adjointe. *Noter $M^{\mathrm{ad}} = M^{\mathrm{tr}}$ si $M \in \mathrm{Mat}_{n\times n}(\mathbb{R})$.*

(b) On appelle matrice symétrique *ou* autoadjointe *toute matrice $M \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ telle que $M^{\mathrm{ad}} = M^{\mathrm{tr}} = M$.*

(c) On appelle matrice hermitienne *ou* autoadjointe *toute matrice $M \in \mathrm{Mat}_{n\times n}(\mathbb{C})$ telle que $M^{\mathrm{ad}} = M$. Noter que M est aujoadjointe si et seulement si $M^{\mathrm{tr}} = \overline{M}$. Noter aussi qu'une matrice symétrique n'est rien d'autre qu'une matrice hermitienne à coefficients réels.*

(d) On appelle matrice orthogonale *ou* isométrie *toute matrice $M \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ telle que $M^{\mathrm{ad}}M = M^{\mathrm{tr}}M = \mathrm{id}$.*

(e) On appelle matrice unitaire *ou* isométrie *toute matrice $M \in \mathrm{Mat}_{n\times n}(\mathbb{C})$ telle que $M^{\mathrm{ad}}M = \mathrm{id}$. Noter que M est unitaire si et seulement si $M^{\mathrm{tr}}\overline{M} = \mathrm{id}$. Noter aussi qu'une matrice orthognale n'est rien d'autre qu'une matrice unitaire à coefficients réels.*

Définition 10.2. *Nous définissons les groupes de matrices suivants où la loi de multiplication est la composition de matrices :*

(a) $\mathrm{GL}_n(K) = \{M \in \mathrm{Mat}_{n\times n}(K) \mid \det(M) \neq 0\}$*, le* groupe linéaire général *sur K,*

(b) $\mathrm{SL}_n(K) = \{M \in \mathrm{Mat}_{n\times n}(K) \mid \det(M) = 1\}$*, le* groupe linéaire spécial *sur K,*

(c) $O_n = \{M \in \mathrm{GL}_n(\mathbb{R}) \mid M^{\mathrm{tr}}M = \mathrm{id}\}$*, le* groupe orthogonal*;*

(d) $\mathrm{SO}_n = \{M \in \mathrm{SL}_n(\mathbb{R}) \mid M^{\mathrm{tr}}M = \mathrm{id}\}$*, le* groupe orthogonal spécial*,*

(e) $U_n = \{M \in \mathrm{GL}_n(\mathbb{C}) \mid M^{\mathrm{tr}}\overline{M} = \mathrm{id}\}$*, le* groupe unitaire*,*

(f) $\mathrm{SU}_n = \{M \in \mathrm{SL}_n(\mathbb{C}) \mid M^{\mathrm{tr}}\overline{M} = \mathrm{id}\}$*, le* groupe unitaire spécial.

Noter qu'il s'agit effectivement de groupes comme des calculs simple montrent ; par exemple pour (e), on a $(MN)^{\mathrm{tr}}\overline{(MN)} = N^{\mathrm{tr}}M^{\mathrm{tr}}\overline{M}\ \overline{N} = 1$ si $M^{\mathrm{tr}}\overline{M} = 1$ et $N^{\mathrm{tr}}\overline{N} = 1$.

Lemme 10.3. *Soit $M \in \mathrm{Mat}_{n\times n}(K)$ une matrice carrée.*

(a) Les assertions suivantes sont équivalentes :

  (i) M est autoadjoint.

  (ii) Pour tout $v, w \in K^n$ on a : $v^{\mathrm{tr}}M^{\mathrm{tr}}\overline{w} = v^{\mathrm{tr}}\overline{M}\overline{w}$.
  Noter qu'en termes du produit scalaire canonique, l'assertion se réécrit comme :
  $\langle Mv, w\rangle = \langle v, Mw\rangle$.

(b) Les assertions suivantes sont équivalentes :

  (i) M est une isométrie.

  (ii) Pour tout $v, w \in K^n$ on a : $v^{\mathrm{tr}}M^{\mathrm{tr}}\overline{M}\overline{w} = v^{\mathrm{tr}}\overline{w}$.
  Noter qu'en termes du produit scalaire canonique, l'assertion se réécrit comme :
  $\langle Mv, Mw\rangle = \langle v, w\rangle$.

Démonstration. Nous avons démontré la partie (a) au début de la section 9. La démonstration de la partie (b) procède par exactement les mêmes arguements. Plus précisement, elle est immédiate de la formule $e_i^{\mathrm{tr}} M e_j = m_{i,j}$ pour toute matrice carrée $M = (m_{i,j})$. □

Il est très facile de donner des exemples de matrices symétriques ou hermitiennes (choisir n'importe quels coefficients réels sur la diagonale, écrire n'importe quels coefficients réels ou complexes (selon la situation) dans la partie sous la diagonale principale, remplir la partie au dessus de la diagonale principales par les valeurs correspondantes).

Lemme 10.4. *Soit $M \in \mathrm{Mat}_{n\times n}(K)$ une matrice carrée. Les assertions suivantes sont équivalentes :*

(i) M est une isométrie (i.e. unitaire ou orthogonal) ;

(ii) les colonnes de M formes une base orthonormale de K^n (pour le produit scalaire canonique) ;

(iii) les lignes de M formes une base orthonormale de K^n (pour le produit scalaire canonique).

Démonstration. Par la définition de la multiplication de deux matrices, l'assertion (ii) est précisement l'égalité $M^{\mathrm{tr}}\overline{M} = \mathrm{id}$, donc l'assertion (i). L'assertion (iii) est l'assertion (ii) pour la matrice M^{tr}. Donc l'équivalence entre (iii) et (i) revient à l'équivalence

$$M^{\mathrm{ad}} M = \mathrm{id} \Leftrightarrow M^{-1} = \overline{M}^{\mathrm{tr}} \Leftrightarrow M\overline{M^{\mathrm{tr}}} = \mathrm{id}.$$

□

Lemme 10.5. *On a*

$$O_2 = \{ \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix} \in \mathrm{GL}_2(\mathbb{R}) \mid 0 \le \alpha < 2\pi \} \cup \{ \begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ \sin(\alpha) & -\cos(\alpha) \end{pmatrix} \in \mathrm{GL}_2(\mathbb{R}) \mid 0 \le \alpha < 2\pi \}.$$

Démonstration. Notons d'abord que la matrice $M = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}$ est bien orthogonale :

$$M^{\mathrm{tr}} M = \begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ -\sin(\alpha) & \cos(\alpha) \end{pmatrix} \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix} = \begin{pmatrix} \cos^2(\alpha)+\sin^2(\alpha) & 0 \\ 0 & \cos^2(\alpha)+\sin^2(\alpha) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Le calcul pour la matrice $\begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ \sin(\alpha) & -\cos(\alpha) \end{pmatrix}$ est similair.
Soit maintenant $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ une matrice orthogonale, donc

$$M^{\mathrm{tr}} M = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a^2+c^2 & ab+cd \\ ab+cd & b^2+d^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Des égalités $a^2 + c^2 = 1$ et $b^2 + d^2 = 1$, on obtient $0 \le \alpha, \beta < 2\pi$ tels que

$$a = \cos(\alpha), c = \sin(\alpha), d = \cos(\beta), b = \sin(\beta).$$

L'égalité $ab + cd = 0$ donne donc

$$0 = \cos(\alpha)\sin(\beta) + \sin(\alpha)\cos(\beta) = \sin(\alpha + \beta).$$

Nous en concluons

$$\alpha + \beta = m\pi$$

pour un $m \in \mathbb{Z}$. Si m est pair, nous trouvons :

$$\cos(\beta) = \cos(m\pi - \alpha) = \cos(m\pi)\cos(\alpha) + \sin(m\pi)\sin(\alpha) = \cos(\alpha)$$

et

$$\sin(\beta) = \sin(m\pi - \alpha) = \sin(m\pi)\cos(\alpha) - \cos(m\pi)\sin(\alpha) = -\sin(\alpha)$$

ce qui donne

$$\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}.$$

Si m est impair, nous trouvons :

$$\cos(\beta) = \cos(m\pi - \alpha) = \cos(m\pi)\cos(\alpha) + \sin(m\pi)\sin(\alpha) = -\cos(\alpha)$$

et

$$\sin(\beta) = \sin(m\pi - \alpha) = \sin(m\pi)\cos(\alpha) - \cos(m\pi)\sin(\alpha) = +\sin(\alpha)$$

ce qui donne

$$\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) = \begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ \sin(\alpha) & -\cos(\alpha) \end{pmatrix},$$

comme voulu. □

Nous changeons maintenant de point de vue : au lieu de matrices, nous regardons des applications linéaires entre espaces hermitiens.

Proposition 10.6. *Soient V et W deux K-espaces hermitiens de dimensions n et m et $\varphi : V \to W$ une application K-linéaire.*

(a) Il existe une unique application K-linéaire $\varphi^{\mathrm{ad}} : W \to V$ telle que pour tout $v \in V$ et tout $w \in W$

$$\langle \varphi(v), w \rangle = \langle v, \varphi^{\mathrm{ad}}(w) \rangle.$$

Noter que le produit scalaire à gauche est celui de W, et le produit à droite celui de V.

L'application φ^{ad} est appelée l'adjointe de φ.

(b) Soient S une K-base orthonormale de V et T une K-base orthonormale de W. Alors

$$M_{S,T}(\varphi^{\mathrm{ad}}) = \overline{M_{T,S}(\varphi)^{\mathrm{tr}}}$$

(la matrice obtenue de la transposée par la conjugaison complexe).

Donc $M_{S,T}(\varphi^{\mathrm{ad}})$ est la matrice adjointe de $M_{T,S}(\varphi)$.

Démonstration. Soient $S = s_1, \ldots, s_n$ et $T = t_1, \ldots, t_m$ les deux bases orthonormales. Soit

$$M_{T,S}(\varphi) = (a_{i,j})_{1 \le i \le m, 1 \le j \le n},$$

c'est-à-dire $\varphi(s_i) = \sum_{k=1}^{m} a_{k,i} t_k$. Nous allons prendre (b) comme la définition de φ^{ad} : c'est l'application K-linéaire représentée par $\overline{M_{T,S}(\varphi)^{\mathrm{tr}}}$. Concrètement, nous avons $\varphi^{\mathrm{ad}}(t_j) = \sum_{k=1}^{n} \overline{a_{j,k}} s_k$.

Nous vérifions d'abord :

$$\begin{aligned}\langle\varphi(s_i),t_j\rangle &= \langle\sum_{k=1}^m a_{k,i}t_k,t_j\rangle = \sum_{k=1}^m a_{k,i}\langle t_k,t_j\rangle &&= a_{j,i}\\ \langle s_i,\varphi^{\mathrm{ad}}(t_j)\rangle &= \langle s_i,\sum_{k=1}^m \overline{a_{j,k}}s_k\rangle = \sum_{k=1}^m a_{j,k}\langle s_i,s_k\rangle &&= a_{j,i}\end{aligned}$$

Nous pouvons maintenant obtenir (a) par la linéarité : soient $v=\sum_{i=1}^n b_is_i$ et $w=\sum_{j=1}^m c_jt_j$; nous avons

$$\begin{aligned}\langle\varphi(v),w\rangle &= \langle\varphi(\sum_{i=1}^n b_is_i),\sum_{j=1}^m c_jt_j\rangle\\ &= \langle\sum_{i=1}^n b_i\varphi(s_i),\sum_{j=1}^m c_jt_j\rangle\\ &= \sum_{i=1}^n b_i\sum_{j=1}^m \overline{c_j}\langle\varphi(s_i),t_j\rangle\\ &= \sum_{i=1}^n b_i\sum_{j=1}^m \overline{c_j}\langle s_i,\varphi^{\mathrm{ad}}(t_j)\rangle\\ &= \langle\sum_{i=1}^n b_is_i,\sum_{j=1}^m c_j\varphi^{\mathrm{ad}}(t_j)\rangle\\ &= \langle\sum_{i=1}^n b_is_i,\varphi^{\mathrm{ad}}(\sum_{j=1}^m c_jt_j)\rangle\\ &= \langle v,\varphi^{\mathrm{ad}}(w)\rangle.\end{aligned}$$

Pour l'unicité de φ^{ad}, écrivons $\varphi^{\mathrm{ad}}(t_j)=\sum_{k=1}^n d_{k,j}s_k$, et calculons

$$a_{j,i}=\langle\varphi(s_i),t_j\rangle=\langle s_i,\varphi^{\mathrm{ad}}(t_j)\rangle=\langle s_i,\sum_{k=1}^n d_{k,j}s_k\rangle=\sum_{k=1}^n\overline{d_{k,j}}\langle s_i,s_k\rangle=\overline{d_{i,j}}.$$

On obtient donc $d_{i,j}=\overline{a_{j,i}}$, l'unicité. □

Noter que si $K=\mathbb{R}$, l'adjointe d'une matrice M est la matrice transposée.

Proposition 10.7. *Soient U,V,W des K-espaces hermitiens de dimensions finies et $U\xrightarrow{\varphi,\psi}V\xrightarrow{\eta}W$ des applications K-linéaires. Alors :*

(a) $\mathrm{id}_V^{\mathrm{ad}}=\mathrm{id}_V$,

(b) $(\varphi+\psi)^{\mathrm{ad}}=\varphi^{\mathrm{ad}}+\psi^{\mathrm{ad}}$,

(c) $\forall\,x\in K:(x\varphi)^{\mathrm{ad}}=\overline{x}\varphi^{\mathrm{ad}}$,

(d) $(\eta\circ\varphi)^{\mathrm{ad}}=\varphi^{\mathrm{ad}}\circ\eta^{\mathrm{ad}}$ *et*

(e) $(\varphi^{\mathrm{ad}})^{\mathrm{ad}} = \varphi$.

Les mêmes assertions sont vraies pour des matrices.

Démonstration. Les assertions pour les matrices sont facilement vérifiées. Le seul endroit où il faut faire attention est $(M \circ N)^{\mathrm{tr}} = N^{\mathrm{tr}} \circ M^{\mathrm{tr}}$, c'est le lemme 9.1. □

Définition 10.8. *Soient V un K-espace hermitien de dimension finie et $\varphi : V \to V$ un K-endomorphisme.*
On dit que φ est autoadjoint (selbstadjungiert, self adjoint) *si $\varphi = \varphi^{\mathrm{ad}}$.*

En vue de la proposition 10.6, nous avons donc

$$\varphi \text{ est autoadjoint } \Leftrightarrow M_{S,S}(\varphi) \text{ est autoadjoint,}$$

pour S une base orthnormale de V.
Pour la démonstration de la proposition suivante, nous avons besoin d'un petit lemme.

Lemme 10.9. *Soit $(V, \langle, \rangle)$ un espace hermitien. Alors, si $v \perp V$ pour $v \in V$, alors $v = 0$.*

Démonstration. Si $v \perp V$, on a en particulier, $v \perp v$, donc $0 = \langle v, v\rangle = |v|^2$ ce qui implique bien $v = 0$. □

Proposition 10.10. *Soient V un K-espace hermitien de dimension finie et $\varphi : V \to V$ un K-endomorphisme.*

(a) Les assertions suivantes sont équivalentes.

(i) φ est autoadjoint ($\varphi = \varphi^{\mathrm{ad}}$).

(ii) $\langle v, w\rangle_\varphi := \langle \varphi(v), w\rangle$ pour $v \in V$ et $w \in V$ est une forme hermitienne.

(b) Si φ est autoadjoint, alors : $\varphi = 0 \Leftrightarrow \forall\, v \in V : \langle v, v\rangle_\varphi = 0$.

Démonstration. (a) Il est toujours vrai (même si φ n'est pas autoadjoint) que $\langle \cdot, \cdot\rangle_\varphi$ est linéaire dans la première variable et sesquilinéaire dans la deuxième. Il faut donc regarder la troisième propriété dans la définition des formes hermitiennes 9.2. Soient $v, w \in V$. D'abord on fait le calcul

$$\overline{\langle v, w\rangle_\varphi} = \overline{\langle \varphi(v), w\rangle} = \overline{\langle v, \varphi^{\mathrm{ad}}(w)\rangle} = \langle \varphi^{\mathrm{ad}}(w), v\rangle = \langle w, v\rangle_{\varphi^{\mathrm{ad}}}.$$

Nous avons donc

$$\begin{aligned}
&\forall\, v, w \in V : \overline{\langle v, w\rangle_\varphi} = \langle w, v\rangle_\varphi \\
\Leftrightarrow &\forall\, v, w \in V : \langle \varphi^{\mathrm{ad}}(w), v\rangle = \langle \varphi(w), v\rangle \\
\Leftrightarrow &\forall\, v, w \in V : \langle (\varphi^{\mathrm{ad}} - \varphi)(w), v\rangle = 0 \\
\Leftrightarrow &\forall\, w \in V : (\varphi^{\mathrm{ad}} - \varphi)(w) \perp V = 0 \\
\Leftrightarrow &\forall\, w \in V : (\varphi^{\mathrm{ad}} - \varphi)(w) = 0 \\
\Leftrightarrow &\varphi^{\mathrm{ad}} = \varphi
\end{aligned}$$

par le lemme 10.9.
(b) Si $\varphi = 0$, il en suit trivialement

$$\langle v, v\rangle_\varphi = \langle \varphi(v), v\rangle = \langle 0, v\rangle = 0.$$

Supposons maintenant $\langle v, v\rangle_\varphi = 0$ pour tout $v \in V$. Soient $v, w \in V$ et soit $a \in K$. On calcule

$$\begin{aligned}
0 &= \langle v + aw, v + aw\rangle_\varphi \\
&= \langle \varphi(v + aw), v + aw\rangle \\
&= \underbrace{\langle \varphi(v), v\rangle}_{=0} + \langle \varphi(v), aw\rangle + \langle \varphi(aw), v\rangle + \underbrace{\langle \varphi(aw), aw\rangle}_{=0} \\
&= \overline{a}\langle \varphi(v), w\rangle + a\langle \varphi(w), v\rangle \\
&= \overline{a}\langle \varphi(v), w\rangle + a\langle w, \varphi(v)\rangle \\
&= \overline{a}\langle \varphi(v), w\rangle + a\overline{\langle \varphi(v), w\rangle} \\
&= 2 \cdot \mathrm{Re}(\overline{a}\langle \varphi(v), w\rangle).
\end{aligned}$$

Avec $a = 1$, nous obtenons $0 = \mathrm{Re}(\langle \varphi(v), w\rangle)$, et avec $a = i$ on trouve $0 = \mathrm{Im}(\langle \varphi(v), w\rangle)$. En conséquence, nous avons pour tout $v, w \in V$

$$0 = \langle \varphi(v), w\rangle.$$

Pour tout $v \in V$, nous trouvons donc $\varphi(v) \perp V$, d'où le résultat voulu $\varphi(v) = 0$ par le lemme 10.9. $\square$

Si l'on applique la proposition précédente avec φ_M pour une matrice carrée M, on retrouve le résultat de la discussion au début de la section 9. Alors :

(a) $M = M^{\mathrm{ad}} \Leftrightarrow M$ est autoadjoint $\Leftrightarrow (v, w) \mapsto v^{\mathrm{tr}}Aw$ est une forme hermitienne.

(b) Si M est autoadjoint, alors : $M = 0 \Leftrightarrow \forall\, v \in K^n : v^{\mathrm{tr}}Mv = 0$.

Nous introduisons maintenant les applications qui ne changent pas les longueurs : ce sont les « isométries ».

Définition 10.11. *Soit V un espace hermitien. On appelle* isométrie *tout $\varphi \in \mathrm{End}_K(V)$ tel que pour tout $v \in V$*

$$|\varphi(v)| = |v|.$$

Lemme 10.12. *Soient V un espace hermitien et $\varphi \in \mathrm{End}_K(V)$. Les assertions suivantes sont équivalentes :*

(i) φ est une isométrie.

(ii) $\varphi^{\mathrm{ad}} \circ \varphi = \mathrm{id}_V$ (en particulier, φ est un isomorphisme).

(iii) Pour tout $v, w \in W$: $\langle \varphi(v), \varphi(w)\rangle = \langle v, w\rangle$.

Démonstration. « (i) $\Rightarrow$ (ii) » : Nous avons pour tout $v \in V$:

$$\langle v, v\rangle = \langle \varphi(v), \varphi(v)\rangle = \langle v, \varphi^{\mathrm{ad}}(\varphi(v))\rangle,$$

donc

$$\langle v, (\varphi^{\mathrm{ad}} \circ \varphi - \mathrm{id}_V)(v)\rangle = 0 \text{ et, alors, } \langle (\varphi^{\mathrm{ad}} \circ \varphi - \mathrm{id}_V)(v), v\rangle = 0.$$

Noter que $\varphi^{\mathrm{ad}} \circ \varphi - \mathrm{id}_V$ est autoadjoint, donc la proposition 10.10(b) implique $\varphi^{\mathrm{ad}} \circ \varphi - \mathrm{id}_V = 0$, alors $\varphi^{\mathrm{ad}} \circ \varphi - \mathrm{id}_V$.
« (ii) $\Rightarrow$ (iii) » : Soient $v, w \in V$, alors

$$\langle \varphi(v), \varphi(w)\rangle = \langle v, \varphi^{\mathrm{ad}}(\varphi(w))\rangle = \langle v, w\rangle.$$

« (iii) $\Rightarrow$ (i) » : Soit $v \in V$. Alors,

$$|\varphi(v)|^2 = \langle \varphi(v), \varphi(v)\rangle = \langle v, v\rangle = |v|^2.$$

$\square$

Par ce lemme, nous avons

$$\varphi \text{ est une isométrie } \Leftrightarrow M_{S,S}(\varphi) \text{ est une isométrie (i.e. orthogonal ou unitaire)}$$

pour S une base orthonormale de V.
Jusqu'ici nous avons considéré deux types d'endomorphismes/matrices : autoadjoints et isométries. On aimerait pouvoir traiter quelques-unes de leurs propriétés en parallèle. On cherche donc une généralisation commune. Les opérateurs normaux sont une telle généralisation. Nous donnons d'abord la définition de façon « métrique ».

Définition 10.13. *Soit V un espace hermitien. On appelle* opérateur normal *tout $\varphi \in \mathrm{End}_K(V)$ tel que pour tout $v \in V$*

$$|\varphi(v)| = |\varphi^{\mathrm{ad}}(v)|.$$

Exemple 10.14. *(a) Si φ est autoadjoint, on a $\varphi^{\mathrm{ad}} = \varphi$, donc, φ est normal.*

(b) Si φ est une isométrie, on a que φ est un isomorphisme et $\varphi^{\mathrm{ad}} = \varphi^{-1}$. Comme $|\varphi(v)| = |v|$ pour tout $v \in V$, en particulier aussi pour $\varphi^{-1}(v)$, on trouve $|\varphi^{\mathrm{ad}}(v)| = |\varphi^{-1}(v)| = |\varphi(\varphi^{-1}(v))| = |v|$, donc, φ est normal.

Proposition 10.15. *Soient V un espace hermitien et $\varphi \in \mathrm{End}_K(V)$. Les assertions suivantes sont équivalentes :*

(i) φ est normal.

(ii) $\varphi^{\mathrm{ad}} \circ \varphi = \varphi \circ \varphi^{\mathrm{ad}}$.

Démonstration. Faisons d'abord le calcul

$$\begin{aligned}|\varphi(v)|^2 - |\varphi^{\mathrm{ad}}(v)|^2 &= \langle \varphi(v), \varphi(v)\rangle - \langle \varphi^{\mathrm{ad}}(v), \varphi^{\mathrm{ad}}(v)\rangle \\ &= \langle \varphi(v), (\varphi^{\mathrm{ad}})^{\mathrm{ad}}(v)\rangle - \langle \varphi^{\mathrm{ad}}(v), \varphi^{\mathrm{ad}}(v)\rangle \\ &= \langle \varphi^{\mathrm{ad}} \circ \varphi(v), v\rangle - \langle \varphi \circ \varphi^{\mathrm{ad}}(v), v\rangle \\ &= \langle (\varphi^{\mathrm{ad}} \circ \varphi - \varphi \circ \varphi^{\mathrm{ad}})(v), v\rangle.\end{aligned}$$

Notons que $\varphi \circ \varphi^{\mathrm{ad}} - \varphi^{\mathrm{ad}} \circ \varphi$ est autoadjoint. En conséquence (proposition 10.10(b)) nous avons

$$\big(\forall\, v \in V : |\varphi(v)|^2 = |\varphi^{\mathrm{ad}}(v)|^2\big) \Leftrightarrow \varphi^{\mathrm{ad}} \circ \varphi = \varphi \circ \varphi^{\mathrm{ad}}.$$

□

En termes de matrices, nous avons donc :

$$\varphi \text{ est normal } \iff \overline{M}^{\mathrm{tr}} M = M \overline{M}^{\mathrm{tr}} \overset{\text{définition}}{\iff} M \text{ est normal}$$

où $M = M_{S,S}(\varphi)$ pour S une base orthonormale de V.

Lemme 10.16. *Soient V un espace hermitien et $\varphi \in \mathrm{End}_K(V)$ normal. Soit $a \in \mathrm{Spec}(\varphi)$ une valeur propre de φ.*

(a) $E_\varphi(a) = E_{\varphi^{\mathrm{ad}}}(\overline{a})$.

(b) Si φ est autoadjoint, alors $a \in \mathbb{R}$.

(c) Si φ est une isométrie, alors $|a| = 1$

Démonstration. (a) Nous démontrons d'abord $\ker(\varphi) = \ker(\varphi^{\mathrm{ad}})$ pour tout opérateur normal. Soit $v \in V$, alors,

$$v \in \ker(\varphi) \Leftrightarrow \varphi(v) = 0 \Leftrightarrow |\varphi(v)| = 0 \overset{\text{déf. normalité}}{\Leftrightarrow} |\varphi^{\mathrm{ad}}(v)| = 0 \Leftrightarrow \varphi^{\mathrm{ad}}(v) = 0 \Leftrightarrow v \in \ker(\varphi^{\mathrm{ad}}).$$

Posons maintenant $\psi := \varphi - a \cdot \mathrm{id}_V$. C'est aussi un opérateur normal :

$$\begin{aligned}\psi\circ\psi^{\mathrm{ad}} &= (\varphi - a\cdot\mathrm{id}_V)\circ(\varphi - a\cdot\mathrm{id}_V)^{\mathrm{ad}} = (\varphi - a\cdot\mathrm{id}_V)\circ(\varphi^{\mathrm{ad}} - \overline{a}\cdot\mathrm{id}_V) = \varphi\circ\varphi^{\mathrm{ad}} - a\cdot\varphi^{\mathrm{ad}} - \overline{a}\cdot\varphi + a\cdot\overline{a}\cdot\mathrm{id}_V \\ &= \varphi^{\mathrm{ad}} \circ \varphi - a \cdot \varphi^{\mathrm{ad}} - \overline{a} \cdot \varphi + a \cdot \overline{a} \cdot \mathrm{id}_V = (\varphi - a \cdot \mathrm{id}_V)^{\mathrm{ad}} \circ (\varphi - a \cdot \mathrm{id}_V) = \psi^{\mathrm{ad}} \circ \psi.\end{aligned}$$

Le calcul précédent nous donne donc

$$E_\varphi(a) = \ker(\varphi - a \cdot \mathrm{id}_V) = \ker(\psi) = \ker(\psi^{\mathrm{ad}}) = \ker(\varphi^{\mathrm{ad}} - \overline{a} \cdot \mathrm{id}_V) = E_{\varphi^{\mathrm{ad}}}(\overline{a}).$$

(b) Pour tout $v \in E_\varphi(a)$ nous avons $v \in E_\varphi(\overline{a})$, donc $a \cdot v = \varphi(v) = \varphi^{\mathrm{ad}}(v) = \overline{a} \cdot v$, alors $a = \overline{a}$ et en conséquence $a \in \mathbb{R}$.
(c) Pour tout $v \in E_\varphi(a)$ nous avons $v = \varphi^{-1}(\varphi(v)) = \varphi^{-1}(a \cdot v) = a \cdot \varphi^{-1}(v) = a \cdot \overline{a} \cdot v = |a|^2 \cdot v$, donc $|a|^2 = 1$. □

Exemple 10.17. *Cet exemple nous donne un avant-goût du théorème spectral.*

(a) D'abord nous continuous l'analyse de O_2 du lemme 10.5.

(1) Soit $M = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}$. Son polynôme caractéristique est

$$(X - \cos(\alpha))^2 + \sin^2(\alpha) = X^2 - 2\cos(\alpha)X + 1$$

dont le discriminant est $4\cos^2(\alpha) - 4 \leq 0$ avec égalité si et seulement si $|\cos(\alpha)| = 1$, si et seulement si $\alpha \in \pi\mathbb{Z}$.

En conséquence, si $\alpha \notin \pi\mathbb{Z}$, alors M n'admet pas de vecteur propre et n'est donc pas diagonalisable. Cela est aussi géométriquement clair car M représente la rotation par l'angle α qui ne fixe aucun vecteur sauf si l'angle est un multiple de π.

Si α est un multiple pair de π, alors $M = \mathrm{id}$. Si α est un multiple impair de π, alors $M = -\mathrm{id}$.

(2) Soit $M = \begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ \sin(\alpha) & -\cos(\alpha) \end{pmatrix}$. Son polynôme caractéristique est

$$X^2 - \cos^2(\alpha) - \sin^2(\alpha) = X^2 - 1 = (X-1)(X+1).$$

La matrice M est donc diagonalisable avec valeurs propres -1 et 1. Un vecteur propre pour la valeur propre 1 est donné par $\begin{pmatrix} \cos(\alpha/2) \\ \sin(\alpha/2) \end{pmatrix}$ et le vecteur $\begin{pmatrix} \sin(\alpha/2) \\ -\cos(\alpha/2) \end{pmatrix}$ est propre pour la valeur propre -1.

Il s'agit géométriquement d'une réflection à une axe (vecteur propre pour la valeur propre 1).

(b) Soit $M \in \mathrm{Mat}_{3\times 3}(\mathbb{R})$ une matrice orthogonale. Son polynôme caractéristique est unitaire de degré 3 et admet donc un zéro réel λ_1. Par le lemme 10.16, ce zéro est soit 1, soit -1. Il existe donc un vecteur propre v_1 pour la valeur propre λ_1. Nous pouvons le normaliser tel que $|v_1| = 1$.

Par Gram-Schmidt, nous pouvons trouver des vecteurs v_2, v_3 tels que v_1, v_2, v_3 forment une base orthonormale de $\mathbb{R}^3$. En plus, comme M est une isométrie, pour $i = 1, 2$, nous avons

$$0 = \langle v_i, v_1 \rangle = \langle Mv_i, Mv_1 \rangle = \lambda_1 \langle Mv_i, v_1 \rangle.$$

Cela veut dire que M envoie le sous-espace $W \leq \mathbb{R}^3$ engendré par v_2, v_3 dans lui même.

Si l'on écrit les vecteurs v_1, v_2, v_3 comme colonnes dans une matrice C (qui est orthogonale !), nous obtenons alors

$$C^{\mathrm{tr}} M C = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & a & b \\ 0 & c & d \end{pmatrix}.$$

La matrice $A := \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ est orthogonale et appartient à O_2.

Si $\det(A) = \det(M)/\lambda_1 = 1$, nous avons que $A = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}$ pour un $0 \leq \alpha < 2\pi$. Si $\det(A) = -1$, nous pouvons trouver une base w_2, w_3 de W consistant en vecteurs propres normalisés : $|w_i| = 1$ pour $i = 2, 3$ pour les valeurs propres $1, -1$. En conséquence, v_1, w_2, w_3 est une base orthonormale de $\mathbb{R}^3$. Si D est la matrice (orthogonale !) dont les colonnes sont ces vecteurs, nous avons finalement

$$D^{\mathrm{tr}} M D = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

pour $\lambda_2 \in \{1, -1\}$.

11 Théorème spectral

Objectifs :

- Connaître et maîtriser les théorèmes spectraux ;
- savoir calculer la diagonalisation de matrices normales complexes, des matrices autoadjointes ;
- savoir calculer la forme normale des matrices orthonormales ;
- connaître des exemples et savoir démontrer des propriétés simples.

Soit V un espace hermitien et soient $U, W \leq V$ deux sous-espaces vectoriel. Nous écrivons $U \textcircled{\perp} W$ pour $U + W$ si la somme est directe ($U \oplus W$) et les deux espaces sont orthogonaux ($U \perp W$).

Lemme 11.1. *Soient V un espace hermitien et $\varphi \in \mathrm{End}_K(V)$ normal. Alors, pour tout $a_1, \dots, a_n \in \mathrm{Spec}(\varphi)$ distincts, nous avons*

$$E_\varphi(a_1) + E_\varphi(a_2) + \cdots + E_\varphi(a_n) = E_\varphi(a_1) \textcircled{\perp} E_\varphi(a_2) \textcircled{\perp} \cdots \textcircled{\perp} E_\varphi(a_n).$$

Démonstration. Nous avons déjà vu dans le lemme 3.10 que la somme d'espaces propres est directe. Soient $0 \neq v \in E_\varphi(a_i)$ et $0 \neq w \in E_\varphi(a_j)$ avec $i \neq j$ (donc $w \in E_{\varphi^{\mathrm{ad}}}(\overline{a_j})$ par le lemme 10.16). Nous avons

$$\langle \varphi(v), w \rangle = \langle a_i v, w \rangle = a_i \langle v, w \rangle,$$

mais aussi

$$\langle \varphi(v), w \rangle = \langle v, \varphi^{\mathrm{ad}}(w) \rangle = \langle v, \overline{a_j} w \rangle = a_j \langle v, w \rangle,$$

donc $\langle v, w \rangle = 0$. □

Nous démontrons d'abord le théorème spectral pour les opérateurs normaux à coefficients complexes. La raison en est que dans ce cas nous disposons du théorème suivant.

Théorème 11.2 (Théorème fondamental de l'algèbre)**.** *Tout polynôme $f \in \mathbb{C}[X]$ de degré ≥ 1 possède un zéro.*

Démonstration. Cours d'analyse. □

Corollaire 11.3. *Soient $0 \subsetneq W \subseteq V$ un sous-espace et $\varphi : V \to V$ une application K-linéaire telle que $\varphi(W) \subseteq W$. Alors il existe $0 \neq w \in W$ et $a \in K$ tels que $\varphi(w) = aw$. En plus, toute valeur propre de la restriction de $\varphi|_W$ à W est aussi une valeur propre de φ.*

Démonstration. Soit $f = \mathrm{car}_{\varphi|_W} \in \mathbb{C}[X]$ le polynôme caractéristique de la restriction de φ à W. Comme $W \neq 0$, on a $\deg(f) \geq 1$. En plus, f divise car_φ (cela suit, par exemple, par le lemme 7.1 (a) et les règles pour les déterminants des matrices en blocs). Cela montre déjà que le spectre de $\varphi|_W$ est contenu dans le spectre de φ. En plus, par le théorème fondamental de l'algèbre 11.2, le polynôme f possède un zéro $a \in \mathbb{C}$. Donc il existe un vecteur propre non-zéro $w \in W$ pour la valeur propre a. □

Théorème 11.4 (Théorème spectral pour opérateurs normaux)**.** *Soient V un $\mathbb{C}$-espace hermitien de dimension finie et $\varphi \in \mathrm{End}_K(V)$. Les assertions suivantes sont équivalentes :*

(i) φ *est normal.*

(ii) $V = \bigoplus_{a \in \mathrm{Spec}(\varphi)} E_\varphi(a)$ *(en particulier,* φ *est diagonalisable).*

(iii) V *possède une base orthonormale de vecteurs propres pour* φ*.*

Démonstration. « (i) $\Rightarrow$ (ii) » : Nous savons déjà que $W := \bigoplus_{a \in \mathrm{Spec}(\varphi)} E_\varphi(a)$ est un sous-espace de V et nous savons que la somme est orthogonale à cause du lemme 11.1. Le corollaire 9.13(b) nous renseigne sur l'existence d'un complément orthogonal $V = W \oplus W^\perp$. Le but est de montrer $W^\perp = 0$. Le lemme 10.16 implique que $W = \bigoplus_{a \in \mathrm{Spec}(\varphi)} E_{\varphi^{\mathrm{ad}}}(\overline{a})$, donc $\varphi^{\mathrm{ad}}(W) \subseteq W$. Soit maintenant $v \in W^\perp$. Alors pour tout $w \in W$,

$$\langle \varphi(v), w \rangle = \langle v, \varphi^{\mathrm{ad}}(w) \rangle = 0,$$

montrant $\varphi(v) \in W^\perp$. Donc nous pouvons restreindre φ sur $W^\perp$. Si $W^\perp$ était non-zéro, par le corollaire 11.3 il contiendrait un vecteur propre non-zéro pour une valeur propre de φ. Comme ce vecteur serait aussi dans W, on obtiendrait que l'intersection de W et $W^\perp$ est non-zéro, une contradiction. Donc $W^\perp = 0$, comme énoncé.
« (ii) $\Rightarrow$ (iii) » : Il suffit de choisir une base orthonormale de chaque $E_\varphi(a)$ et prendre la réunion ; on aura automatiquement une base orthonormale de V à cause de l'orthogonalité des espaces propres.
« (iii) $\Rightarrow$ (i) » : Soit $S = s_1, \dots, s_n$ une base orthonormale de V de vecteurs propres. Soit a_i la valeur propre associée à s_i (nous ne demandons pas que les a_i soient deux-à-deux distincts). Nous avons donc $\varphi(s_i) = a_i \cdot s_i$. Soit $1 \leq j \leq n$. On a

$$\langle s_j, \varphi^{\mathrm{ad}}(s_i) - \overline{a_i} s_i \rangle = \langle s_j, \varphi^{\mathrm{ad}}(s_i) \rangle - \langle s_j, \overline{a_i} s_i \rangle = \langle \varphi(s_j), s_i \rangle - a_i \langle s_j, s_i \rangle = (a_j - a_i)\langle s_j, s_i \rangle = 0.$$

Donc $(\varphi^{\mathrm{ad}}(s_i) - \overline{a_i} s_i) \perp V$, alors $\varphi^{\mathrm{ad}}(s_i) = \overline{a_i} \cdot s_i$ par le lemme 10.9. Le calcul

$$\varphi(\varphi^{\mathrm{ad}}(s_i)) = \varphi(\overline{a_i} \cdot s_i) = \overline{a_i} \cdot \varphi(s_i) = \overline{a_i} \cdot a_i \cdot s_i$$
$$\varphi^{\mathrm{ad}}(\varphi(s_i)) = \varphi^{\mathrm{ad}}(a_i \cdot s_i) = a_i \cdot \varphi^{\mathrm{ad}}(s_i) = a_i \cdot \overline{a_i} \cdot s_i,$$

implique $\varphi \circ \varphi^{\mathrm{ad}} = \varphi^{\mathrm{ad}} \circ \varphi$, la normalité de φ. □

Donnons maintenant la traduction en termes de matrices du théorème spectral 11.4.

Corollaire 11.5. *Soit* $M \in \mathrm{Mat}_{n \times n}(\mathbb{C})$ *une matrice. Alors les assertions suivantes sont équivalentes :*

(i) M *est normal, c'est-à-dire* $\overline{M}^{\mathrm{tr}} \cdot M = M \cdot \overline{M}^{\mathrm{tr}}$*.*

(ii) Il existe une matrice unitaire $C \in \mathrm{Mat}_{n \times n}(\mathbb{C})$ *telle que* $\overline{C}^{\mathrm{tr}} \cdot M \cdot C$ *est une matrice diagonale.*

Démonstration. « (i) $\Rightarrow$ (ii) » : Soit $\varphi = \varphi_M$ l'endomorphisme de $\mathbb{C}^n$ tel que $M_{S,S}(\varphi) = M$ où S est la base canonique (qui est orthonormale pour le produit scalaire canonique !). Selon la proposition 10.6 nous avons $\overline{M}^{\mathrm{tr}} = M_{S,S}(\varphi^{\mathrm{ad}})$. Donc l'hypothèse que M est normal revient à dire que φ est normal. Nous utilisons le théorème 11.4 pour obtenir une base orthonormale T de vecteurs propres. Donc $C_{S,T}^{-1} \cdot M_{S,S}(\varphi) \cdot C_{S,T} = M_{T,T}(\varphi)$ est une matrice diagonale. Rappelons maintenant que les colonnes de $C := C_{S,T}$ sont les vecteurs de la base T. Comme T est orthonormal pour le produit scalaire canonique, nous avons $C \cdot \overline{C}^{\mathrm{tr}} = \mathrm{id}_n$ et l'assertion est démontrée.

« (ii) $\Rightarrow$ (i) » : Soit $\overline{C}^{\mathrm{tr}} \cdot M \cdot C = \mathrm{diag}(a_1, \dots, a_n)$, la matrice diagonale avec $a_1, \dots, a_n$ sur la diagonale. Notons d'abord

$$\overline{(\overline{C}^{\mathrm{tr}} \cdot MC)}^{\mathrm{tr}} = \overline{C}^{\mathrm{tr}} \cdot \overline{M}^{\mathrm{tr}} \cdot C = \mathrm{diag}(\overline{a_1}, \dots, \overline{a_n}).$$

Comme des matrices diagonales commutent, nous trouvons

$$\begin{aligned}\overline{(\overline{C}^{\mathrm{tr}} \cdot MC)}^{\mathrm{tr}} \cdot (\overline{C}^{\mathrm{tr}} \cdot MC) &= \overline{C}^{\mathrm{tr}} \cdot \overline{M}^{\mathrm{tr}} \cdot C \cdot \overline{C}^{\mathrm{tr}} \cdot MC = \overline{C}^{\mathrm{tr}} \cdot \overline{M}^{\mathrm{tr}} \cdot MC \\ &= (\overline{C}^{\mathrm{tr}} \cdot MC) \cdot \overline{(\overline{C}^{\mathrm{tr}} \cdot MC)}^{\mathrm{tr}} = \overline{C}^{\mathrm{tr}} \cdot M \cdot C \cdot \overline{C}^{\mathrm{tr}} \overline{M}^{\mathrm{tr}} \cdot C = \overline{C}^{\mathrm{tr}} \cdot M \cdot \overline{M}^{\mathrm{tr}} \cdot C,\end{aligned}$$

donc $\overline{M}^{\mathrm{tr}} \cdot M = M \cdot \overline{M}^{\mathrm{tr}}$. □

Regardons maintenant le cas $K = \mathbb{R}$.

Lemme 11.6. *Soit $M \in \mathrm{Mat}_{n \times n}(\mathbb{R})$ une matrice que nous considérons sur $\mathbb{C}$.*

(a) Pour tout $\mu \in \mathbb{C}$ et tout $v \in \mathbb{C}^n$ on a l'équivalence : $v \in E_M(\mu) \Longleftrightarrow \overline{v} \in E_M(\overline{\mu})$.

(b) Pour $\mu \in \mathbb{C}$ on a l'équivalence : $\mu \in \mathrm{Spec}(M) \Longleftrightarrow \overline{\mu} \in \mathrm{Spec}(M)$.

(c) Pour $\mu \in \mathbb{R}$, l'espace propre $E_M(\mu) \subseteq \mathbb{C}^n$ possède une base dans $\mathbb{R}^n$.

(d) Soit $\mu \in \mathrm{Spec}(M)$ tel que $\mu \in \mathbb{C} \setminus \mathbb{R}$ et soit $v \in E_M(\mu)$ tel que $|v| = 1$.
Posons $x := \frac{1}{\sqrt{2}}(v + \overline{v}), y := \frac{1}{i\sqrt{2}}(v - \overline{v}) \in E_M(\mu) \oplus E_M(\overline{\mu})$.
Alors $|x| = 1$, $|y| = 1$, $x \perp y$, $Mx = \mathrm{Re}(\mu) \cdot x - \mathrm{Im}(\mu) \cdot y$ et $My = \mathrm{Re}(\mu) \cdot y + \mathrm{Im}(\mu) \cdot x$.
Donc, M agit sur $\langle x, y \rangle$ par la matrice $\begin{pmatrix} \mathrm{Re}(\mu) & \mathrm{Im}(\mu) \\ -\mathrm{Im}(\mu) & \mathrm{Re}(\mu) \end{pmatrix}$ pour la base $\{x, y\}$.

Démonstration. (a) On observe : $Mv = \mu \cdot v \Longleftrightarrow \overline{Mv} = M\overline{v} = \overline{\mu \cdot v} = \overline{\mu} \cdot \overline{v}$ ce qui implique le résultat. (b) en est une conséquence directe.
(c) Il suffit de montrer que $E_M(\mu)$ admet un système de générateurs dans $\mathbb{R}^n$. Soit $v_1, \dots, v_r \in \mathbb{C}^n$ une $\mathbb{C}$-base de $E_M(\mu)$. Posons $x_j = \mathrm{Re}(v_j) = \frac{1}{2}(v_j + \overline{v_j})$ et $y_j = \mathrm{Im}(v_j) = \frac{1}{2i}(v_j - \overline{v_j})$ pour $j = 1, \dots, r$. Ces vecteurs appartiennent à $E_M(\mu)$ car $\overline{v_j}$ y appartient pour tout j. Comme $v_j = x_j + iy_j$, les vecteurs $x_1, \dots, x_r, y_1, \dots, y_r$ engendrent $E_M(\mu)$.
(d) Notons d'abord que $v \perp \overline{v}$ car $E_M(\mu) \perp E_M(\overline{\mu})$ comme $\mu \neq \overline{\mu}$. On a

$$|x|^2 = \langle x, x \rangle = (\frac{1}{\sqrt{2}})^2 \langle v + \overline{v}, v + \overline{v} \rangle = \frac{1}{2}\big(\langle v, v \rangle + \langle \overline{v}, \overline{v} \rangle + \langle v, \overline{v} \rangle + \langle \overline{v}, v \rangle\big) = 1.$$

Le calcul de $|y|$ est similair :

$$|y|^2 = \langle y, y \rangle = (\frac{1}{\sqrt{2}})^2 \langle v - \overline{v}, v - \overline{v} \rangle = \frac{1}{2}\big(\langle v, v \rangle + \langle \overline{v}, \overline{v} \rangle - \langle v, \overline{v} \rangle - \langle \overline{v}, v \rangle\big) = 1.$$

On a aussi :

$$\langle x, y \rangle = \frac{i}{2} \langle v + \overline{v}, v - \overline{v} \rangle = \frac{i}{2}\big(\langle v, v \rangle - \langle \overline{v}, \overline{v} \rangle + \langle \overline{v}, v \rangle - \langle v, \overline{v} \rangle\big) = 0.$$

Calculons maintenant l'action de M :

$$\begin{aligned}
Mx &= \frac{1}{\sqrt{2}}(Mv + M\overline{v}) = \frac{1}{\sqrt{2}}(\mu v + \overline{\mu v}) = \frac{1}{2\sqrt{2}}\big((\mu+\overline{\mu})(v+\overline{v}) + (\mu-\overline{\mu})(v-\overline{v})\big) \\
&= \frac{1}{2}(\mu+\overline{\mu})x - \frac{1}{2i}(\mu-\overline{\mu})y = \mathrm{Re}(\mu)\cdot x - \mathrm{Im}(\mu)\cdot y \\
My &= \frac{1}{i\sqrt{2}}(Mv - M\overline{v}) = \frac{1}{\sqrt{2}i}(\mu v - \overline{\mu v}) = \frac{1}{2i\sqrt{2}}\big((\mu+\overline{\mu})(v-\overline{v}) + (\mu-\overline{\mu})(v+\overline{v})\big) \\
&= \frac{1}{2}(\mu+\overline{\mu})y + \frac{1}{2i}(\mu-\overline{\mu})x = \mathrm{Re}(\mu)\cdot y + \mathrm{Im}(\mu)\cdot x.
\end{aligned}$$

□

Corollaire 11.7. *Soit* $M \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ *une matrice normale, c'est-à-dire* $M^{\mathrm{tr}} \cdot M = M \cdot M^{\mathrm{tr}}$. *Soient* $\lambda_1, \dots, \lambda_r, \mu_1, \dots, \mu_s, \overline{\mu_1}, \dots, \overline{\mu_s}$ *pour* $n = r + 2s$ *et* $\lambda_1, \dots, \lambda_r \in \mathbb{R}$ *et* $\mu_1, \dots, \mu_s \in \mathbb{C} \setminus \mathbb{R}$ *les coefficients diagonaux de la matrice du corollaire 11.5. On pose* $\alpha_i = \mathrm{Re}(\mu_i)$ *et* $\beta_i = \mathrm{Im}(\mu_i)$ *pour* $1 \le i \le s$.
Alors, il existe une matrice orthogonale $C \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ *telle que*

$$C^{\mathrm{tr}} \cdot M \cdot C = \begin{pmatrix}
\lambda_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & \lambda_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\
0 & \dots & 0 & \lambda_r & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & \dots & \dots & 0 & \alpha_1 & \beta_1 & 0 & 0 & 0 & 0 \\
0 & \dots & \dots & 0 & -\beta_1 & \alpha_1 & 0 & 0 & 0 & 0 \\
0 & \dots & \dots & 0 & 0 & 0 & \ddots & \ddots & 0 & 0 \\
0 & \dots & \dots & 0 & 0 & 0 & \ddots & \ddots & 0 & 0 \\
0 & \dots & \dots & 0 & 0 & 0 & 0 & 0 & \alpha_s & \beta_s \\
0 & \dots & \dots & 0 & 0 & 0 & 0 & 0 & -\beta_s & \alpha_s
\end{pmatrix}.$$

Démonstration. Selon le corollaire 11.5 et le lemme 11.6, nous avons une base orthonormale

$$w_1, w_2, \dots, w_r, v_1, \overline{v_1}, v_2, \overline{v_2}, \dots v_s, \overline{v_s}$$

de $\mathbb{C}^n$ consistant en vecteurs propres pour les valeurs propres

$$\lambda_1, \lambda_2, \dots, \lambda_r, \mu_1, \overline{\mu_1}, \mu_2, \overline{\mu_2}, \dots, \mu_s, \overline{\mu_s}$$

où $n = r + 2s$ et la propriété $w_i \in \mathbb{R}^n$ pour $1 \le i \le r$ est satisfaite. Comme dans le lemme 11.6, posons $x_j = \frac{1}{\sqrt{2}}(v_j + \overline{v_j})$ et $y_j = \frac{1}{i\sqrt{2}}(v_j - \overline{v_j})$.
Alors, $w_1, w_2, \dots, w_r, x_1, y_1, x_2, y_2, \dots, x_s, y_s$ forment une base orthonormale de $\mathbb{R}^n$. Si cette base orthonormale est écrite dans les colonnes d'une matrice C (qui est alors orthogonale), alors $C^{-1}MC$ est de la forme énoncée. Cela suit des calculs dans le lemme 11.6. □

Remarque 11.8. *Soit* $M \in \mathrm{Mat}_{n\times n}(K)$ *pour* $K \in \{\mathbb{R}, \mathbb{C}\}$. *Pour calculer la matrice* C *des corollaires 11.5 et 11.7, on peut utiliser les techniques déjà appris. On procède ainsi :*

(1) Calculer le polynôme caractéristique.

(2) Calculer les valeurs propres dans $\mathbb{C}$ (comme zéros du polynôme caractéristique).

(3) Si $M \in \mathrm{Mat}_{n\times n}(\mathbb{C})$, pour tout $a \in \mathrm{Spec}(M)$, calculer une $\mathbb{C}$-base de $E_M(a)$.

(4) Si $M \in \mathrm{Mat}_{n\times n}(\mathbb{R})$, pour tout $a \in \mathrm{Spec}(M)$ réel, calculer une $\mathbb{R}$-base de $E_M(a)$, et pour tout $a \in \mathrm{Spec}(M)$ non-réel, calculer une $\mathbb{C}$-base de $E_M(a)$.

(5) En utilisant Gram-Schmidt, calculer une base orthonormale de $E_M(a)$ (sur $\mathbb{R}$ si la base d'origine est sur $\mathbb{R}$) pour tout $a \in \mathrm{Spec}(M)$.

Noter que si $a \in \mathbb{C} \setminus \mathbb{R}$ et $M \in \mathrm{Mat}_{n\times n}(\mathbb{R})$, alors on obtient une base orthonormale de $E_M(\overline{a})$ en appliquant la conjugaison complexe à la base orthonormale de $E_M(a)$.

(6) Si $M \in \mathrm{Mat}_{n\times n}(\mathbb{C})$, écrire les vecteurs des bases orthonormales dans les colonnes de la matrice C.

(7) Si $M \in \mathrm{Mat}_{n\times n}(\mathbb{R})$, trier les valeurs propres de M (vu comme matrice à coefficients complexes) ainsi : d'abord les valeurs propres réelles $\lambda_1, \dots, \lambda_r$, puis $\mu_1, \dots, \mu_s, \overline{\mu_1}, \dots, \overline{\mu_s} \in \mathbb{C} \setminus \mathbb{R}$.

Pour tout vecteur v de la base orthonormale d'un espace propre $E_M(\mu_i)$ pour $i = 1, \dots, s$, calculer les vecteurs x, y comme dans le corollaire 11.7 et ainsi obtenir une base orthonormale à coefficients réels de $E_M(\mu_i) \oplus E_M(\overline{\mu_i})$.

Ecrire les vecteurs des bases orthonormales réelles des $E_M(\lambda_i)$ pour $i = 1, \dots, r$ et de $E_M(\mu_i) \oplus E_M(\overline{\mu_i})$ dans les colonnes de la matrice C.

Exemple 11.9. *Nous faisons un exemple concret pour une matrice symétrique. Soit*

$$M = \begin{pmatrix} 14 & 38 & -40 \\ 38 & 71 & 20 \\ -40 & 20 & 5 \end{pmatrix}.$$

Son polynôme caractéristique est $(X+45)(X-45)(X-90)$.
Calculons les espaces propres :

$$E_M(-45) = \ker \begin{pmatrix} 59 & 38 & -40 \\ 38 & 116 & 20 \\ -40 & 20 & 50 \end{pmatrix} = \langle \begin{pmatrix} 2 \\ -1 \\ 2 \end{pmatrix} \rangle,$$

$$E_M(45) = \ker \begin{pmatrix} -31 & 38 & -40 \\ 38 & 26 & 20 \\ -40 & 20 & -40 \end{pmatrix} = \langle \begin{pmatrix} 4 \\ -2 \\ -5 \end{pmatrix} \rangle$$

et

$$E_M(90) = \ker \begin{pmatrix} -76 & 38 & -40 \\ 38 & -19 & 20 \\ -40 & 20 & -85 \end{pmatrix} = \langle \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} \rangle.$$

Ces vecteurs sont déjà orthogonaux par le lemme 11.1. On peut le vérifier facilement pour s'en convaincre. Il suffit donc de les normaliser et de les écrire dans les colonnes d'une matrices :

$$C = \begin{pmatrix} \frac{2}{3} & \frac{4}{3\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{-1}{3} & \frac{-2}{3\sqrt{5}} & \frac{2}{\sqrt{5}} \\ \frac{2}{3} & \frac{-\sqrt{5}}{3} & 0 \end{pmatrix}.$$

Par construction, C est orthogonal, ce qu'on vérifie aussi par un calcul direct. Nous obtenons par construction (à vérifier par un calcul) :

$$C^{\mathrm{tr}} M C = \begin{pmatrix} -45 & 0 & 0 \\ 0 & 45 & 0 \\ 0 & 0 & 90 \end{pmatrix}.$$

Nous pouvons maintenant déduire un résultat plus fort si φ est autoadjoint.

Corollaire 11.10. *Soit $K \in \{\mathbb{R}, \mathbb{C}\}$. Soit $M \in \mathrm{Mat}_{n\times n}(K)$ une matrice. Alors les assertions suivantes sont équivalentes :*

(i) M est autoadjoint (symétrique/hermitien).

(ii) Il existe une isométrie (matrice unitaire/orthogonale) $C \in \mathrm{Mat}_{n\times n}(K)$ telle que $\overline{C}^{\mathrm{tr}} \cdot M \cdot C = \mathrm{diag}(a_1, \dots, a_n)$ avec $a_1, \dots, a_n \in \mathbb{R}$.

Démonstration. « (i) ⇒ (ii) » : Comme M est autoadjoint, il est normal. Nous pouvons donc appliquer le corollaire 11.7. En plus, on obtient $r = n$ et $s = 0$ dans la notation du corollaire, car $\mathrm{Spec}(M) \subset \mathbb{R}$ par le lemme 10.16.
« (ii) ⇒ (i) » : Soit $\overline{C}^{\mathrm{tr}} \cdot M \cdot C = \mathrm{diag}(a_1, \dots, a_n)$, la matrice diagonale avec $a_1, \dots, a_n \in \mathbb{R}$ sur la diagonale. En prenant encore l'adjointe des deux côtés, nous avons $\overline{C}^{\mathrm{tr}} \cdot M \cdot C = \overline{C}^{\mathrm{tr}} \cdot \overline{M}^{\mathrm{tr}} \cdot C$ car la matrice diagonale est invariante. Donc, $M = \overline{M}^{\mathrm{tr}}$. □

Corollaire 11.11. *Soit $K \in \{\mathbb{R}, \mathbb{C}\}$. Soient V un K-espace hermitien de dimension finie et $\varphi \in \mathrm{End}_K(V)$. Alors les assertions suivantes sont équivalentes :*

(i) φ est autoadjoint.

(ii) $V = \bigoplus_{a\in\mathrm{Spec}(\varphi)} E_\varphi(a)$ (en particulier, φ est diagonalisable) et $\mathrm{Spec}(\varphi) \subset \mathbb{R}$.

(iii) V possède une base orthonormale de vecteurs propres pour des valeurs propres réelles pour φ.

Démonstration. Nous allons déduire ce théorème du corollaire 11.10. Pour cela, soit S une base orthonormale de V. Alors, φ est normal/autoadjoint si et seulement si $M := M_{S,S}(\varphi)$ est normal/autoadjoint (cela provient de la proposition 10.6).
« (i) ⇒ (ii) » : Il suffit d'appliquer le corollaire 11.10 à la matrice M.
« (ii) ⇒ (iii) » : Il suffit encore une fois de choisir une base orthonormale dans chaque espace propre.
« (iii) ⇒ (i) » : Soit T la base orthonormale dans l'hypothèse. Soit C la matrice dont les colonnes sont les vecteurs de la base T. Alors, $\overline{C}^{\mathrm{tr}} \cdot M_{S,S}(\varphi) \cdot C$ est diagonal avec des coefficients réels, donc le corollaire 11.10 nous dit que $M_{S,S}(\varphi)$ est autoadjoint, alors φ l'est aussi. □

Corollaire 11.12. *(a) Soit $M \in \mathrm{Mat}_{n\times n}(\mathbb{C})$ une isométrie. Alors il existe une matrice* unitaire $C \in \mathrm{Mat}_{n\times n}(\mathbb{C})$ *telle que $\overline{C}^{\mathrm{tr}} MC$ est diagonal et tous les coefficients sur la diagonale sont de valeur absolue* 1.

(b) Soit $M \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ une isométrie. Alors Il existe une matrice orthogonale $C \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ *telle que*

$$C^{\mathrm{tr}} \cdot M \cdot C = \begin{pmatrix}
\lambda_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & \lambda_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\
0 & \dots & 0 & \lambda_r & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & \dots & \dots & 0 & \cos(\alpha_1) & \sin(\alpha_1) & 0 & 0 & 0 & 0 \\
0 & \dots & \dots & 0 & -\sin(\alpha_1) & \cos(\alpha_1) & 0 & 0 & 0 & 0 \\
0 & \dots & \dots & 0 & 0 & 0 & \ddots & \ddots & 0 & 0 \\
0 & \dots & \dots & 0 & 0 & 0 & \ddots & \ddots & 0 & 0 \\
0 & \dots & \dots & 0 & 0 & 0 & 0 & 0 & \cos(\alpha_s) & \sin(\alpha_s) \\
0 & \dots & \dots & 0 & 0 & 0 & 0 & 0 & -\sin(\alpha_s) & \cos(\alpha_s)
\end{pmatrix}$$

où $\lambda_1, \dots, \lambda_r \in \{-1, 1\}$.

Démonstration. (a) Cela est une conséquence directe du corollaire 11.5 et du lemme 10.16.
(b) Cela suit du corollaire 11.7 et du lemme 10.16 car pour $z \in \mathbb{C}$ de valeur absolue 1 on a $\mathrm{Re}(z) = \cos(\alpha)$ et $\mathrm{Im}(z) = \sin(\alpha)$ si l'on écrit $z = \exp(i\alpha)$. $\square$

La partie (b) est une généralisation de l'exemple 10.17.

Corollaire 11.13. *Soit $K \in \{\mathbb{R}, \mathbb{C}\}$. Soient V un K-espace hermitien de dimension finie et soit $\varphi \in \mathrm{End}_K(V)$ une isométrie.*

(a) Si $K = \mathbb{C}$, alors il existe une $\mathbb{C}$-base orthonormale S de V telle que $M_{S,S}(\varphi)$ est diagonal et tous les coefficients sur la diagonale sont de valeur absolue 1.

(b) Si $K = \mathbb{R}$, alors il existe une $\mathbb{C}$-base orthonormale S de V telle que $M_{S,S}(\varphi)$ est comme dans la partie (b) du corollaire 11.12.

Démonstration. C'est la traduction du corollaire 11.12 au cas des endomorphismes. $\square$

Définition 11.14. *(a) Soit V un K-espace hermitien de dimension finie et soit $\varphi \in \mathrm{End}_K(V)$ autoadjoint. On dit que φ est* (défini) positif *si la forme hermitienne $\langle , \rangle_\varphi$ de la proposition 10.10 est (définie) positive.*

(b) Soit $M \in \mathrm{Mat}_{n\times n}(K)$ une matrice autoadjointe (symétrique (si $K = \mathbb{R}$) ou hermitienne (si $K = \mathbb{C}$)). On dit que M est (défini) positif *si la forme hermitienne $\langle v, w\rangle_M := \overline{v}^{\mathrm{tr}} Mw$ est (définie) positive.*

Lemme 11.15. *Soit V un K-espace hermitien de dimension finie avec base orthonormale S et soit $\varphi \in \mathrm{End}_K(V)$ autoadjoint. Alors :*

(a) *φ est (defini) positif* $\Longleftrightarrow M_{S,S}(\varphi)$ *est (defini) positif.*

(b) *φ est positif* $\Longleftrightarrow \mathrm{Spec}(\varphi) \subseteq \mathbb{R}_{\geq 0}$.

(c) *φ est défini positif* $\Longleftrightarrow \mathrm{Spec}(\varphi) \subseteq \mathbb{R}_{>0}$.

Démonstration. Exercice. □

Lemme 11.16. *Soit $M \in \mathrm{Mat}_{n\times n}(K)$ une matrice autoadjointe (symétrique (si $K = \mathbb{R}$) ou hermitienne (si $K = \mathbb{C}$)) et positive. Alors il existe une matrice positive $N \in \mathrm{Mat}_{n\times n}(K)$ telle que $N^2 = M$ et $NM = MN$. En plus, M est defini positif si et seulement si N l'est aussi.*

Démonstration. Exercice. □

Théorème 11.17 (Décomposition polaire)**.** *Soit V un K-espace hermitien de dimension finie et soit $\varphi \in \mathrm{End}_K(V)$ un isomorphisme (c'est-à-dire un endomorphisme inversible).*
Alors il existe un unique $\psi \in \mathrm{End}_K(V)$ autoadjoint et positif et une unique isométrie $\chi \in \mathrm{End}_K(V)$ tels que $\varphi = \chi \circ \psi$.

Démonstration. Existence : Par un exercice, φ^{ad} est aussi un isomorphisme. Définissons l'isomorphisme $\theta := \varphi^{\mathrm{ad}} \circ \varphi$. Il est autoadjoint :

$$\theta^{\mathrm{ad}} = (\varphi^{\mathrm{ad}} \circ \varphi)^{\mathrm{ad}} = \varphi^{\mathrm{ad}} \circ (\varphi^{\mathrm{ad}})^{\mathrm{ad}} = \varphi^{\mathrm{ad}} \circ \varphi = \theta,$$

donc $\mathrm{Spec}(\theta) \subseteq \mathbb{R}$ par le lemme 10.16. Montrons maintenant qu'il est défini positif :

$$\langle v, v\rangle_\theta = \langle \theta(v), v\rangle = \langle \varphi^{\mathrm{ad}}(\varphi(v)), v\rangle = \langle \varphi(v), \varphi(v)\rangle = |\varphi(v)|^2 > 0$$

pour tout $0 \neq v \in V$. Par le lemme 11.16 il existe donc $\psi \in \mathrm{End}_K(V)$ défini positif tel que $\psi^2 = \theta$. Posons $\chi := \varphi \circ \psi^{-1}$. Pour terminer la démonstration de l'existence il suffit de montrer que χ est une isométrie :

$$\begin{aligned}\chi^{-1} = \psi \circ \varphi^{-1} = \psi^{-1} \circ \psi^2 \circ \varphi^{-1} &= \psi^{-1} \circ \theta \circ \varphi^{-1} \\ &= \psi^{-1} \circ \varphi^{\mathrm{ad}} \circ \varphi \circ \varphi^{-1} = \psi^{-1} \circ \varphi^{\mathrm{ad}} = (\varphi \circ \psi^{-1})^{\mathrm{ad}} = \chi^{\mathrm{ad}}\end{aligned}$$

où nous avons utilisé $(\psi^{-1})^{\mathrm{ad}} = (\psi^{\mathrm{ad}})^{-1} = \psi^{-1}$ car ψ est autoadjoint.
Unicité : Supposons $\varphi = \chi_1 \circ \psi_1 = \chi_2 \circ \psi_2$ pour des isométries χ_1, χ_2 et des isomorphismes autoadjoints définis positifs ψ_1, ψ_2. On obtient

$$\chi_2^{-1} \circ \chi_1 = \psi_2 \circ \psi_1^{-1} =: \beta.$$

A gauche nous avons une isométrie et à droite un endomorphisme autoadjoint defini positif. Il existe donc une base orthogonale S telle que $M_{S,S}(\beta)$ est diagonal, et les coefficients sur la diagonale sont d'un part réels et positifs (car β est autoadjoint positif) et de valeur absolue 1 (car β est une isométrie). C'est donc l'identité $\beta = \mathrm{id}$ d'où $\chi_1 = \chi_2$ et $\psi_1 = \psi_2$. □

12 Quadriques

Objectifs :

- Savoir faire des opérations simultanées sur les lignes et les colonnes ;
- connaître le liens avec les matrices élémentaires ;
- savoir calculer une matrice diagonale en utilisant des opérations simultanées sur les lignes et les colonnes ;
- connaître la définition des quadriques ;
- connaître la définition de l'équivalence de quadriques ;
- connaître la classification des quadriques ;
- savoir calculer le type dans la classification pour quadrique donnée ;
- connaître des exemples et savoir démontrer des propriétés simples.

Le premier but est d'obtenir une matrice diagonale par opérations simultanées sur les lignes et les colonnes d'une matrice donnée (attention : cela n'équivaut pas à la diagonalisation des sections précédentes). Puis, on appliquera les résultats à la classification des « quadriques ».

Opérations simultanées sur les lignes et les colonnes

On reprend l'étude des opérations élémentaires (de l'algorithme de Gauß) sur les lignes et les colonnes (voir la définition 1.39 et la suite), sauf que nous faison maintenant des opérations simultanées sur les lignes et les colonnes, c'est-à-dire que toute opération qu'on fait sur les lignes doit aussi être faite sur les colonnes. Par exemple, si l'on additionne la 3-ème ligne sur la 5-ème ligne, il faut aussi additionner la 3-ème colonne sur la 5-ème colonne. L'avantage est qu'une matrice symétrique restera symétrique. En parallèle avec le lemme 1.40, on a le lemme suivant.

Lemme 12.1. *Soient $\lambda \in K$, $i, j, n \in \mathbb{N}_{>0}$, $i \neq j$ et $M \in \mathrm{Mat}_{n\times n}(K)$.*

(a) $P_{i,j}^{\mathrm{tr}} M P_{i,j}$ est la matrice obtenue à partir de M en échangeant la i-ième et la j-ième ligne ainsi que la i-ième et la j-ième colonne.

(b) $S_i(\lambda)^{\mathrm{tr}} M S_i(\lambda)$ est la matrice obtenue à partir de M en multipliant la i-ième ligne ainsi que la i-ème colonne par λ. En particulier, le coefficient à (i,i) est multiplié par λ^2.

(c) $Q_{i,j}(\lambda)^{\mathrm{tr}} M Q_{i,j}(\lambda)$ est la matrice obtenue de M en additionnant λ fois la i-ième ligne sur la j-ième ligne, ainsi que λ fois la i-ième colonne sur la j-ième colonne.

Démonstration. Il suffit d'appliquer le lemme 1.40. □

Exemple 12.2. *Soit* $M = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{pmatrix}$*. C'est une matrice symétrique. On écrit la matrice augmentée et on fait les opérations sur les lignes et sur les colonnes (seulement sur la moitié droite). On n'a besoin de la moitié gauche que si on veut une matrice réelle* C *telle* CMC^{tr} *(Attention : dans les considérations précédentes, nous avions la transposée à gauche, ici à droite) est égal à la matrice obtenue par les transformation simultanées sur les lignes et les colonnes.*

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

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

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

$$\mapsto \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ -3 & 0 & 1 & 0 & -3 & -1 \\ -1 & 1 & -1/3 & 0 & 0 & 1/3 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ -3 & 0 & 1 & 0 & -3 & 0 \\ -1 & 1 & -1/3 & 0 & 0 & 1/3 \end{pmatrix}$$

$$\mapsto \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ -\sqrt{3} & 0 & 1/\sqrt{3} & 0 & -\sqrt{3} & 0 \\ -\sqrt{3} & \sqrt{3} & -1/\sqrt{3} & 0 & 0 & 1/\sqrt{3} \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ -\sqrt{3} & 0 & 1/\sqrt{3} & 0 & -1 & 0 \\ -\sqrt{3} & \sqrt{3} & -1/\sqrt{3} & 0 & 0 & 1 \end{pmatrix}$$

Notons que le -1 *au milieu de la partie à droite ne peut pas être transformé en* 1 *car on peut seulement multiplier/diviser par des carrés. Soit* C *la moitié gauche de la matrice finale :* $C = \begin{pmatrix} 1 & 0 & 0 \\ -\sqrt{3} & 0 & 1/\sqrt{3} \\ -\sqrt{3} & \sqrt{3} & -1/\sqrt{3} \end{pmatrix}$*. La moitié droite est la matrice obtenue par des opérations simultanées sur les lignes et les colonnes. Par le lemme 12.1, on a l'égalité suivante (pour s'en convaincre, on peut la vérfier par un petit calcul) :*

$$CMC^{\mathrm{tr}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

En posant $D = C^{\mathrm{tr}}$ *on a la transposée à gauche :* $D^{\mathrm{tr}}MD = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

Nous généralisons maintenant ce que nous avons vu dans l'exemple.

Proposition 12.3. *Soit* K *un corps tel que* $1+1 \neq 0$ *et soit* $M \in \mathrm{Mat}_{n\times n}(K)$ *une matrice symétrique. Alors il existe une matrice* $C \in \mathrm{GL}_n(K)$ *tel que* $C^{\mathrm{tr}}MC$ *est une matrice diagonale.*

Démonstration. La démonstration se fait par récurrence sur n. Le cas $n = 1$ est trivial (rien à faire). Supposons la proposition démontrée pour les matrices de taille $n - 1$.

Soit $M = \begin{pmatrix} m_{1,1} & m_{1,2} & \dots & m_{1,n} \\ m_{2,1} & m_{2,2} & \dots & m_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ m_{n,1} & m_{n,2} & \dots & m_{n,n} \end{pmatrix}$. Si M est la matrice nulle, il n'y a rien à faire. Donc supposons M non-nul. On utlisera des transformations simultanées des lignes et des colonnes. On procède en 2 étapes.

(1) Transformer la matrice à ce que $m_{1,1} \neq 0$.

Cas 1 : il existe i tel que $m_{i,i} \neq 0$: Dans ce cas, on échange la i-ème et la première ligne et la i-ème et la première colonne.

Cas 2 : $m_{i,i} = 0$ pour tout $i = 1, \dots, n$: Comme M n'est pas la matrice 0, il existe $i \neq j$ tels que $m_{i,j} \neq 0$. On additionne la i-ème sur la j-ème ligne et la i-ème sur la j-ème colonne. Cela donne $m_{i,j} + m_{j,i} = 2m_{i,j}$ à la place (j, j) et on s'est donc ramené au cas 1.

(2) Par (1), nous avons $m_{1,1} \neq 0$. Pour tout $i = 2, \dots, n$, on additionne $-m_{1,i}/m_{1,1}$ fois la première ligne sur la i-ème ligne et $-m_{1,i}/m_{1,1}$ fois la première colonne sur la i-ème colonne.

Ainsi on obtient une matrice de la forme $\begin{pmatrix} m_{1,1} & 0 & \dots & 0 \\ 0 & m_{2,2} & \dots & m_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & m_{n,2} & \dots & m_{n,n} \end{pmatrix}$.

L'hypothèse de récurrence appliquée au bloc de taille $n - 1$ qui reste finit la démonstration. □

Corollaire 12.4. *Le rang d'une matrice est invariant par des opérations simultanées sur les lignes et les colonnes.*

Démonstration. Supposons que N est obtenue de M par des opérations simultanées sur les lignes et les colonnes. Par la proposition 12.3 on a $C^{\mathrm{tr}}MC = N$ pour une matrice inversible C. Comme C^{tr} est aussi inversible (par exemple, car $0 \neq \det(C) = \det(C^{\mathrm{tr}})$), nous avons $\mathrm{rk}(N) = \mathrm{rk}(C^{\mathrm{tr}}MC) = \dim(\mathrm{im}(C^{\mathrm{tr}}MC)) = \dim(\mathrm{im}(C^{\mathrm{tr}}M)) = \dim(C^{\mathrm{tr}}(\mathrm{im}(M)) = \dim(\mathrm{im}(M)) = \mathrm{rk}(M)$. □

Quadriques

Dans toute cette section, soit K un corps tel que $1+1 \neq 0$, par exemple $K = \mathbb{R}$ ou $K = \mathbb{C}$. Rappelons d'abord que $K[X_1, X_2, \dots, X_n]$ note l'anneau des polynômes dans les variables $X_1, X_2, \dots, X_n$ à coefficients dans K. Un élément de $K[X_1, X_2, \dots, X_n]$ est de la forme

$$\sum_{i_1=0}^{d_1} \sum_{i_2=0}^{d_2} \dots \sum_{i_n=0}^{d_n} a_{i_1,i_2,\dots,i_n} X_1^{i_1} X_2^{i_2} \dots X_n^{i_n}.$$

Dans la suite, nous ne regardons que des polynômes quadratiques.

Définition 12.5. *On appelle* polynôme quadratique (à n variable et à coefficients dans K) *tout élément de* $K[X_1, X_2, \dots, X_n]$ *de la forme*

$$q(X_1, X_2, \dots, X_n) = \sum_{1 \le i \le j \le n} a_{i,j} X_i X_j + \sum_{i=1}^{n} a_{0,i} X_i + a_{0,0}.$$

Exemple 12.6. *(a) Soit* $n = 1$*. On nomme la variable* X*. Tout polynôme quadratique est de la forme*

$$a_{1,1} X^2 + a_{0,1} X + a_{0,0} = a_2 X^2 + a_1 X + a_0$$

où nous avons rénumeroté les coefficients de façon habituelle.

(b) Soit $n = 2$*. On nomme les variables* X, Y*. Tout polynôme quadratique est de la forme*

$$a_{1,1} X^2 + a_{1,2} XY + a_{2,2} Y^2 + a_{0,1} X + a_{0,2} Y + a_{0,0}.$$

En particulier, on a les exemples :

(1) $\frac{X^2}{a^2} + \frac{Y^2}{b^2} - 1$
(2) $\frac{X^2}{a^2} - \frac{Y^2}{b^2} - 1$
(3) $\frac{X^2}{a^2} - Y$

Lemme 12.7. *Soit* $n \in \mathbb{N}$ *et soit* $A \in \mathrm{Mat}_{(n+1)\times(n+1)}(K)$ *une matrice symétrique. Ses coefficients seront appelés* $a_{i,j}$ *pour* $0 \le i, j \le n$ *(noter qu'on commence la numérotation à* 0 *!). Soit* $\tilde{X}$ *le vecteur des variables précedé de* 1 *:*

$$A = \begin{pmatrix} a_{0,0} & a_{0,1} & \dots & a_{0,n} \\ a_{0,1} & a_{1,1} & \dots & a_{1,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{0,n} & a_{1,n} & \dots & a_{n,n} \end{pmatrix}, \quad \tilde{X} = \begin{pmatrix} 1 \\ X_1 \\ \vdots \\ X_n \end{pmatrix}.$$

Alors le polynôme

$$q_A(X_1, \dots, X_n) = \tilde{X}^{\mathrm{tr}} A \tilde{X} = 2 \sum_{1 \le i < j \le n} a_{i,j} X_i X_j + \sum_{i=1}^{n} a_{i,i} X_i^2 + 2 \sum_{i=1}^{n} a_{0,i} X_i + a_{0,0}$$

est quadratique et tout polynôme quadratique provient d'une unique matrice symétrique A *par cette formule.*

Démonstration. C'est clair. $\square$

Comme dans le lemme précédent, pour $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in K^n$, on note $\tilde{x} = \begin{pmatrix} 1 \\ x_1 \\ \vdots \\ x_n \end{pmatrix}$, le vecteur x précedé par 1.

Définition 12.8. *On appelle* quadrique (en dimension n) *tout ensemble*

$$Q_A := Q_A(K) := \{x \in K^n \mid \tilde{x}^{\mathrm{tr}} A \tilde{x} = 0\} = \{x \in K^n \mid q_A(x) = 0\}$$

où A *est une matrice symétrique dans* $\mathrm{Mat}_{(n+1)\times(n+1)}(K)$.

Exemple 12.9. *On considère $n = 2$.*

(1) Soit $A = \begin{pmatrix} -1 & 0 & 0 \\ 0 & \frac{1}{a^2} & 0 \\ 0 & 0 & \frac{1}{b^2} \end{pmatrix}$*. On a* $Q_A = \{x \in \mathbb{R}^2 \mid \frac{X^2}{a^2} + \frac{Y^2}{b^2} - 1 = 0\}$*. Géométriquement, il s'agit d'une ellipse.*

(2) Soit $A = \begin{pmatrix} -1 & 0 & 0 \\ 0 & \frac{1}{a^2} & 0 \\ 0 & 0 & \frac{-1}{b^2} \end{pmatrix}$*. On a* $Q_A = \{x \in \mathbb{R}^2 \mid \frac{X^2}{a^2} - \frac{Y^2}{b^2} - 1 = 0\}$*. Géométriquement, il s'agit d'une hyperbole.*

(3) Soit $A = \begin{pmatrix} 0 & 0 & \frac{-1}{2} \\ 0 & \frac{1}{a^2} & 0 \\ \frac{-1}{2} & 0 & 0 \end{pmatrix}$*. On a* $Q_A = \{x \in \mathbb{R}^2 \mid \frac{X^2}{a^2} - Y = 0\}$*. Géométriquement, il s'agit d'une parabole.*

Nous définissons aussi une matrice augmentée : soit $C = (c_{i,j}) \in \mathrm{Mat}_{n\times n}(K)$ une matrice et $y \in K^n$ un vecteur. Nous posons :

$$\widetilde{C_y} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ y_1 & c_{1,1} & \dots & c_{1,n} \\ \vdots & \vdots & \ddots & \vdots \\ y_n & c_{n,1} & \dots & c_{n,n} \end{pmatrix}.$$

Lemme 12.10. *Soit $A \in \mathrm{Mat}_{(n+1)\times(n+1)}(K)$ une matrice symétrique et Q_A la quadrique associée. Soit $\varphi : K^n \to K^n$ une* affinité, *c'est-à-dire, une application de la forme*

$$\varphi(v) = Bv + By$$

où $B \in \mathrm{GL}_n(K)$ et $y \in K^n$. Soit $\tilde{C} := \widetilde{(B^{-1})}_{-y} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ -y_1 & c_{1,1} & \dots & c_{1,n} \\ \vdots & \vdots & \ddots & \vdots \\ -y_n & c_{n,1} & \dots & c_{n,n} \end{pmatrix}.$

Alors $\varphi(Q_A) = Q_{\tilde{C}^{\mathrm{tr}} A \tilde{C}}$. L'image d'une quadrique par une affinité est donc aussi une quadrique.

Démonstration. Tout repose sur l'égalité

$$\tilde{C}\widetilde{\varphi(x)} = \tilde{C}(\widetilde{Bx + By}) = (-\widetilde{y + x} + y) = \tilde{x}.$$

Nous avons donc l'égalité

$$\tilde{x}^{\mathrm{tr}} A \tilde{x} = (\tilde{C}\widetilde{\varphi(x)})^{\mathrm{tr}} A (\tilde{C}\widetilde{\varphi(x)}) = \widetilde{\varphi(x)}^{\mathrm{tr}} \big(\tilde{C}^{\mathrm{tr}} A \tilde{C}\big) \widetilde{\varphi(x)},$$

d'où $x \in Q_A \Leftrightarrow \varphi(x) \in Q_{\tilde{C}^{\mathrm{tr}} A \tilde{C}}$. Cela implique le résultat. □

Définition 12.11. *Soient $q_1(X_1, \dots, X_n)$ et $q_2(X_1, \dots, X_n)$ des polynômes quadratiques provenant de matrices symétriques $A, B \in \mathrm{Mat}_{(n+1)\times(n+1)}(K)$, c'est-à-dire $q_1 = q_A$, $q_2 = q_B$.*
On dit que $q_1(X_1, \dots, X_n)$ et $q_2(X_1, \dots, X_n)$ sont équivalents *s'il existe $C \in \mathrm{GL}_n(K)$, $y \in K^n$ et $0 \neq x \in K$ tels que $\widetilde{C_y}^{\mathrm{tr}} A \widetilde{C_y} = xB$. Dans ce cas on dit aussi que Q_A et Q_B sont équivalents.*

On a donc par le lemme 12.10 que si $q_A(X_1,\dots,X_n)$ et $q_B(X_1,\dots,X_n)$ sont équivalents, alors existe une affinité $\varphi : K^n \to K^n$ telle que $\varphi(Q_A) = Q_B$, notamment $\varphi(v) = C^{-1}v + C^{-1}y$ dans la notation de la définition précédente.
Notre prochain but est de caractériser les quadriques à équivalence près. Pour cela, nous avons besoin de la définition suivante.

Définition 12.12. *On appelle* système de représentants de $K^\times$ modulo les carrés *tout ensemble* $R \in K \setminus \{0\}$ *qui satisfait que pour tout* $x \in K^\times$ *il existe un unique* $r \in R$ *et un* $y \in K$ *tels que* $x = r \cdot y^2$.

Exemple 12.13. *(a) Si* $K = \mathbb{C}$, *alors* $R = \{1\}$ *est un système de représentants de* $\mathbb{C}^\times$ *modulo les carrés. En effet, tout élément de* $\mathbb{C}$ *est un carré.*

(b) Si $K = \mathbb{R}$, *alors* $R = \{-1, 1\}$ *est un système de représentants de* $\mathbb{R}^\times$ *modulo les carrés. En effet, tout élément positif de* $\mathbb{R}$ *est un carré, et tout élément négatif est moins un carré.*

(c) On appelle sans carré *tout entier* $m \in \mathbb{Z}$ *tel qu'il n'est divisible par aucun carré d'un nombre premier. Soit* $R = \{m \in \mathbb{Z} \mid m \text{ est sans carré }\}$.

Si $K = \mathbb{Q}$, *alors* R *est un système de représentants de* $\mathbb{Q}^\times$ *modulo les carrés. En effet, on peut écrire*

$$\frac{a}{b} = ab\frac{1}{b^2} = m\left(\frac{q}{b}\right)^2$$

où $ab = mq^2$ *pour* $m \in \mathbb{Z}$ *sans carré. En plus, si* $m = m'\left(\frac{r}{s}\right)^2$ *pour* m, m' *sans carré, alors* $m' \mid m$ *; de la même façon,* $m \mid m'$ *; comme* m *et* m' *ont le même signe, on obtient* $m = m'$, *l'unicité.*

Dans le théorème de la classification des quadriques, nous utiliserons les notations suivantes : Pour $n \in \mathbb{N}$, les coefficients des matrices symétriques $A \in \mathrm{Mat}_{(n+1)\times(n+1)}(K)$ seront numérotés ainsi :

$$A = \begin{pmatrix} a_{0,0} & a_{0,1} & \dots & a_{0,n} \\ a_{0,1} & a_{1,1} & \dots & a_{1,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{0,n} & a_{1,n} & \dots & a_{n,n} \end{pmatrix}.$$

Soit A_n le bloc de taille $n \times n$ de A en bas à droite :

$$A_n = \begin{pmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{1,n} & \dots & a_{n,n} \end{pmatrix}.$$

En vue de la définition de l'équivalence de deux polynômes quadratiques (et de l'équivalence de deux quadriques), nous avons le droit d'appliquer les opérations suivantes :

- Additionner la i-ème ligne ($i \geq 2$) sur toute autre ligne (attention : pas valable pour $i = 1$).
- Echanger la i-ème et la j-ème ligne pour $i, j \geq 2$ (attention : pas valable si $i = 1$ ou $j = 1$).
- Multiplier la i-ème ligne par un scalaire non-zéro (attention : pas valable pour $i = 1$).

- Multiplier toute la matrice par un scalaire non-zéro.

Lemme 12.14. *Soient $A \in \mathrm{Mat}_{(n+1)\times(n+1)}(K)$ symétrique, $C \in \mathrm{GL}_n(K)$ et $y \in K^n$. Alors*

$$\left(\widetilde{C_y}^{\mathrm{tr}} A \widetilde{C_y}\right)_n = C^{\mathrm{tr}} A_n C.$$

En particulier, le rang de A_n est égal au rang de $\left(\widetilde{C_y}^{\mathrm{tr}} A \widetilde{C_y}\right)_n$. Donc le rang de A_n est invariant par équivalence de polynômes quadratiques.

Démonstration. Les faits que la première colonne de $\widetilde{C_y}^{\mathrm{tr}}$ est le vecteur $\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ et que la première ligne de $\widetilde{C_y}$ est le vecteur $\begin{pmatrix} 1 & 0 & \dots & 0 \end{pmatrix}$ montrent le résultat. □

Théorème 12.15 (Classification des quadriques)**.** *Soit R un système de représentants de $K^\times$ modulo les carrés. Soit $q_A(X_1, \dots, X_n)$ le polynôme quadratique associé à la matrice symétrique $A \in \mathrm{Mat}_{(n+1)\times(n+1)}(K)$. Soit r le rang de la matrice A_n.*
Nous avons les trois cas :

(I) Si $\mathrm{rk}(A) = r$, alors il existe des $a_2, a_3, \dots, a_r \in R$ tels que $q_A(X_1, \dots, X_n)$ est équivalent à $X_1^2 + a_2X_2^2 + a_3X_3^2 + \dots + a_rX_r^2$.

(II) Si $\mathrm{rk}(A) = r + 1$, alors il existe des $a_1, a_2, \dots, a_r \in R$ tels que $q_A(X_1, \dots, X_n)$ est équivalent à $a_1X_1^2 + a_2X_2^2 + \dots + a_rX_r^2 + 1$.

(III) Si $\mathrm{rk}(A) = r + 2$, alors $r \le n - 1$ et il existe des $a_1, a_2, \dots, a_r \in R$ tels que $q_A(X_1, \dots, X_n)$ est équivalent à $a_1X_1^2 + a_2X_2^2 + \dots + a_rX_r^2 + 2X_{r+1}$.

Démonstration. Pour obtenir ces formes spéciales, nous avons le droit d'utiliser seulement ces opérations simultanées sur les lignes et les colonnes qui correspondent aux matrices $\widetilde{C_y}$ avec C une des matrices de la définition 1.39 et $y \in K^n$ n'importe quel vecteur.
On procède en plusieurs étapes :

(1) A cause du lemme 12.14, la proposition 12.3 montre qu'en utilisant des matrices $\widetilde{C_0}$ la matrice A peut être transformée en

$$B = \begin{pmatrix} b_{0,0} & b_{0,1} & \dots & b_{0,r} & b_{0,r+1} & \dots & b_{0,n} \\ b_{0,1} & b_{1,1} & 0 & 0 & 0 & \dots & 0 \\ \vdots & 0 & \ddots & 0 & 0 & \ddots & 0 \\ b_{0,r} & 0 & \dots & b_{r,r} & 0 & \dots & 0 \\ b_{0,r+1} & 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \ddots & \vdots \\ b_{0,n} & 0 & \dots & 0 & 0 & \dots & 0 \end{pmatrix}$$

pour $b_{i,i} \neq 0$ pour $1 \le i \le r$ de façon à ce que q_A et q_B soient équivalents.

(2) Noter qu'additionner la i-ème ligne (pour $i > 1$) sur la première correspond à la matrice $\widetilde{\mathrm{id}_{e_{i-1}}}^{\mathrm{tr}}$ où e_{i-1} est le $(i-1)$-ème vecteur canonique. Nous pouvons donc transformer notre matrice plus pour obtenir

$$B = \begin{pmatrix} b_{0,0} & 0 & \dots & 0 & b_{0,r+1} & \dots & b_{0,n} \\ 0 & b_{1,1} & 0 & 0 & 0 & \dots & 0 \\ \vdots & 0 & \ddots & 0 & 0 & \ddots & 0 \\ 0 & 0 & \dots & b_{r,r} & 0 & \dots & 0 \\ b_{0,r+1} & 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \ddots & \vdots \\ b_{0,n} & 0 & \dots & 0 & 0 & \dots & 0 \end{pmatrix}.$$

(3) C'est ici où la distinction des cas intervient.

(I) Supposons $b_{0,0} = b_{0,r+1} = b_{0,r+2} = \cdots = b_{0,n} = 0$. Dans ce cas le rang de B (qui est égal au rang de A) est égal à r. On pourra encore diviser par $b_{1,1}$ (à cause de l'élément $0 \neq x \in K$ dans la définition de l'équivalence) pour obtenir

$$B = \begin{pmatrix} 0 & 0 & \dots & \dots & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & 0 & b_{2,2} & \ddots & \vdots & 0 & \ddots & 0 \\ 0 & 0 & 0 & \ddots & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & b_{r,r} & 0 & \dots & 0 \\ 0 & 0 & \dots & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \end{pmatrix}.$$

Finalement, par la muliplication de la i-ème colonne et la i-ème ligne pour $2 \leq i \leq r$ par un élément convenable a dans K (ce qui revient à multiplier $b_{i,i}$ par a^2) on peut choisir $b_{i,i}$ dans R. Maintenant q_B est précisement de la forme de (I) dans l'assertion.

(II) Supposons $b_{0,r+1} = b_{0,r+2} = \cdots = b_{0,n} = 0$, mais $b_{0,0} \neq 0$. Dans ce cas le rang de B (qui est égal au rang de A) est égal à $r+1$. Après division par $b_{0,0}$, on obtient

$$B = \begin{pmatrix} 1 & 0 & \dots & \dots & 0 & 0 & \dots & 0 \\ 0 & b_{1,1} & 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & 0 & b_{2,2} & \ddots & \vdots & 0 & \ddots & 0 \\ 0 & 0 & 0 & \ddots & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & b_{r,r} & 0 & \dots & 0 \\ 0 & 0 & \dots & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \end{pmatrix}.$$

Comme dans (I), on peut aussi forcer $b_{i,i} \in R$ pour $1 \leq i \leq r$. Maintenant q_B est précisement de la forme de (II) dans l'assertion.

(III) Supposons qu'il existe $r+1 \leq i \leq n$ tel que $b_{0,i} \neq 0$. En échangeant des lignes et des colonnes simultanément, on peut d'abord obtenir $b_{0,r+1} \neq 0$. En divisant la matrice par $b_{0,r+1}$, on peut donc mettre ce coefficient égal à 1. En additionnant $-b_{0,j}$ fois la $(r+1)$-ème ligne sur la j-ème pour $r+2 \leq j \leq n$ (ce qui correspond à la matrice $(\widetilde{Q_{r,j-1}})_0^{\mathrm{tr}}$) on arrive à annuler $b_{0,j}$ pour ces j. On a donc la matrice

$$B = \begin{pmatrix} 0 & 0 & \dots & 0 & 0 & 1 & 0 & \dots & 0 \\ 0 & b_{1,1} & 0 & 0 & 0 & 0 & \dots & \dots & 0 \\ \vdots & 0 & b_{2,2} & 0 & \dots & 0 & \ddots & \ddots & 0 \\ 0 & 0 & 0 & \ddots & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & b_{r,r} & 0 & 0 & \dots & 0 \\ 1 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & \dots & 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & 0 & \ddots & \vdots \\ 0 & 0 & \dots & \dots & 0 & 0 & 0 & \dots & 0 \end{pmatrix}.$$

On voit que le rang de B est égal à $r+2$. Comme dans (I) et (II), on peut aussi forcer $b_{i,i} \in R$ pour $1 \leq i \leq r$. Maintenant q_B est précisement de la forme de (III) dans l'assertion.

Cela finit la démonstration. $\square$

Corollaire 12.16. *Soit $K = \mathbb{C}$. Soit $q(X_1, \dots, X_n) \in \mathbb{C}[X_1, \dots, X_n]$ un polynôme quadratique non-nul. Alors il est équivalent à un unique polynôme parmi les $3n-1$ polynômes suivants :*

(I) $X_1^2 + \dots + X_r^2$ *pour* $1 \leq r \leq n$ *;*

(II) $X_1^2 + \dots + X_r^2 + 1$ *pour* $1 \leq r \leq n$ *;*

(III) $X_1^2 + \dots + X_r^2 + 2X_{r+1}$ *pour* $1 \leq r \leq n-1$.

Démonstration. On sait que $R = \{1\}$ est un système de représentants pour $\mathbb{C}^\times$ modulo les carrés. Donc le théorème 12.15 implique que q est équivalent à un des polynômes dans la liste. L'unicité provient du fait que dans ce cas le rang ensemble avec le type ((I), (II), (III)) est assez pour uniquement caractériser le polynôme. $\square$

Notre prochain but est une classification explicite des quadriques réelles. Pour cela, il nous faut encore montrer le théorème suivant de Sylvester. D'abord nous avons besoin d'un lemme.

Lemme 12.17. *Soit $A \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ une matrice symétrique et soit $\langle v, w\rangle_A := \langle Av, w\rangle$ la forme symétrique définie par A sur $\mathbb{R}^n$.*

(a) Il existe des sous-espaces $V_+, V_-, V_0 \leq \mathbb{R}^n$ tels que

- $\mathbb{R}^n = V_+ \bigoplus V_- \bigoplus V_0$,
- *pour tout* $0 \neq v \in V_+$*, on a* $\langle v, v\rangle_A > 0$,
- *pour tout* $0 \neq v \in V_-$*, on a* $\langle v, v\rangle_A < 0$ *et*

- *pour tout* $0 \neq v \in V_0$*, on a* $\langle v, v\rangle_A = 0$*.*

(b) Si V_+, V_-, V_0 *sont des sous-espaces ayant les propriétés dans (a), alors*

- $\dim V_+$ *est le nombre de valeurs propres positives de* A*,*
- $\dim V_-$ *est le nombre de valeurs propres négatives de* A *et*
- $\dim V_0$ *est le nombre de valeurs propres* 0 *de* A*.*

On doit compter les valeurs propres avec multiplicités, c'est-à-dire le nombre de fois la valeur propre apparaît sur la diagonale après diagonalisation.

Démonstration. Par le théorème spectral, nous avons une base orthonormale

$$v_1, \dots, v_s, v_{s+1}, \dots, v_r, v_{r+1}, \dots, v_n$$

de $\mathbb{R}^n$ telle que v_i pour $1 \leq i \leq s$ sont des vecteurs propres pour une valeur propre positive, v_i pour $s+1 \leq i \leq r$ sont des vecteurs propres pour une valeur propre négative et v_i pour $s+1 \leq i \leq r$ sont des vecteurs propres pour la valeur propre 0. On prend V_+ comme le sous-espace engendré par $v_1, \dots, v_s$ et V_- comme celui engendré par $v_{s+1}, \dots, v_r$ et V_0 comme celui engendré par $v_{r+1}, \dots, v_n$. Il est clair que toutes les propriétés de (a) et (b) sont satisfaites pour ces espaces.
Soient maintenant V'_+, V'_-, V'_0 d'autres espaces ayant les propriétés de (a). On montre que $V_+ \cap (V'_- \oplus V'_0) = 0$: si $0 \neq v = w_- + w_0$ pour $w_- \in V'_-$ et $w_0 \in V'_0$ était un vecteur dans l'intersection, on aurait d'un part $\langle v, v\rangle_A > 0$ et d'autre part $\langle w_- + w_0, w_- + w_0\rangle_A = \langle w_-, w_-\rangle_A + \langle w_0, w_0\rangle_A \leq 0$. Donc $v = 0$. Cela montre que $V_+ \oplus V'_- \oplus V'_0$ est un sous-espace de $\mathbb{R}^n$, donc $\dim V_+ \leq \dim V'_+$. Par symétrie, on a aussi $\dim V'_+ \leq \dim V_+$, donc l'égalité. Les arguments pour les deux autres égalités sont similairs. □

Théorème 12.18 (Sylvester)**.** *Soit* $A \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ *une matrice symétrique et soit* $C \in \mathrm{GL}_n(\mathbb{R})$*. Alors,* A *et* $C^{\mathrm{tr}}AC$ *possèdent le même nombre de valeurs propres positives. La même assertion est vraie pour les valeurs propres négatives.*

Démonstration. On utilise la notation du lemme 12.17 pour la forme bilinéaire $\langle,\rangle_A$. Faisons d'abord le calcul général :

$$\langle Cv, Cw\rangle_A = \langle ACv, Cw\rangle = \langle C^{\mathrm{tr}}ACv, w\rangle = \langle v, w\rangle_{C^{\mathrm{tr}}AC}$$

pour $v, w \in \mathbb{R}^n$. Considérons $C^{-1}V_+$. Si $0 \neq v \in C^{-1}V_+$ (donc $Cv \in V_+$), alors le calcul ci-dessus donne $0 < \langle v, v\rangle_{C^{\mathrm{tr}}AC}$. En plus, si $w \in C^{-1}V_-$, alors on obtient de la même manière $0 = \langle v, w\rangle_{C^{\mathrm{tr}}AC}$. Faisant des arguments similaires, on obtient que $C^{-1}V_+$, $C^{-1}V_-$, $C^{-1}V_0$ sont des sous-espaces qui satisfont aux propriétés dans (a) du lemme 12.17 pour la forme bilinéaire $\langle,\rangle_{C^{\mathrm{tr}}AC}$. Donc la dimension de V_+ (qui est le nombre de valeurs propres positives pour A) est égal au nombre de valeurs propres positives pour $C^{\mathrm{tr}}AC$. L'argument pour les valeurs propres négatives est le même. □

Corollaire 12.19. *Soit* $K = \mathbb{R}$*. Soit* $q(X_1, \dots, X_n) \in \mathbb{R}[X_1, \dots, X_n]$ *un polynôme quadratique non-nul. Alors il est équivalent à un unique polynôme parmi les* $\frac{3n^2+5n}{2} - 1$ *polynômes suivants :*

(I) $X_1^2 + \dots + X_s^2 - X_{s+1}^2 - \dots - X_r^2$ *pour* $1 \leq s \leq r \leq n$ *;*

(II) $X_1^2 + \cdots + X_s^2 - X_{s+1}^2 - \cdots - X_r^2 + 1$ *pour* $0 \leq s \leq r \leq n$, $1 \leq r$ *;*

(III) $X_1^2 + \cdots + X_s^2 - X_{s+1}^2 - \cdots - X_r^2 + 2X_{r+1}$ *pour* $0 \leq s \leq r \leq n-1$, $1 \leq r$.

Démonstration. On sait que $R = \{-1, 1\}$ est un système de représentants pour $\mathbb{R}^\times$ modulo les carrés. Donc le théorème 12.15 implique que q est équivalent à un des polynômes dans la liste. L'unicité provient du fait que la différence entre le rang de la grande matrice et celui du bloc de taille n en bas à droite détermine le type ((I), (II), (III)). Donc il suffit de connaîtr le nombre de valeurs propres positives (et négatives), à cause du théorème de Sylvester 12.18.
Le nombre de polynômes de type (I) de rang r est égal à r (le signe devant X_1 est toujours $+$), donc il existe $1 + 2 + \cdots + n = \frac{n(n+1)}{2}$ polynômes de type (I). Le nombre de polynômes de type (II) de rang r est égal à $r + 1$ (le signe devant X_1 peut être 1 ou -1), donc il existe $2 + 3 + \cdots + (n + 1) = \frac{(n+1)(n+2)}{2} - 1$ polynômes de type (II). Similairement, Le nombre de polynômes de type (III) de rang r est égal à $r + 1$, mais r est borné par $n - 1$, donc il existe $2 + 3 + \cdots + n = \frac{n(n+1)}{2} - 1$ polynômes de type (III). On obtient donc

$$\frac{n(n+1)}{2} + \frac{(n+1)(n+2)}{2} - 1 + \frac{n(n+1)}{2} - 1 = \frac{3n^2 + 5n}{2} - 1,$$

le nombre dans l'assertion. □

13 Dualité

Objectifs :

- Maîtriser les concepts d'espace dual et d'application duale ;
- connaître la relation avec les matrices transposées ;
- connaître la définition et les propriétés fondamentales des forme bilinéaires ;
- connaître l'application aux rangs des lignes et des colonnes de matrices ;
- connaître des exemples et savoir démontrer des propriétés simples.

Dans cette section, nous introduisons une théorie de dualité valable pour tout corps K (non seulement pour $\mathbb{R}$ et $\mathbb{C}$).
Les résultats principaux de cette section sont

- l'interprétation des matrices transposées comme les matrices qui représentent des applications « duales » ;
- le rang des colonnes d'une matrice est égale au rang des lignes ; cela est parfois utile pour des calculs.

Nous commençons par l'interprétation des matrices transposées comme les matrices représentant les applications duales. Pour cela nous introduisons d'abord l'espace vectoriel dual V^* d'un espace vectoriel V.

Lemme 13.1. *Soient V, W deux K-espaces vectoriels.*

(a) L'ensemble des applications K-linéaires

$$\mathrm{Hom}_K(V, W) := \{f : V \to W \mid f \text{ est } K\text{-linéaire } \}$$

est un K-espace vectoriel pour l'addition

$$(f + g)(v) := f(v) + g(v) \text{ pour } f, g \in \mathrm{Hom}_K(V, W) \text{ et } v \in V$$

et la multiplication scalaire

$$(x.f)(v) := x.(f(v)) = f(x.v) \text{ pour } f \in \mathrm{Hom}_K(V, W),\ x \in K \text{ et } v \in V.$$

(b) Soient S une K-base de V et $f : S \to W$ une application. Alors, il existe un unique $F \in \mathrm{Hom}_K(V, W)$ tel que $F|_S = f$, notamment $F(\sum_{s\in S} a_s s) = \sum_{s\in S} a_s f(s)$.

Démonstration. Calculs simples. □

Définition 13.2. *Soit V un K-espace vectoriel. Le K-espace vectoriel (voir le lemme 13.1(a))*

$$V^* := \mathrm{Hom}_K(V, K)$$

est appelé l'espace dual de V.

Proposition 13.3. *Soit V un K-espace vectoriel de dimension finie n.*

(a) Soit $S = \{s_1, \ldots, s_n\}$ une K-base de V. Pour tout $1 \le i \le n$, soit s_i^* l'unique (par le lemme 13.1(b)) élément dans V^* tel que pour tout $1 \le j \le n$ on a $s_i^*(s_j) = \delta_{i,j} = \begin{cases} 1 & \text{si } i = j, \\ 0 & \text{si } i \neq j. \end{cases}$

Alors, $S^* := \{s_1^*, \ldots, s_n^*\}$ est une K-base de V^*, appelé la base duale.

(b) Si V est de K-dimension finie, alors $\dim_K(V^*) = \dim_K(V)$.

Démonstration. (a) Indépendance K-linéaire : Soit $0 = \sum_{i=1}^n a_i s_i^*$ avec $a_1, \ldots, a_n \in K$. Alors, pour tout $1 \le j \le n$ nous avons

$$0 = \sum_{i=1}^n a_i s_i^*(s_j) = \sum_{i=1}^n a_i \delta_{i,j} = a_j.$$

Génération : Soit $f \in V^*$. Pour $1 \le j \le n$, posons $a_j := f(s_j)$ et $g := \sum_{i=1}^n a_i s_i^* \in V^*$. Nous avons

$$g(s_j) = \sum_{i=1}^n a_i s_i^*(s_j) = a_j = f(s_j)$$

pour tout $1 \le j \le n$, donc $f = g$.

(b) La dimension de V est le cardinal de toute base pour V. Par (a), la base duale possède le même cardinal que toute base de V, donc la dimension de V^* est égale à celle de V. □

Définition-Lemme 13.4. *Soient V, W deux K-espaces vectoriels et $\varphi : V \to W$ une application K-linéaire. Alors, l'application*

$$\varphi^* : W^* \to V^*, \quad f \mapsto \varphi^*(f) = f \circ \varphi$$

est K-linéaire. Elle est appelée l'application duale de φ.

Démonstration. D'abord il faut observer que $\varphi \circ f$ est K-linéaire ; mais, cela résulte du fait que la composition de deux applications linéaires est linéaire. Soient $f, g \in W^*$ et $x \in K$. Nous concluons la preuve par le calcul

$$\begin{aligned}\varphi^*(x \cdot f + g)(v) = ((x \cdot f + g) \circ \varphi)(v) = (x \cdot f + g)(\varphi(v))\\ = xf(\varphi(v)) + g(\varphi(v)) = (x\varphi^*(f) + \varphi^*(g))(v).\end{aligned}$$

pour tout $v \in V$, donc $\varphi^*(x \cdot f + g) = x\varphi^*(f) + \varphi^*(g)$. $\square$

Proposition 13.5. *Soient V, W deux K-espaces vectoriels et $\varphi : V \to W$ une application K-linéaire. Soient en plus $S = \{s_1, \ldots, s_n\}$ une K-base de V et $T = \{t_1, \ldots, t_m\}$ une K-base de W. Alors,*

$$\big(M_{T,S}(\varphi)\big)^{\mathrm{tr}} = M_{S^*,T^*}(\varphi^*).$$

Donc, la matrice qui représente φ^* pour les bases duales est la matrice transposée de la matrice qui représente φ.

Démonstration. Nous écrivons

$$M_{T,S}(\varphi) = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{pmatrix} \text{ et } M_{S^*,T^*}(\varphi^*) = \begin{pmatrix} b_{1,1} & b_{1,2} & \cdots & b_{1,m} \\ b_{2,1} & b_{2,2} & \cdots & b_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n,1} & b_{n,2} & \cdots & b_{n,m} \end{pmatrix}.$$

Cela signifie

$$\varphi(s_j) = \sum_{i=1}^{m} a_{i,j} t_i \text{ et } \varphi^*(t_k^*) = \sum_{i=1}^{n} b_{i,k} s_i^*$$

pour tout $1 \leq j \leq n$ et $1 \leq k \leq m$. Donc, d'un côté

$$(\varphi^*(t_k^*))(s_j) = t_k^*(\varphi(s_j)) = t_k^*(\sum_{i=1}^{m} a_{i,j} t_i) = \sum_{i=1}^{m} a_{i,j} t_k^*(t_i) = a_{k,j}$$

et de l'autre côté

$$(\varphi^*(t_k^*))(s_j) = \sum_{i=1}^{n} b_{i,k} s_i^*(s_j) = b_{j,k},$$

d'où $a_{k,j} = b_{j,k}$, comme énoncé. $\square$

L'espace dual donne lieu à une forme bilinéaire naturelle, comme nous allons le voir dans l'exemple 13.8(b) ; d'abord nous faisons les définitions nécessaires.

Définition 13.6. *Soient V, W deux K-espaces vectoriels. On appelle* forme bilinéaire *toute application*

$$\langle \cdot, \cdot \rangle : V \times W \to K$$

telle que

- $\forall\, a \in K \ \forall\, v_1, v_2 \in V \ \forall\, w \in W : \langle av_1 + v_2, w\rangle = a\langle v_1, w\rangle + \langle v_2, w\rangle$ *(linéarité dans la première variable) et*
- $\forall\, b \in K \ \forall\, v \in V \ \forall\, w_1, w_2 \in W : \langle v, bw_1 + w_2\rangle = b\langle v, w_1\rangle + \langle v, w_2\rangle$ *(linéarité dans la deuxième variable).*

Soit $\langle \cdot, \cdot \rangle : V \times W \to K$ une forme bilinéaire. Pour un sous-espace $V_1 \leq V$, on appelle

$$V_1^\perp := \{w \in W \mid \forall v \in V_1 : \langle v, w\rangle = 0\} \leq W$$

le complément orthogonal de V_1 dans W.
Pour un sous-espace $W_1 \leq W$, on appelle

$$W_1^\perp := \{v \in V \mid \forall w \in W_1 : \langle v, w\rangle = 0\} \leq V$$

le complément orthogonal de W_1 dans V.
On dit que la forme bilinéaire est non-dégénérée *si*

- $\forall\, 0 \neq v \in V \ \exists\, w \in W : \langle v, w\rangle \neq 0$ *et*
- $\forall\, 0 \neq w \in W \ \exists\, v \in V : \langle v, w\rangle \neq 0.$

Dans la suite, nous allons écrire $\langle v, W_1\rangle = 0$ pour $\forall\, w \in W_1 : \langle v, w\rangle = 0$ (et aussi dans l'autre sens).

Lemme 13.7. *Soient V, W deux K-espaces vectoriels et $\langle \cdot, \cdot \rangle : V \times W \to K$ une forme bilinéaire.*

(a) Pour tout sous-espace $V_1 \leq V$, le complément orthogonal de V_1 dans W est un sous-espace de W et pour tout sous-espace $W_1 \leq W$, le complément orthogonal de W_1 dans V est un sous-espace de V.

(b) Soient $W_1 \leq W_2 \leq W$ des sous-espaces. Alors, $W_2^\perp \leq W_1^\perp$.
Également : $V_2^\perp \leq V_1^\perp$ pour tous sous-espaces $V_1 \leq V_2 \leq V$.

(c) La forme bilinéaire est non-dégénérée si et seulement si $W^\perp = 0$ et $V^\perp = 0$.

Démonstration. (a) Soit $V_1 \leq V$ un sous-espace. Soient $w_1, w_2 \in V_1^\perp$, c'est-à-dire, $\langle v, w_i\rangle = 0$ pour $i = 1, 2$ et tout $v \in V_1$. On a donc pour tout $a \in K$ l'égalité

$$\langle v, aw_1 + w_2\rangle = a\langle v, w_1\rangle + \langle v, w_2\rangle = 0,$$

d'où $aw_1 + w_2 \in V_1^\perp$. L'argument pour $W_1^\perp$ est le même.
(b) Soit $v \in W_2^\perp$. Par définition $\langle v, W_2\rangle = 0$, alors en particulier $\langle v, W_1\rangle = 0$, donc $v \in W_1^\perp$. La deuxème assertion suit par le même argument.
(c) Cela n'est qu'une autre manière d'écrire la définition. $\square$

Exemple 13.8. *(a) L'application*

$$\langle\cdot,\cdot\rangle : K^n \times K^n \to K, \quad \langle \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}, \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} \rangle = \begin{pmatrix} a_1 & a_2 & \dots & a_n \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = \sum_{i=1}^{n} a_i b_i$$

est bilinéaire et non-dégénérée.

(b) Soit V un K-espace vectoriel de dimension finie. L'application

$$\langle\cdot,\cdot\rangle : V^* \times V \to K, \quad \langle f, v\rangle := f(v)$$

est bilinéaire et non-dégénérée.

Soient $S = \{s_1, \dots, s_n\}$ une K-base de V et S^* la base duale. Soient $f = \sum_{i=1}^{n} a_i s_i^* \in V^*$ et $v = \sum_{i=1}^{n} b_i s_i \in V$. Alors

$$\langle f, v\rangle = \langle \sum_{i=1}^{n} a_i s_i^*, \sum_{j=1}^{n} b_j s_j \rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} a_i b_j \langle s_i^*, s_j\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} a_i b_j s_i^*(s_j)$$

$$= \sum_{i=1}^{n} a_i b_i = \begin{pmatrix} a_1 & a_2 & \dots & a_n \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}.$$

Nous avons retrouvé la forme bilinéaire de (a).

Proposition 13.9. *Soient V, W deux K-espaces vectoriels de dimensions finies et $\langle\cdot,\cdot\rangle : V \times W \to K$ une forme bilinéaire non-dégénérée.*

(a) Les applications

$$\varphi : V \to W^*, \;\; v \mapsto \varphi(v) =: \varphi_v \text{ avec } \varphi_v(w) := \langle v, w\rangle,$$

et

$$\psi : W \to V^*, \;\; w \mapsto \psi(w) =: \psi_w \text{ avec } \psi_w(v) := \langle v, w\rangle$$

sont des isomorphismes K-linéaires.

(b) $\dim_K(V) = \dim_K(W)$.

Démonstration. La K-linéarité de φ et ψ est claire. Nous montrons l'injectivité de φ. Pour cela soit $v \in \ker(\varphi)$, c'est-à-dire, $\varphi_v(w) = \langle v, w\rangle = 0$ pour tout $w \in W$. Le fait que la forme bilinéaire soit non-dégénérée implique donc $v = 0$, d'où l'injectivité. Nous en déduisons $\dim_K(V) \leq \dim_K(W^*) = \dim_K(W)$.
Les mêmes arguments appliqués à ψ donnent que ψ est injectif et donc $\dim_K(W) \leq \dim_K(V^*) = \dim_K(V)$, d'où $\dim_K(V) = \dim_K(W)$. En conséquence, φ et ψ sont des isomorphismes (car la dimension de l'image est égale à la dimension de l'espace d'arrivée qui sont donc égaux). $\square$

Corollaire 13.10. *Soient V, W deux K-espaces vectoriels de dimension finie.*

(a) Alors, l'application

$$\psi : V \to (V^*)^*, \quad v \mapsto \psi_v = \mathrm{ev}_v : V^* \to K \text{ où } \psi_v(f) = \mathrm{ev}_v(f) = f(v) \text{ pour } f \in V^*$$

est un isomorphisme K-linéaire.

(b) Soit $\alpha : V \to W$ une application K-linéaire. Alors, le diagramme

$$\begin{tikzcd} V \arrow[r, "\alpha"] \arrow[d, "\psi_1"] & W \arrow[d, "\psi_2"'] \\ (V^*)^* \arrow[r, "(\alpha^*)^*"] & (W^*)^*. \end{tikzcd}$$

est commutatif où ψ_1 et ψ_2 sont les isomorphismes de (a), c'est-à-dire $\psi_2 \circ \alpha = (\alpha^*)^* \circ \psi_1$.

(c) Soit $t_1, \ldots, t_n$ une K-base de V^*. Alors, il existe une K-base $s_1, \ldots, s_n$ de V telle que $t_i(s_j) = \delta_{i,j}$ pour tout $1 \le i, j \le n$.

Démonstration. (a) La forme bilinéaire $V^* \times V \to K$, donnée par $\langle f, v \rangle \mapsto f(v)$ de l'exemple 13.8(b) est non-dégénérée. L'application ψ est le ψ de la proposition 13.9.
(b) Soit $v \in V$. Nous avons d'un côté $(\alpha^*)^*(\psi_1(v)) = (\alpha^*)^*(\mathrm{ev}_v) = \mathrm{ev}_v \circ \alpha^*$ et de l'autre côté $\psi_2(\alpha(v)) = \mathrm{ev}_{\alpha(v)}$ avec les notations de (a). Pour vérifier que les deux sont égaux, soit $f \in W^*$. Nous avons

$$\mathrm{ev}_v(\alpha^*(f)) = \mathrm{ev}_v(f \circ \alpha) = f(\alpha(v)) \text{ et } \mathrm{ev}_{\alpha(v)}(f) = f(\alpha(v)),$$

donc l'égalité demandée.
(c) Soit $t_1^*, \ldots, t_n^* \in (V^*)^*$ la base duale, c'est-à-dire $t_j^*(t_i) = \delta_{i,j}$ pour tout $1 \le i, j \le n$. Comme ψ de (a) est un isomorphisme, il existe $s_1, \ldots, s_n$ (automatiquement une K-base de V car c'est l'image d'une base par un isomorphisme) tels que $\psi(s_j) = \mathrm{ev}_{s_j} = t_j^*$, donc $t_j^*(f) = f(s_j)$ pour tout $f \in V^*$. En particulier, nous avons $t_j^*(t_i) = t_i(s_j) = \delta_{i,j}$. $\square$

Proposition 13.11. *Soient V, W deux K-espaces vectoriels de dimensions finies et $\langle \cdot, \cdot \rangle : V \times W \to K$ une forme bilinéaire non-dégénérée.*

(a) Soit $S = \{s_1, \ldots, s_n\}$ une K-base de V. Alors, il existe une K-base $T = \{t_1, \ldots, t_n\}$ de W telle que $\langle s_i, t_j \rangle = \delta_{i,j}$ pour tout $1 \le i, j \le n$.

(b) Pour tout sous-espace $V_1 \le V$ nous avons $(V_1^\perp)^\perp = V_1$.

Également : pour tout sous-espace $W_1 \le W$ nous avons $(W_1^\perp)^\perp = W_1$.

(c) Pour tout sous-espace $V_1 \le V$ nous avons $\dim_K(V_1^\perp) = \dim_K(V) - \dim_K(V_1)$.

Également : pour tout sous-espace $W_1 \le W$ nous avons $\dim_K(W_1^\perp) = \dim_K(W) - \dim_K(W_1)$.

Démonstration. (a) Nous considérons le K-isomorphisme $\varphi : V \to W^*$ de la proposition 13.9 et nous posons $f_i := \varphi(s_i) = \varphi_{s_i}$ pour tout $1 \le i \le n$. Le corollaire 13.10 nous permet de choisir une K-base $t_1, \ldots, t_n$ de W telle que $f_i(t_j) = \delta_{i,j}$ pour tout $1 \le i, j \le n$. Finalement, nous avons $\langle s_i, t_j\rangle = \varphi_{s_i}(t_j) = f_i(t_j) = \delta_{i,j}$, comme exigé.
(b,c) Nous choisissons une K-base $s_1, \ldots, s_d$ de V_1 que nous étendons en une K-base

$$s_1, \ldots, s_d, s_{d+1}, \ldots, s_n$$

de V par la proposition 1.30. En utilisant (a) nous obtenons une K-base $t_1, \ldots, t_n$ de W telle que $\langle s_i, t_j\rangle = \delta_{i,j}$ pour tout $1 \le i, j \le n$.
Nous montrons d'abord $V_1^{\perp} = \langle t_{d+1}, \ldots, t_n\rangle$. L'inclusion « $\supseteq$ » est claire. Soit donc $w = \sum_{i=1}^n a_i t_i \in V_1^{\perp}$, c'est-à-dire $\langle V_1, w\rangle = 0$, donc pour tout $1 \le j \le d$ nous avons

$$0 = \langle s_j, w\rangle = \langle s_j, \sum_{i=1}^n a_i t_i\rangle = \sum_{i=1}^n a_i \langle s_j, t_i\rangle = a_j,$$

et donc $w \in \langle t_{d+1}, \ldots, t_n\rangle$. En conséquence, $\dim_K(V_1^{\perp}) = n - d = \dim_K(V) - \dim_K(V_1)$.
Le même argument utilisé pour $V_1^{\perp}$ montre que $\langle s_1, \ldots, s_d\rangle$ est une K-base pour $(V_1^{\perp})^{\perp}$ qui est donc égal à V_1. $\square$

Corollaire 13.12. *Soient V, W deux K-espaces vectoriels et $\varphi : V \to W$ une application K-linéaire. Nous avons les égalités*

(1) $\operatorname{im}(\varphi)^{\perp} = \ker(\varphi^*)$ *(où $\perp$ provient de la forme bilinéaire naturelle $W^* \times W \to K$),*

(2) $\ker(\varphi)^{\perp} = \operatorname{im}(\varphi^*)$ *(où $\perp$ provient de la forme bilinéaire naturelle $V^* \times V \to K$),*

(3) $\dim_K(\operatorname{im}(\varphi)) = \dim_K(\operatorname{im}(\varphi^*))$ *et*

(4) $\dim_K(\ker(\varphi)) = \dim_K(\ker(\varphi^*))$.

Démonstration. On montre d'abord (1). Soit $f \in W^*$. Alors

$$f \in \operatorname{im}(\varphi)^{\perp} \Leftrightarrow \forall\, v \in V : 0 = \langle f, \varphi(v)\rangle = f(\varphi(v)) \Leftrightarrow f \circ \varphi = 0 \Leftrightarrow f \in \ker(\varphi^*),$$

donc (1).
Nous adaptons légèrement les arguments pour obtenir (2) ainsi. Soit $v \in V$. Alors

$$\begin{aligned} v \in \operatorname{im}(\varphi^*)^{\perp} \Leftrightarrow \forall\, f \in W^* : 0 = \langle \varphi^*(f), v\rangle = \langle f \circ \varphi, v\rangle = f(\varphi(v)) = \langle f, \varphi(v)\rangle \\ \Leftrightarrow \varphi(v) \in W^{\perp} \Leftrightarrow \varphi(v) = 0 \Leftrightarrow v \in \ker(\varphi), \end{aligned}$$

d'où $\operatorname{im}(\varphi^*)^{\perp} = \ker(\varphi)$. En appliquant, la proposition 13.11 nous obtenons $\operatorname{im}(\varphi^*) = \ker(\varphi)^{\perp}$; c'est (2).
Par le corollaire 1.38, nous avons $\dim_K(V) = \dim_K(\operatorname{im}(\varphi)) + \dim_K(\ker(\varphi))$. La proposition 13.11 nous donne

$$\dim_K(\operatorname{im}(\varphi)) = \dim_K(V) - \dim_K(\ker(\varphi)) = \dim_K(\ker(\varphi)^{\perp}) = \dim_K(\operatorname{im}(\varphi^*)),$$

d'où (3). L'argument pour obtenir (4) est similaire :

$$\dim_K(\ker(\varphi)) = \dim_K(V) - \dim_K(\mathrm{im}(\varphi)) = \dim_K(\mathrm{im}(\varphi)^\perp) = \dim_K(\ker(\varphi^*)),$$

ce qui conclut la démonstration. $\square$

Définition 13.13. *Soit $M \in \mathrm{Mat}_{m\times n}(K)$ une matrice.*
Le rang des colonnes *de M est défini comme la dimension du sous-espace de K^m engendré par les colonnes de M (vues comme éléments de K^m).*
Le rang des lignes *de M est défini comme la dimension du sous-espace de K^n engendré par les lignes de M (vues comme éléments de K^n).*

Corollaire 13.14. *Soit $M \in \mathrm{Mat}_{m\times n}(K)$. Alors, le rang des colonnes de M est égal au rang des lignes de M. On parle simplement du* rang *de M.*

Démonstration. Le rang de M est la dimension de l'image de φ_M, l'application K-linéaire $K^n \to K^m$ associée à M (qui envoie $v \in K^n$ sur $Mv \in K^m$). La matrice représentant φ_M^* pour la base duale est M^{tr}. Donc le corollaire suit directement du corollaire 13.12 car le rang des colonnes de M^{tr} est égal au rang des lignes de M. $\square$

Exemple 13.15. *Considérons la matrice $\begin{pmatrix} 3 & 5 & 1 \\ 1 & 2 & 3 \\ 4 & 7 & 4 \end{pmatrix}$. On s'intéresse à son rang (de colonnes). Il est évident que la troisième ligne est la somme des deux premières lignes (qui sont linéairement indépendantes). Donc le rang de M est 2. Il me semble plus difficile de « voir » une combinaison non-triviale des colonnes, mais nous savons qu'il en existe une.*

Nous finissons cette section par des propriétés utiles.

Proposition 13.16. *Soient V, W deux K-espaces vectoriels de dimensions finies et $\langle \cdot, \cdot \rangle : V \times W \to K$ une forme bilinéaire non-dégénérée. Soient $W_1 \le W$ et $W_2 \le W$ des sous-espaces. Alors, nous avons*

(a) $(W_1 \cap W_2)^\perp = W_1^\perp + W_2^\perp$ *et*

(b) $(W_1 + W_2)^\perp = W_1^\perp \cap W_2^\perp$.

Également avec V au lieu de W.

Démonstration. (a) « $\supseteq$ » : Comme $W_1 \cap W_2 \le W_i$ est un sous-espace pour $i = 1, 2$, nous avons $W_i^\perp \le (W_1 \cap W_2)^\perp$, donc $W_1^\perp + W_2^\perp \le (W_1 \cap W_2)^\perp$ car $(W_1 \cap W_2)^\perp$ est un sous-espace.
(b) « $\subseteq$ » : Pour $i = 1, 2$ nous avons $W_i \le W_1 + W_2$, donc nous obtenons $(W_1 + W_2)^\perp \le W_i^\perp$ ce qui implique $(W_1 + W_2)^\perp \le W_1^\perp \cap W_2^\perp$.
(a) « $\subseteq$ » : En combinant les inclusions montrées, nous avons

$$W_1 \cap W_2 = ((W_1 \cap W_2)^\perp)^\perp \le (W_1^\perp + W_2^\perp)^\perp \le (W_1^\perp)^\perp \cap (W_2^\perp)^\perp = W_1 \cap W_2,$$

donc nous avons l'égalité partout et, en particulier, $(W_1 \cap W_2)^\perp = W_1^\perp + W_2^\perp$.
(b) Il suffit d'utiliser (a) avec $W_1^\perp$ et $W_2^\perp$ au lieu de W_1 et W_2 pour obtenir $(W_1^\perp \cap W_2^\perp)^\perp = (W_1^\perp)^\perp + (W_2^\perp)^\perp$ et donc $W_1^\perp \cap W_2^\perp = (W_1 + W_2)^\perp$. $\square$

14 Quotients

Objectifs :

- Connaître et maîtriser la définition de quotient d'espaces vectoriels ;
- connaître les théorèmes d'isomorphisme et d'autres résultats importants ;
- savoir calculer dans des quotients d'espaces vectoriels ;
- connaître des exemples et savoir démontrer des propriétés simples.

Définition 14.1. *Soit V un K-espace vectoriel et $W \leq V$ un sous-espace.*
Tout ensemble de la forme

$$v + W = \{v + w \mid w \in W\}$$

avec $v \in V$ s'appelle sous-espace affine.
Deux sous-espaces $v_1 + W$ et $v_2 + W$ sont appelés parallèles. *Ils sont donc tous les deux parallèles à W.*

Pour comprendre la suite, il est utile de rappeler la définition de congruences modulo n, c'est-à-dire l'ensemble $\mathbb{Z}/n\mathbb{Z}$ (pour $n \in \mathbb{N}_{\geq 1}$), apprise dans le cours *Structures mathématiques*. Pour souligner l'analogie, on peut écrire $V = \mathbb{Z}$ et $W = n\mathbb{Z} = \{nm \mid m \in \mathbb{Z}\}$.
On rappelle que l'ensemble

$$a + n\mathbb{Z} = \{a + mn \mid m \in \mathbb{Z}\} = \{\dots, a - 2n, a - n, a, a + n, a + 2n, \dots\}$$

est la classe d'équivalence de $a \in \mathbb{Z}$ pour la relation d'équivalence définie sur $\mathbb{Z}$ par

$$a \sim_{n\mathbb{Z}} a' \quad \Leftrightarrow \quad a \equiv a' \mod n \quad \Leftrightarrow \quad n \mid (a - a') \quad \Leftrightarrow \quad a - a' \in n\mathbb{Z} \quad \Leftrightarrow \quad a + n\mathbb{Z} = a' + n\mathbb{Z}.$$

On fera essentiellement la même définition dans le cas des espaces vectoriels.

Définition 14.2. *Soit V un K-espace vectoriel et $W \subseteq V$ un sous-espace vectoriel. La relation binaire sur V donnée par*

$$v_1 \sim_W v_2 \quad \stackrel{\text{définition}}{\Longleftrightarrow} \quad v_1 - v_2 \in W$$

pour $v_1, v_2 \in V$ définit une relation d'équivalence.
Les classes d'équivalence sont les sous-espaces affines de la forme

$$v + W = \{v + w \mid w \in W\}.$$

L'ensemble de ces classes est noté V/W et appelé l'ensemble des classes suivant W. *C'est l'ensemble de tous les sous-espaces affines parallèles à W.*

Rappelons également l'addition 'modulaire', c'est-à-dire l'addition de $\mathbb{Z}/n\mathbb{Z}$. La somme de $a + n\mathbb{Z}$ et $b + n\mathbb{Z}$ est définie comme

$$(a + n\mathbb{Z}) + (b + n\mathbb{Z}) := (a + b) + n\mathbb{Z}.$$

Pour voir que cette addition est bien définié, on fait **l'observation fondamentale** : soient $a, a', b, b' \in \mathbb{Z}$ tels que

$$a \equiv a' \mod n \quad \text{et} \quad b \equiv b' \mod n,$$

c'est-à-dire,

$$a + n\mathbb{Z} = a' + n\mathbb{Z} \quad \text{et} \quad b + n\mathbb{Z} = b' + n\mathbb{Z}$$

alors,

$$a + b \equiv a' + b' \mod n,$$

c'est-à-dire,

$$(a + b) + n\mathbb{Z} = (a' + b') + n\mathbb{Z}.$$

La preuve en est très facile : comme $n \mid (a' - a)$ et $n \mid (b' - b)$, il existe $c, d \in \mathbb{Z}$ tels que $a' = a + cn$ et $b' = b + dn$; donc

$$a' + b' = (a + cn) + (b + dn) = (a + b) + n(c + d)$$

ainsi, n divise $(a' + b') - (a + b)$, donc $(a' + b') + n\mathbb{Z} = (a + b) + n\mathbb{Z}$. Un petit exemple :

$$(3 \equiv 13 \mod 10 \quad \text{et} \quad 6 \equiv -24 \mod 10) \quad \Rightarrow \quad 9 \equiv -11 \mod 10.$$

Voici la généralisation aux espaces vectoriels. Noter qu'il ne suffit pas de définir une addition, mais il faut aussi définir une multiplication scalaire.

Proposition 14.3. *Soient K un corps, V un K-espace vectoriel, $W \leq V$ un K-sous-espace vectoriel et V/W l'ensemble des classes suivant W.*

(a) *Pour tout $v_1, v_2 \in V$ la classe $(v_1 + v_2) + W$ ne dépend que des classes $v_1 + W$ et $v_2 + W$.*

Donc, on peut définir l'application, appelée addition*,*

$$+ : V/W \times V/W \to V/W, \quad (v_1 + W, v_2 + W) \mapsto (v_1 + W) + (v_2 + W) := (v_1 + v_2) + W.$$

(b) *Pour tout $a \in K$ et tout $v \in V$, la classe $a.v + W$ ne dépend que de la classe $v + W$. Donc, on peut définir l'application, appelée* multiplication scalaire*,*

$$. : K \times V/W \to V/W, \quad (a, v + W) \mapsto a.(v + W) := a.v + W.$$

(c) $(V/W, +, ., 0 + W)$ *est un K-espace vectoriel, appelé* quotient de V par W.

(d) *L'application*

$$\pi : V \to V/W, \quad v \mapsto v + W$$

est K-linéaire et surjective de noyau $\ker(\pi) = W$ *; elle est appelée* projection naturelle.

Démonstration. (a) Supposons $v_1+W = v_1'+W$ et $v_2+W = v_2'+W$. Il existe donc $w_1, w_2 \in W$ tels que $v_1 = v_1'+w_1$ et $v_2 = v_2'+w_2$. Alors $v_1+v_2 = v_1'+v_2'+(w_1+w_2)$ d'où $(v_1+v_2)-(v_1'-v_2') \in W$ et alors $(v_1+v_2)+W = (v_1'+v_2')+W$.
(b) Supposons $v+W = v'+W$. Il existe donc $w \in W$ tel que $v = v'+w$. Alors $av = a(v'+w) = av'+aw$ d'où $av - av' = aw \in W$ et alors $av+W = av'+W$.
(c) Vérification standard des axiomes qui définissent un espace vectoriel (voir la définition 1.1).
(d) Linéarité : Soient $v_1, v_2 \in V$ et $a \in K$, alors $\pi(av_1+v_2) = (av_1+v_2)+W = a(v_1+W)+(v_2+W) = a\pi(v_1)+\pi(v_2)$.
Surjectivité : La classe $v+W$ est l'image de v sous π.
Calcul du noyau : Soit $v \in V$. Alors $v \in \ker(\pi)$ si et seulement si $v+W = 0+W = W$ et cela est le cas si et seulement si $v \in W$. $\square$

Théorème 14.4 (1er théorème d'isomorphisme/Homomorphiesatz)**.** *Soient K un corps et $\varphi : V \to Y$ une application K-linéaire. Soit $W := \ker(\varphi)$ son noyau.*

(a) Pour $v \in V$, l'image $\varphi(v)$ ne dépend que de la classe $v+W$.

(b) La partie (a) nous permet de définir $\overline{\varphi}(v+W) := \varphi(v)$ pour $v \in V$. Cela résulte en une application

$$\overline{\varphi} : V/W \to Y, \quad v+W \mapsto \overline{\varphi}(v+W) := \varphi(v)$$

qui est injective et K-linéaire. Elle donne lieu à un isomorphisme K-linéaire

$$\overline{\varphi} : V/W \to \mathrm{im}(\varphi).$$

Démonstration. (a) Soient $v, v' \in V$ tels que $v+W = v'+W$. Il existe donc $w \in W$ tel que $v = v'+w$. Nous avons $\varphi(v) = \varphi(v'+w) = \varphi(v')+\varphi(w) = \varphi(v')$ car $\varphi(w) = 0$ comme $w \in W = \ker(\varphi)$.
(b) Linéarité : Soient $v_1, v_2 \in V$ et $a \in K$. Nous avons $\overline{\varphi}(a(v_1+W)+(v_2+W)) = \overline{\varphi}((av_1+v_2)+W) = \varphi(av_1+v_2) = a\varphi(v_1)+\varphi(v_2) = a\overline{\varphi}(v_1)+\overline{\varphi}(v_2)$.
Injectivité : Soit $v+W \in \ker(\overline{\varphi})$. Donc $\overline{\varphi}(v+W) = \varphi(v) = 0$ d'où $v \in \ker(\varphi) = W$, alors $v+W = 0+W$. Cela montre $\ker(\overline{\varphi}) = \{0+W\}$, alors $\overline{\varphi}$ est injectif. $\square$

La proposition suivante est importante car elle décrit les sous-espaces vectoriels des espaces vectoriels quotients.

Proposition 14.5. *Soient K un corps, V un K-espace vectoriel, $W \leq V$ un sous-espace vectoriel, et $\pi : V \to V/W$ la projection naturelle.*

(a) L'application

$$\Phi : \{\text{sous-espaces vectoriels de } V/W\} \longrightarrow \{\text{sous-espaces vectoriels de } V \text{ qui contiennent } W\},$$

donnée par $X \mapsto \pi^{-1}(X)$ est bijective. L'inverse Ψ de Φ est $Y \mapsto \pi(Y)$.

(b) Soient $X_1, X_2 \leq V/W$ deux sous-espaces vectoriels. Alors

$$X_1 \subseteq X_2 \quad \Leftrightarrow \quad \Phi(X_1) \subseteq \Phi(X_2).$$

Démonstration. (a)

- Pour un sous-espace $X \leq V/W$ l'image réciproque $\Phi(X) = \pi^{-1}(X)$ est en effet un sous-espace vectoriel : soient $v_1, v_2 \in V$ tels que $v_1 \in \pi^{-1}(X)$ et $v_2 \in \pi^{-1}(X)$, donc $\pi(v_1) = v_1 + W \in X$ et $\pi(v_2) = v_2 + W \in X$. Alors pour $a \in K$, on a $a\pi(av_1 + v_2) = \pi(v_1) + \pi(v_2) \in X$, d'où $av_1 + v_2 \in \pi^{-1}(X)$.

  En plus $\pi^{-1}(W) \supseteq \pi^{-1}(\{0\}) = \ker(\pi) = W$.

- Nous savons par la proposition 1.36 que les images des applications linéaires entre espaces vectoriels sont des sous-esapces vectoriels, donc $\Psi(Y) = \pi(Y)$ est un sous-espace vectoriel de V/W.

- Voici une assertion auxiliaire :

  Soient $\pi : V \to V'$ un homomorphisme K-linéaire entre espaces vectoriels et $Y \leq V$ un sous-espace vectoriel qui contient $\ker(\pi)$. Alors $\pi^{-1}(\pi(Y)) = Y$.

  On vérifie cette égalité :
  « $\subseteq$ » : Soit $x \in \pi^{-1}(\pi(Y))$, donc $\pi(x) \in \pi(Y)$, donc $\pi(x) = \pi(y)$ pour un $y \in Y$. Donc $0 = \pi(x) - \pi(y) = \pi(x - y)$, donc $x - y \in \ker(\pi) \subseteq Y$, donc $x - y = y' \in Y$, donc $x = y + y' \in Y$.

  « $\supseteq$ » : Soit $y \in Y$, donc $\pi(y) \in \pi(Y)$, donc $y \in \pi^{-1}(\pi(Y))$.

- Soit $Y \leq V$ un sous-espace vectoriel tel que $W \subseteq Y$.

  Par l'assertion auxiliaire on a : $\Phi(\Psi(Y)) = \pi^{-1}(\pi(Y)) = Y$.

- Voici une autre assertion auxiliaire :

  Soient $\pi : V \to V'$ une application surjective (pas nécessairement entre espaces vectoriels) et $X \subseteq V'$ un sous-ensemble. Alors $X = \pi(\pi^{-1}(X))$.

  On vérifie cette égalité.
  « $\subseteq$ » : Soit $x \in X$. Comme π est surjectif, il existe $v \in V$ tel que $\pi(v) = x$. Donc $v \in \pi^{-1}(X)$ et $x = \pi(v) \in \pi(\pi^{-1}(X))$.

  « $\supseteq$ » : Soit $v' \in \pi(\pi^{-1}(X))$. Donc, il existe $v \in \pi^{-1}(X)$ tel que $v' = \pi(v)$. Mais, $v' = \pi(v)$ appartient à X car $v \in \pi^{-1}(X)$.

- Soit $X \leq V/W$ un sous-espace vectoriel.

  Par l'assertion auxiliaire on a : $\Psi(\Phi(X)) = \pi(\pi^{-1}(X)) = X$.

(b) est clair. □

Proposition 14.6 (Deuxième théorème d'isomorphisme)**.** *Soient K un corps, V un K-espace vectoriel et $X, W \subseteq V$ des sous-espaces vectoriels. Alors, l'homomorphisme K-linéaire*

$$\varphi : X \to (X + W)/W, \quad x \mapsto x + W,$$

« induit » (par le théorème d'isomorphisme 14.4) l'isomorphisme K-linéaire

$$\overline{\varphi} : X/(X \cap W) \to (X + W)/W, \quad x + (X \cap W) \mapsto x + W.$$

Démonstration. L'homomorphisme φ est visiblement surjectif et son noyau est composé des éléments $x \in X$ tels que $x + W = W$, donc $x \in X \cap W$, montrant $\ker(\varphi) = X \cap W$. L'existence de $\overline{\varphi}$ résulte donc d'une application directe du théorème d'isomorphisme 14.4. $\square$

Proposition 14.7 (Troisième théorème d'isomorphisme)**.** *Soient K un corps, V un K-espace vectoriel et $W_1 \subseteq W_2$ deux sous-espaces vectoriels de V. Alors, l'homomorphisme K-linéaire*

$$\varphi : V/W_1 \to V/W_2, \quad v + W_1 \mapsto v + W_2$$

« induit » (par le théorème d'isomorphisme 14.4) l'isomorphisme K-linéaire

$$\overline{\varphi} : (V/W_1)/(W_2/W_1) \to V/W_2, \quad v + W_1 + (W_2/W_1) \mapsto v + W_2.$$

Démonstration. L'homomorphisme φ est visiblement surjectif et son noyau est composé des éléments $v + W_1 \in V/W_1$ tels que $v + W_2 = W_2$ ce qui est équivalent à $v + W_1 \in W_2/W_1$. Donc $\ker(\varphi) = W_2/W_1$. L'existence de $\overline{\varphi}$ résulte donc d'une application directe du théorème d'isomorphisme 14.4. $\square$