Algèbre linéaire 2

Christophe Reutenauer
Laboratoire de combinatoire et d'informatique mathématique,
Université du Québec à Montréal

5 septembre 2024

Table des matières

Remerciements : Sylvain St-Amand, Geneviève Lefebvre pour la référence [3].

1 Introduction

Ce cours fait suite au cours d'algèbre linéaire 1, dont on pourra lire les notes [1]. Nous notons $\mathbb{R}$ l'ensemble des nombres réels.

2 Corps des scalaires

En algèbre linéaire 1, nous avons étudié les espaces vectoriels sur les réels : l'opération externe, ou produit externe, associe à tout réel a et à tout vecteur v un vecteur noté av. Nous avons étudié ce qu'on appelle les *espaces vectoriels sur* $\mathbb{R}$.

On peut définir la notion d'espace vectoriel sur d'autres "corps" que $\mathbb{R}$. Un *corps* est un ensemble qui, comme $\mathbb{R}$, possède un 0 (*zéro*) et un 1 (*un*), et qui a quatre opérations : addition, soustraction, multiplication, division, avec les propriétés usuelles ; en particulier, on peut toujours diviser un élément par un autre élément, sauf si ce dernier est nul.

Nous n'irons pas dans les détails ici. Regardons plutôt des exemples :

— Le corps des nombres réels $\mathbb{R}$;
— Le corps des nombres rationnels $\mathbb{Q}$;
— Le corps des nombres complexes $\mathbb{C}$;
— Pour p un nombre premier, le corps $\mathbb{F}_p$ dont les éléments sont les entiers modulo p ; autrement dit $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$.

Le corps $\mathbb{Q}$ est un *sous-corps* de $\mathbb{R}$, lequel est un sous-corps de $\mathbb{C}$.

3 Polynômes, racines et factorisation, division euclidienne et Bezout

Soit $\mathbb{K}$ un corps. On note $\mathbb{K}[x]$ l'ensemble des polynômes en la variable x à coefficients dans $\mathbb{K}$. Un polynôme s'écrit $P(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0$, avec les *coefficients* a_i dans $\mathbb{K}$. Une *racine* (ou un *zéro*) de $P(x)$ dans $\mathbb{L}$, où $\mathbb{L}$

est un sur-corps de $\mathbb{K}$ (c'est-à-dire $\mathbb{K}$ est un sous-corps de L) est un élément α de $\mathbb{L}$ tel que $P(\alpha) = 0$, où $P(\alpha) = a_n\alpha^n + a_{n-1}\alpha^{n-1} + \cdots + a_0$. Remarquez que $P(x) \in \mathbb{K}[x]$ et aussi $P(x) \in \mathbb{L}[x]$.

Par exemple, 1 est une racine de $x^2 - 1$ dans $\mathbb{R}$, et i (le nombre complexe"racine de -1") est une racine de $x^2 + 1$ dans $\mathbb{C}$.

Le polynôme $P(x) \in \mathbb{K}[x]$ est dit *scindé sur* $\mathbb{K}$ s'il s'écrit $P(x) = a_n \prod_{i=1}^{i=n}(x - \alpha_i)$, où les α_i sont dans $\mathbb{K}$. Notez que ses racines sont alors les α_i et qu'il n'y en a pas d'autres, même dans un sur-corps. C'est pourquoi on dit aussi que P *a toutes ses racines dans* $\mathbb{K}$.

Par exemple, le polynôme $x^2 - 1 = (x - 1)(x + 1)$, ses racines sont $1, -1 \in \mathbb{Q}$ et il n'y a pas d'autre racine, que ce soit dans $\mathbb{Q}$, $\mathbb{R}$ ou $\mathbb{C}$.

Si dans la factorisation ci-dessus de $P(x)$, plusieurs des α_i sont égaux, on appelle *multiplicité* de α_i le nombre de α_j égaux à α_i.

Par exemple, la racine 1 est de multiplicité 1 dans le polynôme $x^2 - 1 \in \mathbb{Q}[x]$ (on dit aussi qu'elle est *simple*), alors qu'elle est de multiplicité 3 dans $(x - 1)^3(x + 3)^2$.

Un théorème important est le *théorème fondamental de l'algèbre* : tout polynôme à coefficients sur $\mathbb{C}$ est scindé.

Une propriété importante des polynômes est la suivante ; elle est analogue au "théorème fondamental de l'arithmétique", lequel se formule comme suit : tout entier $n \in \mathbb{Z}$, non nul, a une unique factorisation $n = \epsilon p_1 \dots p_k$, où les p_i sont des nombres premiers et où $\epsilon = 1$ ou -1 (l'unicité est à l'ordre près). Pour les polynômes, le théorème analogue est : tout polynôme $\in \mathbb{K}[x]$, non nul, a une unique factorisation $P = \epsilon P_1 \cdots P_k$, où ϵ est un scalaire non nul, et où les polynômes P_i sont unitaires irréductibles ; ici, *unitaire* signifie que le polynôme a un coefficient dominant égal à 1 ; et *irréductible* signifie que c'est un polynôme P de degré au moins 1, qu'on ne peut pas factoriser le polynôme comme un produit de deux polynômes qui sont tous deux de degré plus petit que P.

Notez que les polynômes de degré 1 sont tous irréductibles.

Une propriété importante des polynômes est la *division euclidienne* : si A, B sont dans $\mathbb{K}[x]$, avec $B \neq 0$, alors il existe $Q, R \in \mathbb{K}[x]$ tels que $A = BQ + R$ et $\deg(R) < \deg(B)$.

On dit que deux polynômes dans $\mathbb{K}[x]$ sont *premiers entre eux* s'ils n'ont pas de diviseur commun dans $\mathbb{K}[x]$ (sauf les polynômes constants). Une autre dont nous aurons besoin est que si $A, B \in \mathbb{K}[x]$ n'ont pas de diviseur commun, alors il existe $P, Q \in \mathbb{K}[x]$ tels que $AP + BQ = 1$ (*théorème de Bezout*).

Si deux polynômes $P, Q \in \mathbb{K}[x]$ sont scindés sur $\mathbb{K}$, alors ils n'ont pas de diviseur commun si et seulement s'ils n'ont pas de racine commune. Autrement dit, si $P(x) = a \prod_{i=1}^{i=n}(x - \alpha_i)$ et $Q(x) = b \prod_{j=1}^{i=m}(x - \beta_j)$

$(a, b, \alpha_i, \beta_j \in \mathbb{K},\ a, b \neq 0)$, alors P, Q sont premiers entre eux si et seulement si $\{\alpha_i \mid i = 1, \ldots, n\} \cap \{\beta_j \mid j = 1, \ldots, m\} = \emptyset$. Le cas particulier dont nous aurons besoin est quand $P = a(x-\alpha)^k, Q = b(x-\beta)^\ell$: ils sont premiers entre eux si et seulement si $\alpha \neq \beta$).

Exercice 3.1. *a) Montrer que si P, Q sont deux polynômes dans $\mathbb{R}[x]$ qui sont premiers entre eux, alors ils n'ont pas de racine commune dans $\mathbb{C}$ (utiliser le théorème de Bezout).*

b) Montrer que si P, Q sont deux polynômes dans $\mathbb{R}[x]$ qui ne sont pas premiers entre eux, alors ils ont une racine commune dans $\mathbb{C}$ (utiliser le théorème fondamental de l'algèbre).

4 Rappels d'algèbre linéaire 1 : espaces vectoriels, sous-espaces, bases

Dans toutes ces notes $\mathbb{K}$ est un corps.

4.1 Les huit axiomes d'un espace vectoriel

Définition 4.1. *Un* espace vectoriel sur $\mathbb{K}$ *(ou* $\mathbb{K}$-espace vectoriel*) est un ensemble E qui a deux opérations. La première est appelée* addition*, ou* somme*, et la seconde est appelée* produit externe *(ou* multiplication scalaire*). L'addition associe à deux éléments quelconques x, y de E un élément de E noté $x + y$. Le produit externe associe à un élément $a \in \mathbb{K}$ et à un élément x de E un élément de E noté ax. Ces deux opérations jouissent des propriétés suivantes (appelées* axiomes des espaces vectoriels*) : quels que soient les éléments x, y, z de E et les éléments a, b de $\mathbb{K}$, on a :*

1. $x + y = y + x$ *(commutativité) ;*
2. $(x + y) + z = (x + y) + z$ *(associativité ;*
3. *Il existe un élément* 0 *de E tel que $x + 0 = x$ (existence de l'élément neutre) ;*
4. *Il existe un élément x' de E tel que $x + x' = 0$ (existence de l'opposé, et ce x' est noté $-x$) ;*
5. $1x = x$ *(le produit externe de $1 \in \mathbb{K}$ avec x est égal à x) ;*
6. $(a + b)x = ax + bx$ *(distributivité) ;*
7. $a(x + y) = ax + ay$ *(distributivité) ;*
8. $(ab)x = a(bx)$ *(associativité).*

Les éléments de E sont appelés des *vecteurs* et les éléments de $\mathbb{K}$ sont appelés des *scalaires*.

Abus de notation : on utilise la même notation 0 pour le zéro de l'espace vectoriel E (appelé le *vecteur nul*) et pour le zéro de $\mathbb{K}$. S'il y a un risque de confusion, on peut écrire 0_E et $0_{\mathbb{K}}$.

Exemple d'espace vectoriel : $\mathbb{K}^n$.

4.2 Sous-espaces vectoriels

Définition 4.2. *Un* sous-espace vectoriel *d'un espace vectoriel V est un sous-ensemble E non vide de V qui est fermé sous les deux opérations de V ; ce qui signifie que quels que soient les vecteurs u, v dans E et le scalaire a, on a $u + v \in E$ et av dans E.*

Pour vérifier la non-vacuité de E, on peut vérifier que $0 \in E$.

Exercice 4.1. *Montrer que $\{(a, b), b = a^2\}$ n'est pas un sous-espace de $\mathbb{R}^2$.*

Exercice 4.2. *Même chose pour $\{(a, b), a, b \geq 0\}$.*

4.3 Produits d'espaces vectoriels

Le *produit des espaces vectoriels* $E_1, \ldots, E_n$ est le produit cartésien $E_1 \times \cdots \times E_n$ de ces espace vectoriels, avec somme et produit externe définis par

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

$$a(x_1, \ldots, x_n) = (ax_1, \ldots, ax_n).$$

Le cas particulier où $E_i = \mathbb{K}$ pour tout i est particulièrement important : $\mathbb{K}^n$ est un espace vectoriel.

4.4 Combinaisons linéaires et bases

Définition 4.3. *Soit E un $\mathbb{K}$-espace vectoriel. Une* combinaison linéaire *dans E est une expression de la forme $a_1e_1 + \ldots + a_ne_n$ où $a_1, \ldots, a_n$ sont dans $\mathbb{K}$ et $e_1, \ldots, e_n$ sont dans E. On appelle* coefficients *de la combinaison linéaire les éléments $a_1, \ldots, a_n$.*

Toute combinaison linéaire représente un vecteur de E.

Définition 4.4. *Une base d'un espace vectoriel E est une famille $e_1, \ldots, e_n$ de vecteurs dans E telle que tout vecteur dans E est une combinaison linéaire de ces vecteurs, et ceci de manière unique.*

Théorème 4.1. *Deux bases quelconques de E ont le même nombre d'éléments.*

Cette cardinalité commune des bases de E s'appelle la *dimension* de E. A strictement parler, nous ne nous occupons ici que des espaces de dimension finie. Il existe aussi une théorie des espaces de dimension infinie.

Théorème 4.2. *Si des vecteurs $v_1, \ldots, v_n$ engendrent V (c'est-à-dire, si tout vecteur dans V est combinaison linéaire de ces vecteurs), alors il existe une sous-famille de ces vecteurs qui forme une base de V. En particulier la dimension de V est $\leq n$.*

Comme conséquence très importante, on a que *si un espace vectoriel peut être engendré par un nombre fini de vecteurs, alors il a une base finie.* Il est vrai aussi que tout espace vectoriel a une base, non nécessairement finie, mais nous n'irons pas trop dans cette direction.

Théorème 4.3. *Soit V est un espace de dimension finie. Si W est un sous-espace de V, alors* $\dim(W) \leq dim(V)$ *et l'inégalité est stricte si $W \neq V$.*

Il est équivalent de dire que si W est un sous-espace de V, de dimension finie, et que si $\dim(W) = \dim(V)$, alors $W = V$.

On dit que des vecteurs $v_1, \ldots, v_n$ d'un espace vectoriel V sont *linéairement indépendants* si pour tous scalaires $a_1, \ldots, a_n$ tels que $a_1 v_1 + \cdots + a_n v_n = 0$, on $a_1 = \ldots = a_n = 0$.

Théorème 4.4. *Si des vecteurs $v_1, \ldots, v_n$ d'un espace vectoriel V sont linéairement indépendants, alors il existe une base de V qui les contient. En particulier* $\dim(V) \geq n$.

Théorème 4.5. *Soit V un espace vectoriel et $v_1, \ldots, v_n$ des vecteurs dans V. Deux quelconques des trois conditions suivantes impliquent la troisième :*
(i) $\dim(V) = n$ *;*
(ii) $v_1, \ldots, v_n$ engendrent V ;
(iii) $v_1, \ldots, v_n$ sont linéairement indépendants.

Notez que (ii) et (iii) signifient que ces vecteurs forment une base de V.

Exercice 4.3. *Les vecteurs* $(1,2), (2,3), (3,4)$ *forment-ils une base de* $\mathbb{K}^2$ *?*

Exercice 4.4. *Montrer que* $(1,0,0), (0,1,0), (0,0,1)$ *forment une base de* $\mathbb{K}^3$.

Exercice 4.5. *Les vecteurs* $(1,1,1), (1,-1,0), (2,0,1)$ *forment-ils une base de* $\mathbb{K}^3$ *?*

Exercice 4.6. *Montrez que les vecteurs* $(1,-1,0),(0,1,-1),(-1,0,1) \in \mathbb{K}^3$ *sont linéairement dépendants.*

Exercice 4.7. *Donnez plusieurs bases de l'espace vectoriel* $\{(a,b,c) \in \mathbb{K}^3, a+b+c=0\}$.

Exercice 4.8. *Montrer que pour* $\mathbb{K}=\mathbb{R}$*, l'espace vectoriel* $\mathbb{K}^3$ *a pour base* $(1,2,0),(0,1,2),(1,1,1)$.

Exercice 4.9. *Montrer que pour* $\mathbb{K}=\mathbb{F}_3$*, l'exercice précédent n'est pas vrai. Indication : le corps* $\mathbb{F}_3$ *est l'ensemble* $\{0,1,2\}$ *avec l'addition* $1+1=2, 1+2=0, 2+2=1$ *et la multiplication par* $2\cdot 2=1$*, avec les propriétés uselles pour* 0 *et* 1*. Remarque :* $\mathbb{F}_3=\mathbb{Z}/3\mathbb{Z}$*, et calculer dans* $\mathbb{F}_3$*, c'est calculer avec les entiers, modulo 3.*

Exercice 4.10. *On utilise ici la notion de base infinie, et de dimension infinie. Montrer que le* $\mathbb{K}$*-espace vectoriel* $\mathbb{K}[x]$ *a pour base* $1,x,x^2,,\ldots,x^n,\ldots$*; et aussi pour base* $1,1+x,1+x+x^2,\ldots,1+x+\cdots+x^n,\ldots$.

4.5 Sous-espace engendré et rang

Le *sous-espace engendré par les vecteurs* $v_1,\ldots,v_n$ est l'ensemble des combinaisons linéaires de ces vecteurs. On le note $\mathrm{Vect}(v_1,\ldots,v_n)$ ou $\langle v_1,\ldots,v_n\rangle$.

Le *rang* d'un ensemble de vecteurs est par définition la dimension du sous-espace qu'ils engendrent.

Exercice 4.11. *Montrer que* $Vect(v_1,\ldots,v_n) = Vect(v_1,\ldots,v_{n-1})$ *si et seulement si* v_n *est combinaison linéaire de* $v_1,\ldots,v_{n-1}$.

5 Rappel d'algèbre linéaire 1 : applications linéaires

Définition 5.1. *Soient* E,F *des espaces vectoriels. Une fonction* $f: E\to F$ *(c'est-à-dire une fonction de* E *vers* F*) est appelée une* application linéaire *si pour tous vecteurs* x,y *dans* E *et tout scalaire* a *on* $f(x+y)=f(x)+f(y)$ *et* $f(ax)=af(x)$.

Lorsque $E=F$, on parle d'*endomorphisme de* E, et on note $\mathrm{End}(E)$ l'ensemble des endomorphismes de E. Lorsque $F=\mathbb{K}$, on parle de *forme linéaire sur* E.

Définition 5.2. *Le* noyau *d'une application linéaire est l'ensemble des vecteurs qu'elle envoie sur* 0.

Si f est une application linéaire de E vers F, on note $\mathrm{Ker}(f)$ son noyau. On a donc $\mathrm{Ker}(f) = \{x \in E | f(x) = 0\} = f^{-1}(\{0\}) = f^{-1}(0)$.

Proposition 5.1. *Une application linéaire est injective si et seulement si son noyau est (le sous-espace) nul.*

Le noyau et l'image d'une application linéaire sont des sous-espaces.

Théorème 5.1. *(théorème du rang) Soient E, F des espaces vectoriels de dimensions finies et $f : E \to F$ une application linéaire. On a* $\dim(E) = \dim(\mathrm{Ker}(f)) + \dim(\mathrm{Im}(f))$.

Ce théorème s'appelle théorème du rang car on appelle *rang* de f la dimension de $\mathrm{Im}(f)$. Notation $\mathrm{rg}(f)$.

Proposition 5.2. *La composée de deux applications linéaires est une application linéaire.*

Lorsqu'on compose plusieurs fois un endomorphisme f avec lui-même, on parle de *puissance* et on le note f^n. On a donc $f^1 = f$, $f^n = f^{n-1} \circ f$ si $n \geq 2$.

Définition 5.3. *Un* isomorphisme *est une application linéaire bijective. S'il existe un isomorphisme de E vers F, on dit que E est F sont* isomorphes.

Proposition 5.3. *Si f est un isomorphisme, alors la fonction réciproque f^{-1} est aussi un isomorphisme.*

Théorème 5.2. *Si $f : E \to F$ est un isomorphisme, alors* $\dim(E) = \dim(F)$. *Réciproquement, si* $\dim(E) = \dim(F)$, *alors il existe des isomorphismes de E vers F.*

Théorème 5.3. *Soit $f : E \to F$ une application linéaire et B une base de E.*

(i) f est injective si et seulement si $f(B)$ est linéairement indépendant.
(ii) f est surjective si et seulement si $f(B)$ engendre F.
(iii) f est bijective si et seulement si $f(B)$ est une base de F.

Définition 5.4. *Soit $f : E \to F$ une application linéaire, $u_1, \ldots, u_n$ une base de E et $v_1, \ldots, v_p$ une base de F. La* matrice de f dans ces bases *est la matrice $M(f) = (a_{ij})$ de taille $p \times n$ telle que pour tout $j = 1, \ldots, n$, on a $f(u_j) = \sum_i a_{ij} v_i$.*

Si x est un vecteur dans E, avec $f(x) = y \in F$, alors on peut écrire dans les bases $x = \sum_j \alpha_j u_j$ et $y = \sum_i \beta_j v_i$. Les coefficients α_j et β_i, qui sont dans $\mathbb{K}$, sont liés par le produit matriciel

$$\begin{bmatrix} \beta_1 \\ \vdots \\ \beta_p \end{bmatrix} = M(f) \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}$$

On notera la cohérence de ce produit matriciel : à gauche une colonne, c'est-à-dire une matrice $p \times 1$, et à droite le produit d'une matrice $p \times n$ par une matrice-colonne $n \times 1$. Ce genre de vérification est souvent utile, et permet de vérifier les formules.

Dans le cas particulier où f un endomorphisme de E et $u_1, \ldots, u_n$ une base de E, on parle de *matrice de f dans cette base* : c'est le cas particulier de la définition précédente où les deux espaces sont égaux, et où les deux bases ont égales. On a donc $f(u_j) = \sum_i a_{ij} u_i$.

La matrice d'un endomorphisme est toujours une matrice carrée. Rappelons que l'*ordre* d'une matrice carrée est le nombre de ses lignes (= le nombre de ses colonnes).

Proposition 5.4. *Soient f, g des endomorphismes de E ; alors on a : $M(g \circ f) = M(g)M(f)$, $M(f^n) = M(f)^n$. La fonction $f \mapsto M(f)$ est un isomorphisme de l'espace des endomorphismes de E vers l'espace des matrices carrées d'ordre* $\dim(E)$.

Rappelons que la somme et le produit externe de l'espace des endomorphismes de E sont définis pat $(f + g)(x) = f(x) + g(x)$ et $(af)(x) = af(x)$.

Théorème 5.4. *Soit f un endomorphisme de E, avec* $\dim(E) < \infty$. *Alors f est un isomorphisme $\Leftrightarrow$ f est injectif $\Leftrightarrow$ f est surjectif $\Leftrightarrow$ $M(f)$ est une matrice inversible.*

Un endomorphisme bijectif est appelé un *automorphisme*. Pour un automorphisme f, on définit la *puissance* f^n pour $n \in \mathbb{Z}$.

Concluons cette section avec les changements de bases. Soient f un endomorphisme de E, $U = (u_1, \ldots, u_n)$, $V = (v_1, \ldots, v_n)$ des bases de E. Soit M_U la matrice de f dans la base U et de même pour M_V. On appelle *matrice de passage de U à V* la matrice $P_{U,V} = (b_{ij})$ la matrice $n \times n$ définie par

$$v_j = \sum_i b_{ij} u_i.$$

Autrement dit la j-ème colonne de $P_{U,V}$ comporte les coefficients de la combinaion linéaire de $u_1, \ldots, u_n$ qui représente v_j.

Alors $P_{U,V}$ est l'inverse de la matrice $P_{V,U}$. De plus

$$M_V = P_{V,U} M_U P_{U,V} = P_{U,V}^{-1} M_U P_{U,V}.$$

La *classe de conjugaison* d'une matrice carrée M d'ordre k est l'ensemble des matrices de la formePMP^{-1}, P carrée d'ordre n et inversible.

Corollaire 5.1. *L'ensemble des matrices d'un endomorphisme $f : E \to E$ dans les bases de E forme une classe de conjugaison .*

Exercice 5.1. *Montrer que $(a, b, c, d) \mapsto (a, b, c)$ de $\mathbb{K}^4$ dans $\mathbb{K}^3$ est une application linéaire.*

Exercice 5.2. *Montrer que $x \mapsto x+1$ de $\mathbb{K}$ dans lui-même n'est pas linéaire.*

Exercice 5.3. *Même chose avec $x \mapsto x^2$, sauf si $1 + 1 = 0$ dans $\mathbb{K}$.*

Exercice 5.4. *L'application linéaire de l'exercice 5.1 est-elle injective ? surjective ? quel est son rang ?*

Exercice 5.5. *Soit $f(a, b, c, d) = (a + b, b)$, $f : \mathbb{K}^4 \to \mathbb{K}^2$. Déterminer le noyau, l'image et le rang de f.*

Exercice 5.6. *Montrer que $(a, b, c, d) \mapsto (d, a, c, b)$ est un isomorphisme de $\mathbb{K}^4$ dans lui-même.*

Exercice 5.7. *Soit f l'endomorphisme de $\mathbb{K}^2$ tel que $f(x, y) = (ax+by, cx+dy$. Quelle est sa matrice dans la base $e_1 = (1, 0)$, $e_2 = (0, 1)$. A quelle condition f est-il un isomorphisme ?*

Exercice 5.8. *Soit $f(x, y) = (2x, y)$ l'endomorphisme de $\mathbb{K}^2$. Déterminer sa matrice dans la base de l'exercice 5.7. Puis dans la base $(2, 1), (1, 1)$.*

Exercice 5.9. *Soit f l'endomorphisme de $\mathbb{K}^2$ telle que $f(x, y) = (x + y, y)$. Calculer f^n.*

Exercice 5.10. *Même chose avec $f(x, y) = (y, 0)$.*

Exercice 5.11. *Calculer le noyau et l'image de l'endomorphime de l'exercice 5.10*

6 Rappel d'algèbre linéaire 1 : déterminants

6.1 Développement de Laplace

Soit $A = (a_{ij})_{1 \leq i,j \leq n}$ une matrice $n \times n$ sur $\mathbb{K}$. Le *ij-mineur* de cette matrice (ou par abus, le mineur de a_{ij}) est le déterminant de la matrice $(n-1) \times (n-1)$ obtenue en supprimant dans A la i-ème ligne et la j-ème colonne ; cette matrice est notée A^{ij}. Le mineur en question est donc $\det(A^{ij})$.

Le *développement de Laplace du déterminant selon la ligne i* est

$$\det(A) = \sum_{1 \leq j \leq n} (-1)^{i+j} a_{ij} \det(A^{ij}).$$

Le *développement de Laplace du déterminant selon la colonne j* est

$$\det(A) = \sum_{1 \leq i \leq n} (-1)^{i+j} a_{ij} \det(A^{ij}).$$

6.2 Produit et matrices par blocs

Théorème 6.1. $\det(AB) = \det(A)\det(B)$.

Corollaire 6.1. *Deux matrices carrées conjuguées ont même déterminant.*

Théorème 6.2. *Si la matrice carrée A a une décomposition par blocs $A = \begin{bmatrix} A_1 & B \\ 0 & A_2 \end{bmatrix}$, où A_1, A_2 sont des matrices carrées, alors* $\det(A) = \det(A_1)\det(A_2)$.

6.3 Matrice adjointe

On appelle *ij-cofacteur* de A (ou par abus, cofacteur de a_{ij}) le terme $(-1)^{i+j}\det(A^{ij})$.

La *matrice adjointe* de A, ou *comatrice*, est la matrice

$$\operatorname{adj}(A) = ((-1)^{i+j}\det(A^{ji}))_{1 \leq i,j \leq n}$$

Attention à la transposition : en position ij, on met le cofacteur de a_{ji}.

Théorème 6.3. $A\operatorname{adj}(A) = \operatorname{adj}(A)A = \det(A)I_n$.

Théorème 6.4. *A inversible* $\Leftrightarrow$ $\det(A) \neq 0$. *Dans ce cas* $A^{-1} = \frac{1}{\det(A)}\operatorname{adj}(A)$.

6.4 Rang

Le *rang* d'une matrice (rectangulaire)A est par définition le plus grand r tel que A ait une sous-matrice de taille $r \times r$ de déterminant non nul.

Théorème 6.5. *Le rang d'une matrice est égal au rang de ses vecteurs colonnes, au rang de ses vecteurs lignes, au rang de toute application linéaire dont elle est la matrice, et au nombre de lignes non nulles de sa matrice réduite échelonnée.*

Exercice 6.1. *(questions de cours de cegep-algèbre linéaire 1) Compléter chaque phrase :*

- *Si dans M, deux lignes sont égales, alors* $\det(M) = \ldots$.
- *Si dans M, une ligne est combinaison linéaire des autres lignes, alors* $\det(M) = \ldots$.
- *Si dans M on échange deux lignes, alors son déterminant est* $\ldots$.
- *Si M est carrée d'ordre n et $a \in \mathbb{K}$, alors* $\det(aM) = \ldots$.
- *Le déterminant de l'inverse de M est* $\ldots$.
- *Le déterminant de la transposée de M est* $\ldots$.
- *Le déterminant d'une matrice triangulaire est* $\ldots$.

Exercice 6.2. *Calculer le déterminant des matrices suivantes :* $\begin{bmatrix} a & b \\ -b & a \end{bmatrix}$, $\begin{bmatrix} a & b & c \\ b & c & a \\ c & a & b \end{bmatrix}$, $\begin{bmatrix} 1 & 1 & 1 \\ a & b & c \\ a^2 & b^2 & c^2 \end{bmatrix}$.

Exercice 6.3. *Calculer l'inverse de la matrice* $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

Exercice 6.4. *Soient A, B des matrices de taille $n \times p$ et $p \times n$. Montrer que* $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$.

Exercice 6.5. *Calculer le déterminant de la matrice* $\begin{bmatrix} 0 & a & b & c \\ a & 0 & c & b \\ b & c & 0 & a \\ c & b & a & 0 \end{bmatrix}$.

Exercice 6.6. *Soit M la matrice carrée ayant la décomposition par blocs* $M = \begin{bmatrix} 0 & A \\ C & B \end{bmatrix}$, *où A, C sont des matrices carrées. Quel est son déterminant ?*

7 Intersection et somme directe de sous-espaces

7.1 Intersection

Théorème 7.1. *Soit $\mathcal{E}$ un ensemble de sous-espaces d'un espace vectoriel V. Alors leur intersection $\bigcap_{E\in\mathcal{E}} E$ est un sous-espace de V.*

La preuve est tout-à-fait analogue au cas de deux sous-espaces (voir [1]).

Corollaire 7.1. *Soient $v_1, \dots, v_n$ des vecteurs dans V. Alors le sous-espace $\mathrm{Vect}(v_1, \dots, v_n)$ engendré par ces vecteurs est égal à l'intersection de tous les sous-espaces de V qui contiennent ces vecteurs. C'est aussi le plus petit sous-espace qui contient ces vecteurs.*

Autrement dit, soit $\mathcal{E}$ l'ensemble des sous-espaces E de V tels que $v_1, \dots, v_n \in E$. Alors $\bigcap_{E\in\mathcal{E}} E$ est le sous-espace engendré par $v_1, \dots, v_n$.

Démonstration. 1. L'intersection $\bigcap_{E\in\mathcal{E}} E$ est un sous-espace qui contient $v_1, \dots, v_n$. Elle contient donc toutes les combinaisons linéaires de ces vecteurs, donc contient $\mathrm{Vect}(v_1, \dots, v_n)$.

2. Réciproquement, celui-ci est un sous-espace qui contient $v_1, \dots, v_n$. Donc l'intersection est contenue dans $\mathrm{Vect}(v_1, \dots, v_n)$. □

7.2 Somme

Définition 7.1. *Soit $\mathcal{E}$ un ensemble de sous-espaces d'un espace vectoriel V. La* somme *de ces sous-espaces est l'intersection de tous les sous-espaces qui les contiennent tous.*

D'après le théorème 7.1, la somme des $E \in \mathcal{E}$ est un sous-espace de V.

Lorsque ces sous-espaces sont en nombre fini, disons $E_1, \dots, E_n$, on note aussi $E_1 + \cdots + E_n$ leur somme.

Proposition 7.1. *Soient $E_1, \dots, E_n$ des sous-espaces de V. On a $E_1 + \cdots + E_n = \{v_1 + \cdots + v_n, v_1 \in E_1, \dots, v_n \in E_n\}$.*

Démonstration. On vérifie sans peine que le membre droit est un sous-espace vectoriel F. Il contient chaque E_i. Par conséquent $E_1 + \cdots + E_n \subset F$, par définition de la somme. Réciproquement, si $v \in F$, alors $v = \sum_i v_i$, $v_i \in E_i$; si W est un sous-espace qui sontient tous les E_i, alors $v \in W$; par suite F est contenu dans l'intersection de tous ces W, donc $F \subset E_1 + \cdots + E_n \subset F$. □

Définition 7.2. *Soient $E_1, \dots E_n$ des sous-espaces de V. On dit que la somme des sous-espaces est* directe *si quels que soient les vecteurs $v_1 \in E_1, \dots, v_n \in E_n$, on a $v_1 + \dots + v_n = 0$ seulement si chaque $v_i = 0$* [1].

Exemple 7.1. *Soient $\{(a, b, 0), a, b \in \mathbb{K}\}$ et $\{0, c, d), c, d \in \mathbb{K}\}$, qui sont deux sous-espaces de $\mathbb{K}^3$. Leur somme n'est pas directe.*

Proposition 7.2. *La somme $E_1 + E_2$ est directe si et seulement si $E_1 \cap E_2 = 0$.*

Attention : la généralisation de cette proposition à plus de deux sous-espaces n'est pas que les sous-espaces soient deux à deux d'intersection nulle ; cette dernière condition est insuffisante. Voir l'exercice 7.4.

Démonstration. Si la somme est directe, soit $x \in E_1 \cap E_2$. Notons que $x + (-x) = 0$; comme $x \in E_1$ et $-x \in E_2$, la définition de somme directe implque que $x = 0 = -x$. Donc $E_1 \cap E_2 = 0$.

Réciproquement, si cette intersection est nulle, soit $x_i \in E_i$ tels que $x_1 + x_2 = 0$. Alors $x_1 = -x_2 \in E_2$, donc $x_1 \in E_1 \cap E_2$, et par suite $x_1 = 0$ et enfin $x_2 = 0$. □

Notons la terminologie : on dit que deux sous-espaces E_1, E_1 de V sont *supplémentaires* si leur somme est directe et si cette somme est égale à V.

Proposition 7.3. *Soient $E_1, \dots E_n$ des sous-espaces de V et soit B_i une base de E_i. La somme $E = E_1 + \dots + E_n$ est directe si et seulement si les bases B_i sont disjointes et si $B_1 \cup \dots \cup B_n$ est une base de E.*

Une telle base de V s'appelle une *base adaptée à la somme directe* $E_1 + \dots + E_n$.

Démonstration. Il est clair que la réunion des B_i engendre la somme des sous-espaces E_i.

Supposons que la somme soit directe. Alors les bases doivent être disjointes (sinon il existe $i \neq j$ et un vecteur v non nul dans $E_i \cap E_j$, et ce vecteur a deux écritures distinctes comme somme de vecteurs dans E_i, contradiction). De plus, les vecteurs dans $B_1 \cup \dots \cup B_n$ sont linéairement indépendants : en effet, un relation de dépendance linéaire entre ces vecteurs se réécrit, à cause de la définition 7.2, en n relations de dépendance linéaire entre les vecteurs de chaque B_i ; les coefficients sont donc tous nuls.

1. On appelle souvent ce type de somme directe une *somme directe interne* , le mot "interne" indiquant que les espaces qu'on additionne sont tous sous-espaces d'un même espace.

Réciproquement, supposons que $B_1 \cup \ldots \cup B_n$ est une base de la somme des sous-espaces ; si $\sum_i v_i = 0$, avec $v_i \in E_i$, alors écrivons chaque v_i comme combinaison linéaire de v_i ; nous obtenons alors une relation de dépendance linéaire des vecteurs de $\cup_i B_i$, laquelle doit avoir tous ses coefficients nuls, et donc chaque $v_i = 0$. La somme est donc directe. □

Corollaire 7.2. *On suppose que V est de dimension finie. La somme des sous-espaces $E_1, \ldots E_n$ est directe si et seulement si* $\dim(E_1 + \cdots + E_n) = \dim(E_1) + \cdots + \dim(E_n)$.

Exercice 7.1. *Soient $E_1, \ldots E_n, F_1, \ldots F_n$ des sous-espaces de l'espace vectoriel V tels que V est la somme directe de $E_1, \ldots E_n$, et aussi de $F_1, \ldots, F_n$. On suppose que $E_i \subset F_i$ pour tout i. Montrer que $E_i = F_i$ pour tout i.*

7.3 Sous-espace stables, somme directe et restriction

On dit qu'un sous-espace W est *stable* sous l'endomorphisme f si $f(W) \subset W$.

Définition 7.3. *Soit f un endomorphisme de V et W un sous-espace stable sous f. La* restriction de f *à W est l'endomorphisme de W, noté $f|W$, défini par $(f|W)(w) = f(w)$ pour tout w dans W.*

On a donc $f|W \in \mathrm{End}(W)$.

Proposition 7.4. *Soit f un endomorphisme de V, et supposons que V soit somme directe des sous-espaces $E_1, \ldots, E_n$. Soit B une base adaptée à cette somme directe. Supposons aussi que chaque E_i soit stable sous f. Alors la matrice de f dans cette base a la forme diagonale par blocs*

$$\begin{bmatrix} M_1 & 0 & \cdots & 0 \\ 0 & M_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & M_n \end{bmatrix}, \tag{1}$$

où M_i est la matrice, dans la base B_i, de la restriction de f à E_i.

Une matrice de cette forme est appelée *somme diagonale* des matrices $M_1, \ldots, M_n$. Attention : une somme diagonale de matrices *n'est pas une somme de matrices* au sens usuel. Du reste, on ne pourrait pas additionner des matrices carrées de taille différentes, mais on peut toujours faire leur somme diagonale.

Démonstration. Prouvons-le dans un cas particulier. Supposons que V soit de dimension 4, et somme directe de E_1 et E_2. Supposons que E_1 ait pour base e_1, e_2 et E_2 ait pour base e_3, e_4. Alors e_1, e_2, e_3, e_4 est une base de V. On a par hypothèse $f(e_1) = ae_1 + be_2$, $f(e_2) = ce_1 + de_2$. Donc les deux premières colonnes de la matrice de f dans la base ci-dessus sont $\begin{bmatrix} a \\ b \\ 0 \\ 0 \end{bmatrix}$ et $\begin{bmatrix} c \\ d \\ 0 \\ 0 \end{bmatrix}$. En raisonnant de même avec e_3 et e_4, on trouve que la matrice de f est de la forme $\begin{bmatrix} a & b & 0 & 0 \\ c & d & 0 & 0 \\ 0 & 0 & e & g \\ 0 & 0 & h & i \end{bmatrix}$, qui est une somme diagonale de deux matrices carrées. □

Proposition 7.5. *Soit g, f des endomorphismes de V, W un sous-espace stable sous f et g, et a un scalaire. Alors W est stable sous $af, f+g, g\circ f, f^n$. De plus, $(af)|W = a(f|W)$, $(f+g)|W = f|W + g|W$, $(g\circ f)|W = (g|W)\circ (f|W)$ et $(f|W)^n = f^n|W$.*

Démonstration. □

Exercice 7.2. *Soient $v_i \in V$, $i = 1, \ldots, n$. Soit $\mathbb{K}v_i = \{av_i, a \in \mathbb{K}\}$; montrer que ceci est le sous-espace engendré par v_i. Montrer que $\mathbb{K}v_1 + \cdots + \mathbb{K}v_n$ est le sous-espace engendré par $v_1, \ldots, v_n$. Montrer que cette somme est directe si et seulement si les v_i sont linéairement indépendants.*

Exercice 7.3. *Soit V un espace vectoriel et $p : V \to V$ un endomorphsime* idempotent*, c'est-à-dire $p^2 = p$. Montrer que V est somme directe de* $\mathrm{Ker}(p)$ *et de* $\mathrm{Im}(p)$.

Exercice 7.4. *Soient $E_1, \ldots E_n$ des sous-espaces de V. Montrer que la somme de ces sous-espaces est directe si et seulement si quel que soit $i = 1, \ldots, n-1$, on a $(E_1 + \cdots + E_i) \cap E_{i+1} = 0$.*

Exercice 7.5. *Montrer que si $E_1, \ldots, E_k$ sont des sous-espaces de V tels que la somme $V_1 + \cdots + V_{k-1}$ est directe (égale à E disons), et si la somme $E + E_k$ est directe, alors la somme $E_1 + \cdots + E_k$ est directe.*

Exercice 7.6. *Soient $W_1, \dots, W_k$ des sous-espaces de V dont la somme W est directe. Montrer que la dimension de W est la somme des dimensions des W_i.*

Exercice 7.7. *Enoncer et démontrer une réciproque de la proposition 7.4.*

8 Base du produit d'espaces vectoriels

Théorème 8.1. *Soient $E_1, \dots, E_n$ des espaces vectoriels sur $\mathbb{K}$. Soit B_i une base de E_i. Alors $E_1 \times \cdots \times E_n$ a pour base $\bigcup_{1 \leq i \leq n} B'_i$, où $B'_i = 0 \times \cdots \times 0 \times B_i \times 0 \times \cdots \times 0$ (le B_i en position i). En particulier,* $\dim(E_1 \times \cdots \times E_n) = \sum_{1 \leq i \leq n} \dim(B_i)$.

Exemple 8.1. *Si X a pour base $x_1, x_2, \dots$, et Y a pour base $y_1, y_2, \dots$, alors $X \times Y$ a pour base les vecteurs $(x_1, 0), (x_2, 0), \dots, (0, y_1), (0, y_2), \dots$.*

Démonstration. □

Exercice 8.1. *Soient deux sous-espaces vectoriels V et W d'un espace vectoriel E de dimension finie, et $\varphi : V \times W \to E$, $(v, w) \mapsto v + w$. Montrer que φ est une application linéaire, que son image est la somme $V + W$ et que son noyau est* $\mathrm{Ker}(\varphi) = \{(h, -h), h \in V \cap W\}$. *Montrer que* $\mathrm{Ker}(\varphi)$ *et $V \cap W$ sont isomorphes. En déduire que* $\dim(V) + \dim(W) = \dim(V + W) + \dim(V \cap W)$.

9 Polynôme caractéristique

9.1 Substitution d'une matrice ou d'un endomorphisme dans un polynôme

Si P est un polynôme, $P = a_n x^n + \cdots + a_1 x + a_0 \in \mathbb{K}[x]$, M est une matrice carrée d'ordre k sur $\mathbb{K}$, et f un endomorphisme du $\mathbb{K}$-espace vectoriel V, on définit

$$P(M) = a_n M^n + \cdots + a_1 M + a_0 I_k,$$

qui est une matrice carrée de même taille que M, et on définit aussi

$$P(f) = a_n x^n + \cdots + a_1 x + a_0 \,\mathrm{id}_V$$

(où les puissances de f sont ses puissance pour la composition, $f^2 = f \circ f$ etc...), qui est un endomorphisme de V.

Proposition 9.1. *1. Soit M une matrice carrée d'ordre k sur $\mathbb{K}$. La fonction $\mathbb{K}[x] \to M_{kk}(\mathbb{K}), P(x) \mapsto P(M)$, est une application linéaire qui préserve la multiplication (c'est-à-dire $\forall P, Q \in \mathbb{K}[x], (PQ)(M) = P(M)Q(M)$).*

2. Soit $f \in \mathrm{End}(V)$. La fonction $\mathbb{K}[x] \to \mathrm{End}(V), P(x) \mapsto P(f)$, est une application linéaire qui transforme la multiplication des polynômes en la composition des endomorphismes.

Cette dernière propriété signifie que si P, Q sont des polynômes, alors $(PQ)(f) = P(f) \circ Q(f)$.

Démonstration. 1. On prend $P = \sum_i a_i x^i, Q = \sum_j b_j x^j \in \mathbb{K}[x]$ et $\alpha \in \mathbb{K}$. On a $(P+Q)(M) = (\sum_i (a_i + b_i)x^i)(M) = \sum_i (a_i + b_i)M^i = \sum_i a_i M^i = \sum_i b_i M^i = P(M) + Q(M)$. De plus, $(\alpha P)(M) = \sum_i \alpha a_i x^i)(M) = \sum_i \alpha a_i M^i = \alpha \sum_i a_i M^i = \alpha P(M)$. Enfin $(PQ)(M) = (\sum_i a_i b_j x^{i+j})(M) = \sum_i a_i b_j M^{i+j} = (\sum_i a_i M^i)(\sum_j b_j M^j) = P(M)Q(M)$.

2. On utilise la proposition 5.4, qui implique que si M est la matrice de f dans une base, alors $P(M)$ est la matrice de $P(f)$ dans cette base. □

Corollaire 9.1. *Les matrices (resp. les endomorphismes) de la forme $P(M)$ (resp. $P(f)$), $P \in \mathbb{K}[x]$, commutent entre elles (resp. eux) pour la multiplication (resp. la composition).*

Exercice 9.1. *a) On suppose qu'on a une matrice carrée M satisfaisant $M^2 = 0$. Montrer que les matrices $I - M$ et $I + M$ sont inverses l'une de l'autre. Indication : utiliser l'identité $(1-x)(1+x) = 1-x^2$ et la proposition 9.1.*

b) Quel est l'énoncé analogue pour les endomorphismes ?

c) Comment peut-on généraliser a) si on a $M^n = 0$?

9.2 Polynôme caractéristique

Définition 9.1. *Le* polynôme caractéristique *d'une matrice carrée M d'ordre n est* $\det(xI_n - M) \in \mathbb{K}[x]$.

Théorème 9.1. *(Cayley-Hamilton) Soit $P(x)$ le polynôme caractéristique de M. On a $P(M) = 0$.*

Ecrivons $P(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$. Alors $P(M) = a_n M^n + a_{n-1}M^{n-1} + \cdots + a_1 M + a_0 I_n$.

Voir l'exercice 9.6 pour la détermination de certains coefficients du polynôme caractéristique ; en particulier, son *coefficient dominant* a_n vaut 1.

Exemple 9.1. $n = 2$, $M = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, $P(x) = x^2 - (a+d)x + ad - bc = x^2 - \mathrm{Tr}(M)x + \det(M)$. *Donc* $\begin{bmatrix} a & b \\ c & d \end{bmatrix}^2 - (a+d)\begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} ad-bc & 0 \\ 0 & ad-bc \end{bmatrix} = 0$, *ce qu'on peut vérifier directement.*

Lemme 9.1. *Si une matrice carrée d'ordre h a des coefficients qui sont des polynômes de degré ≤ 1, alors son déterminant est un polynôme de degré $\leq h$.*

Démonstration. Ça se démontre par une récurrence évidente, en utilisant le développement de Laplace. □

Preuve du théorème 9.1. Soit $N = xI_n - M$ et A la matrice adjointe de N. Alors par le lemme chaque coefficient de A est de degré au plus $n-1$. D'après le théorème 6.3, on a $NA = P(x)I_n$, où $P(x) = \det(N)$, le polynôme caractéristique de M. On peut écrire $A = A_{n-1}x^{n-1} + \cdots + A_1x + A_0$, où les matrices A_i sont carrées d'ordre n et à coefficients dans $\mathbb{K}$.

On a donc $(xI_n - M)\sum_i A_i x^i = P(x)I_n = (a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0)I_n$. On en déduit les équations, par identification des monômes en x :

$$\begin{aligned} A_{n-1} &= a_nI_n \\ A_{n-2} - MA_{n-1} &= a_{n-1}I_n \\ &\vdots \\ A_0 - MA_1 &= a_1I_n \\ -MA_0 &= a_0I_n \end{aligned}$$

On en déduit

$$\begin{aligned} P(M) &= a_nM^n + a_{n-1}M^{n-1} + \cdots + a_1M + a_0I_n \\ &= M^n(a_nI_n) + M^{n-1}(a_{n-1}I_n) + \cdots + M(a_1I_n) + a_0I_n \\ &= M^nA_{n-1} + M^{n-1}(A_{n-2} - MA_{n-1}) + \cdots + M(A_0 - MA_1) - MA_0 \\ &= M^nA_{n-1} + M^{n-1}A_{n-2} - M^nA_{n-1} + \cdots + MA_0 - M^2A_1 - MA_0 = 0. \end{aligned}$$

□

Définition 9.2. *Le* polynôme caractéristique d'un endomorphisme *de V est le polynôme caractéristique de sa matrice dans une base de V.*

Ceci est bien défini car on a la

Proposition 9.2. *Si deux matrices carrées A, B sont conjuguées, elles ont le même polynôme caractéristique.*

Démonstration. On a $B = P^{-1}AP$. Alors $xI - B = P^{-1}(xI - A)P$. On prend les déterminants, on utilise la formule du produit, en remarquant que $det(P^{-1}) = det(P)^{-1}$, et on en déduit le résultat. □

Corollaire 9.2. *Soit $P(x)$ le polynôme caractéristique d'un endomorphisme. Alors $P(f) = 0$.*

Ici $P(f)$ est l'endomorphisme $f^n + a_{n-1}f^{n-1} + \cdots + a_1 f + a_0 \,\mathrm{id}$.

Démonstration. On choisit une base, et soit M la matrice de f dans cette base. Soit $P(x) = \det(xI_n - M)$. Alors $P(x)$ est le polynôme caractéristique à la fois de M et de f. De plus, $P(M)$ est la matrice de $P(f)$ dans cette base (proposition 5.4) ; comme $P(M) = 0$, on $P(f) = 0$. □

Exercice 9.2. *Calculer le polynôme caractéristique de la matrice* $\begin{bmatrix} 2 & -1 & 2 \\ 10 & -5 & 7 \\ 4 & -2 & 2 \end{bmatrix}$.

Exercice 9.3. *On suppose que M est la matrice (1). Montrer que son polynôme caractéristique est le produit des polynômes caractéristiques des M_i.*

Exercice 9.4. *Montrer que si A, B sont des matrices carrées, alors AB et BA ont le même polynôme caractéristique. Indication : supposer B inversible.*

Exercice 9.5. ** Soit $f \in \mathrm{End}(V)$. On dit que V est* irréductible sous l'action de f *si V ne possède pas de sous-espace stable sous f, sauf V et $\{0\}$ (ces deux sous-espaces sont toujours trivialement stables). Montrer que si le polynôme caractéristique de f est irréductible, alors V est irréductible sous l'action de f.*

Exercice 9.6. ** Montrer que le polynôme caractéristique de M, matrice $n \times n$, est de la forme $x^n - \mathrm{Tr}(M)x^{n-1} + \cdots + (-1)^n \det(M)$. Utiliser le développement de Laplace et le lemme 9.1.*

Exercice 9.7. ** Une preuve fausse classique du théorème de Cayley-Hamilton est comme suit : la substitution de x par M dans le polynôme caractéristique* $\det(xI_n - M)$ *est égale à* $\det(MI_n - M)$*, donc à* $\det(M - M) =$

$\det(0) = 0$. *Montrer que la même "preuve" démontre que si on pose* $Q(x) = \mathrm{Tr}(xI_n - M)$, *alors* $Q(M) = 0$ *(on peut aussi remplacer* Tr *par une autre fonction des matrices vers les scalaires). Puis vérifier que cette dernière identité n'est pas vraie.*

10 Valeurs, vecteurs et sous-espaces propres

Définition 10.1. *Soit* f *un endomorphisme de l'espace vectoriel* V *sur* $\mathbb{K}$. *On dit que* $\lambda \in \mathbb{K}$ *est une* valeur propre *de* f *s'il existe un vecteur* v *non nul dans* V *tel que* $f(v) = \lambda v$. *Un tel vecteur est alors appelé* vecteur propre, *et on dit qu'il est* associé à la valeur propre λ.

Définition 10.2. *Soit* A *une matrice carrée d'ordre* n. *Soit* f *l'endomorphisme de l'espace vectoriel des colonnes* $V = M_{n1}(\mathbb{K})$ *dont* A *est la matrice. Autrement dit, si* $v \in V$, *alors* $f(v) = Av$. *Les* valeurs propres *(resp.* vecteurs propres*) de* A *sont les valeurs propres (resp. vecteurs propres) de* f.

Remarquez qu'un vecteur propre est, par définition, toujours non nul.

Proposition 10.1. *Un endomorphisme de* V *a la valeur propre* λ *si et seulement si* $f - \lambda\,\mathrm{id}$ *n'est pas un automorphisme de* V.

Démonstration. Si λ est une valeur propre de f, alors il existe v non nul tel que $f(v) = \lambda v$. Donc $(f - \lambda\,\mathrm{id})(v) = 0$. Donc $f - \lambda\,\mathrm{id}$ n'est pas un automorphisme de V, car non injectif (proposition 5.1).

Réciproquement, si $f - \lambda\,\mathrm{id}$ n'est pas un automorphisme de V, alors $f - \lambda\,\mathrm{id}$ n'est pas injectif (proposition 5.1). Donc $Ker(f - \lambda\,\mathrm{id}) \neq 0$. Il existe donc $v \neq 0$ dans V tel que $(f - \lambda\,\mathrm{id})(v) = 0$. C'est-à-dire : $f(v) = \lambda v$. Donc λ est une valeur propre de f. □

Le *déterminant* d'un endomorphisme de V est par définition le déterminant de sa matrice dans une base de V. Ceci est bien défini, car si on change de base, alors on conjugue la matrice (corollaire 9.2), et par suite le déterminant reste le même.

Il découle des théorèmes 5.4 et 6.4 qu'un endomorphisme est un automorphisme si et seulement si son déterminant est non nul. On obtient donc le

Corollaire 10.1. *Un endomorphisme a la valeur propre* λ *si et seulement sile déterminant de* $f - \lambda\,\mathrm{id}$ *est nul.*

Corollaire 10.2. *Les valeurs propres d'une matrice (resp. d'un endomorphisme) sont les racines du polynôme caractéristique.*

Définition 10.3. *Soit λ une valeur propre de l'endomorphisme f de V. Le* sous-espace propre *associé à λ est* $\mathrm{Ker}(f - \lambda Id)$.

Autrement dit, c'est l'ensemble des $v \in V$ tels que $f(v) = \lambda v$. Ce sous-espace n'est jamais nul, λ étant supposé être une valeur propre. Remarquez que si 0 est une valeur propre de f (c'est-à-dire, f est non injectif), alors le sous-espace propre associé est le noyau de f.

Exemple 10.1. *Soit $f(x, y) = (x, 2y)$. La matrice de cet endomorphisme de $\mathbb{K}^2$ dans la base canonique est $\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$. Son polynôme caractéristique est $(x-1)(x-2)$ et les valeurs propres sont donc $1, 2$. Les sous-espaces propres associés sont $\mathbb{K}e_1$ et $\mathbb{K}e_2$. L'espace a une base constitué de vecteurs propres, donc f est diagonalisable (voir plus loin pour la définition).*

Exemple 10.2. *Cette fois, $f(x, y) = (x + y, y)$. La matrice est $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$. Son polynôme caractéristique est $(x-1)^2$. L'unique valeur propre est 1, et le sous-espace propre associé est le noyau de l'endomorphisme $(x, y) \mapsto (y, 0)$, c'est-à-dire Ke_1.*

Exemple 10.3. *Soit $\mathbb{C}$ vu comme un espace vectoriel sur $\mathbb{R}$. Il est de dimension 2, avec base 1 et i. Soit f l'endomorphisme de $\mathbb{C}$ défini par $f(z) = iz$. Sa matrice dans la base ci-dessus est $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$. Son polynôme caractéristique est $x^2 + 1$, il n'y a pas de valeurs propres (sur $\mathbb{R}$).*

Proposition 10.2. *La somme des sous-espaces propres d'un endomorphisme est une somme directe.*

Nous prouvons cette proposition plus loin, après quelques préparations.

Lemme 10.1. *Soient $P_1, \ldots, P_k$ des polynômes premiers entre eux deux à deux, et P leur produit. Soit f un endomorphisme de V. Alors* $\mathrm{Ker}(P(f))$ *est la somme directe des sous-espaces* $\mathrm{Ker}(P_i(f))$, $i = 1, \ldots, k$.

Démonstration. 1. Cas $k = 2$. On suppose donc $P = P_1P_2$. Par le théorème de Bezout, il existe des polynômes Q_1, Q_2 tels que $P_1Q_1 + P_2Q_2 = 1$. On note $P_i(f) = g_i, Q_i(f) = h_i$. Alors $g_1 \circ h_1 + g_2 \circ h_2 = \mathrm{id}_V$. Soit $v \in \mathrm{Ker}(P(f))$. Alors $v = \mathrm{id}(v) = (g_1 \circ h_1 + g_2 \circ h_2)(v) = g_1 \circ h_1(v) + g_2 \circ h_2(v) = u_2 + u_1$, où l'on pose $u_2 = g_1 \circ h_1(v)$ et $u_1 = g_2 \circ h_2(v)$. On a $g_2(u_2) = g_2 \circ g_1 \circ h_1(v)$. D'après le corollaire 9.1, les endomorphismes g_2, g_2, h_1 commutent entre eux ; donc $g_2(u_2) = h_1(g_2 \circ g_1(v))$. De plus, $P = P_2P_1$, donc par la proposition 9.1,

$P(f) = g_2 \circ g_1$, et par suite $g_2 \circ g_1(v) = 0$. Donc $g_2(u_2) = 0$ et $u_2 \in \mathrm{Ker}(g_2) = \mathrm{Ker}(P_2(f))$. De manière analogue, $u_1 \in \mathrm{Ker}(P_1(f))$. Comme $v = u_1 + u_2$, on a prouvé que $\mathrm{Ker}(P(f))$ est la somme des sous-espaces $\mathrm{Ker}(P_i(f))$, $i = 1, 2$.

Reste à montrer que la somme est directe : prenons $v \in \mathrm{Ker}(P_1(f)) \cap \mathrm{Ker}(P_2(f)) = \mathrm{Ker}(g_1) \cap \mathrm{Ker}(g_2)$. On a $Q_1P_1 + Q_2P_2 = 1$, donc par la proposition 9.1, $v = h_1 \circ g_1(v) + h_2 \circ g_2(v) = 0 + 0 = 0$.

2. Cas général : récurrence sur k. Supposons donc $k \geq 3$ est le résultat prouvé pour $k-1$ polynômes $P-1, \ldots, P_{k-1}$. Posons $Q = P_1 \cdots P_{k-1}$. Par le cas $k = 2$, $\mathrm{Ker}(P(f))$ est la somme directe de $\mathrm{Ker}(Q(f))$ et de $\mathrm{Ker}(P_k(f))$. Par hypothèse de récurrence, $\mathrm{Ker}(Q(f))$ est la somme directe des $\mathrm{Ker}(P_i(f))$, $i = 1, \ldots, k-1$. Donc, par l'exercice 7.5, $\mathrm{Ker}(P(f))$ est la somme directe des $\mathrm{Ker}(P_i(f))$, $i = 1, \ldots, k$. □

Preuve de la proposition 10.2. On applique le lemme aux polynômes $x - \lambda$, λ valeur propre de l'endomorphisme : ces polynômes sont premiers entre eux deux à deux.

Donc le noyau de $\prod_\lambda (f - \lambda \,\mathrm{id})$ (où le produit est sur toutes les valeurs propres de f) est égal à la somme directe des noyaux des $f - \lambda \,\mathrm{id}$, c'est-à-dire, des sous-espaces propres de f. □

Exercice 10.1. *Montrer que f est un automorphisme si et seulement si 0 n'est pas vecteur propre de f.*

Exercice 10.2. *Montrer que si f est un automorphisme et λ une valeur propre de f, alors λ^{-1} est une valeur propre de f^{-1}. Comparer les vecteurs propres de f et f^{-1}.*

Exercice 10.3. *Si λ est une valeur propre de f, et $P(x)$ un polynôme, montrer que $P(\lambda)$ est une valeur propre de $P(f)$ (commencer par $P(x) = x^n$).*

Exercice 10.4. *soit f un endomorphisme de V, v un vecteur non nul et D le sous-espace $D = Vect(v)$. Montrer que D est stable sous f si et seulement si v est un vecteur propre de f.*

11 Diagonalisation

Définition 11.1. *Un endomorphisme de V est* diagonalisable *si V a une base constituée de vecteurs propres de f.*

Proposition 11.1. *Un endomorphisme f de V est diagonalisable si et seulement si V a une base où la matrice de f est diagonale.*

Démonstration. C'est direct, en utilisant la définition de la matrice d'un endomorphisme. □

Corollaire 11.1. *f est diagonalisable si seulement sa matrice est conjuguée à une matrice diagonale.*

Démonstration. Si f est diagonalisable, on utilise le corollaire 5.1. Réciproquement, si la matrice de f dans une base est conjuguée à une matrice diagonale, alors par la section 5 (matrice de changement de base), il existe une base où la matrice de f est diagonale ; donc f est diagonalisable par la proposition précédente. □

Par définition, une matrice est dite *diagonalisable* si elle est conjuguée à une matrice diagonale. De manière équivalente (par le corollaire), l'endomorphisme qu'elle représente est diagonalisable.

Proposition 11.2. *Un endomorphisme de V est diagonalisable si et seulement si la somme de ses sous-espaces propres est V.*

Démonstration. Si f est diagonalisable, V a une base constitué de vecteurs propres. Tout vecteur v est combinaison linéaire de ces vecteurs. Comme toute combinaison linéaire de vecteurs propres correspondant à la valeur propre λ est un vecteur propre pour λ, qui est dans $\mathrm{Ker}(f - \lambda\,\mathrm{id})$, on en déduit que v est dans la somme des $\mathrm{Ker}(f - \lambda\,\mathrm{id})$, les sous-espaces propres de f.

Réciproquement, si V est somme des sous-espaces propres, c'est une somme directe (proposition 10.2). Une base de V adaptée à cette somme directe nous fournit donc une base de vecteurs propres de f. □

Théorème 11.1. *Si le polynôme caractéristique de f est égal à $P(x) = \prod_{1 \leq i \leq n}(x - \lambda_i)$, avec $\lambda_1, \ldots, \lambda_n$ distincts, alors f est diagonalisable.*

La réciproque du théorème n'est pas vraie cependant, comme le montre l'endomorphisme identité.

On peut exprimer l'hypothèse du théorème de la manière suivante : le polynôme caractéristique de f est scindé et n'a que des racines simples.

Démonstration. D'après le théorème de Cayley-Hamilton, $P(f) = 0$. Donc $V = \mathrm{Ker}(P(f))$.

D'après le lemme 10.1, $\mathrm{Ker}(P(f))$ est somme des $\mathrm{Ker}(f - \lambda_i\,\mathrm{id})$, et on applique la proposition 11.2. □

Exercice 11.1. *On suppose que dans la base $v_1, \dots, v_n$ la matrice de f est diagonale. Indiquer quelles sont les valeurs propres, et les sous-espaces propres de f.*

Exercice 11.2. *Pour l'endomorphisme dont la matrice est donnée en l'exercice 9.2, calculer les valeurs propres et les sous-espaces propres. Est-il diagonalisable ?*

Exercice 11.3. *Montrer qu'un endomorphisme idempotent est diagonalisable (voir l'exercice 7.3).*

Exercice 11.4. *Calculer les valeurs propre de la matrice* $\begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$. *Montrer qu'elle n'est pas diagonalisable.*

Exercice 11.5. *Calculer les valeurs propre de la matrice* $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 0 & 1 & 3 \end{bmatrix}$. *Montrer qu'elle est diagonalisable.*

Exercice 11.6. *On considère un endomorphisme f dont la matrice dans la base $v_1, \dots, v_n$ est diagonale, avec des éléments diagonaux λ_i distincts. Soit g un endomorphisme qui commute avec f. Montrer que $f(g(v_i)) = \lambda_i g(v_i)$. En déduire qu'il existe $a_i \in \mathbb{K}$ tel que $g(v_i) = a_i v_i$ et que g est diagonalisable. Indication : le sous-espace propre de f pour λ_i est de dimension 1.*

Exercice 11.7. *Soit V un espace vectoriel de dimension finie, f un endomorphisme de V, Λ l'ensemble des valeurs propres de f, et pour $\lambda \in \Lambda$, n_λ la dimension du sous-espace propre de V. Montrer que f est diagonalisable si et seulement si $\sum_{\lambda \in \Lambda} n_\lambda = \dim(V)$. Indication : utiliser le corollaire 7.2.*

12 Polynôme minimal

Définition 12.1. *Le* polynôme minimal *d'une matrice carrée M (resp. d'un endomorphisme f) est le polynôme unitaire $P(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 \in \mathbb{K}[x]$, de degré n minimum, tel que $P(M) = M^n + a_1 M^{n-1} + \cdots + a_1 M + a_0 I = 0$) (resp. $P(f) = f^n + a_1 f^{n-1} + \cdots + a_1 f + a_0 \,\mathrm{id} = 0$).*

Le polynôme minimal est unique : car si $P(x) = x^n + \cdots$ et $Q(x) = x^n + \cdots$ sont deux polynômes distincts tels que $P(M) = Q(M) = 0$, on trouve que $(P - Q)(M) = 0$, $P - Q \neq 0$, et on en déduit un polynôme de degré plus petit.

Le polynôme minimal existe, car le polynôme caractéristique, disons $P(x)$, est un polynôme unitaire et le théorème de Cayley-Hamilton nous assure que $P(f) = 0$.

Remarquons aussi que si un polynôme non nul Q satisfait $Q(f) = 0$ (ou $Q(M) = 0$), alors $\deg(Q)$ est $\geq$ au degré du polynôme minimal (diviser Q par son coefficient dominant).

Exemple 12.1. *Le polynôme minimal de la matrice identité est $x-1$. Celui de la matrice $\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$ est $(x-1)(x-2)$. Celui de la matrice $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ est $(x-1)^2$. On vérifie en effet pour ces deux matrices que le polynôme minimal ne peut être de degré 1 (donc de la forme $x-a, a \in \mathbb{K}$).*

Proposition 12.1. *Le polynôme minimal divise le polynôme caractéristique.*

Démonstration. Soit A le polynôme caractéristique et B le polynôme minimal. On fait la division euclidienne du premier par le second : $A = BQ + R$, où $\deg(R) < \deg(B)$. On a alors, par la proposition 9.1, le théorème 9.1, et par définition du polynôme minimal : $0 = A(f) = B(f)Q(f) + R(f) = R(f)$ et on doit avoir $R = 0$ par minimalité du degré du polynôme minimal. □

La même preuve démontre le

Corollaire 12.1. *Si M est une matrice (resp. si f est un endomorphisme) et $A(x)$ un polynôme telle que $A(M) = 0$ (resp. tel que $A(f) = 0$), alors A est un multiple du polynôme minimal de M (resp. de f).*

Ce corollaire permet de calculer le polynôme minimal de f, connaissant un tel polynôme $A(x)$. En effet, le polynôme minimal sera forcément un diviseur de $A(x)$, et comme le polynôme minimal est unitaire, il n'y a qu'un nombre fini de diviseurs. Pour montrer que A est le polynôme minimal, il suffit de vérifier que pour tout diviseur unitaire B de A, de degré strictement plus petit que A, on a $B(f) \neq 0$.

Par exemple (sur le corps des réels) : si on sait que $f^3 = \mathrm{id}$, le polynôme minimal est un diviseur de $x^3 - 1$. On a $x^3 - 1 = (x-1)(x^2+x+1)$ (factorisation en polynômes irréductibles dans $\mathbb{R}[x]$). Le polynôme minimal de f est donc soit x^2+x+1, soit $x-1$, soit x^3-1. Et c'est x^3-1 si et seulement si $f - \mathrm{id} \neq 0$ et $f^2 + f + \mathrm{id} \neq 0$.

Corollaire 12.2. *Si le polynôme caractéristique est scindé sur $\mathbb{K}$, le polynôme minimal aussi.*

Démonstration. En effet un polynôme qui divise un polynôme scindé est scindé. □

Proposition 12.2. *Les valeurs propres sont les racines du polynôme minimal.*

Démonstration. Toute racine du polynôme du polynôme minimal est une racine du polynôme caractéristique, car celui-là divise celui-ci, d'après la proposition précédente. Donc toute racine du polynôme minimal est une valeur propre (corollaire 10.2).

Réciproquement, soit λ une valeur propre de f, et v une valeur propre associée. Alors $f(v) = \lambda v$. Il existe une base $v = v_1, v_2, \ldots, v_n$ de V. La matrice de f dans cette base a la forme triangulaire par blocs $M = \begin{bmatrix} \lambda & B \\ 0 & D \end{bmatrix}$. Si $P(x)$ est le polynome minimal, on a $0 = P(M) = \begin{bmatrix} P(\lambda) & B' \\ 0 & P(D) \end{bmatrix}$. Donc $P(\lambda) = 0$. □

Théorème 12.1. *Un endomorphisme (resp. une matrice) est diagonalisable si et seulement si son polynôme minimal est scindé et n'a que des racines simples.*

Démonstration. Le polynôme minimal d'une matrice diagonale est $\prod_\lambda (x - \lambda)$, où le produit est sur tous les λ sur la diagonale, sans répétition. Ce polynôme est scindé et n'a que des racines simples. On en déduit une des implications du théorème.

Réciproquement, supposons que le polynôme minimal de $f \in \mathrm{End}(V)$ est scindé et n'a que des racines simples. On applique alors le lemme 10.1 et la proposition 11.2. □

Exercice 12.1. *Quel est le polynôme minimal de la matrice diagonale dont la diagonale est* $1, 1, 1, 1, 2, 2$ *?*

Exercice 12.2. *Plus généralement : une matrice diagonale dont la diagonale est* $a_1, \ldots, a_k$ *?*

Exercice 12.3. *Calculer le polynôme minimal de la matrice* $\begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$.

Exercice 12.4. *On suppose que M est la matrice (1). Montrer que son polynôme minimal est le plus petit multiple commun des polynômes minimaux des M_i.*

13 Triangularisation

Définition 13.1. *Un endomorphisme est* triangularisable *s'il existe une base où sa matrice est triangulaire. Une matrice est* triangularisable *si elle est conjuguée à une matrice triangulaire.*

De manière analogue, une matrice est triangularisable si et seulement si elle est la matrice d'un endomorphisme triangularisable.

Théorème 13.1. *Une matrice (resp. un endomorphisme) est triangularisable si et seulement si son polynôme minimal est scindé.*

Démonstration. Soit f un endomorphisme de V, espace de dimension n.

Si la matrice de f dans une base de V est triangulaire, alors son polynôme caractéristique est scindé, car c'est le produit de $x-a$ pour tous les éléments diagonaux a de la matrice. Donc le polynôme minimal aussi (corollaire 12.2).

Réciproquement, si le polynôme minimal est scindé, alors f a au moins une valeur propre λ (proposition 12.2). Soit v un vecteur propre associé et $v = v_1, v_2, \ldots, v_n$ une vase de V. Dans cette base, la matrice de f a la forme triangulaire par blocs $M = \begin{bmatrix} \lambda & l \\ 0 & D \end{bmatrix}$, où l est une matrice-ligne de longeur $n-1$, et D une matrice carrée d'ordre $n-1$. En raisonnant comme dans la preuve de la proposition 12.2, on voit que $P(D) = 0$, où P est le polynôme minimal de f. Donc le polynôme minimal de D est scindé, car il divise P (corollaire 12.1). En raisonnant par récurrence sur la dimension, on obtient que D est conjuguée à une matrice triangulaire supérieure : $GDG^{-1} = T$. Alors $\begin{bmatrix} 1 & 0 \\ 0 & G \end{bmatrix} \begin{bmatrix} \lambda & l \\ 0 & D \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & G^{-1} \end{bmatrix} = \begin{bmatrix} \lambda & l \\ 0 & GD \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & G^{-1} \end{bmatrix} = \begin{bmatrix} \lambda & lG^{-1} \\ 0 & GDG^{-1} \end{bmatrix} = \begin{bmatrix} \lambda & lG^{-1} \\ 0 & T \end{bmatrix}$, qui est triangulaire supérieure. Comme $\begin{bmatrix} 1 & 0 \\ 0 & G \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & G^{-1} \end{bmatrix}$, on voit que M est conjuguée à une matrice triangulaire supérieure. Donc par changement de base, f a une matrice triangulaire supérieure.

□

Corollaire 13.1. *Le polynôme minimal est scindé si et seulement si le polynôme caractéristique est scindé.*

Démonstration. Si le polynôme caractéristique est scindé, le polynôme minimal aussi (corollaire 12.2). Réciproquement, si le polynôme minimal est scindé, nous pouvons supposer que f a, dans une certaine base, une matrice

triangulaire. Le polynôme caractéristique de cette matrice est scindé, car c'est le produit de $x - a$ pour tous les éléments diagonaux a de la matrice ; celui de f est donc aussi scindé. □

Le théorème fondamental de l'algèbre implique le

Corollaire 13.2. *Si le corps de base est $\mathbb{C}$, tout endomorphisme, et toute matrice carrée, est triangularisable.*

Exercice 13.1. *Montrer que la matrice* $\begin{bmatrix} 0 & 1 \\ 2 & 0 \end{bmatrix}$ *est triangularisable sur $\mathbb{R}$ mais pas sur $\mathbb{Q}$.*

Exercice 13.2. *Montrer que la matrice de f dans la base $v_1, \ldots, v_n$ est triangulaire supérieure si et seulement si elle est triangulaire inférieure dans la base $v_n, \ldots, v_1$.*

Exercice 13.3. *Montrer que la matrice de f dans la base $v_1, \ldots, v_n$ est triangulaire supérieure si seulement si pour tout $j = 1, \ldots, n$, $f(v_j) \in \mathrm{Vect}(v_1, \ldots, v_j)$.*

Exercice 13.4. *Quelle est la condition dans l'exercice précédent si on remplace "triangulaire supérieure" par "triangulaire supérieure avec des 0 sur la diagonale" ? Montrer qu'alors $f^n = 0$.*

Exercice 13.5. *Utiliser l'exercice précédent pour montrer qu'une matrice strictement triangulaire (= triangulaire avec des 0 sur la diagonale) est nilpotente.*

Exercice 13.6. *Soit M une matrice* unipotente*, c'est-à-dire $M = I_n + N$, où N est une matrice triangulaire supérieure avec des 0 sur la diagonale. Montrer que M est inversible, d'inverse $I_n - N + N^2 - N^3 + \cdots + (-1)^{n-1} N^{n-1}$.*

Exercice 13.7. *On considère l'espace vectoriel des polynômes en la variable x et de degré au plus $n - 1$. Montrer que l'opération de dérivation est un endomorphisme dont la matrice dans la base $1, x, \ldots, x^{n-1}$ est triangulaire.*

Exercice 13.8. *On considère l'espace vectoriel des matrices 2 fois 2, avec sa base des matrices élémentaires E_{ij}, $i, j = 1, 2$. Soit f l'endomorphisme de cet espace défini par $f(M) = ME_{12}$. Montrer qu'en ordonnant la base comme il faut, la matrice de f est triangulaire.*

14 Sous-espaces caractéristiques

Définition 14.1. *Soit f un endomorphisme dont le polynôme minimal est scindé, de la forme*

$$\prod_{1\leq i\leq k}(x-a_i)^{n_i} \tag{2}$$

$n_i \geq 1$, a_i racines distinctes. Le sous-espace caractéristique de f associé à la valeur propre a_i *est* $\operatorname{Ker}(f - a_i \operatorname{id})^{n_i}$.

Proposition 14.1. *Le sous-espace propre associé à la valeur propre a_i est contenu dans le sous-espace caractéristique associé à cette valeur propre*

Démonstration. □

Théorème 14.1. *Soit f un endomorphisme de V dont le polynôme minimal est scindé. Alors V est somme directe des sous-espaces caractéristiques de f.*

Démonstration. C'est une conséquence du lemme 10.2, car les polynômes $(x-a_i)^{n_i}$ sont deux à deux premiers entre eux. □

Proposition 14.2. *Tout sous-espace caractéristique f est stable sous f.*

Démonstration. On va montrer plus généralement que si $P(x)$ est un polynôme, alors $\operatorname{Ker}(P(f))$ est stable sous f. Prenons en effet un $v \in \operatorname{Ker}(P(f))$. Alors $P(f)(v) = 0$. Or $f \circ P(f) = P(f) \circ f$ (corollaire 9.1). Donc $0 = f(P(f)(v)) = f \circ P(f)(v) = P(f) \circ f(v) = P(f)(f(v))$ et par suite $f(v) \in \operatorname{Ker}(P(f))$. □

En appliquant la proposition 7.4, on obtient le

Corollaire 14.1. *Dans une base adaptée à la somme directe du théorème 14.1, la matrice de f est diagonale par blocs.*

Théorème 14.2. *Soit f un endomorphisme dont le polynôme caractéristique $P(x)$ est scindé : $P(x) = \prod_{i=1}^{i=p}(x-\alpha_i)^{d_i}$, où les α_i sont distincts. Alors la dimension du sous-espace caractéristique correspondant à α_i est d_i.*

Notons que d_i est la multiplicité de α_i en tant que racine du polynôme caractéristique P. On l'appelle la *multiplicité* de la valeur propre α_i.

Démonstration. Soit $\prod_{i=1}^{i=p}(x-\alpha_i)^{m_i}$ le polynôme minimal de f. Soit $V_i = \mathrm{Ker}((f-\alpha_i\,\mathrm{id})^{m_i})$ ls sous-espace caractéristique correspondant à la valeur propre α_i. Soit f_i la restriction de f à V_i : elle envoie V_i dans lui-même (proposition 14.2). On a $(f_i-\alpha_i\,\mathrm{id}_{V_i})^{m_i}=0$ (appliquer proposition 7.5), donc le polynôme minimal de f_i divise $(x-\alpha_i)^{m_i}$ (corollaire 12.1). Par suite la seule valeur propre de f_i est α_i. Posons $e_i = \dim(V_i)$. Alors le polynôme caractéristique de f_i est $(x-\alpha_i)^{e_i}$. Il s'ensuit, par le corollaire 14.1 et l'exercice 9.3, que le polynôme caractéristique de f est $\prod_{i=1}^{i=p}(x-\alpha_i)^{e_i}$, et donc $e_i = \dim(V_i) = d_i$. □

Théorème 14.3. *Soit f un endomorphisme dont le polynôme minimal est scindé. Alors f est somme d'un endomorphisme diagonalisable et d'un endomorphisme nilpotent qui commutent.*

Démonstration. Soit V_i le sous-espace caractéristique correspondant à la valeur propre α_i. Définissons f_i comme dans la preuve précédente, et aussi les endomorphismes g_i, h_i de V_i par : $g_i(v)=\alpha_i v$, $h_i(v)=f_i(v)-g_i(v)$. Puisque $(f_i-\alpha_i\,\mathrm{id}_{V_i})^{m_i}=0$, on a $h_i^{m_i}=0$. De plus g_i et h_i commutent, car g_i est un multiple scalaire de l'identité de V_i.

Définissons maintenant g,h par $g(v_1+\cdots+v_p)=g_1(v_1)+\cdots+g_p(v_p)$ et $h(v_1+\cdots+v_p)=h_1(v_1)+\cdots+h_p(v_p)$ pour tous $v_i\in V_i$. Comme la somme des V_i est directe, ceci est bien défini. On vérifie alors que h nilpotent, $gh=hg$ et $f=g+h$. □

Proposition 14.3. *Avec un polynôme minimal comme en (2), soient m_i des entiers naturels avec $m_i\geq n_i$. Alors* $\mathrm{Ker}(f-a_i\,\mathrm{id})^{n_i}=\mathrm{Ker}(f-a_i\,\mathrm{id})^{m_i}$.

C'est vrai en particulier si on prend pour m_i la multiplicité de la valeur propre a_i (= sa multiplicité dans le polynôme caractéristique).

Démonstration. On montre que V est somme directe des $\mathrm{Ker}(f-a_i\,\mathrm{id})^{m_i}$, en procédant comme dans la preuve du théorème 14.1. Puis on montre que $\mathrm{Ker}(f-a_i\,\mathrm{id})^{n_i}\subset\mathrm{Ker}(f-a_i\,\mathrm{id})^{m_i}$, en procédant comme dans la preuve de la proposition 14.1. Puis on utilise l'exercice 7.1. □

Exercice 14.1. *Montrer que si f et g commutent, alors* $\mathrm{Ker}(g)$ *est stable sous f.*

Exercice 14.2. *On considère une* filtration *de V, c'est-à-dire une suite croissante de sous-espaces $V_1\subset\ldots\subset V_p=V$. Soit $d_i=\dim(V_i)-\dim(V_{i-1})$, où l'on pose $V_0=0$. Une* base adaptée *à cette filtration est une base*

$v_1, \dots, v_n$ de V telle que les $d_1 + \cdots + d_i$ éléments de cette base forment une base de V_i. Montrer qu'une telle base existe. Montrer que si f est un endomorphisme stabilisant chaque V_i (c'est-à-dire $f(V_i) \subset V_i$)), alors la matrice de f dans une base adaptée est triangulaire par blocs.

Exercice 14.3. *On suppose que le polynôme minimal est $\prod_{i=1}^{i=p}(x-\alpha_i)^{m_i}$, et que le polynôme caractéristique est $\prod_{i=1}^{i=p}(x-\alpha_i)^{d_i}$, avec des α_i distincts. Montrer que $m_i \leq d_i$ et que $\operatorname{Ker}(f - a_i \operatorname{id})^{m_i} = \operatorname{Ker}(f - a_i \operatorname{id})^{d_i}$.*

Exercice 14.4. *A quelle condition les sous-espaces caractéristiques sont-il égaux aux sous-espaces propres ?*

15 Espaces cycliques et matrice compagne

Définition 15.1. *Soit $f \in \operatorname{End}(V)$. On dit que V est* cyclique sous l'action de f *s'il existe un vecteur v tel que $V = \operatorname{Vect}(v, f(v), f^2(v), \dots, f^p(v), \dots)$.*

Autrement dit, tout vecteur w dans V est combinaison linéaire des vecteurs $f^i(v)$, $i \in \mathbb{N}$. Attention : cela signifie que w est combinaison linéaire d'un *ensemble fini* de vecteurs de la forme $f^i(v)$.

Prenons un exemple : V avec base e_1, e_2 et endomorphisme $f(e_1) = e_2, f(e_2) = e_1 + e_2$. Alors $v, f(v)$ engendrent V (car $v = e_2$, $f(v) = e_1 + e_2$) ; donc tout vecteur w dans V est combinaison linéaire des vecteurs $f^i(v)$, $i \in \mathbb{N}$.

Un contre-exemple : $\dim(V) = 2$, $f = \operatorname{id}$.

Proposition 15.1. *Dans les conditions de la définition ci-dessus, si $n = \dim(V)$, alors $v, f(v), \dots, f^{n-1}(v)$ est une base de V.*

Lemme 15.1. *Soient $v_0, v_1, v_2, \dots, v_N$ des vecteurs linéairement dépendants d'un espace vectoriel. Il existe alors k tel que v_k soit linéairement dépendant de $v_0, v_1, \dots, v_{k-1}$.*

Démonstration. On choisit $k = 1+$ le maximum des i tels que $v_1, \dots, v_i$ soient linéairement indépendants. □

Preuve de la proposition 15.1. Il est certain que les vecteurs $v, f(v), \dots, f^n(v)$ sont linéairement dépendants : car ce sont $n+1$ vecteurs dans un espace de dimension n. Donc il existe par le lemme $k = 0, \dots, n$ tel que $f^k(v)$ est combinaison linéaire de $v, f(v), \dots, f^{k-1}(v)$.

Nous montrons plus loin que tout vecteur $f^i(v)$ ($i \in \mathbb{N}$) est combinaison linéaire de $v, f(v), \dots, f^{k-1}(v)$. Donc ces k vecteurs engendrent

$\mathrm{Vect}(v, f(v), f^2(v), \ldots, f^p(v), \ldots)$; cet espace est par hypothèse égal à V. Donc ces k vecteurs engendrent V. Comme $k \leq n = \dim(V)$, on doit avoir $k = n$. Donc $v, f(v), \ldots, f^{n-1}(v)$ est une base de V.

Il reste à montrer que tout vecteur $f^i(v)$ est combinaison linéaire de $v, f(v), \ldots, f^{k-1}(v)$. C'est clair pour $i = 0, \ldots, k$. Supposons que ce soit vrai pour $i \geq k$. Alors $f^i(v)$ est combinaison linéaire de $v, f(v), \ldots, f^{k-1}(v)$; en appliquant f, on trouve que $f^{i+1}(v)$ est combinaison linéaire de $f(v), \ldots, f^k(v)$; comme $f^k(v)$ est combinaison linéaire de $v, f(v), \ldots, f^{k-1}(v)$, on en déduit que $f^{i+1}(v)$ est combinaison linéaire de $v, f(v), \ldots, f^{k-1}(v)$, et on conclut par récurrence sur i. □

Théorème 15.1. *Dans la base de la proposition précédente, la matrice de f est de la forme*

$$\begin{bmatrix} 0 & \cdots & \cdots & 0 & \alpha_0 \\ 1 & \ddots & \ddots & \vdots & \alpha_1 \\ 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & \alpha_{n-2} \\ 0 & \cdots & 0 & 1 & \alpha_{n-1} \end{bmatrix} \tag{3}$$

où $x^n - \alpha_{n-1}x^{n-1} - \cdots - \alpha_1 x - \alpha_0$ est le polynôme caractéristique et minimal de f.

Par exemple, pour $n = 1$, $n = 2$ et $n = 3$, les matrices sont $[\alpha_0]$, $\begin{bmatrix} 0 & \alpha_0 \\ 1 & \alpha_1 \end{bmatrix}$ et $\begin{bmatrix} 0 & 0 & \alpha_0 \\ 1 & 0 & \alpha_1 \\ 0 & 1 & \alpha_2 \end{bmatrix}$.

Démonstration. Posons $e_i = f^{i-1}(v)$. Alors $e_1, \ldots, e_n$ est une base de V. De plus pour $i < n$, on a $f(e_i) = f(f^{i-1}(v)) = f^i(v) = e_{i+1}$, et $f(e_n)$ est combinaison linéaire de $e_1, \ldots, e_n$: $f(e_n) = \alpha_0 e_1 + \ldots \alpha_{n-1} e_n$. On en déduit la forme de la matrice dans la base.

Pour le reste, voir l'exercice 15.3. □

On appelle *matrice compagne* du polynôme $x^n - \alpha_{n-1}x^{n-1} - \cdots - \alpha_1 x - \alpha_0$ la matrice (3). Réciproquement, ce polynôme s'appelle le *polynôme compagnon* de la matrice ci-dessus.

Le théorème qui suit ne sera pas démontré ici. On peut le déduire de [2] section 9.7.

Théorème 15.2. *1) Etant donné un endomorphisme $f \in \mathrm{End}(V)$, V a une base où la matrice de f est la matrice compagne d'un polynôme P si*

et seulement si le polynôme caractéristique de f est égal à son polynôme minimal.

2) Etant donnée une matrice carrée $M \in M_{n,n}(\mathbb{K})$), M est conjuguée à la matrice compagne d'un polynôme P si et seulement si le polynôme caractéristique de M est égal à son polynôme minimal.

Notons que dans les deux cas, P est égal au polynôme caractéristique.

Exercice 15.1. *Soit M une matrice et P son polynôme caractéristique. A quelle condition M est-elle égale à la matrice compagne de P ? Vérifier que l'égalité n'est pas vraie pour $M = I_2$.*

Exercice 15.2. *Déterminer la matrice compagne du polynôme $x^2 - x - 1$. Calculer les puissances de cette matrice en fonction des* nombres de Fibonacci *F_n définis récursivement par $F_0 = F_1 = 1$ et $F_{n+2} = F_{n+1} + F_n$ pour $n \in \mathbb{N}$.*

Exercice 15.3. *Soit M la matrice (3) et f l'endomorphisme associé de $\mathbb{K}^n$. Soit $e_1, \dots, e_n$ la base canonique, et $v = e_1$. Montrer que $f^i(v) = e_{i+1}$ pour $i = 0, \dots, n-1$. En déduire que les endomorphismes* $\mathrm{id}, f, \dots, f^{n-1}$ *sont linéairement indépendants. En déduire que le polynôme minimal est de degré $\geq n$. En déduire qu'il est égal au polynôme caractéristique, lequel est $x^n - \alpha_{n-1}x^{n-1} - \cdots - \alpha_1 x - \alpha_0$.*

16 Forme de Jordan

16.1 Cas nilpotent

Soit f un endomorphisme *nilpotent*, c'est-à-dire dont une certaine puissance est nulle.

Définition 16.1. *Une* base de Jordan *pour f est une base de la forme $v_{1,1}, \dots, v_{1,n_1}, v_{2,1}, \dots, v_{2,n_2}, \dots, v_{k,1}, \dots, v_{k,n_k}$, telle que $f(v_{i,j}) = v_{i,j+1}$, avec la convention $v_{i,n_{i+1}} = 0$.*

Par exemple, $v_{1,1}, v_{1,2}, v_{1,3}, v_{2,1}, v_{3,1}, v_{3,2}$ (avec $n_1 = 3, n_2 = 1, n_3 = 2$) et $f(v_{1,1}) = v_{1,2}, f(v_{1,2}) = v_{1,3}, f(v_{1,3}) = 0, f(v_{2,1}) = 0, f(v_{3,1}) = v_{3,2}, f(v_{3,2}) = 0$.

Il est commode de représenter ceci par un *graphe orienté*, où les flèches indiquent l'action de f.

$$v_{1,1} \mapsto v_{1,2} \mapsto v_{1,3} \mapsto 0$$
$$v_{2,1} \mapsto 0$$
$$v_{3,1} \mapsto v_{3,2} \mapsto 0$$

Ce graphe est une réunion de *chaînes*. Appelons *bloc de Jordan nilpotent* une matrice carrée d'ordre n de la forme

$$\begin{bmatrix} 0 & 0 & \cdots & \cdots & 0 \\ 1 & \ddots & \ddots & & \vdots \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 & 0 \\ 0 & \cdots & 0 & 1 & 0 \end{bmatrix}$$

Par exemple, pour $n = 1$, $n = 2$ et $n = 3$, les blocs de Jordan nilpotents sont $\begin{bmatrix} 0 \end{bmatrix}$, $\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$ et $\begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$.

Lemme 16.1. *Soit f un endomorphisme nilpotent. Dans une base de Jordan pour f, la matrice de f est une somme diagonale de blocs de Jordan nilpotents. Réciproquement, si la matrice de f dans une certaine base est une somme diagonale de blocs de Jordan nilpotents, alors cette base est une base de Jordan pour f.*

Démonstration. Un exemple suffira. La matrice de l'endomorphisme de l'exemple ci-dessus est

$$\begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} & 0 & 0 & 0 \\ \mathbf{1} & \mathbf{0} & \mathbf{0} & 0 & 0 & 0 \\ \mathbf{0} & \mathbf{1} & \mathbf{0} & 0 & 0 & 0 \\ 0 & 0 & 0 & \mathbf{0} & 0 & 0 \\ 0 & 0 & 0 & 0 & \mathbf{0} & \mathbf{0} \\ 0 & 0 & 0 & 0 & \mathbf{1} & \mathbf{0} \end{bmatrix}$$

□

Nous somme sur le chemin qui va démontrer le

Théorème 16.1. *Soit $f \in \mathrm{End}(V)$ nilpotent. Il existe une base de Jordan pour f, dans laquelle sa matrice est une somme diagonale de blocs de Jordan nilpotents. Cette matrice est unique à l'ordre près des blocs.*

Remarquons que l'unicité dans le théorème est équivalente à ceci : le *multi-ensemble* des tailles des blocs de Jordan est unique. Dans notre exemple, le multi-ensemble est $\{1, 2, 3\}$ (qui est un ensemble). Un autre

exemple est $\{1, 1, 2, 3, 3\}$, ce qui signifie que la matrice a deux blocs de Jordan de taille 1, un seul de taille 2, et 2 blocs de taille 3. Ce multi-ensemble est appelé le *type* de la base de Jordan ; avec les notations de la définition 16.1, le type est $\{n_1, n_2, \ldots, n_k\}$.

Pour prouver l'unicité nous avons besoin de plusieurs lemmes.

Lemme 16.2. *Soit B la base de Jordan de la définition 16.1. Alors* $\mathrm{Ker}(f) = \mathrm{Vect}(v_{1,n_1}, v_{2,n_2}, \ldots, v_{k,n_k})$. *En particulier* $\dim(\mathrm{Ker}(f)) = k =$ *le nombre de chaînes dans la base de Jordan.*

Démonstration. Il est clair que les v_{i,n_i} sont dans le noyau, donc une inclusion est claire. Une fois n'est pas coutume, nous allons faire la preuve de l'inclusion réciproque sur l'exemple courant. La généralisation devrait être évidente. Soit $v \in \mathrm{Ker}(f)$. Alors $v = av_{1,1} + bv_{1,2} + cv_{1,3} + dv_{2,1} + ev_{3,1} + gv_{3,2}$. Alors $0 = f(v) = av_{1,2} + bv_{1,3} + ev_{3,2}$, ce qui implique que $a = b = e = 0$. Donc $v = cv_{1,3} + dv_{2,1} + gv_{3,2} \in \mathrm{Vect}(v_{1,3}, v_{2,1}, v_{3,2})$, ce qu'on voulait démontrer. □

Lemme 16.3. *Soit p un entier ≥ 1. Alors* $\dim(\mathrm{Ker}(f^p)) = a_1 + \cdots + a_p$, *où a_i est le nombre de n_i qui sont plus grand ou égal à i.*

Démonstration. 1. Commençons par supposer que $k = 1$, c'est-à-dire : il y a une seule chaîne dans la base de Jordan. Supposons que la chaîne soit de longueur n : $v_1 \mapsto v_2 \mapsto \ldots \mapsto v_n \mapsto 0$. Le noyau de f est le sous-espace engendré par v_n, le noyau de f^2 est engendré par v_n, v_{n-1} ; si $p \leq n$, le noyau de f^p est engendré par $v_n, v_{n-1}, \ldots, v_{n-p+1}$, et si $p > n$, il est engendré par tous les v_i. On voit donc que $\mathrm{Ker}(f^p)$ est de dimension $a_1 + \cdots + a_p$, car $a_p = 1$ si $p \leq n$, et $a_p = 0$ si $p > n$.

2. Supposons maintenant k quelconque. Chaque chaîne de la base de Jordan définit un sous-espace stable sous f, et donc sous f^p. L'espace V est somme directe de ces sous-espaces.

Le lemme se déduit alors du résultat suivant : soit g un endomorphisme de V et supposons que V soit somme directe des V_i, $i = 1, \ldots, k$, qui sont chacun stable sous g. Alors $Ker(g)$ est somme directe des noyaux des restrictions de g aux sous-espaces V_i. □

Lemme 16.4. *Deux bases de Jordan pour f ont toujours le même type.*

Démonstration. Il suffit de montrer que le type d'une base de Jordan de f ne dépend que de f. Mais f détermine entièrement les nombres $\mathrm{Ker}(f^p)$ pour tout $p \geq 1$. Donc f détermine entièrement les nombres $a_1 + \cdots + a_p$ du lemme précédent, et par suite les nombres a_p.

Mais on a $k =$ le nombres de chaînes dans la base de Jordan $= a_1$. De plus $a_2 =$ le nombre de n_i qui sont ≥ 2, donc $a_1 - a_2 =$ le nombre de n_i qui sont égaux à 1. De plus, a_3 est le nombre de n_i qui sont égaux à ≥ 3, donc $a_2 - a_3 =$ le nombre de n_i qui sont égaux à 2 etc...

Donc la donnée de f détermine entièrement le nombre de n_i qui sont égaux à un entier donné, donc le type de la base de Jordan. □

Lemme 16.5. *Soit $V = E + L$, où E, L sont des sous-espaces de V et $e_1, \ldots, e_n$ une base de E. Il existe des vecteurs $l_1, \ldots, l_p$ dans L tels que $e_1, \ldots, e_n, l_1, \ldots, l_p$ forment une base de V.*

Pour comprendre ce lemme, notez que sous l'hypothèse supplémentaire $E \cap L = 0$, la somme $V = E + L$ est directe, et le lemme découle du corollaire 7.3.

Démonstration. Soit $l_1, \ldots, l_p, l_{p+1}, \ldots, l_q$ une base de L telle que $l_{p+1}, \ldots, l_q$ soit une base de $E \cap L$. Alors $e_1, \ldots, e_n, l_1, \ldots, l_p$ forment une base de V. La vérification est laissée en exercice. □

Preuve du théorème 16.1. Le lemme 16.4 implique l'unicité, car si dans une certaine base, f a une matrice qui est une somme diagonale de blocs de Jordan nilpotent, alors cette base est une base de Jordan pour f (lemme 16.1).

Pour l'existence d'une base de Jordan, nous raisonnons sur l'*exposant de nilpotence*, qui le plus petit $r \geq 1$ tel que $f^r = 0$. Le cas $r = 1$ est évident : n'importe quelle base est une base de Jordan, avec tous les $n_i = 1$.

Pour le cas général, supposons $r \geq 2$. Soit $W = \mathrm{Im}(f)$. Alors W est stable sous f. Soit g la restriction de f à W : on a $g \in \mathrm{End}(W)$ et $g^{r-1} = 0$. En effet, si $w \in W$, alors $w = f(v)$ et l'on a $g^{r-1}(w) = f^{r-1}(w) = f^r(v) = 0$. Donc l'exposant de nilpotence de g est $< r$. Par hypothèse de récurrence, W possède une base de Jordan pour g :

$$\begin{array}{c} w_{1,1} \mapsto w_{1,2} \mapsto \ldots \mapsto w_{1,n_1} \\ \ldots \\ w_{k,1} \mapsto w_{k,2} \mapsto \ldots \mapsto w_{k,n_k} \end{array}$$

où les flèches indiquent l'action de g, donc de f. On va compléter cette base de W en une base de V, qui sera une base de Jordan pour f. Les vecteurs $w_{1,1}, \ldots, w_{k,1}$ sont dans $W = \mathrm{Im}(f)$. Il existe donc des vecteurs $v_1, \ldots, v_k$ dans V tels que $f(v_1) = w_{1,1}, \ldots, f(v_k) = w_{k,1}$. Nous obtenons le graphe, où

les flèches indiquent l'action de f :

$$\begin{gathered} v_1 \mapsto w_{1,1} \mapsto w_{1,2} \mapsto \ldots \mapsto w_{1,n_1} \\ \ldots \\ v_k \mapsto w_{k,1} \mapsto w_{k,2} \mapsto \ldots \mapsto w_{k,n_k} \end{gathered}$$

Cela ressemble fort à une base de Jordan pour f, mais il manque quelque chose. Nous prouvons plus bas que tous ces vecteurs v_i et $w_{i,j}$ sont linéairement indépendants. Notons E l'espace qu'ils engendrent. Nous montrons plus bas aussi que $V = E + \mathrm{Ker}(f)$. Le lemme implique qu'il existe des vecteurs $l_1, \ldots, l_p$ dans $\mathrm{Ker}(f)$ tels que ceux-ci, joints à tous les v_i et w_{ij}, forment une base de V. C'est une base de Jordan pour f, car $f(l_i) = 0$ pour tout $i = 1, \ldots, p$.

1. Montrons que les vecteurs v_i et w_{ij} sont linéairement indépendants. Une relation de dépendance linéaire s'écrit $\sum_i a_i v_i + \sum_{i,j} b_{i,j} w_{i,j} = 0$. Appliquons f : $\sum_i a_i w_{i,1} + \sum_{i,j} b_{i,j} w_{i,j+1} = 0$, avec la convention $w_{i,n_i+1} = 0$. Il s'ensuit que tous les a_i sont nuls, et revenant à l'égalité précédente, que tous les b_{ij} sont nuls, car les w_{ij} sont linéairement indépendants.

2. Montrons que $f(E) = W$. Comme W est l'image de f, on a $f(E) \subset W$. Réciproquement, on sait que W est engendré par les w_{ij}. Donc si $w \in W$, $w = \sum_{ij} b_{ij} w_{i,j} = f(\sum_{ij} b_{ij} w_{i,j-1})$, en posant $w_{i,0} = v_i$. Donc $w \in f(E)$ et donc $W \subset f(E)$.

3. Montrons que $V = E + \mathrm{Ker}(f)$. Soit donc $v \in V$. Alors $f(v) \in \mathrm{Im}(f) = W = f(E)$, donc il existe $e \in E$ tel que $f(v) = f(e)$. Par suite $v - e \in \mathrm{Ker}(f)$ et donc $v = e + (v - e) \in E + \mathrm{Ker}(f)$. □

Exercice 16.1. *On considère un graphe orienté $G = (S, A)$, où S est l'ensemble fini des* sommets *de G et $A \subset S \times S$ l'ensemble de ses arêtes. Un* chemin *de longueur k dans G est une suite d'arêtes $a_1, \ldots, a_k$ telle que $a_i = (s_i, t_i)$ et $t_i = s_{i+1}$ pour $i = 1, \ldots, k-1$. Supposons pour simplifier que $S = \{1, \ldots, n\}$. La* matrice d'adjacence *de $M = (m_{ij})$ de G est la matrice définie par $m_{ij} = 1$ si (i, j) est une arête, $= 0$ sinon. Montrer par récurrence sur k que le coefficient ij de M^k est égal au nombre de chemins dans G de longueur k.*

Exercice 16.2. *Montrer que la matrice d'adjacence du graphe associé à une base de Jordan est le bloc de Jordan correspondant, si l'on numérote les sommets de manière adéquate. Calculer ses puissances en utilisant l'exercice précédent.*

Exercice 16.3. *Soit V un espace vectoriel ayant une base de quatre vecteurs u, v, w, x, et f un endomorphisme nilpotent de V ayant la base de*

Jordan avec la chaîne associée $u \to v \to w \to x \to 0$; ce qui signifie que l'endomorphisme f satisfait : $f(u) = v, f(v) = w, f(w) = x, f(x) = 0$.

a) trouver une base du noyau de f, et sa dimension.

b) trouver un base du noyau de f^2 (la composée de f avec elle-même), et sa dimension.

c) trouver une base du noyau de f^3, et sa dimension.

d) trouver une base du noyau de f^4, et sa dimension.

e) qu'en est-il des autres puissances de f ?

Exercice 16.4. *Supposez qu'on vous dit qu'un endomorphisme f nilpotent de l'espace vectoriel V satisfait à : a) $Ker(f)$ est de dimension 8 ; b) $Ker(f^2)$ est de dimension 11 ; c) $f^2 = 0$. Trouver le graphe d'une base de Jordan de f ; c'est-à-dire, déterminer les longueurs des chaînes de cette base, et le nombre de chaînes de chaque longueur.*

Indications : quelle est la dimension de V ? Quelle est la longueur maximum d'une chaîne ?

Exercice 16.5. *On a un endomorphisme nilpotent f de V qui satisfait : $dim(Ker(f)) = 6$; $dim(Ker(f^2)) = 10$; $dim(Ker(f^3)) = 12$; $dim(Ker(f^4)) = 13$; $f^4 = 0$.*

Déterminer le type d'une base de Jordan pour f.

Exercice 16.6. *On considère une base de Jordan de f nilpotent. Calculer les dimensions des images des puissances de f. Quel est le plus petit r tel que $f^r = 0$?*

Exercice 16.7. *Soit $f \in \mathrm{End}(V)$, avec $n = \dim(V)$, et on suppose que le plus petit r tel que $f^r = 0$ est n. Quelle est la forme de Jordan de f ?*

Exercice 16.8. *Soit $f \in \mathrm{End}(V)$, avec $n = \dim(V)$, et on suppose que le plus petit r tel que $f^r = 0$ est $n - 1$. Quelle est la forme de Jordan de f ?*

Exercice 16.9. *On considère l'espace vectoriel V de base a, b, c, d et $f \in End(V)$ tel que $f(a) = b, f(c) = b, f(b) = d, f(d) = 0$. Trouver une base de Jordan pour f.*

Exercice 16.10. *Rappelons que l'exposant de nilpotence d'une matrice carrée M est le plus petit n tel que $M^n = 0$. Déterminer l'exposant de*

nilpotence de la matrice

$$\begin{bmatrix} 0 & 0 & \cdots & \cdots & 0 \\ 1 & 0 & \ddots & & \vdots \\ 1 & 1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 & 0 \\ 1 & \cdots & 1 & 1 & 0 \end{bmatrix}$$

En déduire la forme de Jordan de cette matrice.

Exercice 16.11. *Dans l'exercice précédent, prenons $n = 3$. Pour l'endomorphisme f associé, montrer qu'il existe une base u, v, w de l'espace ambiant tel que $f(u) = v + w, f(v) = w, f(w) = 0$. Montrer que $u, v + w, w$ est une base de Jordan pour f. Généraliser pour $n \geq 4$.*

Exercice 16.12. *Montrer que le type d'un endomorphisme nilpotent d'un espace de dimension 1, 2 ou 3 est entièrement déterminé par son exposant de nilpotence. Appliquer ceci pour déterminer le type des matrices* $\begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}$, $\begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$, *et* $\begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$.

Exercice 16.13. *Montrer par un exemple que le résultat de l'exercice précédent n'est plus vrai en dimension 4 (énumérer les types et déterminer l'exposant de nilpotence dans chaque cas).*

16.2 Forme de Jordan : cas général

Définition 16.2. *Un* bloc de Jordan de valeur propre λ et d'ordre n *est une matrice carrée d'ordre n de la forme*

$$\begin{bmatrix} \lambda & 0 & \cdots & \cdots & 0 \\ 1 & \lambda & \ddots & & \vdots \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \lambda & 0 \\ 0 & \cdots & 0 & 1 & \lambda \end{bmatrix}$$

Autrement dit, on a des λ sur la diagonale, des 1 sur la diagonale juste en-dessous, et des 0 partout ailleurs. Si $\lambda = 0$, c'est un bloc de Jordan nilpotent.

Théorème 16.2. *Pour tout endomorphisme f de V dont le polynôme minimal est scindé, il existe une base où la matrice de f est diagonale par blocs, chaque bloc étant un bloc de Jordan. Cette matrice est unique à l'ordre près des blocs.*

Une matrice qui est somme diagonale de blocs de Jordan est dite *sous forme de Jordan*. Trouver une telle base et calculer la matrice, c'est *mettre f sous forme de Jordan*.

Exemples : si f est diagonalisable, sa forme de Jordan est une matrice diagonale ; les blocs de Jordan sont tous de taille 1.

Lemme 16.6. *Soit M une matrice somme diagonale de blocs de Jordan J_k, $k = 1, \dots, p$, d'ordre n_k et valeur propre λ_k. Alors le polynôme caractéristique de M est $\prod_{1 \leq k \leq p}(x - \lambda_k)^{n_k}$ et son polynôme minimal est $\prod_{\lambda}(x-\lambda)^{m_\lambda}$, où ce produit est sur toutes les valeurs propres distinctes λ de f et où $m_\lambda = \max\{n_k, \lambda_k = \lambda\}$.*

Démonstration. Le polynôme caractéristique de M est le produit des polynômes caractéristiques des J_k. De plus le polynôme caractéristique d'un bloc de Jordan J, d'ordre n avec valeur propre λ, est $(x-\lambda)^n$, puisque c'est une matrice triangulaire.

Quant au polynôme minimal, c'est le plus petit multiple commun des polynômes minimaux des blocs (exercice 12.4). □

Preuve du théorème 16.2. 1. Soit $\prod_{1 \leq i \leq p}(x-\alpha_i)^{a_i}$ le polynôme minimal de f, où les α_i sont distincts et les $a_i \geq 1$. Nous savons que l'espace V est somme directe des sous-espaces caractéristiques $V_i = \mathrm{Ker}((f-\alpha_i I_n)^{a_i})$ et que chaque V_i est stable sous f. Notons f_i la restriction de f à V_i : $f_i \in \mathrm{End}(V_i)$. Posons $f_i = \alpha_i \,\mathrm{id}_{V_i} + h_i$. Alors $h_i = f_i - \alpha_i \,\mathrm{id}_{V_i}$ est un endomorphisme nilpotent de V_i : en effet, $h_i^{a_i} = (f - \alpha_i \,\mathrm{id}_V)^{a_i} \mid V_i$ (par la proposition 7.5) $= 0$. Par le théorème 16.1, il possède donc une base de Jordan v_{rs} : $h_i(v_{rs}) = v_{r,s+1}$ avec $v_{r,n_r+1} = 0$. On a $f_i(v_{rs}) = \alpha_i v_{rs} + v_{r,s+1}$. Donc dans la base des v_{rs} de V_i, la matrice de f_i est une somme diagonale de blocs de Jordan associés à la valeur propre α. L'espace V a pour base la réunion, pour $i = 1, \dots, p$, des bases de Jordan de V_i pour f_i (proposition 7.3). On en déduit par la proposition 7.4 que la matrice de f est une somme diagonale de blocs de Jordan.

2. Unicité (cas d'une seule valeur propre). On suppose que f a une seule valeur propre α. Alors $(f - \alpha \,\mathrm{id})^n = 0$ et donc $f = \alpha \,\mathrm{id} + h$, h nilpotent. On est ramené à l'unicité dans le cas nilpotent (théorème 16.1).

3. Unicité (cas général). On se ramène au cas précédent en se restreignant aux sous-espaces caractéristiques. Les détails sont laissés au lecteur. □

Corollaire 16.1. *Pour toute matrice carrée M dont le polynôme minimal est scindé, il existe une matrice carrée inversible P et une matrice carrée N sous forme de Jordan telle que $N = P^{-1}MP$. La matrice N est unique à l'ordre près des blocs de Jordan.*

Trouver P et N, c'est ce qu'on appelle *mettre M sous forme de Jordan.* Rappelons le corollaire 13.1. Une fois qu'on a calculé les racines du polynôme caractéristique et qu'on a constaté qu'il est scindé (c'est la partie algorithmiquement difficile en général), calculer P et N ne pose pas de problème algorithmique compliqué, puisqu'on est ramené à résoudre des équations linéaires : il faut calculer des bases des sous-espaces caractéristiques, donc calculer des bases de noyaux (voir [1] 8.10), puis trouver des bases de Jordan nilpotentes, en suivant la méthode de la preuve du théorème 16.1.

Exercice 16.14. *Quelle est la forme de Jordan les matrices suivantes ($\mathbb{K} = \mathbb{C}$) ?* $\begin{bmatrix} 1 & 3 \\ 3 & 2 \end{bmatrix}$, $\begin{bmatrix} 1 & 3 \\ -3 & 2 \end{bmatrix}$, $\begin{bmatrix} 2 & 3 \\ 0 & 2 \end{bmatrix}$, $\begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 2 \\ 0 & 0 & 3 \end{bmatrix}$, $\begin{bmatrix} 3 & 1 & 0 \\ 0 & 3 & 1 \\ 0 & 1 & 3 \end{bmatrix}$, $\begin{bmatrix} 3 & 1 & 1 \\ 0 & 3 & 2 \\ 0 & 0 & 3 \end{bmatrix}$ *et* $\begin{bmatrix} 3 & 0 & 1 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix}$.

Exercice 16.15. *Soit M une matrice dont le polynôme caractéristique est $(x-a)(x-b)(x-c)$. Montrer que la connaissance du polynôme minimal de M implique celle de sa forme de Jordan.*

Exercice 16.16. *On trouve dans la littérature des blocs de Jordan de la forme*

$$\begin{bmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 & 0 \\ \vdots & & \ddots & \lambda & 1 \\ 0 & \cdots & \cdots & 0 & \lambda \end{bmatrix}$$

Autrement dit : les 1 sont au-dessus de la diagonale, au lieu d'être en-dessous comme dans la définition 16.2.

1. Montrer que si $e_1 \to e_2 \to \cdots \to e_n \to 0$ est une chaîne représentant une base de Jordan nilpotente pour l'endomorphisme nilpotent f, alors la matrice de de f dans la base $e_n, \ldots, e_2, e_1$ (ordre inversé !) est de la forme précédente, avec $\lambda = 0$.

2. Si la matrice de f, dans une certaine base, est une somme diagonale de blocs de Jordan de la forme de la définition 16.2, montrer qu'on peut permuter les éléments de cette base de manière à obtenir une base où la matrice de f est une somme diagonale de matrices de la forme ci-dessus. On peut s'inspirer de l'exercice 13.2.

17 Espaces euclidiens

17.1 Inégalité de Cauchy-Schwartz

Définition 17.1. *Une* forme bilinéaire *sur un $\mathbb{K}$-espace vectoriel V est une application $B : V \times V \mapsto \mathbb{K}$ telle que pour tout v, les applications $x \mapsto B(x, v)$ et $x \mapsto B(v, x)$ de V dans $\mathbb{K}$ sont linéaires.*

Autrement dit : pour tous $x, y, v \in V$ et pour tout $a \in \mathbb{K}$, on a : $B(x + y, v) = B(x, v)+B(y, v)$, $B(ax, v) = aB(x, v)$, $B(v, x+y) = B(v, x)+B(v, y)$, $B(v, ax) = aB(v, x)$.

Une conséquence entre autres : $B(v, 0) = B(0, v) = 0$.

ATTENTION :B n'est pas en général une application linéaire de l'espace vectoriel produit $V \times V$ dans $\mathbb{K}$. Ceci signifierait que, par exemple, $B(x + y, v + w) = B(x, v) + B(y, w)$. La forme bilinéaire $\mathbb{K} \times K \to \mathbb{K}$, $(x, y) \mapsto xy$ est un contre-exemple explicite.

Exemple : $V = \mathbb{K}^n$, $B((x_1, \ldots, x_n), (y_1, \ldots, y_n)) = \sum_i x_i y_i$. Vérifiez que c'est bel et bien une forme bilinéaire.

Il est utile de savoir calculer avec des formes bilinéaires et des combinaisons linéaires : on a, pour tous les vecteurs $u_1, \ldots, u_n$, $v_1, \ldots, v_p$ et tous les scalaires $a_1, \ldots, a_n, b_1, \ldots, b_p$, la formule

$$B(\sum_i a_i u_i, \sum_j b_j v_j) = \sum_{ij} a_i b_j B(u_i, v_j).$$

Vérifiez cette formule : soit par une double récurrence sur n et p, soit en utilisant le fait qu'une application linéaire préserve les combinaisons linéaires. On peut aussi la vérifier pour des petites valeurs de n, p.

Une forme bilinéaire B est dite *symétrique* si $B(u, v) = B(v, u)$ pour tous les vecteurs u, v.

A partir de maintenant nous prenons $\mathbb{K} = \mathbb{R}$. Un *produit scalaire* sur un espace vectoriel réel V est une forme bilinéaire symétrique telle que pour tout vecteur v non nul on a $B(v, v) > 0$. On note usuellement le produit scalaire $\langle u, v\rangle$ au lieu de $B(u, v)$.

Un espace vectoriel réel de dimension finie muni d'un produit scalaire est appelé un *espace euclidien.*

On note alors $||u|| = \sqrt{\langle u, u\rangle}$, et on appelle ce nombre positif ou nul la *norme* de u. On a la *relation de polarisation*

$$||x+y||^2 = ||x||^2 + ||y||^2 + 2\langle x, y\rangle, \tag{4}$$

pour tous les vecteurs x, y. En effet, $||x+y||^2 = \langle x+y, x+y\rangle = \langle x, x\rangle + \langle x, y\rangle + \langle y, x\rangle + \langle y, y\rangle$, ce qui implique la relation, à cause de la symétrie.

On dit que des vecteurs x, y sont *orthogonaux* si leur produit scalaire est nul.

Lorsque deux vecteurs x, y sont orthogonaux, la relation de polarisation devient la *relation de Pythagore*

$$||x+y||^2 = ||x||^2 + ||y||^2.$$

Exemple 17.1. *1. Sur $\mathbb{R}^n$, on a le produit scalaire $\langle (x_1, \ldots, x_n), (y_1, \ldots, y_n)\rangle = \sum_i x_i y_i$.*

2. Sur l'espace vectoriel des polynômes de degré au plus n, on a le produit scalaire $\langle P, Q\rangle = \int_{-1}^{1} P(t)Q(t)dt$. Vérifier.

Théorème 17.1. *(Inégalité de Cauchy-Schwartz) Pour tous vecteurs x, y, on a $|\langle x, y\rangle| \leq ||x||\,||y||$, avec égalité si et seulement si x, y sont colinéaires.*

Rappelons que x, y *colinéaires* signifie qu'il engendrent un sous-espace de dimension au plus 1. Cela signifie qu'ils sont soit tous les deux nuls, soit que l'un est un multiple scalaire de l'autre.

Démonstration. Si $x = 0$, c'est clair : on a en fait égalité, et x, y sont colinéaires. Supposons donc x non nul. Pour tout t réel, on a $\langle tx+y, tx+y\rangle \geq 0$. Donc $t^2\langle x, x\rangle + 2t\langle x, y\rangle + \langle y, y\rangle \geq 0$. Ceci implique que le discriminant de ce polynôme du second degré en t est ≤ 0. Donc $\langle x, y\rangle^2 - \langle x, x\rangle\langle y, y\rangle \leq 0$. Donc $\langle x, y\rangle^2 \leq \langle x, x\rangle\langle y, y\rangle$. On obtient l'inégalité du théorème en prenant les racines carrées.

Si on a égalité, alors le discriminant est nul et il y a une racine réelle t, c'est-à-dire $\langle tx + y, tx + y\rangle = 0$. Donc $tx + y = 0$ et x, y sont colinéaires.

Si x, y sont colinéaires, alors $y = \alpha x$, et on a : $\langle x, y\rangle = \alpha\langle x, x\rangle$ d'une part, et $\langle x, x\rangle\langle y, y\rangle = \alpha^2\langle x, x\rangle^2$ d'autre part. D'où l'égalité. □

Exercice 17.1. *Montrer que si x est un vecteur non nul, alors la norme de $\frac{1}{||x||}x$ est 1.*

Exercice 17.2. *Montrer que sur* $\mathbb{R}^2$*, on a le produit scalaire* $\langle (a,b),(a',b')\rangle = aa' + bb' + \frac{1}{2}(ab' + ba')$.

Exercice 17.3. ** Montrer que la norme satisfaît les propriétés suivantes :* $||ax|| = |a|\,||x||$ *;* $||x|| = 0 \Leftrightarrow x = 0$ *;* $||x+y|| \leq ||x|| + ||y||$*. Montrer qu'on a égalité dans la dernière inégalité si et seulement* $x = 0$ *ou* $y = ax$ *pour un* $a \geq 0$.

17.2 Orthogonalité et bases orthonormales

Une *base orthonormale* de E, espace euclidien, est une base $e_1, \ldots, e_n$ de E formée de vecteurs de norme 1 et deux à deux orthogonaux.

Une base $e_1, \ldots, e_n$ est orthonormale si et seulement si quels que soient i, j dans $\{1, \ldots, n\}$, on a $\langle e_i, e_j\rangle = \delta_{ij}$, où le symbole δ_{ij} signifie 1 si $i = j$ et 0 si $i \neq j$.

Théorème 17.2. *Soit* $e_1, \ldots, e_n$ *une base orthonormale de* E*. Pour tout* v *dans* E*, on a*

$$v = \sum_i \langle v, e_i\rangle e_i$$

et

$$||v||^2 = \sum_i \langle v, e_i\rangle^2.$$

Donc si $v = \sum_i a_i e_i$ ($a_i \in \mathbb{R}$), alors $||v||^2 = \sum_i a_i^2$.

Démonstration. Ecrivons $v = \sum_j a_j e_j$. Prenons le produit scalaire avec e_i : $\langle v, v_i\rangle = \sum_j a_j \langle e_j, e_i\rangle = a_i$.

De plus $||v||^2 = \langle v, v\rangle = \langle \sum_j a_j e_j, \sum_j a_j e_j\rangle = \sum_{i,j} a_i a_j \langle e_i, e_j\rangle = \sum_{i,j} a_i a_j \delta_{ij} = \sum_i a_i^2$. □

Corollaire 17.1. *Sous les mêmes hypothèses, si* $u = \sum_i a_i e_i$*,* $v = \sum_i b_i e_i$*, alors* $\langle u, v\rangle = \sum_i a_i b_i$.

Démonstration. $\langle u, v\rangle = \sum_{i,j} a_i b_j \langle e_i, e_j\rangle = \sum_{i,j} a_i b_j \delta_{i,j} = \sum_i a_i b_i$. □

Rappelons l'*orthonormalisation* de Gram-Schmidt (voir [1] Théorème 9.1).

Théorème 17.3. *Soit* $e_1, \ldots, e_n$ *une base de* E *espace euclidien. Il existe une unique base orthonormale* $v_1, \ldots, v_n$ *de* E *telle que :* $\forall j = 1, \ldots, n$*,* $Vect(e_1, \ldots, e_j) = Vect(v_1, \ldots, v_j)$ *et que* $\langle e_j, v_j\rangle > 0$.

Ce théorème implique en particulier que tout espace euclidien possède au moins une base orthonormale.

Remarquons aussi que tout sous-espace vectoriel d'un espace euclidien est un espace euclidien, donc admet une base orthonormale.

Exercice 17.4. *Soit E un espace euclidien.*

Soient $e_1, \dots, e_n$ des vecteurs non nuls et deux à deux orthogonaux de E (on a donc $\langle e_i, e_j \rangle = \delta_{ij}$, où le symbole de Kronecker δ_{ij} *signifie 1 si $i = j$ et 0 si $i \neq j$). Montrer que ces vecteurs sont linéairement indépendants (calculer $\langle e_j, \sum_i a_i e_i \rangle$). En déduire que la dimension de l'espace est au moins n.*

En déduire que la dimension de l'espace est égale au nombre maximum de vecteurs deux à deux orthogonaux.

17.3 Projections orthogonales

Définition 17.2. *Si E est un espace euclidien et X un sous-ensemble, on note $X^\perp$ l'ensemble des vecteurs de E qui sont orthogonaux à tous les vecteurs dans X. On l'appelle l'*orthogonal *de X.*

L'orthogonal de X un sous-espace de V (exercice 17.5).

Théorème 17.4. *Soit V un sous-espace de l'espace euclidien E. Alors E est somme directe de V et $V^\perp$.*

Démonstration. Soit $v_1, \dots, v_k$ une base orthonormale de V. Alors un vecteur e dans E est dans $V^\perp$ si et seulement s'il est orthogonal à chaque v_i (exercice 17.6).

Soit e un vecteur quelconque et posons $v = \sum_i \langle e, v_i \rangle v_i \in V$. On a alors $\langle v, v_j \rangle = \langle \sum_i \langle e, v_i \rangle v_i, v_j \rangle = \sum_i \langle e, v_i \rangle \langle v_i, v_j \rangle = \sum_i \langle e, v_i \rangle \delta_{ij} = \langle e, v_j \rangle$. Alors $\langle e - v, v_j \rangle = \langle e, v_j \rangle - \langle v, v_j \rangle = \langle e, v_j \rangle - \langle e, v_j \rangle = 0$. Donc $e - v \in V^\perp$. Comme $e = e - v + v$, on obtient que $E = V^\perp + V$. Le fait que $V \cap V^\perp = 0$ découle de ce que $\langle v, v \rangle = 0$ seulement si $v = 0$. □

On appelle *projection orthogonale* de E sur son sous-espace V la projection p de E dans lui-même qui correspond à la somme directe $E = V \oplus V^\perp$. Autrement dit, si $e = v + w$, $v \in V, v \in V^\perp$, elle envoie e sur v. On a, d'après la preuve ci-dessus,

$$p(e) = \sum_i \langle e, v_i \rangle v_i, \tag{5}$$

si $v_1, \dots, v_k$ est une base orthonormale de V. La projection p est une application linéaire.

Exercice 17.5. *Montrer que $X^\perp$ est l'intersection de tous les $x^\perp$, $x \in X$ (on note $x^\perp$ pour $\{x\}^\perp$). Montrer que $x^\perp$ est un sous-espace, et aussi $X^\perp$.*

Exercice 17.6. *Si X engendre V, montrer que $X^\perp = V^\perp$.*

Exercice 17.7. *Soit V un sous-espace de l'espace euclidien E. Montrer que E a une base orthonormale qui contient une base orthonormale de V. Montrer que tout base orthonormale de V peut se compléter en une base orthonormale de E.*

Exercice 17.8. *On définit la fonction $p : E \to E$ par la formule (5). On suppose que $v_1, \ldots, v_k$ est une base orthonormale d'un sous-espace V de E. Calculer $p(v_j)$. Montrer directement (sans utiliser la notion de projection orthogonale) que p est une application linéaire, d'image V, de noyau $V^\perp$, et que $p \circ p = p$.*

17.4 Isométries

Définition 17.3. *Un endomorphisme u d'un espace euclidien est dit* orthogonal *si pour tous $x, y \in E$, on a $\langle u(x), u(y)\rangle = \langle x, y\rangle$.*

Autrement dit, u préserve le produit scalaire.

Théorème 17.5. *Les conditions suivantes sont équivalentes, pour un endomorphisme u d'un espace euclidien.*

(i) u est orthogonal ;

(ii) u est une isométrie, *c'est-à-dire $||u(x)|| = ||x||$ pour tout vecteur x.*

(iii) il existe une base orthonormale transformée par u en une base orthonormale ;

(iv) u transforme toute base orthonormale en une base orthonormale ;

(v) il existe une base orthonormale où la matrice M de u est orthogonale, *c'est-à-dire M est inversible et $M^{-1} = {}^tM$;*

(vi) la matrice de u dans toute base orthonormale est orthogonale.

Démonstration. (i) implique (iv) ; en effet l'orthonormalité d'une base se vérifie par le produit scalaire.

(iv) implique (iii) : évident.

(iii) implique (ii) : découle du théorème 17.2.

(ii) implique (i) : on utilise la relation de polarisation (4) : $2\langle u(x), u(y)\rangle = ||u(x) + u(y)||^2 - ||u(x)||^2 - ||u(y)||^2 = ||u(x + y)||^2 - ||u(x)||^2 - ||u(y)||^2 = ||x + y||^2 - ||x||^2 - ||y||^2 = 2\langle x, y\rangle$.

(i) implique (vi) : soit $e_1, \ldots, e_n$ une base othonormale et $M = (m_{ij})$. On a donc $u(e_j) = \sum_i m_{ij} e_i$. On a $\langle u(e_j), u(e_k)\rangle = \langle \sum_i m_{ij} e_i, \sum_{i'} m_{i'k} e_{i'}\rangle =$

$\sum_{i,i'} m_{ij} m_{i'k} \langle e_i, e_{i'} \rangle = \sum_i m_{ij} m_{ik}$. Cette dernière somme est le coefficient jk du produit ${}^tMM$.

Si u est orthogonal, on a $\langle u(e_j), u(e_k) \rangle = \langle e_j, e_k \rangle = \delta_{jk}$ et on déduit du calcul précédent que ${}^tMM = I_n$. Par suite ${}^tM$ est l'inverse de M.

(vi) implique (v) : évident.

(v) implique (iii) : même calcul que dans l'implication (i) implique (vi). □

Corollaire 17.2. *Un endomorphisme orthogonal est un automorphisme.*

Exercice 17.9. *Quel est le déterminant d'une matrice orthogonale ?*

Exercice 17.10. *Montrer que si σ est une bijection de $\{1, \ldots, n\}$ dans lui-même, alors la matrice M définie par $m_{ij} = 1$ si $i = \sigma(j)$ et $= 0$ sinon, est une matrice orthogonale.*

Exercice 17.11. *Montrer que la composée de deux isométries est une isométrie, et de même que l'inverse d'une isométrie.*

Exercice 17.12. *Montrer qu'une matrice orthogonale d'ordre 2 est de la forme* $\begin{bmatrix} \cos(t) & \sin(t) \\ -\sin(t) & \cos(t) \end{bmatrix}$ *ou* $\begin{bmatrix} \cos(t) & \sin(t) \\ \sin(t) & -\cos(t) \end{bmatrix}$. *Vérifier aussi directement que ces matrices sont bien orthogonales. Indications : écrire que la matrice multipliée par sa transposée est égale à l'identité ; si $a^2 + b^2 = 0$, il existe t tel que $a = \cos(t)$ et $b = \sin(t)$; si de plus $c^2 + d^2 = 1$ et $ac + bd = 0$, alors $(c, d) = \pm(-b, a)$.*

Exercice 17.13. *Montrer qu'une matrice carrée est orthogonale si et seulement si ses colonnes forment une base orthonormale de l'espace des colonnes, pour le produit scalaire usuel.*

17.5 Décomposition en valeurs singulières d'une matrice

Rappelons que ${}^tA$ désigne la transposée de la matrice A.

Théorème 17.6. *Soit A une matrice réelle de taille $m \times n$ avec $m \leq n$. Il existe une matrice orthogonale P (resp. R), de taille $m \times m$ (resp. $n \times n$) telle que*

$$A = P[D, \mathbf{0}]R$$

(décomposition par blocs) où la matrice D est diagonale, avec des éléments diagonaux ≥ 0, de taille $m \times m$, et $\mathbf{0}$ est une matrice nulle de taille $m \times (n - m)$.

Les éléments diagonaux non nuls de D sont appelés *valeurs singulières* de A. Comme le montrera la preuve, ce sont les racines carrées des valeurs propres non nulles de $A\,{}^tA$, qui sont > 0.

Lemme 17.1. *(i) Si P est une matrice réelle telle que $P\,{}^tP = 0$, alors $P = 0$.*

(ii) Si Q est une matrice réelle telle que ${}^tQQ = I_r$, alors ses r colonnes sont de norme 1 et deux à deux orthogonales pour le produit scalaire usuel.

Preuve du théorème 17.6, tirée de [3]. La matrice $M = A\,{}^tA$ est symétrique (calculer sa transposée). Il découle du corollaire 10.3 dans [1] qu'elle est diagonalisable et qu'elle a des valeurs propres réelles. Il découle même du théorème 10.4 dans [1] qu'il existe une matrice orthogonale P telle que $P^{-1}MP$ soit diagonale.

Montrons que ses valeurs propres sont ≥ 0. Soit donc a valeur propre non nulle de M, et v un vecteur colonne propre associé. On a $Mv = av$. Donc ${}^tvMv = {}^tvA\,{}^tAv = {}^tuu$, où u est le vecteur colonne ${}^tAv$. Ceci est aussi égal ${}^tvav = a\,{}^tvv$. Comme v est non nul, ${}^tvv$ est > 0. De plus, u n'est pas nul (sinon $av = Mv = A\,{}^tAv = Au$ est nul, ce qui contredit $a \neq 0$ et $v \neq 0$), donc ${}^tuu$ est > 0. De l'égalité $a\,{}^tvv = {}^tuu$, on tire donc que $a > 0$.

La matrice A est de taille $m \times n$, donc $M = A\,{}^tA$ est de taille $m \times m$, de même que P. Soit r le rang de M ; donc $r \leq m \leq n$. Comme P est orthogonale, ${}^tP = P^{-1}$. La matrice ${}^tPMP$ est donc diagonale, de rang r aussi, et cette matrice diagonale a les même valeurs propres que M. Les éléments diagonaux d'une matrice diagonale sont ses valeurs propres, avec leurs multiplicités. Le rang d'une matrice diagonale est égal au nombre d'éléments diagonaux non nuls ; ceux de ${}^tPMP$ sont positifs strictement, comme nous avons vu ci-dessus. Nous pouvons donc supposer que

$$ {}^tPMP = D^2, D = \mathrm{Diag}(d_1, \ldots, d_r, 0, \ldots, 0), $$

où il y a $m - r$ zéros, et les d_i non nuls.

Nous partitionnons la matrice P en une deux blocs P_1, P_2 de tailles $m \times r$ et $m \times (m - r)$:

$$ P = [P_1, P_2]. $$

Nous obtenons

$$ {}^tP\,A\,{}^tAP = \begin{bmatrix} {}^tP_1 \\ {}^tP_2 \end{bmatrix} A\,{}^tA\,[P_1, P_2] = \begin{bmatrix} D_1^2 & 0 \\ 0 & 0 \end{bmatrix} $$

où $D_1 = \mathrm{Diag}(d_1, \ldots, d_r)$. Donc ${}^tP_1 A\,{}^tAP_1 = D_1^2$, d'où $D_1^{-1\,t}P_1 A\,{}^tAP_1 D_1^{-1} = I_r$ et

$$ (D_1^{-1}\,{}^tP_1 A)\,{}^t(D_1^{-1\,t}P_1 A) = I_r \tag{6} $$

et

$$ {}^tP_2 A\, {}^tA P_2 = 0, $$

donc

$$ ({}^tP_2 A)\, {}^t({}^tP_2 A) = 0. $$

On en déduit que ${}^tP_2 A = 0$ (lemma 17.1). Posons

$$ Q_1 = {}^tA P_1 D_1^{-1} = {}^t(D_1^{-1\,t} P_1 A). $$

C'est une matrice de taille $n \times r$, et $r \leq n$. On a par (6)

$$ {}^tQ_1 Q_1 = I_r. $$

donc les r colonnes, de taille n, sont orthogonales et de norme 1 (lemme 17.1 (ii)). On peut donc compléter ces r colonnes en une base orthonormale de l'espace des colonnes $\mathbb{R}^{n\times 1}$ (voir exercice 17.7). Il existe donc une matrice Q_2 de taille $n \times (n-r)$ telle que les colonnes de

$$ Q = [Q_1, Q_2] $$

forment un base orthonormale de $\mathbb{R}^{n\times 1}$; donc Q est une matrice orthogonale (exercice 17.13). Notons que par suite

$$ {}^tQ_1 Q_2 = 0. $$

On a

$$ {}^tPAQ = \begin{bmatrix} {}^tP_1 \\ {}^tP_2 \end{bmatrix} A[Q_1 Q_2] = \begin{bmatrix} {}^tP_1 AQ_1 & {}^tP_1 AQ_2 \\ {}^tP_2 AQ_1 & {}^tP_2 AQ_2 \end{bmatrix}. $$

Comme ${}^tP_2 A = 0$ comme nous l'avons vu plus haut, nous avons ${}^tP_2 AQ_1 = 0$ et ${}^tP_2 AQ_2 = 0$. Nous avons aussi

$$ 0 = {}^tQ_1 Q_2 = D_1^{-1}\, {}^tP_1 AQ_2, $$

donc ${}^tP_1 AQ_2 = 0$. On a aussi ${}^tP_1 AQ_1 = {}^tP_1 A\, {}^tA P_1 D_1^{-1} = D_1$, comme il découle d'une équation vue plus haut. On a donc

$$ {}^tPAQ = \begin{bmatrix} D_1 & 0 \\ 0 & 0 \end{bmatrix}, $$

et on en déduit, en comptant le nombre de lignes et colonnes, que

$$ {}^tPAQ = [D, 0], $$

et enfin que

$$ A = P[D, 0]\, {}^tQ. $$

□

18 Espaces hermitiens

Soit V un espace vectoriel sur $\mathbb{C}$. Une forme *sequilinéaire* sur V est une application $S : V \times V \to \mathbb{C}$ telle que, quelque que soient les vecteurs u, v et le scalaire a, on ait

$$\begin{array}{rcl} S(u+v,w) & = & S(u,w)+S(v,w) \\ S(u,v+w) & = & S(u,v)+S(u,w) \\ S(au,v) & = & aS(u,v) \\ S(u,av) & = & \bar{a}S(u,v), \end{array}$$

où $\bar{a}$ est le conjugué du nombre complexe a. On voit donc que la seule différence avec la notion de forme bilinéaire est la dernière équation.

Un exemple : $V = \mathbb{C}^n, S((x_1,\ldots,x_n),(y_1,\ldots,y_n)) = \sum_i x_i\bar{y}_i$.

Une *forme hermitienne* sur V est une forme sequilinéaire S qui satisfait de plus la condition

$$S(v,u) = \overline{S(u,v)}$$

pour tous les vecteurs u, v. Une conséquence de cette condition est que

$$S(v,v) \in \mathbb{R}$$

pour tout vecteur v : en effet un nombre complexe égal à son conjugué est un nombre réel.

Dans l'exemple précédent, S est hermitien, car $S((y_1,\ldots,y_n),(x_1,\ldots,x_n)) = \sum_i y_i\bar{x}_i = \overline{\sum_i \bar{y}_i x_i}$ (car la conjugaison complexe est une involution qui préserve la somme et le produit) $= \overline{S((x_1,\ldots,x_n),(y_1,\ldots,y_n))}$.

Un *produit scalaire* sur un espace vectoriel complexe est une forme hermitienne, qu'on note $\langle u,v\rangle$ plutôt que $S(u,v)$, telle que

$$\langle v,v\rangle \geq 0$$

pour tout vecteur v, avec de plus

$$\langle v,v\rangle = 0 \Rightarrow v = 0.$$

Un *espace hermitien* est un espace vectoriel de dimension finie sur $\mathbb{C}$ muni d'un produit scalaire. Ces espaces ont de nombreux points communs avec les espaces euclidiens.

Dans la suite V est un espace hermitien. La *norme* d'un vecteur v est $||v|| = \sqrt{\langle v,v\rangle}$. Deux vecteurs u, v sont dits *orthogonaux* si $\langle u,v\rangle = 0$; notez que l'orthogonalité est une notion symétrique, car $\langle u,v\rangle = 0 \Leftrightarrow \langle v,u\rangle = 0$.

Le théorème de Pythagore est toujours vrai : si u, v sont orthogonaux, alors $||x+y||^2 = ||x||^2 + ||y||^2$.

De même, l'inégalité de Cauchy-Schwartz : Pour tous vecteurs x, y, on a $|\langle x, y\rangle| \leq ||x||\,||y||$, avec égalité si et seulement si x, y sont colinéaires. La notation $|z|$ désigne le module du nombre complexe z.

19 Formes quadratiques

19.1 Matrice d'une forme bilinéaire

Définition 19.1. *La* matrice d'une forme bilinéaire B *dans une base* $e_1, \ldots, e_n$ *de* E *est la matrice* $(B(e_i, e_j))_{1 \leq i,l \leq n}$.

Proposition 19.1. *Soient* M *la matrice d'une forme bilinéaire dans une base de* E *et soit* u, v *les vecteurs colonnes associés à deux vecteurs de* x, y *de* E. *Alors*

$$B(x, y) = {}^t u M v.$$

Démonstration. Soit $e_1, \ldots, e_n$ la base considérée. Ecrivons $u = {}^t(u_1, \ldots, u_n)$ et $v = {}^t(v_1, \ldots, v_n)$. Alors $B(x, y) = B(\sum_i u_i e_i, \sum_j v_j e_j) = \sum_{ij} u_i B(e_i, e_j) v_j$, qui est bien égal à ${}^t u M v$. □

Le *rang* d'une forme bilinéaire est par définition le rang de sa matrice. Ceci est bien défini, car on a la

Proposition 19.2. *Si* M, M' *sont les matrices de la forme bilinéaire* B *dans deux bases, alors il existe une matrice* P *inversible telle que* $M = {}^t P M' P$.

Démonstration. Soient $U = (u_1, \ldots, u_n)$ et $V = (v_1, \ldots, v_n)$ les deux bases. Soient $M_u = (B(u_i, u_j))$ et $M_v = (B(v_i, v_j))$ les matrices de la forme dans ces deux bases respectivement. On considèrer la matrice de changement de base $P = P_{UV} = (b_{ij})$. On a donc $v_j = \sum_i b_{ij} u_i$. On a alors $B(v_r, v_s) = B(\sum_i b_{ir} u_i, \sum_{js} b_{js} u_j) = \sum_{ij} b_{ir} b_{js} B(u_i, u_j) = \sum_{ij} b_{ir} B(u_i, u_j) b_{js}$: c'est le coefficient rs dans le produit matriciel ${}^t P M_u P$. Donc $M_v = {}^t P M_u P$. □

Exercice 19.1. *Montrer que la relation qui relie deux matrices carrées* M, M' *d'ordre* n, *s'il existe* P *inversible telle que* $M = {}^t P M' P$, *est une relation d'équivalence.*

19.2 Formes bilinéaires symétriques et forme quadratiques

Une forme bilinéaire B est dite *symétrique* si pour tous vecteurs x, y, on a $B(x, y) = B(y, x)$. Une *forme quadratique* est une fonction sur un espace vectoriel E de la forme $x \mapsto B(x, x)$, où B est une forme bilinéaire sur E.

Théorème 19.1. *Il y a une bijection naturelle entre formes bilinéaires symétriques et forme quadratiques sur E. A une forme bilinéaire symétrique B est associée la forme quadratique $Q(x) = B(x, x)$. A la forme quadratique Q est associée la forme bilinéaire symétrique*

$$B(x, y) = \frac{1}{2}(Q(x + y) - Q(x) - Q(y)).$$

Cette dernière équation s'appelle *relation de polarisation.*

Notez que dans ce théorème, on suppose que 2 (c'est-à-dire $1 + 1$) n'est pas égal à 0 dans le corps $\mathbb{K}$: ceci est bien sûr vrai pour les corps usuels $\mathbb{Q}, \mathbb{R}, \mathbb{C}$, mais ce n'est pas vrai dans le corps $\mathbb{F}_2$, ou l'on a $1 + 1 = 0$. On dit alors que la *caractéristique* de $\mathbb{K}$ est différente de 2.

Démonstration. Soit B une forme bilinéaire symétrique et $Q = B(x, x)$ la forme quadratique associée. Alors $\frac{1}{2}(Q(x + y) - Q(x) - Q(y)) = \frac{1}{2}(B(x + y, x + y) - B(x, x) - B(y, y)) = \frac{1}{2}(B(x, x) + B(x, y) + B(y, x) + B(y, y) - B(x, x) - B(y, y)) = B(x, y)$.

Réciproquement soit Q une forme quadratique, donc $Q = H(x, x)$ pour une forme bilinéaire H. Alors posons $B(x, y) = \frac{1}{2}(Q(x + y) - Q(x) - Q(y))$; on a alors $B(x, y) = \frac{1}{2}(H(x + y, x + y) - H(x, x) - H(y, y)) = \frac{1}{2}(H(x, y) + H(y, x))$ (par un calcul semblable au calcul précédent) ; donc B est une forme bilinéaire symétrique. De plus, on a $B(x, x) = \frac{1}{2}(2H(x, x)) = Q(x)$. □

Exercice 19.2. *Montrer qu'une forme bilinéaire est symétrique si et seulement si sa matrice dans une base (resp. dans toute base) est symétrique.*

Exercice 19.3. *Sur l'espace des matrices carrées d'ordre 2, on considère la fonction $(A, B) \mapsto \mathrm{Tr}(AB)$. Montrer que c'est une forme bilinéaire symétrique, calculer sa matrice dans la base canonique, préalablement ordonnée, et calculer la forme quadratique associée.*

Exercice 19.4. *Soit B une forme bilinéaire symétrique et $Q(x) = B(x, x)$ la forme quadratique associée. Soient $v_1, \ldots, v_n$ des vecteurs deux à deux orthogonaux pour B (c'est-à dire $B(v_i, v_j) = 0$ si $i \neq j$). Montrer que $Q(v_1 + \cdots + v_n) = Q(v_1)^2 + \cdots + Q(v_n)^2$.*

Exercice 19.5. *Montrer que Q est une forme quadratique, si et seulement si :*

1. Pour tout vecteur x et tout scalaire a, on a $Q(ax) = a^2Q(x)$;

2. La fonction $E \times E \to \mathbb{K}$, $(x, y) \mapsto \frac{1}{2}(Q(x+y) - Q(x) - Q(y))$ est une forme bilinéaire symétrique.

Exercice 19.6. *Soit E un espace vectoriel avec base $e_1, \ldots, e_n$, et soient $a_1, \ldots, a_k \in \mathbb{K}$, $k \leq n$. On considère la forme quadratique définie par $Q(x_1e_1 + \ldots + x_ne_n) = a_1x_1^2 + \cdots + a_kx_k^2$. Déterminer la forme bilinéaire symétrique associée, ainsi que sa matrice dans la base donnée.*

Exercice 19.7. *Le* noyau *d'une forme bilinéaire symétrique B est $\{x \in E, \forall y \in E, B(x, y) = 0\}$. Montrer que c'est un sous-espace de E dont la dimension est égale au rang de la matrice de B dans une base.*

Exercice 19.8. *Montrer que le noyau de la forme bilinéaire symétrique de l'exercice 19.6 est* $\mathrm{Vect}(e_{k+1}, \ldots, e_n)$.

19.3 Orthogonalisation

Théorème 19.2. *Toute forme bilinéaire symétrique sur un espace vectoriel V admet une base orthogonale.*

Ici, nous appelons orthogonaux deux vecteurs u, v tels que $B(u, v) = 0$. Une *base orthogonale* est une base telle que ses vecteurs soient deux à deux orthogonaux.

Il est clair qu'une base est orthogonale si et seulement si la matrice dans cette base de la forme est diagonale.

Démonstration. On peut supposer B non nulle. Alors la forme quadratique associée est non nulle. Choisissons un vecteur e_1 tel que $B(e_1, e_1) \neq 0$. Soit $H = \{v \in V, B(e_1, v) = 0\}$. Alors H est un sous-espace de V. Soit $v \in V$, et soit $a = B(v, e_1)$ et $u = v - aB(e_1, e_1)^{-1}e_1$. Alors $B(u, e_1) = B(v, e_1) - B(aB(e_1, e_1)^{-1}e_1, e_1) = a - aB(e_1, e_1)^{-1}B(e_1, e_1) = 0$. Donc $u \in H$. Comme $v = u + aB(e_1, , e_1)^{-1}e_1$, on a $V = H + \mathrm{Vect}(e_1)$. La somme est directe, car $e_1 \notin H$, sinon on aurait $B(e_1, e_1) = 0$.

La restriction de B à H est une forme bilinéaire symétrique de H. Par hypothèse de récurrence sur la dimension, il existe une base orthogonale $e_1, \ldots, e_n$ de H. Alors $e_1, \ldots, e_n$ est une base orthogonale de V pour B. □

19.4 Polynôme associé à une forme quadratique et une base

On suppose dans toute la suite de ce chapitre que la caractéristique de $\mathbb{K}$ *est* $\neq 2$.

Rappelons qu'une forme linéaire sur un $\mathbb{K}$-espace vectoriel est une application linéaire $\varphi : E \to \mathbb{K}$. Si $e_1, \ldots, e_n$ est une base de E, il existe des a_i (à savoir $a_i = \varphi(e_i)$) tels que pour tout $e = \sum_i x_i e_i$, on a $\varphi(e) = \sum_i a_i x_i$. Donc, une forme linéaire n'est rien d'autre qu'un polynôme homogène de degré 1 en plusieurs variables $x_1, \ldots, x_n$, une fois qu'on a choisi une base. Les a_i dépendront de cette base.

Pour les formes quadratiques, on a un résultat analogue.

Proposition 19.3. *Etant donnée une forme quadratique* Q *sur* E *et une base* $e_1, \ldots, e_n$ *de* E*, posons* $a_i = Q(e_i)$ *et* $b_{ij} = B(e_i, e_j)$*, où* B *est la forme bilinéaire symétrique associée à* Q*. Alors, pour tout* $e = \sum_i x_i e_i \in E$*,*

$$Q(e) = \sum_i a_i x_i^2 + 2 \sum_{i<j} b_{ij} x_i x_j.$$

Réciproquement, si on a cette formule, alors $a_i = Q(e_i)$ *et* $b_{ij} = B(e_i, e_j)$*.*

Nous appellerons *forme polynomiale* de Q et d'une base le polynôme de la proposition.

On parle souvent, par abus de langage, de *forme linéaire* (resp. *forme quadratique*) pour désigner un polynôme en plusieurs variables, homogène, et de degré 1 (resp. de degré 2).

Démonstration. On a $Q(e) = B(e,e) = B(\sum_i x_i e_i, \sum_i x_i e_i) = \sum_{i,j} B(e_i, e_j) x_i x_j$ (par bilinéarité) $= \sum_i B(e_i, e_i) x_i^2 + \sum_{i \neq j} B(e_i, e_j) x_i x_j$. Cette dernière somme est par symétrie de B égale à $2 \sum_{i<j} B(e_i, e_j) x_i x_j$, et la preuve est finie.

Réciproquement, si on a la formule, alors en utilisant la relation $B(e, f) = \frac{1}{2}(Q(e+f) - Q(e) - Q(f))$, on trouve que $B(e_i, e_i) = a_i$, puisque $Q(e_i) = a_i$ (car pour $e = e_i$, on a $x_j = \delta_{ij}$) ; de plus, si $i < j$, alors $B(e_i, e_j) = b_{ij}$ (car si $e = e_i + e_j$, on a $x_i = x_j = 1$ et $x_k = 0$ pour les autres k, donc $Q(e) = a_i + a_j + 2b_{ij}$). □

19.5 Equivalence des formes quadratiques

Deux formes quadratiques Q, Q' (resp. bilinéaires symétriques B, B') sur E sont dites *équivalentes* s'il existe un automorphisme u de E tel que pour tout vecteur x, on a $Q'(x) = Q(u(x))$ (resp. $B'(x, y) = B(u(x), u(y))$ pour tous vecteurs x, y).

Proposition 19.4. *Soient B, B' des formes bilinéaires symétriques et Q, Q' les formes quadratiques associées respectives. Les conditions suivantes sont équivalentes :*

(i) B, B' sont équivalentes ;

(ii) Q, Q' sont équivalentes ;

(iii) dans des bases appropriées, B et B' ont la même matrice ;

(iv) dans les bases appropriées, Q et Q' ont la même forme polynomiale.

Démonstration. L'équivalence de (i) et (ii) résulte des formules dans le théorème 19.1.

(i) implique (iii) : Si B, B' sont équivalentes, alors $B'(x, y) = B(u(x), u(y))$ pour un automorphisme u de E ; soit $e_1, \ldots, e_n$ une base de E. Donc $u(e_1), \ldots, u(e_n)$ est une base de E ; on a $B'(e_i, e_j) = B(u(e_i), u)e_j)$, donc la matrice de B' dans la première base est égale à la matrice de B dans la deuxième base.

(iii) implique (iv) : découle de la proposition 19.3.

(iv) impliqe (ii) : par hypothèse, il existe une base $e_1, \ldots, e_n$ et une base $e'_1, \ldots, e'_n$ telles que la forme polynomiale de Q' dans la première base est égale à la forme polynomiale de Q dans le deuxième. Soit u l'automorphisme de E qui envoie e_i sur e'_i. On a alors pour tout vecteur $e = \sum_i x_i e_i$: $Q'(e) = Q(\sum_i x_i e'_i)$, par l'hypothèse sur les formes polynomiales. Comme $u(e) = \sum_i x_i e'_i$, on obtient $Q'(e) = Q(u(e))$.

□

Soit $\mathcal{B}$ (resp. $\mathcal{Q}$) l'ensemble des formes bilinéaires symétriques (resp. des formes quadratiques) sur l'espace vectoriel E. Le groupe $Aut(E)$ des automorphismes de E agit à droite sur $\mathcal{B}$ (resp. $\mathcal{Q}$), par la formule $(B.u)(e, f) = B(u(e), u(f)$ (resp. $(Q.u)(e) = Q(u(e))$. On laisse au lecteur le soin de vérifier que $B.u$ (resp. $Q.u$) ainsi défini est bien une forme bilinéaire symétrique (resp. une forme quadratique), et que l'action est bien une action *à droite* (voir l'exercice 19.9). Il est clair alors que Q, Q' sont équivalentes si et seulement si elles ont dans la même orbite sous le groupe ; de même pour les formes bilinéaires symétriques.

Par ailleurs, les deux actions sont compatibles avec la bijection entre formes bilonéaires symétriques et formes quadratiques : si Q, Q' correspondent respectivement à B, B' et si $u \in Aut(E)$, alors : $Q' = Q.u \Leftrightarrow B' = B.u$, comme le lecteur peut le vérifier par les formules dans le théorème 19.1.

Il découle des propriétés des actions de groupes que l'équivalence des formes quadratiques, et des formes bilinéaires symétriques, est bien une relation d'équivalence.

Exercice 19.9. *On suppose qu'un groupe A agit à gauche sur un ensemble E. Soit K un ensemble. On considère l'ensemble $\mathcal{F}$ des fonctions de A dans K. Pour une telle fonction, soit q, et pour a dans A, on définit la fonction $q.a : E \to K$ par $(q.a)(e) = q(a.e)$. Montrer qu'on définit ainsi une action à droite de A sur $\mathcal{F}$.*

19.6 Formes bilinéaires symétriques : classification sur $\mathbb{C}$ et $\mathbb{R}$

Théorème 19.3. *Toute forme quadratique sur $\mathbb{C}$ est équivalente à une forme quadratique de la forme $x_1^2 + x_2^2 + \cdots + x_r^2$, où r est le rang de cette forme. Deux forme quadratiques sont équivalentes si et seulement si elles on t même rang.*

Théorème 19.4. *Toute forme quadratique sur $\mathbb{R}$ est équivalente à une forme quadratique de la forme $x_1^2 + x_2^2 + \cdots + x_s^2 - x_{s+1}^2 - \cdots - x_{s+t}^2$, où le couple (s,t) est unique. De plus $s+t$ est le rang de cette forme. Deux formes quadratiques sont équivalentes si et seulement si elles ont même signature.*

On appelle (s,t) la signature de la forme quadratique. Elle est dite *définie positive* si sa signature est $(n,0)$, où n est la dimension de l'espace.

Preuve du théorème 19.3. 1. Supposons que la forme bilinéaire symétrique associée est B soit de rang r. Il existe une base orthogonale $e_1, \ldots, e_n$ pour B. La matrice de B dans cette base est la matrice diagonale dont les éléments sont les $B(e_i, e_i)$. Le rang de cette matrice est le nombre de e_i non nuls. Quitte à réordonner la base, on peut supposer que $B(e_i, e_i) = 0$ pour $r > i$. Pour $i = 1, \ldots, r$, le nombre complexe $B(e_i, e_i)$ est non nul et nous choisissons une racine carrée de ce nombre, et notons z_i son inverse. Remplaçons e_i par $z_i e_i$. Alors $B(e_i, e_i) = 1$. Il s'ensuit par la proposition 19.3 que la forme quadratique est celle de l'énoncé.

2. Découle de la première partie, car dans deux bases appropriées les deux formes ont la même matrice. □

Preuve du théorème 19.4. 1. Il existe une base orthogonale $e_1, \ldots, e_n$ pour B. La matrice de B dans cette base est la matrice diagonale dont les éléments sont les $B(e_i, e_i)$. Le rang de cette matrice est le nombre de e_i non nuls. Quitte à réordonner la base, on peut supposer que $B(e_i, e_i) = 0$ pour $r > i$. De plus, on peut supposer $B(e_i, e_i) > 0$ pour $i = 1, \ldots, s$ et $B(e_i, e_i) < 0$ pour $i = s+1, \ldots, r$, et on pose $t = r - s$. Nous définissons le réel a_i comme l'inverse de la racine carrée de $B(e_i, e_i)$ si $i = 1, \ldots r$, et comme l'inverse de

la racine carrée de $-B(e_i, e_i)$ si $i = r+1, \ldots, r$. Nous remplaçons e_i par $a_i e_i$ pour tous ces i. Alors $B(e_i, e_i) = 1$ dans le premier cas, et $B(e_i, e_i) = -1$ dans le deuxième. Dans cette base, la forme est celle de l'énoncé.

2. Unicité de s, t. Nous supposons que la forme quadratique Q a, dans la base $e_1, \ldots, e_n$, la forme polynomiale donnée dans le théorème. Le rang r de Q est entièrement déteminé par Q.

Notons $F = \mathrm{Vect}(e_1, \ldots, e_s)$ et $G = \mathrm{Vect}(e_{s+1}, \ldots, e_r)$. La restriction de Q à F est $x_1^2 + \cdots + x_s^2$ et à G, c'est $-(x_{s+1}^2 + \cdots + x_r^2)$. La première est donc définie positive, et la seconde est définie négative. Soit H, un sous-espace de $F + G$, tel que la restriction de Q à H soit définie positive. Alors $H \cap G = 0$. Donc la somme $H + G$ est directe et par suite $\dim(H) + \dim(G) \leq \dim(F + G) = r$. Donc $\dim(H) \leq n - \dim(G) \leq s + t - t = s$. Il s'ensuit que s est le maximum des dimensions des sous-espaces de $F + G$ auxquels la restriction de Q est définie positive. Cela détermine s de manière unique. C'est analogue pour t.

3. Si deux formes ont même signature, il existe deux bases où elles ont la même forme polynomiale. Elles sont donc équivalentes. □

Exercice 19.10. *Soient Q_1, Q_2 des formes quadratiques sur les espaces réels E_1, E_2, de signatures $(p_1, q_1), (p_2, q_2)$. Soit $Q; E_1 \times E_2 \to \mathbb{K}$, $Q(x_1, x_2) = Q_1(x_1) + Q_2(x_2)$. Montrer que Q est une forme quadratique de signature $(p_1 + p_2, q_1 + q_2)$.*

Exercice 19.11. *Montrer que le nombre de classes d'équivalence de formes quadratiques sur un espace vectoriel de dimension finie, sur le corps $\mathbb{C}$ et sur le corps $\mathbb{R}$, est fini.*

19.7 Algorithme de Gauss*

Théorème 19.5. *On suppose que $2 \neq 0$ dans $\mathbb{K}$. Toute forme quadratique peut s'écrire comme une combinaison linéaire de carrés de formes linéaires linéairement indépendantes.*

Cette décomposition n'est pas unique ; et par conséquent, l'algorithme ci-dessous n'est pas *déterministe* ; il y a des choix en cours de route.

Ce théorème est une conséquence du théorème 19.2. Réciproquement, l'algorithme de Gauss permet de déterminer le rang et la signature d'une forme quadratique réelle.

Pour prouver le théorème, nous décrivons cet algorithme : partant d'une forme quadratique donnée sous la forme d'un polynôme homogène de degré

2 (voir la proposition 19.3), il donne sa décomposition en somme de combinaisons linéaires de carrés polynômes homogènes de degré 1 linéairement indépendants.

Une partie de l'algorithme repose sur la *complétion du carré*. Faisons quelques exemples.

Exemple 19.1. *La complétion du carré classique consiste à transformer la forme quadratique* $ax^2 + bxy + cy^2$ *comme suit (on suppose le coefficient* a *non nul) :*

$$ax^2 + bxy + cy^2 = a(x^2 + \frac{b}{a}xy) + cy^2 = a(x + \frac{b}{2a}y)^2 - a\frac{b^2}{4a^2}y^2 + cy^2$$

$$= a(x + \frac{b}{2a}y)^2 - \frac{b^2 - 4ac}{4a}y^2.$$

On a donc écrit la forme quadratique de départ comme combinaison linéaire des carrés des formes linéaires $x + \frac{b}{2a}y$ *et* y*, lesquelles sont bien linéairement indépendantes, car* $a \neq 0$*.*

Exemple 19.2. *On veut écrire la forme quadratique* xy *comme somme de deux carrés de formes linéaires. On s'appuie sur le fait que* $(x + y)^2 + (x - y)^2 = 4xy$*. Donc*

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

Revenons à l'algorithme de Gauss. Soit donc $Q(x_1, \ldots, x_n)$ comme dans la proposition 19.3. Nous commençons par supposer que Q contienne un terme x_i^2 avec coefficient non nul. On peut supposer $i = n$. On va faire la complétion du carré avec x_n. On peut écrire $Q = a_n x_n^2 + f(x_1, \ldots, x_{n-1})x_n + Q'(x_1, \ldots, x_{n-1})$, où f' est une forme linéaire et Q' une forme quadratique, ne comportant pas la variable x_n. Alors

$$Q = a_n(x_n + \frac{1}{2a_n}f)^2 - \frac{1}{4a_n}f^2 + Q'.$$

On conclut par récurrence sur le nombre de variables, appliquée à la forme quadratique $Q' - \frac{1}{4a_1^2}f^2$, laquelle est une forme quadratique en $x_1, \ldots, x_{n-1}$. On obtient que cette dernière est une combinaison linéaire des carrés de formes quadratiques $f_1, \ldots, f_k$ en $x_1, \ldots, x_{n-1}$ linéairement indépendantes. Alors les formes linéaires $f_1, \ldots, f_k, x_n + \frac{1}{2a_1}f$ sont linéairement indépendantes (car seule la dernière comporte x_n), et Q est combinaison linéaire de leurs carrés.

On suppose maintenant qu'aucun monôme x_i^2 n'apparaisse dans Q. On a alors

$$Q = ax_1x_2 + x_1f_1(x_3, \dots, x_n) + x_2f_2(x_3, \dots, x_n) + R(x_3, \dots, x_n),$$

où $a \neq 0$, et où f_1, f_2 sont des formes linéaires et R une forme quadratique. On a

$$Q = a(x_1 + \frac{1}{a}f_2)(x_2 + \frac{1}{a}f_1) + R - \frac{1}{a}f_1f_2$$

$$= \frac{a}{4}[(x_1 + \frac{1}{a}f_2 + x_2 + \frac{1}{a}f_1)^2 - (x_1 + \frac{1}{a}f_2 - x_2 - \frac{1}{a}f_1)^2)] + Q'(x_3, \dots, x_n).$$

On conclut alors par récurrence sur n, en utilisant le lemme suivant (appliqué à $g_1 = \frac{1}{a}f_1 + \frac{1}{a}f_2, g_2 = -\frac{1}{a}f_1 + \frac{1}{a}f_2$, et où Q' est par récurrence une combinaion linéaire des carrés de $g_3, \dots, g_p$, formes linéaires en $x_3, \dots, x_p$, linéairement indépendantes).

Lemme 19.1. *Soient $g_1, \dots, g_p$ des formes linéaires en $x_3, \dots, x_n$, tellles que $g_3, \dots, g_p$ sont linéairement indépendantes. Alors les formes linéaires $x_1 + x_2 + g_1$, $x_1 - x_2 + g_2$, $g_3, \dots, g_p$ sont linéairement indépendantes.*

Démonstration. □

19.8 Applications de l'algorithme de Gauss*

1. L'algorithme de Gauss permet d'orthogonaliser une forme quadratique (ce que nous savons déjà par le théorème 19.2, dont la preuve fournit aussi un algorithme). En effet, soit $Q(x_1, \dots, x_n)$ donnée par sa forme polynomiale dans une certaine base $e_1, \dots, e_n$. L'algorithme permet de l'écrire comme une combinaison linéaire $\sum_i a_i\phi_i(x_1, \dots, x_n)^2$ de carrés de forme linéaires.

Conformément à la Proposition 19.3, la variable x_i représente la forme linéaire "i-ème coordonnée dans la base précédente" ; autrement dit, c'est la fonction, notée e_i^*, qui est définie par $e_i^*(\sum_{1 \leq j \leq n} b_je_b) = b_i$ (les b_j sont les coordonnées du vecteur dans la parenthèse).

Donc chaque forme linéaire ϕ_k est combinaison linéaire des fonctions e_i^* (qui sont aussi des formes linéaires). On applique alors un théorème qu'on verra dans un cours sur la dualité (voir [2]) : il existe une base $f_1, \dots, f_n$ de l'espace telle que $f_i^* = \phi_i$ pour tout i. Il s'ensuit que dans cette nouvellle base, la forme quadratique a la forme polynomiale $\sum_i a_ix_i^2$. D'après la dernière assertion de la proposition 19.3, la matrice dans cette base de la forme biliassociée est diagonale, donc cette base est orthogonale.

2. L'algorithme de Gauss permet le calcul de la signature d'une forme quadratique réelle. La signature est (s, t), où s (resp. t) est le nombre de

coefficients positifs (resp. négatifs) dans l'expression de la forme comme combinaison linéaire de carrés de formes linéaires.

Exercice 19.12. *Sur $\mathbb{R}^{2n}$, on définit $Q(x_1, y_1, x_2, y_2, \ldots, x_n, y_n) = \sum_i x_i y_i$. Montrer que Q est une forme quadratique et déterminer sa signature.*

Exercice 19.13. *Trouver un exemple de forme quadratique qui a deux expressions distinctes comme combinaison linéaire de carrés de formes linéaires linéairement indépendantes.*

Exercice 19.14. *Appliquer la méthode de l'exemple 19.1 au formes x^2+xy+y^2 et x^2+xy-y^2, et dire laquelle est définie positive. Plus généralement, dans cet exemple, à quelle condition la forme quadratique est-elle définie positive ?*

Exercice 19.15. *Appliquer la réduction de Gauss aux formes quadratiques suivantes :*

a) $xy + yz + zx$;
b) $xy + yz + zt + tx$;
c) $x^2 + y^2 + 3z^2 + 4xy + 2xz + 2yz$;
d) $x^2 + 2y^2 + 3z^2 + 2xy + 4xz + 2yz$;
e) $(x + y)^2 - x^2 - y^2$.

Exercice 19.16. *Calculer la signature de la forme quadratique réelle de l'exercice 19.3.*

Exercice 19.17. *Calculer la signature de la forme quadratique réelle $\mathrm{Tr}({}^tMM)$ sur l'espace des matrices carrées réelles d'ordre n.*

Exercice 19.18. *Calculer la signature des forme quadratiques réelles suivantes :*

a) $x^2 + y^2 + z^2 - 4(xy + xz + yz)$;
b) $x^2 - 2yz - xz$;
c) $x^2 + 2xy + y^2$;
d) $x^2 + xy + y^2$;
e) $ax^2 + bxy + cy^2$.

Exercice 19.19. *Soit $Q(M) = -4\det(M) + \mathrm{Tr}(M)^2$ la forme quadratique réelle sur l'espace des matrices carrées d'ordre 2. Déterminer sa signature.*

Exercice 19.20. *(pièges !) Déterminer la signature des deux formes quadratiques réelles $(x-y)^2 - (y-x)^2$ et $(x-y)^2 + (y-z)^2 + (z-x)^2$.*

20 Formes multilinéaires alternées et déterminants

Définition 20.1. *Soit E un espace vectoriel sur $\mathbb{K}$. On appelle* forme n-linéaire *une application $f : E^n \mapsto \mathbb{K}$ qui est n-*linéaire*, c'est-à-dire : quel que soit $i = 1, \ldots, n$, et quels que soient v_j, $j \in \{1, \ldots, n\} \setminus i$, la fonction $E \to$, $v \mapsto f(v_1, \ldots, v_{i-1}, v, v_{i+1}, \ldots, v_n)$ est linéaire.*

Nous connaissons déjà les cas particuliers où $n = 1$ ou $n = 2$.

Définition 20.2. *Soit E un espace vectoriel sur $\mathbb{K}$. Une application n-linéaire $f : E^n \mapsto \mathbb{K}$ est dite* alternée *si quels que soient v_i, $i \in \{1, \ldots, n\}$, tels que $v_j = v_k$ pour deux indices distincts j, k, on a $f(v_1, \ldots, v_n) = 0$.*

Théorème 20.1. *Soit $E = \mathbb{K}^n$. La fonction de E^n dans $\mathbb{K}$, qui à n vecteurs $v_1, \ldots, v_n$ associe le déterminant de la matrice dont la j-ème colonne est v_j, est n-linéaire alternée.*

Démonstration. Si on fixe les colonnes d'une matrice, sauf la j-ème, et si on fait varier la j-ème colonne, on obtient, en prenant le déterminant, une application linéaire de $\mathbb{K}^n$ (vu comme des vecteurs colonnes) dans K (voir [1] Proposition 13.2).

De plus, le déterminant d'une matrice ayant deux colonnes égales, est égal à 0. Ceci prouve le théorème. □

Appelons *permutation* de $\{1, \ldots, n\}$ une bijection de cet ensemble dans lui-même. On appelle *matrice de la permutation* σ la matrice P_σ dont l'élément i, j est 1 si $\sigma(j) = i$, et 0 sinon.

Par exemple, la matrice de la permutation $\sigma(1) = 2, \sigma(2) = 3, \sigma(3) = 1, \sigma(4) = 4$ est la matrice $P_{2314} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$. La signature d'une permutation est par définition le nombre $\operatorname{sgn}(\sigma) = \det(P_\sigma)$. Il découle par récurrence et du développement de Laplace du déterminant que la signature est toujours ± 1, voir l'exercice 20.1.

Théorème 20.2. *Le déterminant d'une matrice $n \times n$ $A = (a_{ij})$ est égal à*

$$\sum_\sigma \operatorname{sgn}(\sigma) a_{\sigma(1),1} \cdots a_{\sigma(n),n},$$

où la somme est sur toutes les permutations de $\{1, \ldots, n\}$.

Démonstration. Soient $v_1, \dots, v_n$ les colonnes de la matrice. Son déterminant est une forme n-linéaire alternée des colonnes, que nous notons $f(v_1, \dots, v_n)$. Nous utilisons la base canonique $e_1, \dots, e_n$ de $\mathbb{K}^n$ (vecteurs colonnes). On a donc $\det(A) = f(v_1, \dots, v_n) = f(\sum_i a_{i1}e_i, \dots, \sum_i a_{in}e_i)$. Par la n-linéarité de f, ceci est égal à (prendre une bonne respiration) $\sum_{i_1,\dots,i_n} a_{i_1,1} \cdots a_{i_n,n} f(e_{i_1}, \dots, e_{i_n})$. Mais pour que $f(e_{i_1}, \dots, e_{i_n})$ soit non nul, il faut que $i_1, \dots, i_n$ soient distincts ; autrement dit, la fonction $1 \mapsto i_1, \dots, n \mapsto i_n$ est une permutation σ de $\{1, \dots, n\}$; et dans ce cas on $i_1 = \sigma(1), \dots, i_n = \sigma(n)$.

Par ailleurs, $f(e_{i_1}, \dots, e_{i_n})$ est alors égal au déterminant de la matrice dont les colonnes sont, dans cet ordre, les vecteurs $e_{i_1}, \dots, e_{i_n}$; cette matrice n'est autre que la matrice de permutation P_σ, et par suite son déterminant est $\det(P_\sigma) = \mathrm{sgn}(\sigma)$. D'où la formule. □

Exercice 20.1. *On rappelle que le déterminant d'une matrice est multiplié par -1 quand on échange deux colonnes. Montrer que la signature d'une permutation est ± 1, et plus précisément, $(-1)^k$, où k est le nombre d'échange de colonnes appliquée à la matrice de cette permutation pour la transformer en la matrice identité.*

Exercice 20.2. *Combien y a-t-il de termes dans la formule du théorème 20.2 ?*

Exercice 20.3. *Ecrire la formule du théorème 20.2 pour une matrice carrée d'ordre 2 et 3, et pour les courageux, aussi d'ordre 4.*

21 Exponentielle d'une matrice réelle ou complexe

Soit M une matrice carrée à coefficients complexes. On appelle *exponentielle* de M, la matrice notée e^M égale à

$$e^M = \sum_{0 \le n} \frac{1}{n!} M^n, \tag{7}$$

où M^0 est à interpréter comme la matrice identité de la taille appropriée.

Mais comment définit-on cette somme infinie ?

D'abord, soit (v_n) une suite de vecteurs dans $\mathbb{C}^d$. Ecrivons $v_n = (v_{n1}, \dots, v_{nd})$. On dira que cettte suite *converge* si pour chaque coordonnée $i = 1, \dots, d$, la suite des i-èmes composantes v_{ni} converge vers une limite l_i ; alors la *limite* de (v_n) est le vecteur $(l_1, \dots, l_d) \in \mathbb{C}^d$. Par exemple, pour

$v_n = (1/n, e^{-n}) \in \mathbb{C}^2$, la suite (v_n) converge, car les deux suites $(1/n)$ et e^{-n} convergent vers 0 et 1 repectivement ; donc $v_n = (1/n, e^{-n})$ tend vers $(0, 1)$.

Pour une somme infinie de vecteurs $\sum_n v_n$, le critère est que chaque composante $\sum_n v_{ni}$ converge. On utilisera la condition suffisante suivante : pour que $\sum_n v_n$ converge, il suffit qu'elle converge absolument, c'est-à-dire que les séries numériques $\sum_n |v_{ni}|$ convergent, où $|z|$ est le module du nombre complexe z.

La convergence de la somme infinie (7) se démontre de la manière suivante. Notons z_{nij} le coefficient i, j de la matrice $\frac{1}{n!}M^n$. D'après ce qui précède, il suffit de montrer que pour chaque couple i, j, la somme $\sum |z_{nij}|$ converge.

Notons $|A|$ le maximum des modules des coefficients de la matrice A. Si $A = (a_{ij})$ et $B = (b_{ij})$ sont des matrices carrées de taille $d \times d$, le coefficient i, k de leur produit est $\sum_j a_{ij}b_{jk}$; donc son module est $\leq \sum_j |a_{ij}b_{jk}| = \sum_j |a_{ij}||b_{jk}| \leq d|A||B|$. Il s'ensuit (récurrence sur n) que $|M^n| \leq d^n|M|^n$. Par conséquent $|z_{nij}| \leq \frac{1}{n!}d^n|M|^n$ et par suite $\sum_n |z_{nij}|$ converge (car elle est majorée par la série convergente $\sum_n \frac{1}{n!}(d|M|)^n = e^{d|M|}$).

Tous ces efforts nous donnent que e^M est bien définie. Calculons-là dans les cas particuliers où M est diagonale, et où M est nilpotente, et puis donnons une méthode générale.

Supposons que M soit une matrice diagonale $M = \mathrm{Diag}(a_1, \ldots, a_n)$. Alors $e^M = \mathrm{Diag}(e^{a_1}, \ldots, e^{a_d})$.

Supposons que M soit une matrice nilpotente, avec $M^r = 0$; alors $e^M = \sum_{0 \leq n < r} \frac{1}{n!}M^n$, qui est une somme finie.

Supposons que M soit conjuguée à une autre matrice : $M = P^{-1}NP$. Alors $e^M = P^{-1}e^N P$; ceci découle de ce que la fonction $N \mapsto P^{-1}NP$ est linéaire, préserve le produit, et qu'elle est continue (j'omets les détails). Ceci s'applique en particulier au cas où N est diagonale, et où on connaît P ; ce qui permet de calculer M.

Enfin supposons que $M = M_1 + M_2$ où M_1 et M_2 commutent (pour le produit !). On a alors $e^M = e^{M_1}e^{M_2}$ (pas vrai sans la commutation !). Pour le démontrer, on tire avantage de la convergence absolue, qui permet de manipuler les sommes infinies comme si elles étaient des sommes finies : associativité et commutativité. On a en effet $e^M = \sum_n \frac{1}{n!}(M_1 + M_2)^n = \sum_n \frac{1}{n!} \sum_{i+j=n} \binom{n}{i} M_1^i M_2^j$, et c'est ici qu'on utilise la commutation. Donc $e^M = \sum_{i,j} \frac{1}{i!}\frac{1}{j!}M_1^i M_2^j = \sum_i \frac{1}{i!}M_1^i \sum_j \frac{1}{j!}M_2^j = e^{M_1}e^{M_2}$.

Tout ceci fournit une méthode théorique, et aussi pratique, pour calculer l'exponentielle d'une matrice M : on écrit $M = M_1 + M_2$ où ces deux matrices commutent, où M_1 est diagonalisable et où M_2 est nilpotente. Donc

$M_1 = P^{-1}DP$, D matrice diagonale, M_2 nilpotente. Et on applique ce qui précède : $e^M = e^{M_1}e^{M_2}$.

Exemple 21.1. *Prenons* $M = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$. *Alors* $M = I_2 + N$, $N = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ *nilpotente, de carré nul ; de plus* I_2 *et* M *commutent. On a* $e^{I_2} = \begin{bmatrix} e & 0 \\ 0 & e \end{bmatrix} = eI_2$, $e^N = \sum_{0 \leq n} \frac{1}{n!} N^n = I_2 + N = M$ *et enfin* $e^M = eI_2 M = \begin{bmatrix} e & e \\ 0 & e \end{bmatrix}$.

Exercice 21.1. *Calculer les puissances de* A, *et puis directement l'exponentielle de la matrice* tA, *avec* $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$, *en tirant avantage des formules* $\sin(t) = t - t^3/6 + t^5/5! - \ldots, \cos(t) = 1 - t^2/2 + t^4/4! - \ldots$.

Exercice 21.2. *1. Montrer que* e^{-M} *est l'inverse de* e^M. *2. Montrer que pour* $k \in \mathbb{Z}$, $e^{kM} = (e^M)^k$.

22 Solutionnaire des exercices (esquisses)

4.8 La matrice sont les lignes sont ces trois vecteurs est $\begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 2 \\ 1 & 1 & 1 \end{pmatrix}$. Son déterminant est $1+4-2$, qui vaut 3 dans $\mathbb{R}$. Comme $3 \neq 0$, les vecteurs forment une base de $\mathbb{R}^3$.

4.9 La même matrice, mais vue dans $\mathbb{F}_3$, a déterminant nul dans $\mathbb{F}_3$. Ces vecteurs ne forment donc pas une base de $\mathbb{F}_3^3$.

4.11 Si $\mathrm{Vect}(v_1, \ldots, v_n) = \mathrm{Vect}(v_1, \ldots, v_{n-1})$, alors $v_n \in \mathrm{Vect}(v_1, \ldots, v_n)$ donc $v_n \in \mathrm{Vect}(v_1, \ldots, v_{n-1})$, donc v_n est combinaison linéaire de $v_1, \ldots, v_{n-1}$.

Réciproquement, si v_n est combinaison linéaire de $v_1, \ldots, v_{n-1}$, on peut écrire $v_n = \sum_{i=1}^{i=n-1} a_i v_i$. Alors tout vecteur v dans $\mathrm{Vect}(v_1, \ldots, v_n)$, qui s'écrit donc $v = \sum_{i=1}^{i=n} b_i v_i$, s'écrit aussi $v = \sum_{i=1}^{i=n-1} b_i v_i + b_n v_n = \sum_{i=1}^{i=n-1} b_i v_i + b_n \sum_{i=1}^{i=n-1} a_i v_i = \sum_{i=1}^{i=n-1} b_i v_i + \sum_{i=1}^{i=n-1} b_n a_i v_i = \sum_{i=1}^{i=n-1} (b_i + b_n a_i) v_i$. Donc $v \in \mathrm{Vect}(V_1, \ldots, v_{n-1})$. On a donc l'inclusion $\mathrm{Vect}(v_1, \ldots, v_n) \subset \mathrm{Vect}(v_1, \ldots, v_{n-1})$, et l'inclusion inverse est évidente.

5.4 Pas injective ($(0,0,0,1)$ est dans le noyau), surjective, le rang est 3.

5.11 Le noyau est $\mathbb{K} \times \{0\}$ et l'image aussi est $\mathbb{K} \times \{0\}$.

6.4 La trace de AB est $\sum_{i=1}^{i=n} \sum_{j=1}^{j=p} a_{ij} b_{ji}$. Celle de BA est $\sum_{i=1}^{i=p} \sum_{j=1}^{j=n} b_{ij} a_{ji}$.

6.5 Le déterminant est $a^4+b^4+c^4-2a^2b^2-2a^2c^2-2b^2c^2$, qu'on peut (mais ne doit pas) factoriser en $(a+b+c)(a-b-c)(-a+b-c)(-a-b+c)$.

6.6 Par échanges successifs de deux colonnes, on peut mettre la matrice sous la forme triangulaire par blocs $N = \begin{bmatrix} A & 0 \\ B & C \end{bmatrix}$, où a, c est l'ordre de A, C. On a donc $\det(N) = (-1)^{ac}\det(M)$, où a, c est l'ordre de A, C. De plus, $\det(N) = \det(A)\det(C)$.

7.2

7.3 Si $y \in \mathrm{Ker}(p) \cap \mathrm{Im}(p)$, alors $y = p(x)$, donc $0 = p(y) = p(p(x)) = p^2(x) = p(x) = y$. Si x est dans V, alors $x = x-p(x)+p(x) \in \mathrm{Ker}(p)\cap\mathrm{Im}(p)$, car $p(x-p(x)) = p(x) = p^2(x) = 0$.

7.4 La condition est nécessaire : en effet, si $(E_1+\cdots+E_i)\cap E_{i+1} \neq 0$, alors il existe un vecteur v_i non nul dans cette intersection. Alors $v_i = v_1 + \ldots + v_{i-1}$, où chaque v_j est dans E_j, et par suite $0 = -v_1 + \ldots + (-v_{i-1}) + v_i$ a deux écritures distinctes comme somme de vecteurs dans E_j, ce qui contredit la définition de somme directe.

La conséquence est suffisante : supposons que $0 = v_1 + \ldots + v_n$, avec chaque v_j dans E_j, et, raisonnant par l 'absurde, qu'ils ne soient pas tous nuls. Soit i le plus grand tel que $v_i \neq 0$. Alors $v_i = -v_1 - \ldots - v_{i-1} \in (E_1+\cdots+E_i)\cap E_{i+1}$; donc $v_i = 0$, contradiction.

7.6 C'est une conséquence directe de la proposition 7.3.

7.7 Enoncé de la réciproque : on suppose que la matrice de l'endomorphisme f est, dans une base de V, sous la forme d'une matrice diagonale par blocs, comme dans l'équation (1), où les M_i sont des matrices carrées d'ordre n_i. Soient E_1 le sous-espace engendré par les n_1 premiers éléments de la base, E_2 celui engendré par les n_2 suivants, etc .... Alors V est somme directe de $E_1, \ldots, E_n$ et ce sont des sous-espaces stables sous f.

8.1 L'image de φ est l'ensemble des $v+w, v \in V, w \in W$, et donc par définition, c'est $V+W$. Le noyau de φ est l'ensemble des (v,w) tels que $v+w = 0$; cette égalité signifie que $w = -v$, ce qui implique que $v \in V\cap W$; il s'ensuit que le noyau est comme indiqué. Ceci implique aussi que la première projection $V \times W \to V$, $(v,w) \mapsto v$, restreinte à $\mathrm{Ker}(\varphi)$, envoie $\mathrm{Ker}(\varphi)$ dans $V \cap W$, et c'est un isomorphisme. Le théorème du rang appliqué à φ prouve alors la formule.

9.1 L'identité $I - M^2 = (I-M)(I+M)$ montre que l'inverse de $I-M$ est $I+M$. Pour n quelconque, on utilise l'identité, $I - M^n = (I-M)(I + M + \cdots + M^{n-1})$.

9.2 C'est x^3+x^2.

9.3 Le déterminant d'une matrice triangulaire par blocs est le produit des déterminants des blocs diagonaux.

9.4 On a $BA = B(AB)B^{-1}$ et on applique la proposition 9.2.

9.5 Si V n'est pas irréductible sous l'action de f, il existe un sous-espace W de V qui est stable sous f, avec $0 < p = \dim(W) < \dim(V)$. Il existe une base $v_1, \dots, v_n$ de V telle que $v_1, \dots, v_p$ est une base de W. Dans cette base, la matrice de f est triangulaire par blocs, de la forme $\begin{pmatrix} A & B \\ 0 & D \end{pmatrix}$, où les matrices A, D sont carrées d'ordre p et $n-p$. Par suite, le polynôme caractéristique de f est le produit de ceux de A et D, qui sont de degrés p et $n-p$, et n'est donc par irréductible.

9.6 Le polynôme caractéristique est $P(x) = \det(xI_n - M)$; si on y fait $x = 0$, on obtient que le terme constant de ce polynôme est $\det(-M) = (-1)^n \det(M)$. Notons $N = xI_n - N$. Pour la trace, il faut développer le déterminant $\det(N)$ selon la première ligne, et on obtient $(x - m_{11}) \det(N^{11}) + m_{12} \det N^{12} - m_{13} \det N^{13} + \cdots$. Les $n-1$ derniers termes sont des polynômes de degrés $\leq n-2$, et ne contribuent donc par au coefficient de x^{n-1} dans $P(x)$. On a $N^{11} = xI_{n-1} - M^{11}$, et par hypothèse de récurrence, $\det(N^{11} = x^{n-1} - \mathrm{Tr}(M^{11}x^{n-2} + \cdots$, donc $P(x) = x^n - (m_{11} + \mathrm{Tr}(N^{11})x^{n-1} + \cdots$, et on peut conclure car $\mathrm{Tr}(M) = m_{11} + \mathrm{Tr}(M^{11})$.

9.7 Preuve fausse : si on substitue M à x dans $\mathrm{Tr}(xI_n - M)$, on obtient $\mathrm{Tr}(MI_n - M) = \mathrm{Tr}(M - M) = \mathrm{Tr}(0) = 0$. Prenons une matrice 2 fois 2 générique M, avec des coefficients a, b, c, d ; on a alors $\mathrm{Tr}(xI_2 - M) = 2x - a - d$; lorsqu'on y substitue M à x, on trouve $2M - (a+d)I_2$, qui est nul si et seulement si M est la matrice aI_2.

10.1 f est un automorphisme de $V \Leftrightarrow f$ a un noyau nul $\Leftrightarrow \forall v \in V \setminus 0, f(v) \neq 0v \Leftrightarrow f$ n'a pas la valeur propre 0.

10.2 Si $f(v) = \lambda v$, alors $v = f^{-1}(f(v)) = \lambda f^{-1}(v)$, donc $f^{-1}(v) = \lambda^{-1}v$. Les valeurs propres de f sont les inverse de celles de f^{-1}.

10.3 Si $f(v) = \lambda v$, alors par récurrence, $f^n(v) = \lambda^n v$. Si $P = \sum_i x^i$, alors $P(f)(v) = \sum_i a_i f^n(v) = \sum_i a_i \lambda^i v = (\sum_i a_i \lambda^i)v = P(\lambda)v$.

10.4 Si $f(v) = \lambda v$, alors $f(av) = af(v) = a\lambda v = (a\lambda)v \in D$; donc D est stable. Réciproquement, si D est stable, $f(v)$ est un mutiple scalaire de v, donc v est un vecteur propre de f.

11.2 Les valeurs propres sont 0 et -1. Les vecteurs colonnes du noyau sont les ${}^t(x, y, z)$ qui satisfont $2x - y + 2z = 0, 10x - 5y + 7z = 0, 4x - 2y + 2z = 0$; par calcul, on trouve donc que le noyau est engendré par le vecteur ${}^t(1, 2, 0)$. Les vecteurs colonnes qui sont dans le noyau de $f + \mathrm{id}_V$ satisfont le système $2x - y + 2z = -x, 10x - 5y + 7z = -y, 4x - 2y + 2z = -z$; le calcul montre que ce sous-espace est la droite engendrée par ${}^t(-1, 1, 2)$. La matrice n'est donc pas diagonalisable, car le sous-espace engendré par les vecteurs propres

n'est pas l'espace tout entier.

11.3 D'après l'exercice mentionné, l'espace est somme directe du noyau de $\mathrm{Ker}(p - \mathrm{id})$ et de $\mathrm{Ker}(p)$. Or ceux-ci sont des sous-espaces propres. Donc on applique la proposition 11.2.

11.4 Le calcul montre que le polynôme caractéristique est x^3. Donc l'unique valeur propre est 0. Si l'endomorphisme était diagonalisable, ce serait l'endomorphisme nul.

11.5 Le calcul montre que le polynôme caractéristique est $(x-1)(x-2)(x-3)$. Donc les valeurs propres sont dictinctes et on applique le théorème 11.1.

11.6 Le noyau de $f - \lambda_i \,\mathrm{id}$ est $\mathbb{K}v_i$. On a $f(g(v_i)) = g(f)v_i) = g(\lambda_i v_i) = \lambda_i g(v_i)$. Donc $g(v_i) \in \mathrm{Ker}(f - \lambda_i \,\mathrm{id})$. Donc $g(v_i) = \mu_i v_i$ pour un certain scalaire μ_i.

12.1 C'est $(x-1)(x-2)$.

12.2 Soit E l'ensemble des éléments diagonaux de la matrice. Le polynôme minimal est alors $\prod_{\lambda \in E}(x - \lambda)$.

12.3 C'est $(x-2)^2(x-3)$.

12.4 Si $P(x)$ est un polynôme, alors $P(M)$ est la matrice diagonale par blocs, dont les blocs diagonaux sont les matrices $P(M_1), \ldots, P(M_k)$. Donc $P(M)$ est nul si et seulment si chaque $P(M_i)$ est nul, ce qui signifie que P est multiple du polynôme minimal P_i de M_i. Donc P est le plus petit multiple commun de $P_1, \ldots, P_k$.

13.2 Soit $(a_{ij}$ la matrice de f dans la base $v_1, \ldots, v_n$. On a alors $f(v_j) = \sum_i a_{ij} v_i$. Définissons $u_1 = v_n, u_2 = v_{n-1}, \ldots, u_n = v_1$ et $(b_{ij}$ la matrice de f dans la base $u_1, \ldots, u_n$. Alors $\sum_i b_{ij} u_i = f(u_j) = f(v_{n+1-j} = \sum_i a_{i,n+1-j} v_i = \sum_i a_{i,n+1-j} u_{n+1-i}$. En posant $i' = n+1-i$, ceci est égal à $= \sum_{i'} a_{n+1-i',n+1-j} u_{i'} = \sum_i a_{n+1-i,n+1-j} u_i$. Donc $b_{ij} = a_{n+1-i,n+1-i}$: les deux matrices s'obteinnetn l'une de l'autre par symétrie par rapport à leur centre. On conclut l'exercice, car si l'une est triangulaire supérieure l'autre est triangulaire inférieure.

15.3 On a $g := f^n - \alpha_{n-1} f^{n-1} - \cdots - \alpha_1 f - \alpha_0 \,\mathrm{id} = 0$. En effet $g(v) = 0$. Par conséquent, puis que f et g commutent, f^i et g commutent aussi, donc $g(f^i(v)) = f^i(g(v)) = 0$. Donc $g = 0$ puisque $g(e_i) = 0$ pour tout $i = 1, \ldots, n$. Donc le polynôme minimal de f divise $x^n - \alpha_{n-1} x^{n-1} - \cdots - \alpha_1 x - \alpha_0$. Les puissances $f^k, \ldots, f, \mathrm{id}$ sont linéairement indépendantes si $k < n$, car les vecteurs $f^k(v) = e_{k-1}, \ldots, f(v) = e_2, v = e_1$ sont linéairement indépendants. Donc le polynôme minimal de f est de degré au moins n. Il s'ensuit que c'est bien $x^n - \alpha_{n-1} x^{n-1} - \cdots - \alpha_1 x - \alpha_0$, et c'est aussi le polynôme caractéristique.

16.3 a) $\{x\}$, dimension 1 ; b)$\{w, x\}$, dimension 2 ; c)$\{v, w, x\}$, dimension 3 ; d)$\{u, v, w, x\}$, dimension 4 ; e) même chose que d).

16.4 La dimension de V est celle du noyau de l'endomorphisme nul, qui est f^2, donc sa dimension est 11. Le maximum des longueurs de chaînes est 2, car l'index de nilpotence de f est 2. Comme $11 = 5 \cdot 1 + 3 \cdot 2$, le type de la base de Jordan est $\{1, 1, 1, 1, 1, 2, 2, 2\}$.

16.5 $\{1, 1, 2, 2, 3, 4\}$. Pour comprendre ceci, on peut disposer des étoiles par colonnes, représentant les dimensions des noyaux : la première colonne pour le noyau de f, les deux premières pour le noyau de f^2, etc...

```
*  *  *  *
*  *  *
*  *
*  *
*
*
```

Les longueurs de lignes sont alors les longueurs des chaines.

16.6 On peut commencer par l'exemple courant. L'image de f est $\mathrm{Vect}(v_{12}, v_{13}, v_{32})$, c'est-à-dire les buts des flèches du graphe. L'image de f^2 est $\mathrm{Vect}(v_{13})$, car v_{13} est l'unique but d'un chemin de longueur 2 dans le graphe. Quant à l'image de f^3, elle est nulle. On a $f^3 = 0$.

Avant de faire le cas général, on peut considérer une base de Jordan avec une seule chaîne, de longueur n, donc avec n vecteurs. L'image de f est de dimension $n - 1$, celle de f^2 est de dimension $n - 2$, etc... On a $f^n = 0$.

Cas général : avec les notations de la définition 16.1, l'image de f est $Vect(v_{ij}, i = 1, \ldots, k, j = 2, \ldots, n_i)$. Ces vecteurs sont linéairement indépendants, donc la dimension de l'image est $\sum_i (n_i - 1)$, autrement dit $n - k$, si on note n la dimension de l'espace ambiant. Du reste, on peut déduire ceci du théorème du rang et du lemme 16.2.

Pour ce qui est de l'image de f^2, elle est engendrée par les vecteurs v_{ij} avec $j \geq 2$. Sa dimension est donc la somme des $n_{ij} - 2$, où la somme est sur tous les i tels que $n_i \geq 2$.

Plus généralement, la dimension de l'image de f^r est la somme de tous les $n_i - r$, où on fait la somme sur les i tels que $n_i \geq r$.

En particulier, l'image de f^r est nulle dès que r est $\geq$ au maximum des n_i. On a donc $f^r = 0$ si r est égal à ce maximum, et c'est l'exposant de nilpotence de f.

16.7 Il y a une seule chaine de longueur n ; autrement dit le type est $\{n\}$.

16.8 Il y a une chaine de longueur $n - 1$ et une chaine de longueur 1 ; autrement dit le type est $\{n - 1, 1\}$.

16.9 On $f(c - a) = 0$. Comme $a, b, c - a, d$ est une base de l'espace ambiant, et que de plus $f(a) = b, f(b) = d, f(d) = 0$, on cette dernière base

est une base de Jordan, avec graphe : $a \to b \to d \to 0, c - a \to 0$. Son type est $\{3, 1\}$.

16.10 L'exposant de nilpotence est n. On peut le montrer de plusieurs manières. Par exemple, on peut trouver une formule pour M^r, $r = 1, 2, \ldots, n$, où interviennent les coefficients binomiaux.

On peut aussi montrer, de manière plus paresseuse, montrer que M^r est une matrice triangulaire inférieure $A = (a_{ij})$, avec $a_{ij} = 0$ pour tous les $j \leq i - r + 1$, alors que les coefficients $a_{i,i-r}$ sont égaux à 1.

Pour la base de Jordan, on peut peut utiliser l'exercice 16.6. Comme l'exposant de nilpotence est égal à la dimension de l'espace ambiant, la base de Jordan ne peut avoir qu'une seule chaîne ; elle donc de type $\{n\}$, et la forme de Jordan de la matrice est un bloc de Jordan nilpotent d'ordre n.

16.11 On a en effet $f(u) = v + w, f(v + w) = w, f(w) = 0$. Le graphe est $u \to v + w \to w \to 0$.

16.12 Les types possibles sont $\{3\}$, $\{1, 2\}$ et $\{1, 1, 1\}$. Les exposants de nilpotence sont respectivement 3, 2 et 1. Pour les 4 matrices indiquées, les exposants de nilpotence sont repectivement 3, 2, 2, 3.

16.13 Les types possibles sont $\{4\}$, $\{1, 3\}$, $\{2, 2\}$, $\{1, 1, 2\}$ et $\{1, 1, 1, 1\}$. Les exposants de nilpotence sont respectivement 4, 3, 2, 2, 1. On ne peut donc déduire le type de l'exposant de nilpotence.

16.14 Les deux premières matrices ont deux valeurs propres distinctes λ, μ disons, car le discriminant de leur polynôme caractéristique est non nul dans les deux cas. Donc ces matrices sont diagonalisables et on a deux blocs de Jordan d'ordre 1.

La troisième matrice M_3 a la valeur propre 2 à la multiplicité 2. Comme $M_3 \neq 2I_2$, la forme de Jordan est un seul bloc de degré 2.

La quatrième matrice est diagonalisable car elle a les valeurs propres distinctes 1, 2, 3.

La cinquième matrice a trois valeurs propres distinctes, elle est donc diagonalisable, donc les blocs de Jordan sont de taille 1.

La sixième matrice M_6 a la valeur propre 3 avec multiplicité 3 ; le carré de $M_6 - 3I_3$ n'est pas nul, donc il y a un seul bloc de Jordan d'odre 3 et de valeur propre 3.

Enfin, la dernière matrice M_7 a la valeur propre 3 avec multiplicité 3 ; $M_7 \neq I_3$, donc M_7 n'est pas diagonalisable ; le carré de $M_7 - 3I_3$ est nul, donc il y un bloc de Jordan d'ordre 2 et un d'ordre 1, tous les deux associés à la valeur propre 3.

16.15 La forme de Jordan dépend des égalités entre les trois racines a, b, c.

Si elles sont distinctes, la matrices est diagonalisable, et il y a donc 3 blocs de Jordan d'ordre 1.

Si $a = b \neq c$, alors on peut avoir deux polynômes minimaux : le polynôme caractéristique, ou $(x-a)(x-c)$. Dans le premier cas il y a un bloc de Jordan d'ordre 2 (valeur propre a) et un bloc de Jordan d'ordre 1 (valeur propre c). Dans le deuxième cas, la matrice est diagonalisable.

Si $a = b = c$, alors il peut y avoir deux polynômes minimaux : le polynôme caractéristique, $(x-a)^2$ ou $x-a$. Dans le premier cas, il y a un bloc de Jordan d'ordre 3. Dans le deuxième cas, il y a un bloc d'ordre 2 et un d'ordre 1. Dans le troisième cas, la matrice est diagonalisable.

17.1 On a $\langle x/||x||, x/||x|| \rangle = (1/||x||^2)\langle x, x\rangle = (1/||x||^2)||x||^2 = 1$. Donc $||x|| = 1$.

17.2 On vérifie sans peine que c'est une forme bilinéaire. Soit $(a, b) \in \mathbb{R}^2$. Alors $\langle (a, b), (a, b)\rangle = a^2 + b^2 + ab = (a + 1/2b)^2 + 3/4b^2 \geq 0$, et ça vaut 0 seulement si $(a, b) = 0$.

17.3 On a $\langle ax, ax\rangle = a^2\langle x, x\rangle = a^2||x||^2$. Donc $||ax|| = \sqrt{\langle ax, ax\rangle} = |a|||x||$.

On a $||x|| = 0 \Leftrightarrow \langle x, x\rangle = 0 \Leftrightarrow x = 0$.

L'inégalité est équivalente à l'inégalité obtenue en élevant les deux côtés au carré. On a $||x + y||^2 = \langle x + y, x + y\rangle = ||x||^2 + ||y||^2 + 2\langle x, y\rangle$ (relation de polarisation) et $(||x|| + ||y||)^2 = ||x||^2 + ||y||^2 + 2||x||\,||y||$. Il s'agit donc de montrer que $\langle x, y\rangle \geq ||x||\,||y||$, ce qui découle de l'inégalité de Cauchy-Schwartz.

Enfin, par le calcul précédent, on a l'égalité $||x + y|| = ||x|| + ||y||$ si et seulement si $\langle x, y\rangle = ||x||\,||y||$. Cette dernière égalité est équivalente à $\langle x, y\rangle \geq 0$ et $|\langle x, y\rangle| = ||x||\,||y||$. Par le théorème 17.1, ceci est équivalent à $\langle x, y\rangle \geq 0$ et $x = 0$ ou $y = ax$ $(a \in \mathbb{R})$. Mais $y = ax$ avec $x \neq 0$ et $a < 0$ implique $\langle x, y\rangle < 0$. d'où la conclusion.

17.5 Un vecteur v est dans $X^\perp$ si et seulement s'il est orthogonal à tout vecteur dans X, et ceci est équivalent à $v \in x^\perp$ pour tout x dans X. D'où l'assertion sur l'intersection.

Comme l'intersection de sous-espaces est un sous-espace, il suffit de montrer que $x^\perp$ est un sous-espace pour tout vecteur x. Or la fonction $v \mapsto \langle v, x\rangle$ est une application linéaire de l'espace euclidien vers $\mathbb{R}$. Son noyau est un sous-espace, et c'est justement $x^\perp$.

17.6 Soit $u \in V^\perp$. Donc par définition, u est orthogonal à tout v dans V. Par suite il est orthogonal à tout x dans X (car $X \subset V$) et donc $V^\perp \subset X^\perp$.

Réciproquement, soit $u \in X^\perp$, donc orthogonal à tout x dans X. Soit $v \in V$. Alors $v = \sum_x a_x x$, donc $\langle u, v\rangle = \langle u, \sum_x a_x x\rangle = \sum_x a_x\langle u, x\rangle = 0$. Donc $u \in V^\perp$.

D'où : $V^\perp = X^\perp$.

17.9 On a ${}^tMM = I_n$, et $\det(M) = \det({}^tM)$. Donc $\det(M)^2 = 1$, et $\det(M) = \pm 1$.

17.10 On note M_σ la matrice de l'énoncé. On montre que pour deux permutations σ et τ, on a toujours $M_\sigma M_\tau = M_{\sigma\circ\tau}$. Ensuite que $M_\sigma^{-1} = M_{\sigma^{-1}}$ et on en déduit que $M_\sigma^{-1} = {}^tM_\sigma$.

Autre solution : soit l'endomorphisme f dont la matrice dans la base orthonormale $e_1, \ldots, e_n$ est M_σ. Alors $f(e_j) = e_{\sigma(j)}$. Donc $\langle f(e_j, f(e_k)\rangle = \langle e_{\sigma(j)}, e_{\sigma(k)}\rangle = \delta_{\sigma(j),\sigma(k)} = \delta_{jk}$. Donc la base $e_{\sigma(1)}, \ldots, e_{\sigma(n)}$ est orthonormale et on applique le théorème.

17.11 Une isométrie préserve le produit scalaire. On en déduit facilement l'exercice.

17.12 Soit $M = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ une matrice orthogonale d'ordre 2. On a donc $\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} a & c \\ b & d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. Ça donne $a^2+b^2 = 0$, $ac+bd = 0$, $c^2+d^2 = 1$. La deuxième équation implique que $(c,d) = e(-b,a)$. La première que $a = \cos(t), b = \sin(t)$. La troisième donne alors $1 = e^2$, donc $e = \pm 1$.

19.19 On a $Q(M) = -4(ad-bc)+(a+d)^2 = (a-d)^2+4ad-4ad+4bc = (a-d)^2+4bc = (a-d)^2+(b+c)^2-(b-c)^2$. C'est une combinaison linéaire de 3 formes linéaires ; comme il y a deux coefficients positifs et un négatif, la signature est $(2,1)$.

20.2 $n!$

20.3 $a_{11}a_{22} - a_{21}a_{12}$; $a_{11}a_{22}a_{33} - a_{21}a_{12}a_{33} - a_{11}a_{32}a_{23} + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{23} - a_{31}a_{22}a_{13}$. Pas courageux.

21.1 Soit A cette matrice. Alors la suite des puissances de A une période égale à 4, avec $A^0 = I_2$, $A^1 = 1$, $A^2 = -I_2$, $A^3 = -A$. On en déduit que le coefficient $1,1$ de $\exp(ta)$ est $1 - t^2/2! + t^4/4! - \ldots = \cos(t)$. Pour les autres coefficients, un calcul analogue conduit à la conclusion $\exp(tA) = \begin{bmatrix} \cos(t) & \sin(t) \\ -\sin(t) & \cos(t) \end{bmatrix}$.

Références

[1] C. Reutenauer, Notes de cours d'algèbre linéaire 1, UQAM. 3, 14, 43, 46, 50, 63

[2] C. Reutenauer, Notes de cours d'algèbre linéaire 3, UQAM. 34, 61

[3] S. R. Searle, A. Khuri, Matrix algebra useful for statistics, J. Wiley and Sons, 2017. 3, 50

Index