Algèbre linéaire 3

Guillaume Legendre, remanié par Benjamin Melinand

(version préliminaire du 25 octobre 2022)

Au regretté logo de l'université (2009-2019)

Table des matières

1 Réduction des endomorphismes **3**
1.1 Conjugaison, similitude . . . 3
1.2 Sous-espaces stables par un endomorphisme . . . 4
1.3 Éléments propres d'un endomorphisme . . . 5
1.4 Polynômes d'endomorphismes, polynômes annulateurs . . . 6
1.5 Polynôme caractéristique . . . 9
1.6 Diagonalisation . . . 12
1.7 Trigonalisation . . . 15
1.8 Sous espaces caractéristiques et multiplicités des valeurs propres (hors programme) . . . 17
1.9 Résumé . . . 20
1.10 Application de la réduction au calcul des puissances d'une matrice . . . 21
1.10.1 Grâce au polynôme minimal . . . 21
1.10.2 Dans le cas où A est diagonalisable . . . 22
1.10.3 Dans le cas où A est trigonalisable . . . 22
1.11 Application de la réduction aux suites récurrentes . . . 23
1.11.1 Suites récurrentes d'ordre 1 à valeur vectorielle . . . 23
1.11.2 Suites récurrentes linéaires d'ordre p . . . 23

2 Formes bilinéaires réelles **27**
2.1 Généralités sur les formes bilinéaires réelles . . . 27
2.2 Représentation matricielle d'une forme bilinéaire . . . 28
Changement de base . . . 29
Restriction à un sous-espace vectoriel . . . 29
2.3 Non dégénérescence d'une forme bilinéaire réelle symétrique . . . 29

3 Formes quadratiques réelles **31**
3.1 Généralités . . . 31
3.2 Orthogonalité . . . 33
3.3 Classification des formes quadratiques réelles . . . 36

4 Espaces euclidiens **41**
4.1 Définitions . . . 41
4.2 Orthogonalité dans les espaces euclidiens . . . 43
4.2.1 Généralités . . . 43
4.2.2 Bases orthonormées . . . 43
4.2.3 Projection orthogonale . . . 46
4.3 Structure du dual d'un espace euclidien . . . 48
4.4 Endomorphismes des espaces euclidiens . . . 48
4.4.1 Adjoint d'un endomorphisme . . . 48
4.4.2 Isométries vectorielles . . . 49
4.4.3 Endomorphismes auto-adjoints . . . 52

Dans tout ce document, on désigne par $\mathbb{K}$ un corps qui peut être $\mathbb{R}$, le corps de nombres réels, ou $\mathbb{C}$, le corps des nombres complexes.

Chapitre 1

Réduction des endomorphismes

Réduire un endomorphisme dans un espace vectoriel de dimension finie, c'est notamment trouver une base de l'espace dans laquelle cet endomorphisme est « aisément » étudiable et manipulable. En pratique, ceci revient à déterminer une base par rapport laquelle la matrice représentative de l'endomorphisme possède une forme particulière : diagonale (dans le meilleur des cas), diagonale par blocs ou triangulaire. La réduction des endomorphismes (et des matrices qui leur sont associées) est pour cette raison un outil incontournable de l'étude de ces derniers et ce chapitre lui est consacré.

1.1 Conjugaison, similitude

Comme nous avons pu le dire en introduction, la réduction d'un endomorphisme revient à trouver une base sur laquelle un endomorphisme agit le plus simplement possible. Cela nous pousse à définir la notion suivante.

Définition 1.1 (Conjugaisons d'endomorphismes) *Soit E un $\mathbb{K}$-espace vectoriel et $u, v \in \mathscr{L}(E)$. On dit que u et v sont conjugués s'il existe un endomorphisme $\phi \in GL(E)$ tel que*

$$u = \phi \circ v \circ \phi^{-1},$$

où $GL(E)$ est le groupe général linéaire de *E, c'est-à-dire l'ensemble des automorphismes de E.*

Deux endomorphismes conjugués auront des propriétés similaires. Illustrons cela avec deux exemples.

Exemple 1.2 *Les endomorphismes u et v ont des propriétés géométriques en commun. Soit $x \in E$. On a*

$$u(x) = x \iff \phi \circ v \circ \phi^{-1}(x) = x \iff v(\phi^{-1}(x)) = \phi^{-1}(x).$$

Ainsi, si on note $F = \{x \in E,\ u(x) = x\}$ l'ensemble des points fixes de u, l'ensemble des points fixes de v est $\phi^{-1}(F)$.

Exemple 1.3 *Les endomorphismes u et v ont la même dynamique. On a $u^2 = \phi \circ v \circ \phi^{-1} \circ \phi \circ v \circ \phi^{-1} = \phi \circ v^2 \circ \phi^{-1}$. Ainsi pour tout $k \in \mathbb{N}$, $u^k = \phi \circ v^k \circ \phi^{-1}$.*

La notion de conjugaison est utilisée pour les endomorphismes. Sa traduction matricielle porte un autre nom : la similitude.

Définition 1.4 (Similitude) *Soit $n \in \mathbb{N}$ et $A, B \in M_n(\mathbb{K})$. On dit que A et B sont semblables s'il existe une matrice $P \in GL_n(\mathbb{K})$ tel que*

$$A = PBP^{-1}$$

où $GL_n(\mathbb{K})$ est le groupe général linéaire de degré n et de corps $\mathbb{K}$, *c'est-à-dire l'ensemble des matrices inversibles de $M_n(\mathbb{K})$.*

L'intérêt principal d'utiliser la notion de similitude plutôt que la notion de conjugaison est qu'elle nous fournit directement une base adaptée. Soit E un $\mathbb{K}$-espace vectoriel de dimension finie n munit d'une base $\mathscr{B} = (e_1, \cdots, e_n)$. Soit $u \in \mathscr{L}(E)$ et notons $A = \mathrm{Mat}_{\mathscr{B}}(u)$. Soit $\mathscr{B}' = (e'_1, \cdots, e'_n)$ une autre base de E où u est supposé être plus simple à étudier. Notons $B = \mathrm{Mat}_{\mathscr{B}'}(u)$. Les matrices A et B sont semblables et il existe $P \in GL_n(\mathbb{K})$ tel que $A = PBP^{-1}$. La matrice P est alors la matrice de passage de $\mathscr{B}$ à $\mathscr{B}'$.

1.2 Sous-espaces stables par un endomorphisme

La réduction d'un endomorphisme va reposer sur une propriété de *stabilité* de certains sous-espaces vectoriels sous l'action de cet endomorphisme, notion que nous allons maintenant introduire.

Définition 1.5 (sous-espace vectoriel stable par un endomorphisme) *Soit E un $\mathbb{K}$-espace vectoriel, u un endomorphisme de E et A un sous-espace vectoriel de E. On dit que A est* ***stable*** *(ou* ***invariant****) par u si et seulement si l'image de A par u est incluse dans A, c'est-à-dire*

$$\forall x \in E,\ x \in A \implies u(x) \in A.$$

Définition 1.6 *Soit E un $\mathbb{K}$-espace vectoriel et u un endomorphisme de E. Soit A est un sous-espace vectoriel stable par u. On appelle endomorphisme induit par u sur A l'endomorphisme $u_A \in \mathscr{L}(A)$ défini par*

$$\begin{array}{rcccc} u_A & : & A & \to & A \\ & & x & \mapsto & u(x). \end{array}$$

Une situation importante impliquant des sous-espaces vectoriels stables est celle dans laquelle on considère deux endomorphismes commutant entre eux.

Théorème 1.7 *Soit E un $\mathbb{K}$-espace vectoriel et u et v deux endomorphismes de E qui commutent entre eux, i.e. $u \circ v = v \circ u$. L'endomorphisme v laisse stable l'image de u, le noyau de u, et plus généralement tout sous-espace $\ker(u - \lambda\, id_E)$, avec λ un scalaire.*

DÉMONSTRATION. Puisque les endomorphismes u et v commutent entre eux, on a, pour tout vecteur x de E,

$$v(u(x)) = u(v(x)),$$

le dernier vecteur appartenant à l'image de u. On en déduit donc que $v(\mathrm{Im}(u))$ est inclus dans $\mathrm{Im}(u)$, qui est alors stable par v.

Considérons à présent un vecteur x du noyau de u. On a $u(v(x)) = v(u(x)) = v(0_E) = 0_E$, d'où $v(\ker(u))$ est inclus dans $\ker(u)$, qui est alors stable par v.

Soit enfin un scalaire λ. On a

$$v \circ (u - \lambda\, id_E) = v \circ u - \lambda\, v = u \circ v - \lambda\, v = (u - \lambda\, id_E) \circ v,$$

et les endomorphismes $u - \lambda\, id_E$ et v commutent donc aussi entre eux, ce qui permet de conclure. □

En dimension finie, il est possible d'interpréter matriciellement la propriété de stabilité d'une sous-espace vectoriel par un endomorphisme. Pour le voir, considérons un $\mathbb{K}$-espace vectoriel E de dimension finie égale à n, u un endomorphisme de E, A un sous-espace vectoriel de E stable par u, de dimension égale à p. Soit B un sous-espace vectoriel de E supplémentaire de A, c'est-à-dire tel que $E = A \oplus B$, et $\mathscr{B}$ une base adaptée [1] à une telle décomposition. Dans ce cas, la matrice représentative de l'endomorphisme u dans la base $\mathscr{B}$ est triangulaire supérieure par blocs,

$$\begin{pmatrix} M_{11} & M_{12} \\ 0_{n-p,p} & M_{22} \end{pmatrix},$$

où M_{11} est une matrice de $M_p(\mathbb{K})$, M_{12} est une matrice de $M_{p,n-p}(\mathbb{K})$ et M_{22} est une matrice de $M_{n-p}(\mathbb{K})$. Si l'on suppose de plus que le sous-espace B soit stable par u, alors la matrice de l'endomorphisme dans la base $\mathscr{B}$ est diagonale par blocs,

$$\begin{pmatrix} M_{11} & 0_{p,n-p} \\ 0_{n-p,p} & M_{22} \end{pmatrix}.$$

Plus généralement, si on a décomposé E en une somme directe de sous-espaces stables par u, $E = \oplus_{i=1}^{m} A_i$ avec $u(A_i) \subset A_i$ pour tout entier i appartenant à $\{1, \dots, m\}$, et si $\mathscr{B}$ est une base de E adaptée à cette décomposition, alors la matrice de l'endomorphisme dans la base $\mathscr{B}$ sera de la forme

$$\begin{pmatrix} M_{11} & 0 & \dots & 0 \\ 0 & M_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & M_{mm} \end{pmatrix}.$$

1. Cette adaptation s'entend dans le sens où les p premiers vecteurs composant la base forment une base du sous-espace A et les $n - p$ suivants forment une base du sous-espace B.

1.3 Éléments propres d'un endomorphisme

Définitions 1.8 (éléments propres d'un endomorphisme) *Soit E un $\mathbb{K}$-espace vectoriel et u un endomorphisme de E. Un scalaire λ est une **valeur propre** de u si et seulement s'il existe un vecteur non nul x de E tel que $u(x) = \lambda x$. Un vecteur non nul x de E tel qu'il existe un scalaire λ vérifiant $u(x) = \lambda x$ est un **vecteur propre** de u associé à la valeur propre λ.*

Il est important de remarquer qu'un vecteur propre est associé à une *unique* valeur propre, mais qu'une valeur propre possède une infinité de vecteurs propres, tout multiple non nul d'un vecteur propre donné étant lui-même un vecteur propre associé à la même valeur propre.

On notera également que la définition précédente est donnée pour un espace vectoriel de dimension quelconque. En considérant des espaces de dimension finie non nulle, il est possible de l'étendre à une matrice carrée en posant que les éléments propres de la matrice sont ceux de l'endomorphisme qui lui est canoniquement associé. Ainsi, en considérant une matrice M d'ordre n, on a que le scalaire λ est une valeur propre de M s'il existe une matrice colonne X non nulle de $M_{n,1}(\mathbb{K})$ telle que

$$MX = \lambda X.$$

De même, une matrice colonne X de $M_{n,1}(\mathbb{K})$ sera un vecteur propre de la matrice M si elle est non nulle et qu'il existe un scalaire λ tel que $MX = \lambda X$.

Tous les résultats énoncés pour des endomorphismes dans un espace vectoriel de dimension finie seront par conséquent valables *mutatis mutandis* pour des matrices carrées.

Définition 1.9 (spectre d'un endomorphisme) *L'ensemble des valeurs propres d'un endomorphisme u d'un $\mathbb{K}$-espace vectoriel s'appelle le **spectre** de u et se note* $\mathrm{Sp}(u)$.

Si cette dernière définition s'étend naturellement à toute matrice carrée M, la notation $\mathrm{Sp}(M)$ peut s'avérer ambigüe lorsque les coefficients de M sont des réels. En effet, suivant que l'on conçoit M comme une matrice à coefficients complexes ou réels, on cherchera ses valeurs dans $\mathbb{C}$ ou dans $\mathbb{R}$. Dans ce cas, on a recours aux notations $\mathrm{Sp}_{\mathbb{C}}(M)$ ou $\mathrm{Sp}_{\mathbb{R}}(M)$, qui sont plus explicites.

Remarque 1.10 *Un endomorphisme peut ne pas avoir de valeur propre. Par exemple l'endomorphisme u sur le $\mathbb{R}$-espace vectoriel $\mathbb{R}^2$ tel que $u(e_1) = -e_2$ et $u(e_2) = e_1$ n'a pas de valeur propre.*

Remarque 1.11 *Pour une matrice $M \in M_n(\mathbb{R})$ on a toujours $\mathrm{Sp}_{\mathbb{R}}(M) \subset \mathrm{Sp}_{\mathbb{C}}(M)$. Par exemple, la matrice dans la base canonique de l'endomorphisme de la remarque précédente est $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ et $\mathrm{Sp}_{\mathbb{R}}(A) = \emptyset$ mais $\mathrm{Sp}_{\mathbb{C}}(A) = \{-i, i\}$.*

Proposition et définition 1.12 (sous-espace propre) *Soit E un $\mathbb{K}$-espace vectoriel, u un endomorphisme de E et λ un scalaire. Les assertions suivantes sont équivalentes.*

(i) Le scalaire λ est une valeur propre de u.

(ii) L'endomorphisme $u - \lambda\, id_E$ n'est pas injectif.

(iii) Le noyau $\ker(u - \lambda\, id_E)$ n'est pas réduit à $\{0_E\}$.

*Dans ce cas, le sous-espace vectoriel $E_\lambda = \ker(u - \lambda\, id_E)$ est appelé **sous-espace propre** de u associé à la valeur propre λ.*

DÉMONSTRATION. Si λ est une valeur propre de l'endomorphisme u, alors, par définition, il existe un vecteur x non nul tel que $u(x) = \lambda x$. Il existe par conséquent un vecteur x non nul tel que $(u - \lambda\, id_E)(x) = 0_E$, ce qui signifie encore que l'endomorphisme $u - \lambda\, id_E$ n'est pas injectif, ce qui signifie encore que le noyau de cet endomorphisme n'est pas réduit au vecteur nul. □

Lorsque E est de dimension finie, on a l'équivalence

$$\lambda \in \mathrm{Sp}(u) \iff u - \lambda\, id_E \notin GL(E).$$

Proposition 1.13 *Soit $n \in \mathbb{N}^*$, $A \in M_n(K)$ et $P \in GL_n(\mathbb{K})$. Alors $\mathrm{Sp}(M) = \mathrm{Sp}(PMP^{-1})$ et $E_\lambda(PMP^{-1}) = PE_\lambda(M)$.*

Définition 1.14 (ordre de multiplicité géométrique d'une valeur propre) *On appelle **ordre de multiplicité géométrique** d'une valeur propre λ l'entier* $\dim(E_\lambda)$.

Proposition 1.15 *Soit E un $\mathbb{K}$-espace vectoriel, u un endomorphisme de E. Soit k un entier naturel non nul et $\lambda_1, \dots, \lambda_k$ des valeurs propres de u deux à deux distinctes. Alors, les sous-epaces propres E_{λ_i}, $i = 1, \dots, k$, sont en somme directe.*

DÉMONSTRATION. On raisonne par récurrence sur l'entier k. Pour $k = 1$, il n'y a rien à montrer. Pour k un entier naturel non nul, on suppose que si $\lambda_1, \dots, \lambda_k$ sont des valeurs propres de u deux à deux distinctes et $x_1, \dots, x_k$ sont des éléments des sous-espaces propres qui leur sont respectivement associés, alors l'égalité $x_1 + \dots + x_k = 0_E$ implique que $x_1 = \dots = x_k = 0_E$. Soit alors λ_{k+1} une valeur propre de u distincte des précédentes et x_{k+1} un vecteur propre associé. Dans ce cas, supposons que $x_1 + \dots + x_{k+1} = 0_E$, ce qui équivaut à $x_{k+1} = -(x_1 + \dots + x_k)$, d'où $u(x_{k+1}) = -u(x_1) - \dots - u(x_k)$, par linéarité de u. On a par conséquent $\lambda_{k+1}\, x_{k+1} = -\lambda_1\, x_1 - \dots - \lambda_k\, x_k$, soit encore $0_E = (\lambda_{k+1} - \lambda_1)\, x_1 + \dots + (\lambda_{k+1} - \lambda_k)\, x_k$. L'hypothèse de récurrence conduit alors à $(\lambda_{k+1} - \lambda_i)\, x_i = 0_E$ pour tout entier i dans $\{1, \dots, k\}$ et, puisque $\lambda_{k+1} - \lambda_i \neq 0$ pour tout entier i dans $\{1, \dots, k\}$, on en déduit que $x_i = 0_E$ pour tout entier i dans $\{1, \dots, k\}$, d'où $x_{k+1} = 0_E$. □

Corollaire 1.16 *Soit E un $\mathbb{K}$-espace vectoriel de dimension finie non nulle égale à n et u un endomorphisme de E. Alors, l'endomorphisme u possède au plus n valeurs propres distinctes.*

1.4 Polynômes d'endomorphismes, polynômes annulateurs

Soit $P(X) = a_0 + a_1 X + a_2 X^2 + \dots$ un polynôme de $\mathbb{K}[X]$. Pour tout endomorphisme u d'un $\mathbb{K}$-espace vectoriel E de dimension finie non nulle égale à n, on définit

$$P(u) = a_0\, id_E + a_1\, u + a_2\, u^2 + \dots, \tag{1.1}$$

où, pour tout entier naturel non nul k, on a posé $u^k = \underbrace{u \circ \dots \circ u}_{k \text{ termes}}$ et par convention $u^0 = id_E$.

La preuve du prochain résultat est laissée en exercice au lecteur.

Proposition 1.17 *Soit E un $\mathbb{K}$-espace vectoriel de dimension finie non nulle et u un endomorphisme de E. L'application de $\mathbb{K}[X]$ dans $\mathscr{L}(E)$ qui a tout polynôme P associe l'application linéaire $P(u)$ définie par* (1.1) *est un **morphisme d'algèbres**, c'est-à-dire qu'elle est linéaire, i.e.*

$$\forall \lambda \in \mathbb{K},\ \forall (P, Q) \in (\mathbb{K}[X])^2,\ (\lambda\, P + Q)(u) = \lambda\, P(u) + Q(u),$$

et vérifie

$$\forall (P, Q) \in (\mathbb{K}[X])^2,\ (PQ)(u) = P(u) \circ Q(u).$$

On remarquera en particulier que, pour tous polynômes P et Q et tout endomorphisme u, les endomorphismes $P(u)$ et $Q(u)$ commutent entre eux :

$$P(u) \circ Q(u) = Q(u) \circ P(u).$$

On peut de la même manière définir un morphisme d'algèbres sur l'ensemble des matrices carrées en posant, pour toute matrice M de $M_n(\mathbb{K})$,

$$P(M) = a_0\, I_n + a_1\, M + a_2\, M^2 + \dots$$

Remarque 1.18 *Attention aux notations. Si $P \in \mathbb{K}[X]$, $u \in \mathscr{L}(E)$ où E est un $\mathbb{K}$-espace vectoriel et $x \in E$, $P(u) \in \mathscr{L}(E)$ de sorte que l'évaluation de cet endomorphisme en x s'écrit $P(u)(x)$. L'écriture $P(u(x))$ n'a aucun sens !*

Définition 1.19 (polynôme annulateur) *Soit E un $\mathbb{K}$-espace vectoriel de dimension finie non nulle et u un endomorphisme de E. On dit qu'un polynôme P est un **polynôme annulateur** de u si l'endomorphisme $P(u)$ est nul.*

Une définition similaire vaut pour le polynôme annulateur d'une matrice carrée.

Exemple 1.20 *Si l'endomorphisme u est une projection, alors X^2-X est un polynôme annulateur de u, puisque $u^2 = u$ par définition d'une projection.*

L'ensemble des polynômes annulateurs d'un même endomorphisme forme un *idéal* [2].

Le résultat suivant établit un lien entre les valeurs propres d'un endomorphisme et les racines de tout polynôme annulateur de cet endomorphisme.

Proposition 1.21 *Soit u un endomorphisme d'un $\mathbb{K}$-espace vectoriel de dimension finie non nulle et P un polynôme annulateur de u. On a l'inclusion*

$$\mathrm{Sp}(u) \subseteq \{\textit{racines de } P\}.$$

DÉMONSTRATION. Si x est un vecteur propre de u associé à une valeur propre λ, on a $u(x) = \lambda\, x$, ce qui implique que, pour tout entier naturel k, $u^k(x) = \lambda^k x$ et, plus généralement que, pour tout polynôme P de $\mathbb{K}[X]$, $P(u)(x) = P(\lambda)x$. En particulier, avoir $P(u)(x) = 0_E$ implique que $P(\lambda)x = 0_E$ et donc que $P(\lambda) = 0$, puisque le vecteur x est non nul. □

Attention : toute racine d'un polynôme annulateur n'est pas nécessairement une valeur propre.

Exemple 1.22 *Si u est une projection sur un $\mathbb{K}$-espace vectoriel, alors $(X-10)(X^2-X)$ annule u. Pourtant 10 n'est pas valeur propre de u car les seuls valeurs propres possibles pour une projection sont 0 et 1.*

Définition 1.23 (polynôme minimal) *Un **polynôme minimal** d'un endomorphisme est un polynôme annulateur de cet endomorphisme, unitaire et de degré minimal.*

Proposition 1.24 *Soit u un endomorphisme d'un $\mathbb{K}$-espace vectoriel de dimension finie non nulle. Il existe un unique polynôme minimal de u, noté μ_u. De plus,*

$$\{P \in \mathbb{K}[X]\,,\ P \textit{ annule } u\} = \{\mu_u Q\,,\ Q \in \mathbb{K}[X]\}.$$

Autrement dit le polynôme minimal de u divise tout polynôme annulateur de u et si un polynôme est divisible par le polynôme minimal de u, alors il annule u. Enfin, si on note $d = \deg \mu_u$

$$d = \max\{k \in \mathbb{N}^*\,,\ \textit{la famille } (id_E, u, \cdots, u^{k-1}) \textit{ est libre dans } \mathscr{L}(E)\}.$$

DÉMONSTRATION. Notons E l'espace vectoriel de l'énoncé et n sa dimension. L'ensemble des polynômes annulateurs est le noyau de l'application linéaire suivante

$$\begin{array}{cccc} \Phi & : & \mathbb{K}[X] & \to & \mathscr{L}(E) \\ & & P & \mapsto & P(u). \end{array}$$

Comme $\mathbb{K}[X]$ est de dimension infinie et $\mathscr{L}(E)$ est de dimension finie (de dimension n^2) l'application Φ n'est pas injective et donc il existe un polynôme annulateur non nul. Considérons maintenant l'ensemble des degrés des polynômes annulateurs de u :

$$\{\deg(P)\,|\,P \in \mathbb{K}[X],\ P \neq 0_{\mathbb{K}[X]},\ P(u) = 0_{\mathscr{L}(E)}\}.$$

C'est une partie non vide de $\mathbb{N}$ (au vu du point précédent). Soit m son plus petit élément. Considérons un polynôme annulateur Π de u de degré m. On va montrer que tout polynôme annulateur P de u est un multiple de Π. Pour cela, on effectue la division euclidienne de P par Π :

$$P = \Pi Q + R,$$

où $\deg(R) \leq m-1$. Si $P(u) = 0_{\mathscr{L}(E)}$ et $\Pi(u) = 0_{\mathscr{L}(E)}$, alors $R(u) = 0_{\mathscr{L}(E)}$, ce qui implique que R est nul (car de degré strictement plus petit que le degré minimal) et le polynôme P est donc multiple de Π.

Si maintenant Π_1 et Π_2 sont deux polynômes annulateurs de u de degré m, ils sont multiples l'un de l'autre et ne diffèrent par conséquent que d'une constante multiplicative. On a donc unicité du polynôme minimal si l'on suppose que le coefficient de son terme de plus haut degré est fixé, ce que l'on a fait en posant qu'il est égal à 1.

2. Un idéal est un sous-ensemble remarquable d'un anneau : c'est un sous-groupe du groupe additif de l'anneau, qui est de plus stable par multiplication par les éléments de l'anneau.

Enfin, notons $d = \deg \mu_u$. Comme μ_u annule u, en écrivant $\mu_u(u) = 0$ on constate que d et donc k pour $k \geq d+1$, n'appartiennent pas à l'ensemble

$$\{k \in \mathbb{N}^* \text{ , la famille } (id_E, u, \cdots, u^{k-1}) \text{ est libre dans } \mathscr{L}(E)\}.$$

De plus cet ensemble est non vide (car 1 est dedans). Il admet donc un plus grand élément que nous noterons d'. On a $d' < d+1$ et donc $d' \leq d$. Supposons par l'absurde que $d' < d$ alors il existe des éléments $a_0, \cdots, a_{d'}$ de $\mathbb{K}$ tels que

$$\sum_{k=0}^{d'} a_k u^k = 0.$$

Donc le polynôme $P(X) = \sum_{k=0}^{d'} a_k X^k$ annule u, il est non nul et de degré strictement plus petit que le degré du polynôme minimal. Ceci est une contradiction. □

Exemple 1.25 *Si E est un $\mathbb{K}$-espace vectoriel non réduit à $\{0\}$ et u un projecteur non nul et différent de l'identité, $\mu_u = X^2 - X$.*

Proposition 1.26 *Soit un endomorphisme d'un $\mathbb{K}$-espace vectoriel de dimension finie non nulle. Les racines du polynôme minimal de cet endomorphisme sont exactement les valeurs propres de cet endomorphisme.*

DÉMONSTRATION. Compte tenu de la proposition 1.21, on a juste à montrer que toute racine du polynôme minimal d'un endomorphisme est une valeur propre de cet endomorphisme. Soit u un endomorphisme d'un $\mathbb{K}$-espace vectoriel E de dimension finie non nulle dont le scalaire λ est racine du polynôme minimal. Dans ce cas, on a $\mu_u(X) = (X-\lambda)Q(X)$, avec $\deg(Q) \leq \deg(\mu_u) - 1$ et on a donc

$$0_{\mathscr{L}(E)} = \mu_u(u) = (u - \lambda\, id_E) \circ Q(u).$$

Puisque $Q(u) \neq 0_{\mathscr{L}(E)}$ par minimalité de μ_u, l'endomorphisme $u - \lambda\, id_E$ n'est pas injectif et λ est donc une valeur propre de u. □

Corollaire 1.27 *Si E est un $\mathbb{C}$-espace vectoriel de dimension finie et $f\, \mathscr{L}(E)$ alors $\mathrm{Sp}(f) \neq \emptyset$.*

DÉMONSTRATION. Le polynôme minimal de f admet au moins une racine dans $\mathbb{C}$ par le théorème de d'Alembert-Gauss. □

On peut de la même manière définir le polynôme minimal d'une matrice $A \in M_n(\mathbb{K})$. Comme pour le spectre, on pourrait être tenter de faire attention à la définition du polynôme minimal suivant si on voit la matrice dans $M_n(\mathbb{R})$ ou $M_n(\mathbb{C})$. La prochaine proposition montre que ce n'est pas nécessaire et que la notion de polynôme minimal est indépendante du corps de base.

Proposition 1.28 *Soit $n \in \mathbb{N}^*$ et $M \in M_n(\mathbb{R})$. On note $\mu_{M,\mathbb{R}}$ son polynôme minimal vu comme matrice dans $M_n(\mathbb{R})$ et $\mu_{u,\mathbb{C}}$ son polynôme minimal vu comme matrice dans $M_n(\mathbb{C})$. Alors $\mu_{M,\mathbb{R}} = \mu_{M,\mathbb{C}}$.*

DÉMONSTRATION. Comme $\mu_{M,\mathbb{R}} \in \mathbb{C}[X]$ et que $\mu_{M,\mathbb{R}}$ annule M on a $\mu_{M,\mathbb{C}} | \mu_{M,\mathbb{R}}$ et $d_{\mathbb{C}} \leq d_{\mathbb{R}}$. Notons $d_{\mathbb{R}} = \deg \mu_{M,\mathbb{R}}$ et $d_{\mathbb{C}} = \deg \mu_{M,\mathbb{C}}$. De plus, par la proposition 1.24, la famille $(I_n, M, \cdots, M^{d_{\mathbb{R}}-1})$ est une famille libre de $M_n(\mathbb{R})$. Montrons que cette famille est aussi libre dans $M_n(\mathbb{C})$. Soit $a_0, \cdots, a_{d_{\mathbb{R}}-1})$ des éléments de $\mathbb{C}$ tels que

$$\sum_{k=0}^{d_{\mathbb{R}}-1} a_k M^k = 0.$$

Alors en prenant la partie réelle et la partie imaginaire, on en déduit facilement que la famille $(I_n, M, \cdots, M^{d_{\mathbb{R}}-1})$ est libre dans $M_n(\mathbb{C})$. Par la proposition 1.24 il vient que $d_{\mathbb{R}} \leq d_{\mathbb{C}}$. □

Nous concluons cette section avec le résultat riche de conséquences. Avant de l'énoncer, rappelons tout d'abord que deux polynômes P et Q de $\mathbb{K}[X]$ sont dits *premiers entre eux* si et seulement si

$$D \text{ divise } P \text{ et } Q \text{ dans } \mathbb{K}[X] \implies D \text{ est un polynôme constant.}$$

En particulier, deux polynômes à coefficients complexes sont premiers entre eux si et seulement s'ils n'ont pas de racines communes.

D'après le théorème de Bézout, le fait que les polynômes P et Q sont premiers entre eux équivaut à ce qu'il existe des polynômes R et S tels que $RP + SQ = 1$.

Proposition 1.29 (« lemme de décomposition des noyaux ») *Soit u un endomorphisme d'un $\mathbb{K}$-espace vectoriel E et P et Q des polynômes premiers entre eux. On a*

$$\ker((PQ)(u)) = \ker(P(u)) \oplus \ker(Q(u)).$$

DÉMONSTRATION. Soit x un vecteur appartenant à $\ker(P(u))$. On a alors

$$P(u)(x) = 0_E \implies Q(u)(P(u)(x)) = 0_E \implies (P(u) \circ Q(u))(x) = 0_E \implies (PQ)(u)(x) = 0_E,$$

d'où x appartient à $\ker((PQ)(u))$. Ainsi, on a prouvé que $\ker(P(u)) \subset \ker((PQ)(u))$. De la même manière, on a $\ker(Q(u)) \subset \ker((PQ)(u))$.

Les polynômes P et Q étant premiers entre eux, on a que

$$R(u) \circ P(u) + S(u) \circ Q(u) = id_E.$$

En évaluant cette dernière identité en un vecteur x de $\ker((PQ)(u))$, on obtient

$$(R(u) \circ P(u))(x) + (S(u) \circ Q(u))(x) = x.$$

Posons alors $y = (S(u) \circ Q(u))(x)$ et $z = (R(u) \circ P(u))(x)$. Il vient

$$P(u)(y) = P(u)((S(u) \circ Q(u))(x)) = S(u)((P(u) \circ Q(u))(x)) = S(u)(0_E) = 0_E,$$

et, de la même façon, $Q(u)(z) = 0_E$. On a ainsi que y appartient à $\ker(P(u))$ et que z appartient à $\ker(Q(u))$,d'où $\ker((PQ)(u)) = \ker(P(u)) + \ker(Q(u))$. Enfin, pour tout vecteur x de $\ker(P(u)) \cap \ker(Q(u))$, on a $(R(u) \circ P(u))(x) = (S(u) \circ Q(u))(x) = 0_E$ donc en utilisant l'identité, il vient que $x = 0_E + 0_E = 0_E$, d'où $\ker(P(u)) \cap \ker(Q(u)) = \{0_E\}$. □

Il est possible de généraliser ce résultat par récurrence.

Corollaire 1.30 (« lemme de décomposition des noyaux généralisé ») *Soit u un endomorphisme d'un $\mathbb{K}$-espace vectoriel, k un entier naturel strictement plus grand que 1 et $P_1, \dots, P_k$ des polynômes deux à deux premiers entre eux. On a*

$$\ker((P_1 \dots P_k)(u)) = \ker(P_1(u)) \oplus \dots \oplus \ker(P_k(u)).$$

Exemple 1.31 *Pour un projecteur u sur un $\mathbb{K}$-espace vectoriel E, on a toujours*

$$\ker(u) \oplus \ker(u - Id_E) = E.$$

Le lemme de décomposition des noyaux est particulière intéressant lorsque que l'on l'applique à un polynôme annulateur.

Corollaire 1.32 *Soit u un endomorphisme d'un $\mathbb{K}$-espace vectoriel E, k un entier naturel strictement plus grand que 1 et $P_1, \dots, P_k$ des polynômes deux à deux premiers entre eux. On suppose que le polynôme $P = P_1 \cdots P_k$ annule u. On a*

$$E = \ker(P_1(u)) \oplus \dots \oplus \ker(P_k(u)).$$

1.5 Polynôme caractéristique

On a vu que le scalaire λ était une valeur propre d'un endomorphisme u d'un espace vectoriel E si et seulement si l'endomophisme $u - \lambda\, id_E$ n'est pas injectif. Si l'espace E est de dimension finie, ceci équivaut encore à dire que le déterminant de $u - \lambda\, id_E$ est nul. Trouver les valeurs propres d'un endomorphisme en dimension finie revient donc à résoudre une équation polynomiale.

Définition 1.33 (polynôme caractéristique) *Soit E un $\mathbb{K}$-espace vectoriel de dimension finie non nulle et u un endomorphisme de E. On appelle* ***polynôme caractéristique*** *de u le polynôme χ_u de $\mathbb{K}[X]$ défini par $\chi_u(X) = \det(X\, id_E - u)$.*

Remarque 1.34 *On trouve parfois la définition $\chi_u(X) = \det(u - X\, id_E)$, qui a comme défaut le fait que le polynôme ainsi défini n'est pas nécessairement unitaire, le coefficient du terme de plus haut degré étant égal à -1 élevé à une puissance égale à la dimension de l'espace.*

Pour une matrice M d'ordre n, on a $\chi_M(X) = \det(X\, I_n - M)$. Compte tenu de cette définition et des propriétés du déterminant, il est clair que deux matrices semblables possèdent le même polynôme caractéristique. La réciproque est fausse, comme le montre l'exemple des matrices I_2 et $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$.

Théorème 1.35 *Soit E un $\mathbb{K}$-espace vectoriel de dimension finie non nulle et u un endomorphisme de E. Un scalaire λ est une valeur propre de u si et seulement si $\chi_u(\lambda) = 0$.*

Remarque 1.36 *Les valeurs propres d'une matrice triangulaire sont ses coefficients diagonaux.*

Exemple 1.37 *Le polynôme caractéristique de la matrice*

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

est $\chi_A(X) = X^2 + X - 8$.

Définition 1.38 (ordre de multiplicité algébrique d'une valeur propre) *On appelle **ordre de multiplicité algébrique** d'une valeur propre son ordre de multiplicité en tant que racine du polynôme caractéristique.*

Il découle des dernières définitions que deux matrices ayant le même polynôme caractéristique ont les mêmes valeurs propres, avec les mêmes ordres de multiplicité algébrique.

Proposition 1.39 (degré et coefficients du polynôme caractéristique d'une matrice) *Soit M une matrice d'ordre n à coefficients dans le corps $\mathbb{K}$. Le polynôme caractéristique de M est de degré égal à n et l'on a*

$$\chi_M(X) = X^n - \operatorname{tr}(M) X^{n-1} + \cdots + (-1)^n \det(M).$$

DÉMONSTRATION. Posons $\alpha = X\, I_n - M$. Par la formule de Leibniz pour le déterminant, on a

$$\chi_M(X) = \det(\alpha) = \sum_{\sigma \in \Sigma_n} \varepsilon(\sigma) \prod_{i=1}^{n} \alpha_{\sigma(i)\, i}.$$

Chacun des $n!$ termes de cette somme est le produit de n termes polynomiaux en X de degré inférieur ou égal à un. Le degré de χ_M est donc inférieur ou égal à n. De plus, un des termes de la somme est exactement de degré n si et seulement si chacun des facteurs le composant est de degré exactement égal à un. Ceci équivaut au fait que $\sigma = id_{\{1,\dots,n\}}$. Ainsi, la somme correspondant au polynôme caractéristique de M contient un unique terme de degré n et des termes de degré inférieur ou égal à $n-1$. C'est ainsi un polynôme de degré n, s'écrivant sous la forme

$$\begin{aligned} \chi_M(X) &= (X - m_{11}) \dots (X - m_{nn}) + \text{termes de degré inférieur ou égal à } n-1 \\ &= X^n + \text{termes de degré inférieur ou égal à } n-1. \end{aligned}$$

C'est donc bien un polynôme unitaire. Le coefficient du terme de degré nul de χ_M est par ailleurs donné par sa valeur en 0, qui est $\det(-M) = (-1)^n \det(M)$.

Enfin, pour déterminer le coefficient du terme de degré $n-1$, on observe que, quand la permutation σ n'est pas égale à l'identité, il existe un entier i de $\{1,\dots,n\}$ tel que $\sigma(i) \neq i$. En posant alors $j = \sigma^{-1}(i)$, on a que $\alpha_{\sigma(i)\, i} = -m_{\sigma(i)\, i}$ et $\alpha_{\sigma(j)\, j} = -m_{\sigma(j)\, j}$, et le terme $\varepsilon(\sigma) \prod_{i=1}^{n} \alpha_{\sigma(i)\, i}$ correspondant à cette permutation est de degré inférieur ou égal à $n-2$. Il en découle que

$$\begin{aligned} \chi_M(X) &= (X - m_{11}) \dots (X - m_{nn}) + \text{termes de degré inférieur ou égal à } n-2 \\ &= X^n - (m_{11} + \cdots + m_{nn}) X^{n-1} + \text{termes de degré inférieur ou égal à } n-2 \\ &= X^n - \operatorname{tr}(M) X^{n-1} + \text{termes de degré inférieur ou égal à } n-2. \end{aligned}$$

□

Remarque 1.40 *Pour une matrice* $M \in M_2(\mathbb{K})$ *on a en particulier* $\chi_M(X) = X^2 - \operatorname{tr}(M)X + \det(M)$.

On retrouve avec ce dernier résultat qu'une matrice M d'ordre n possède au plus n valeurs propres. Par ailleurs, si $\mathbb{K} = \mathbb{C}$, cette matrice admet exactement n valeurs propres comptées avec leurs ordres de multiplicité respectifs. Il en va de même si $\mathbb{K} = \mathbb{R}$ et si le polynôme caractéristique est *scindé*, c'est-à-dire décomposable en un produit de facteurs de degré un. Dans ces deux cas, on peut écrire

$$\chi_M(X) = (X - \lambda_1)\dots(X - \lambda_n)$$

où les scalaires $\lambda_1, \dots, \lambda_n$ sont les valeurs propres, distinctes ou confondues, de la matrice M, ou bien

$$\chi_M(X) = (X - \lambda_1)^{m_{\lambda_1}} \dots (X - \lambda_p)^{m_{\lambda_p}}$$

où l'entier naturel p est inférieur ou égal à n, les scalaires $\lambda_1, \dots, \lambda_p$ sont les valeurs propres, distinctes deux à deux, de M et les entiers naturels $m_{\lambda_1}, \dots, m_{\lambda_p}$ sont leurs ordres de multiplicité respectifs.

Proposition 1.41 (propriétés du polynôme caractéristique d'une matrice) *Soit A et B deux matrices carrées de même ordre. On a*

$$\chi_{A^\top} = \chi_A \text{ et } \chi_{AB} = \chi_{BA}.$$

DÉMONSTRATION. Soit A et B deux matrices d'ordre n. Le déterminant d'une matrice étant égal à celui de sa transposée, on a

$$\chi_{A^\top}(X) = \det\left(X\,I_n - A^\top\right) = \det\left((X\,I_n - A)^\top\right) = \det(X\,I_n - A) = \chi_A(X).$$

Pour la seconde assertion, si l'une des matrices, A par exemple, est inversible, alors $A^{-1}(AB)A = BA$, d'où AB et BA sont semblables. Dans le cas général, on considère les matrices de $M_{2n}(\mathbb{K})$ décomposées par blocs

$$M = \begin{pmatrix} BA & -B \\ 0 & 0 \end{pmatrix},\ N = \begin{pmatrix} 0 & -B \\ 0 & AB \end{pmatrix} \text{ et } P = \begin{pmatrix} I_n & 0 \\ A & I_n \end{pmatrix}.$$

On vérifie que $PN = MP = \left(\begin{smallmatrix} 0 & -B \\ 0 & 0 \end{smallmatrix}\right)$ et que P est une matrice triangulaire inférieure à diagonale non nulle, donc inversible. Les matrice M et N sont donc semblables. Le calcul par blocs de leurs polynômes caractéristiques respectifs laisse alors apparaître que $\chi_M(X) = \chi_{BA}(X)X^n$ et $\chi_N(X) = X^n\,\chi_{AB}(X)$, d'où la conclusion. □

Le théorème suivant est un résultat important de ce chapitre.

Théorème 1.42 (« théorème de Cayley–Hamilton ») *Soit u un endomorphisme d'un* $\mathbb{K}$*-espace vectoriel de dimension finie non nulle. On a*

$$\chi_u(u) = 0_{\mathscr{L}(E)}.$$

DÉMONSTRATION. Notons E le sous-espace de l'énoncé et n la dimension de E. On va montrer que, pour tout vecteur x non nul de E, $\chi_u(u)(x) = 0_E$. Soit $\mathscr{E}$ le sous-ensemble de $\mathbb{N}$ défini par

$$\mathscr{E} = \left\{k \in \mathbb{N}^* \mid \{u^i(x)\}_{0 \leq i \leq k-1} \text{ est une famille libre}\right\}.$$

C'est une partie non vide de $\mathbb{N}$ (puisque x est non nul), majorée par n. Elle admet donc un plus grand élément, que l'on note p. Par définition de p, la famille $\{x, u(x), \dots, u^{p-1}(x)\}$ est libre et on peut la compléter au besoin en une base $\mathscr{B} = \{x, u(x), \dots, u^{p-1}(x), e_{p+1}, \dots, e_n\}$ de E. Toujours par définition de p, la famille $\{x, u(x), \dots, u^p(x)\}$ est liée et l'on peut donc poser $u^p(x) = \sum_{i=0}^{p-1} a_i\,u^i(x)$. La matrice de u dans la base $\mathscr{B}$ s'écrit par blocs

$$\begin{pmatrix} B & C \\ 0 & D \end{pmatrix}$$

avec

$$B = \begin{pmatrix} 0 & \dots & \dots & 0 & a_0 \\ 1 & \ddots & & \vdots & a_1 \\ 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \dots & 0 & 1 & a_{p-1} \end{pmatrix} \tag{1.2}$$

une matrice d'ordre p. Un calcul par blocs du déterminant définissant χ_u conduisant à $\chi_u = \chi_B \chi_D$, un développement par rapport à la dernière colonne du déterminant fournissant χ_B donne pour sa part

$$\chi_B(X) = X^{p-1}(X - a_{p-1}) + \sum_{i=0}^{p-2} (-1)^{p+i+1}(-a_i)\,\Delta_i,$$

où

$$\Delta_i = \begin{vmatrix} X & 0 & \dots & 0 & * & \dots & \dots & * \\ * & \ddots & \ddots & \vdots & \vdots & & & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots & & & \vdots \\ * & \dots & * & X & * & \dots & \dots & * \\ 0 & \dots & \dots & 0 & -1 & * & \dots & * \\ \vdots & & & \vdots & 0 & \ddots & \ddots & \vdots \\ \vdots & & & \vdots & \vdots & \ddots & \ddots & * \\ 0 & \dots & \dots & 0 & 0 & \dots & 0 & -1 \end{vmatrix} = X^i(-1)^{p-1-i}.$$

Par suite, on trouve que

$$\chi_B(X) = X^p - a_{p-1} X^{p-1} - \sum_{i=0}^{p-2} a_i X^i = X^p - \sum_{i=0}^{p-1} a_i X^i,$$

On a finalement

$$\chi_u(u) = (u^p - \sum_{i=0}^{p-1} a_i\, u^i) \circ \chi_D(u).$$

Puisque ces deux polynômes de u commutent entre eux, on a obtenu que

$$\chi_u(u)(x) = \chi_D(u)\left(u^p(x) - \sum_{i=0}^{p-1} a_i\, u^i(x) \right) = \chi_D(u)(0_E) = 0_E.$$

Cette égalité restant vraie pour $x = 0_E$, on a prouvé le résultat. □

Ce résultat montre que le polynôme caractéristique (et tous ses multiples) est un polynôme annulateur de l'endomorphisme auquel il est associé.

Remarque 1.43 *Dans la démonstration ci-dessus, on a cherché le plus petit sous-espace vectoriel stable par u contenant le vecteur x. Il correspond au sous-espace vectoriel*

$$F_x := \{P(u)(x)\,,\ P \in \mathbb{K}[X]\}.$$

Dans la base choisie pour ce sous-espace vectoriel, la matrice de l'endomorphisme induit par u a la forme d'une matrice compagnon *(forme de la matrice* (1.2)*). L'endomorphisme induit par u sur le sous-espace est alors dit* cyclique.

Corollaire 1.44 *Soit u un endomorphisme d'un $\mathbb{K}$-espace vectoriel de dimension finie non nulle n. Alors* $\deg \mu_u \leq n$. *En particulier, la famille $(id_E, u, \cdots, u^n)$ est toujours liée.*

Exemple 1.45 (différence entre polynômes minimal et caractéristique) *Pour une homothétie de rapport λ, $u = \lambda\, id_E$, on a $\chi_u(X) = (X - \lambda)^n$ alors que $\mu_u(X) = X - \lambda$.*

1.6 Diagonalisation

Définition 1.46 (endomorphisme diagonalisable) *Soit E un $\mathbb{K}$-espace vectoriel de dimension finie non nulle et u un endomorphisme de E. On dit que u est **diagonalisable** s'il existe une base de E formée de vecteurs propres de u.*

Il résulte de la définition précédente qu'une matrice carrée est diagonalisable si elle représente un endomorphisme diagonalisable. Le résultat suivant donne une justification au vocabulaire employé.

Proposition 1.47 *Soit E un $\mathbb{K}$-espace vectoriel de dimension finie non nulle et u un endomorphisme de E. Les conditions suivantes sont équivalentes.*

(i) L'endomorphisme u est diagonalisable.

(ii) Il existe une base de E dans laquelle la matrice représentative de u est diagonale.

DÉMONSTRATION. On note n la dimension de E. Si $\mathscr{B}=(e_1,\dots,e_n)$ est une base de E formée de vecteurs propres de u associés aux valeurs propres $\lambda_1,\dots,\lambda_n$, comptées avec leurs ordres de multiplicité algébrique respectifs, la matrice représentative de u dans la base $\mathscr{B}$ est la matrice diagonale

$$\begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{pmatrix}.$$

Réciproquement, si $\mathscr{B}$ est une base de E, dire que la matrice représentative de u dans la base $\mathscr{B}$ est la matrice diagonale

$$\begin{pmatrix} \mu_1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \mu_n \end{pmatrix}$$

revient à dire que $u(e_1)=\mu_1\, e_1,\dots,u(e_n)=\mu_n\, e_n$. Les vecteurs de cette base sont donc des vecteurs propres de u et l'endomorphisme est donc diagonalisable. □

On dira qu'une matrice M d'ordre n est diagonalisable si et seulement si elle est semblable à une matrice diagonale. En notant $\mathscr{U}$ une base de $M_{n,1}(K)$ formée de vecteurs propres associés aux valeurs propres $\lambda_1,\dots,\lambda_n$ de M, $\mathscr{B}$ la base canonique de $M_{n,1}(K)$, P la matrice de passage de la base $\mathscr{B}$ à la base $\mathscr{U}$ et D la matrice diagonale telle que $d_{ii}=\lambda_i$, $i=1,\dots,n$, on a alors l'égalité

$$M=PDP^{-1}.$$

L'exemple d'un homothétie est particulièrement éclairant, car diagonaliser un endomorphisme revient à écrire l'espace comme une somme directe de sous-espaces propres de cet endomorphisme pour chacun desquels l'endomorphisme induit est une homothétie. Cette observation conduit à une nouvelle caractérisation de la diagonalisation.

Théorème 1.48 *Un endomorphisme d'un $\mathbb{K}$-espace vectoriel de dimension finie non nulle est diagonalisable si et seulement s'il existe un polynôme non nul, scindé sur $\mathbb{K}$ et à racines simples, annulateur de cet endomorphisme.*

DÉMONSTRATION. Soit u un endomorphisme d'un $\mathbb{K}$-espace vectoriel E de dimension finie non nulle. On note n la dimension de E. Supposons que u est diagonalisable. Soit $\{\lambda_1,\dots,\lambda_r\}$ l'ensemble des valeurs propres de u. Soit $\mathscr{B}=(e_1,\dots,e_n)$ est une base de E formée de vecteurs propres de u associés aux valeurs propres. La matrice représentative de u dans la base $\mathscr{B}$ est une matrice diagonale D et on remarque que $\prod_{i=1}^{r}(X-\lambda_i)$ annule D et est scindé à racines simples. Pour montrer l'implication réciproque, on suppose qu'il existe un polynôme P annulateur de u, de la forme

$$P(X)=\prod_{i=1}^{p}(X-\mu_i),$$

les scalaires μ_i étant deux à deux distincts. Les monômes $X-\mu_1,\dots,X-\mu_p$ étant deux à deux premiers entre eux, le corollaire 1.30 permet d'écrire que

$$E=\bigoplus_{i=1}^{p}\ker(u-\mu_i\, id_E).$$

Une base de E adaptée [3] à cette décomposition en somme directe est une base de vecteurs propres de u et l'endomorphisme est donc diagonalisable. □

3. Si $\ker(u-\mu_i\, id_E)$ se trouve réduit au vecteur nul, on considère qu'une « base » de ce noyau est l'ensemble vide.

Corollaire 1.49 *Un endomorphisme d'un $\mathbb{K}$-espace vectoriel de dimension finie non nulle est diagonalisable si et seulement si son polynôme minimal est scindé à racines simples.*

Enfin, donnons une caractérisation reliée au polynôme caractéristique.

Théorème 1.50 (condition nécessaire et suffisante de diagonalisabilité) *Soit un endomorphisme u d'un $\mathbb{K}$-espace vectoriel E de dimension finie non nulle n. Cet endomorphisme est diagonalisable si et seulement si son polynôme caractéristique est scindé sur $\mathbb{K}$ et l'ordre de multiplicité algébrique de chaque valeur propre est égal à la dimension du sous-espace propre correspondant.*

DÉMONSTRATION. Supposons tout d'abord que le polynôme caractéristique χ_u soit scindé sur $\mathbb{K}$ et que l'ordre de multiplicité algébrique de chaque valeur propre soit égal à la dimension du sous-espace propre correspondant. Posons $\chi_u(X) = \prod_{i=1}^{p}(X-\lambda_i)^{m_{\lambda_i}}$, où les scalaires λ_i, $i = 1,\dots,p$, sont les valeurs propres deux à deux distinctes de u et les entiers m_{λ_i}, $i = 1,\dots,p$, sont leurs ordres de multiplicité respectifs. Pour tout entier i de $\{1,\dots,p\}$, posons par ailleurs $n_{\lambda_i} = \dim(E_{\lambda_i})$. Par hypothèse, on a, pour tout entier i de $\{1,\dots,p\}$, $n_{\lambda_i} = m_{\lambda_i}$, ce qui implique alors que

$$n = \deg(\chi_u) = \sum_{i=1}^{p} m_{\lambda_i} = \sum_{i=1}^{p} n_{\lambda_i}.$$

Les sous-espaces propres étant en somme directe en vertu de la proposition 1.15, il existe donc une base de E formée de vecteurs propres de u et la proposition 1.47 permet alors d'affirmer que l'endomorphisme est diagonalisable.

Réciproquement, supposons que l'endomorphisme u soit diagonalisable. L'espace vectoriel E est somme directe des sous-espaces propres de u, c'est-à-dire

$$E = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_p},$$

où les scalaires λ_i, $i = 1,\dots,p$, sont les valeurs propres deux à deux distinctes de u. Dans une base de E associée à cette décomposition en somme directe, la matrice représentative de u s'écrit

$$D = \begin{pmatrix} \lambda_1 & & & & & & \\ & \ddots & & & & & \\ & & \lambda_1 & & & & \\ & & & \ddots & & & \\ & & & & \lambda_p & & \\ & & & & & \ddots & \\ & & & & & & \lambda_p \end{pmatrix},$$
$$\underbrace{\qquad}_{\dim(E_{\lambda_1})\text{ fois}} \qquad \underbrace{\qquad}_{\dim(E_{\lambda_p})\text{ fois}}$$

ce qui implique que

$$\chi_u(X) = \chi_D(X) = \prod_{i=1}^{p}(X-\lambda_i)^{\dim(E_{\lambda_i})}.$$

Le polynôme caractéristique de u est donc scindé sur $\mathbb{K}$ et l'ordre de multiplicité de chaque valeur propre correspond à la dimension du sous-espace propre associé. □

On a évidemment un énoncé similaire pour une matrice carrée. La preuve du résultat suivant est laissée au lecteur.

Corollaire 1.51 (condition suffisante de diagonalisabilité) *Soit un endomorphisme d'un $\mathbb{K}$-espace vectoriel de dimension finie non nulle égale à n. Si cet endomorphisme possède n valeurs propres distinctes, alors il est diagonalisable et ses sous-espaces propres sont des droites vectorielles.*

Exemple 1.52 (Projection) *Soit E un $\mathbb{K}$-espace vectoriel de dimension finie et $u \in \mathscr{L}(E)$ une projection. Alors u est diagonalisable car $X(X-1)$ annule u.*

Exemple 1.53 *La matrice*

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

n'est pas diagonalisable car $\chi_A(X) = (X-3)^2$ et $E_3(A)$ est de dimension 1.

Exemple 1.54 *Si $A \in M_n(\mathbb{K})$ est une matrice nilpotente non nulle, A n'est pas diagonalisable. En effet, $\chi_A(X) = X^n$ et par le théorème de cayley-Hamilton, comme A est non nulle, $\mu_A(X) = X^k$ avec $k > 1$.*

La diagonalisation d'une matrice en pratique

Soit M une matrice d'ordre n à coefficients dans le corps $\mathbb{K}$.

1. Calculer le polynôme caractéristique χ_M, trouver ses racines et une factorisation associée du polynôme pour obtenir les valeurs propres avec leurs ordres de multiplicité respectifs. Si le polynôme n'est pas scindé, la matrice n'est pas diagonalisable.
2. Si le polynôme caractéristique est scindé, décrire pour chaque valeur propre le sous-espace propre qui lui est associé.
3. Pour chacune des valeurs propres, comparer l'ordre de multiplicité algébrique de la valeur propre avec la dimension du sous-espace propre correspondant. Si ces deux nombres sont égaux pour *toutes* les valeurs propres, alors la matrice est diagonalisable. S'il existe au moins une valeur propre pour laquelle ce n'est pas le cas, la matrice n'est pas diagonalisable.
4. Si la matrice est diagonalisable et si $\{V_1, \dots, V_n\}$ est une base de $M_{n,1}(\mathbb{K})$ formée de vecteurs propres associés aux valeurs propres $\lambda_1, \dots, \lambda_n$ de M comptées avec leur ordre de multiplicité, alors la matrice P dont les colonnes sont les vecteurs V_i, $i = 1, \dots, n$, est telle que
$$M = P\,\mathrm{diag}(\lambda_1, \dots, \lambda_n)P^{-1}.$$

1.7 Trigonalisation

On a vu dans la section précédente que tout endomorphisme n'est pas nécessairement diagonalisable. On va maintenant montrer que l'on peut en revanche toujours trouver une base de l'espace par rapport à laquelle la matrice d'un endomorphisme défini sur un $\mathbb{C}$-espace vectoriel est une matrice triangulaire.

Définition 1.55 (endomorphisme trigonalisable) *Soit un $\mathbb{K}$-espace vectoriel de dimension finie non nulle et un endomorphisme de E. On dit que cet endomorphisme est **trigonalisable** si et seulement s'il existe une base de l'espace dans laquelle sa matrice est triangulaire supérieure.*

On notera que, dans la définition précédente, on aurait pu tout aussi bien remplacer « triangulaire supérieure » par « triangulaire inférieure ».

Par extension, une matrice carrée sera dite trigonalisable si et seulement si elle est semblable à une matrice triangulaire supérieure. Les valeurs propres d'une matrice trigonalisable sont ainsi données par les coefficients diagonaux d'une matrice triangulaire qui lui est semblable.

Théorème 1.56 (condition nécessaire et suffisante de trigonalisabilité) *Un endomorphisme d'un $\mathbb{K}$-espace vectoriel de dimension finie non nulle est trigonalisable si et seulement si son polynôme caractéristique est scindé sur $\mathbb{K}$. En particulier, tout endomorphisme d'un $\mathbb{C}$-espace vectoriel est trigonalisable.*

DÉMONSTRATION. Considérons que l'espace est de dimension n et soit M une matrice de $M_n(\mathbb{K})$ représentant l'endomorphisme. Si l'endomorphisme est trigonalisable, il existe une matrice P de $GL_n(\mathbb{K})$ et une matrice T triangulaire supérieure telles que $M = PTP^{-1}$. En posant
$$T = \begin{pmatrix} \lambda_1 & * & \dots & * \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \dots & 0 & \lambda_n \end{pmatrix},$$
les coefficients λ_i, $i = 1, \dots, n$, appartenant à $\mathbb{K}$, on trouve alors que $\chi_M(X) = \chi_T(X) = \prod_{i=1}^{n}(X - \lambda_i)$. Le polynôme caractéristique de l'endomorphisme est donc scindé sur $\mathbb{K}$.

Montrons à présent la réciproque en raisonnant par récurrence sur l'ordre de la matrice. Si $n = 1$, tout élément de $M_n(\mathbb{K})$ est triangulaire et donc trigonalisable. Supposons à présent que l'entier n soit supérieur ou égal à 1 et que tout élément de $M_n(\mathbb{K})$ dont le polynôme caractéristique est scindé sur $\mathbb{K}$ soit trigonalisable. Soit une matrice M de $M_{n+1}(\mathbb{K})$, telle que χ_M est scindé sur $\mathbb{K}$. L'endomorphisme de $\mathbb{K}^{n+1}$ canoniquement associé à M admet au moins une valeur propre λ_1 appartenant à $\mathbb{K}$. Soit v_1 un vecteur propre associé à λ_1. La famille $\{v_1\}$ étant libre, on peut la compléter en une base de $\mathbb{K}^{n+1}$, la matrice de l'endomorphisme dans cette base s'écrivant alors sous la forme par blocs

$$M' = \begin{pmatrix} \lambda_1 & L \\ 0 & M_1 \end{pmatrix},$$

avec L appartenant à $M_{1,n}(\mathbb{K})$ et M_1 appartenant à $M_n(\mathbb{K})$. Les matrices M et M' étant semblables, il existe une matrice P' de $GL_{n+1}(\mathbb{K})$ telle que

$$M = P'M'(P')^{-1},$$

et un calcul par blocs du déterminant montre que $\chi_{M'}(X) = (X - \lambda_1)\det(X\,I_n - M_1) = (X - \lambda_1)\chi_{M_1}(X)$. D'autre part, χ_M étant scindé sur $\mathbb{K}$, on peut poser $\chi_M(X) = \prod_{i=1}^{n+1}(X - \lambda_i)$, dont on déduit que $\chi_{M_1}(X) = \prod_{i=2}^{n+1}(X - \lambda_i)$. L'hypothèse de récurrence assure alors l'existence d'une matrice P_1 de $GL_n(\mathbb{K})$ et d'une matrice triangulaire supérieure T_1 telles que $M_1 = P_1 T_1 {P_1}^{-1}$.

Il suffit ensuite de poser

$$P'' = \begin{pmatrix} 1 & 0 \\ 0 & P_1 \end{pmatrix}.$$

Par un calcul par blocs, on a $\det(P'') = \det(P_1)$, d'où P'' est inversible et, toujours en calculant par blocs, il vient

$$(P'')^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & {P_1}^{-1} \end{pmatrix}$$

et

$$(P'')^{-1}M'P'' = \begin{pmatrix} \lambda_1 & LP_1 \\ 0 & T_1 \end{pmatrix} = T.$$

Posons finalement $P = P'P''$. Cette dernière matrice est inversible et l'on a

$$P^{-1}MP = (P'')^{-1}(P')^{-1}MP'P'' = (P'')^{-1}M'P'' = T.$$

L'endomorphisme représenté par M est donc bien trigonalisable. □

Le dernier théorème est vrai pour une matrice carrée à coefficient réels pour peu que l'on factorise son polynôme caractéristique sur $\mathbb{C}$ et non sur $\mathbb{R}$. C'est en ce sens que l'on peut dire, comme on l'a écrit en début de section, que l'on peut toujours trouver une base dans laquelle la matrice d'un endomorphisme défini sur un $\mathbb{C}$-espace vectoriel est triangulaire.

Exemple 1.57 *La matrice*

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

est trigonalisable car $\chi_A(X) = X^2 - 1 = (X - 1)(X + 1)$.

Exemple 1.58 *La matrice*

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

n'est pas trigonalisable sur $\mathbb{R}$ *car* $\chi_A(X) = X^2 + 1$ *n'est pas scindé dans* $\mathbb{R}[X]$.

La preuve du résultat suivant est laissée au lecteur.

Corollaire 1.59 *Soit E un $\mathbb{K}$-espace vectoriel de dimension finie non nulle et u un endomorphisme de E. On suppose que le polynôme caractéristique est scindé et qu'il s'écrit*

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

avec les λ_i deux à deux distincts. On a alors

$$\operatorname{tr}(u) = \sum_{i=1}^{p} m_{\lambda_i} \lambda_i \text{ et } \det(u) = \prod_{i=1}^{p} {\lambda_i}^{m_{\lambda_i}}.$$

La trigonalisation d'une matrice en pratique

Soit M une matrice d'ordre n à coefficients dans le corps $\mathbb{K}$.

1. Calculer le polynôme caractéristique χ_M, trouver ses racines et une factorisation associée du polynôme pour obtenir les valeurs propres. Si le polynôme caractéristique n'est pas scindé, la matrice n'est pas trigonalisable.
2. Si le polynôme caractéristique est scindé, il s'agit à présent de trouver une base dans laquelle la matrice de l'endomorphisme représenté par M sera triangulaire supérieure.

   On a intérêt à placer dans cette base le plus grand nombre possible de vecteurs propres de la matrice, en déterminant pour cela une base de chaque sous-espace propre. La réunion de ces bases constitue bien une partie de la base recherchée, puisque les sous-espaces propres sont en somme directe. Le choix de ces premiers vecteurs (ici au nombre de p) impose que la matrice triangulaire sera de la forme

$$T = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 & * & \dots & \dots & * \\ 0 & \ddots & \ddots & \vdots & \vdots & & & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots & & & \vdots \\ \vdots & & \ddots & \lambda_p & * & & & \vdots \\ \vdots & & & \ddots & \lambda_{p+1} & \ddots & & \vdots \\ \vdots & & & & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & & & \ddots & \ddots & * \\ 0 & \dots & \dots & \dots & \dots & \dots & 0 & \lambda_n \end{pmatrix}.$$

   Il reste à compléter la famille obtenue pour arriver à une base. On peut pour cela s'y prendre de la façon suivante.

   Procédons itérativement en supposant que l'on a déjà déterminé les vecteurs $V_1, \dots, V_j$, avec $p \leq j \leq n-1$. Afin de choisir le vecteur suivant, on commence par compléter la famille $\{V_1, \dots, V_j\}$ en une base $\{V_1, \dots, V_j, U_{j+1}, \dots, U_n\}$ de $M_{n,1}(\mathbb{K})$, puis on cherche V_{j+1} sous la forme $V_{j+1} = \sum_{i=j+1}^{n} \alpha_i U_i$, la forme de la matrice T imposant que $MV_{j+1} = \sum_{i=1}^{j} t_{i\,j+1} V_i + \lambda_{j+1} V_{j+1}$. En explicitant cette relation, on obtient n équations linéaires dont les inconnues sont les n coefficients $t_{1\,j+1}, \dots, t_{j\,j+1}, \alpha_{j+1}, \dots, \alpha_n$:

$$\sum_{i=1}^{j} t_{i\,j+1} V_i + \sum_{i=j+1}^{n} \alpha_i (\lambda_{j+1} U_i - MU_i) = 0.$$

   Toute solution non nulle de ce système fournit un vecteur V_{j+1} possible et les coefficients possiblement non nuls correspondants de la $j+1^{\text{e}}$ colonne de T.

1.8 Sous espaces caractéristiques et multiplicités des valeurs propres (hors programme)

On peut aller plus loin que la simple trigonalisation en construisant une base de l'espace ambiant dans laquelle la matrice de l'endomorphisme est diagonale par blocs, chacun des blocs étant une matrice triangulaire supérieure dont les coefficients diagonaux sont identiques.

Soit un $\mathbb{K}$-espace vectoriel E de dimension finie n et $f \in \mathscr{L}(E)$. Soit λ une valeur propre de f. On peut alors écrire

$$\chi_f(X) = (X-\lambda)^m Q(X)$$

avec Q et $(X-\lambda)$ premiers entre eux. On rappelle que m s'appelle la multiplicité algébrique de la valeur propre λ.

De plus on a aussi

$$\mu_f(X) = (X-\lambda)^l R(X)$$

avec R et $(X-\lambda)$ premiers entre eux. Nous allons essayer de comprendre ce que représente ces multiplicités. Commençons par la multiplicité algébrique.

Définition 1.60 (Sous-espace caractéristique) *Soit un $\mathbb{K}$-espace vectoriel E de dimension finie et $f \in \mathscr{L}(E)$. Pour une valeur propre λ de f de multiplicité algébrique m on définit $C_\lambda := \ker(f - \lambda id_E)^m$ que l'on nomme sous-espace caractéristique de f associé à la valeur propre λ.*

On a alors le résultat suivant.

Proposition 1.61 *Soit un $\mathbb{K}$-espace vectoriel E de dimension finie n et $f \in \mathscr{L}(E)$. Soit λ une valeur propre de multiplicité algébrique m. Alors C_λ est stable par f et $\dim C_\lambda = m$. De plus, on a $E_\lambda \subset C_\lambda$ et en particulier la multiplicité géométrique est toujours plus petite que la multiplicité algébrique.*

DÉMONSTRATION. Le but de la preuve va être d'exprimer χ_f de deux manières. Nous savons déjà que

$$\chi_f(X) = (X-\lambda)^m Q(X)$$

avec $(X-\lambda)$ et Q premiers entre eux. Les endomorphismes f et $(f-\lambda id_E)^m$ commutent donc nous avons bien C_λ stable par f. Notons f_λ l'endomorphisme induit par f sur C_λ. Similairement $\ker Q(f)$ est stable par f et nous notons f_Q l'endomorphisme induit par f sur ce noyau. Par le lemme de décomposition des noyaux, nous avons

$$E = \ker(f-\lambda id_E)^m \oplus \ker Q(f).$$

En prenant une base adaptée à cette décomposition nous obtenons la matrice

$$\begin{pmatrix} M_\lambda & 0 \\ 0 & M_Q \end{pmatrix}.$$

La matrice M_λ est la matrice de l'endomorphisme f_λ dans une certaine base et M_Q la matrice de f_Q. En calculant le polynôme caractéristique de la matrice précédente nous obtenons

$$\chi_f = \chi_{f_\lambda} \chi_{f_Q}.$$

Au vu du point précédent nous avons $\chi_{f_\lambda} | \chi_f$. Par définition de C_λ le polynôme $(X-\lambda)^m$ annule f_λ. Ainsi le polynôme minimal de f_λ est scindé et $\mathrm{Sp}(f_\lambda) = \{\lambda\}$. On peut donc trigonaliser f_λ et en calculant le polynôme caractéristique il vient que $\chi_{f_\lambda}(X) = (X-\lambda)^{\dim(C_\lambda)}$. De plus, $\chi_{f_Q} | \chi_f$ et par définition de f_Q, le polynôme Q annule f_Q. Ainsi, λ ne peut être valeur propre de f_Q car sinon elle serait racine de Q (ce qui n'est pas possible car $(X-\lambda)$ et Q sont premiers entre eux). Nous tirons de cela une autre manière d'exprimer χ_f

$$\chi_f(X) = (X-\lambda)^{\dim(C_\lambda)} \chi_{f_Q}(X)$$

avec χ_{f_Q} et $(X-\lambda)$ premiers entre eux. Par unicité de la factorisation en polynômes irréductibles, il vient que $\dim(C_\lambda) = m$. □

Intéressons nous maintenant à la multiplicité dans le polynôme minimal. Nous avons le résultat suivant.

Proposition 1.62 *Soit un $\mathbb{K}$-espace vectoriel E de dimension finie n et $f \in \mathscr{L}(E)$. Soit λ une valeur propre de multiplicité l dans le polynôme minimal et m dans le polynôme caractéristique. Alors $l \leq m$ et $(\ker(f-\lambda id_E)^k)_{k\in\mathbb{N}}$ stationne exactement à partir du rang l, autrement dit*

$$\ker(f-\lambda id_E)^{l-1} \subsetneq \ker(f-\lambda id_E)^l = \ker(f-\lambda id_E)^{l+1}.$$

De plus, on a

$$C_\lambda = \ker(f - \lambda Id_E)^l.$$

DÉMONSTRATION. On note $\mu_f(X) = (X-\lambda)^l R(X)$ avec R et $(X-\lambda)$ premiers entre eux et $\chi_f(X) = (X-\lambda)^m Q(X)$ avec Q et $(X-\lambda)$ premier entre eux. Le théorème de Cayley-Hamilton nous donne que $l \leq m$ et que $R|Q$ (on utilise le lemme de Gauss). Puis on applique le lemme de décomposition des noyaux aux deux polynômes

$$E = \ker(f-\lambda id_E)^l \oplus \ker R(f) \text{ et } E = \ker(f-\lambda id_E)^m \oplus \ker Q(f).$$

On constate alors $\ker(f-\lambda id_E)^l \subset \ker(f-\lambda id_E)^m$ et $\ker R(f) \subset \ker Q(f)$ (car $l \leq m$ et $R|Q$). Cela oblige $\ker(f-\lambda id_E)^l = \ker(f-\lambda id_E)^m$ et donc que la suite des noyaux stationne au maximum à partir du rang l.

Supposons par l'absurde que $\ker(f-\lambda id_E)^{l-1} = \ker(f-\lambda id_E)^l$ (à savoir que la suite stationne avant le rang l). Posons $P(X) = (X-\lambda)^{l-1}R$. On rappelle que

$$E = \ker(f-\lambda id_E)^l \oplus \ker R(f)$$

et nous notons f_λ l'endomorphisme induit par f sur $\ker(f-\lambda id_E)^l$ et f_R l'endomorphisme induit par f sur $\ker R(f)$. On a alors $P(f) = 0$ car P annule f_R (comme $R|P$) et car P annule f_λ comme $\ker(f-\lambda id_E)^{l-1} = \ker(f-\lambda id_E)^l$. Mais ceci est absurde car $\deg P < \deg \mu_f$. □

Corollaire 1.63 *Soit un $\mathbb{K}$-espace vectoriel E de dimension finie n et $f \in \mathscr{L}(E)$. f est diagonalisable si et seulement si χ_f est scindé et pour toute valeur propre λ de f on a* $\ker(f-\lambda id_E) = \ker(f-\lambda id_E)^2$.

On va maintenant s'intéresser plus en détail aux sous-espaces caractéristiques.

Proposition 1.64 *Soit un $\mathbb{K}$-espace vectoriel E de dimension finie n et $f \in \mathscr{L}(E)$. Soit λ une valeur propre de f, notons l la multiplicité de λ dans le polynôme minimal de f et m sa multiplicité algébrique. Notons f_λ l'endomorphisme induit par f sur C_λ. Alors*

$$f_\lambda = \lambda id_{C_\lambda} + n_\lambda$$

et n_λ est un endomorphisme nilpotent d'indice de nilpotence l. De plus, $\chi_{f_\lambda}(X) = (X-\lambda)^m$ et $\mu_{f_\lambda}(X) = (X-\lambda)^l$.

DÉMONSTRATION. On reprend les résultats des preuves de la proposition 1.61 et de la proposition 1.62. On voit que $\chi_{f_\lambda}(X) = (X-\lambda)^m$ et que $\mu_{f_\lambda}(X) = (X-\lambda)^l$. En particulier, X^l est le polynôme minimal de $n_\lambda := f_\lambda - \lambda id_{C_\lambda}$. Donc n_λ est un endomorphisme nilpotent d'indice de nilpotence l. □

On peut maintenant introduire la décomposition en sous-espace caractéristique d'un endomorphisme dont le polynôme caractéristique est scindé. Cette décomposition est intéressante car elle généralise la décomposition en sous-espaces propres à des endomorphisme qui ne pas nécessairement diagonalisables.

Théorème 1.65 (Décomposition en sous-espaces caractéristiques) *Soit un $\mathbb{K}$-espace vectoriel E de dimension finie n et $f \in \mathscr{L}(E)$. On suppose que χ_f est scindé et qu'il s'écrit*

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

avec les λ_i deux à deux distincts. Alors

$$E = \oplus_{i=1}^r C_{\lambda_i}.$$

De plus, il existe une base adaptée à cette décomposition telle que la matrice de f dans cette base a la forme

$$\begin{pmatrix} C_{11} & 0 & \dots & 0 \\ 0 & C_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & C_{rr} \end{pmatrix}$$

et, pour tout $i \in [\![1, r]\!]$, $C_{ii} \in M_{m_{\lambda_i}}(\mathbb{K})$ est la matrice de l'endomorphisme induit par f sur C_{λ_i} dans une certaine base et a la forme

$$C_{ii} = \begin{pmatrix} \lambda_i & * & \dots & * \\ 0 & \lambda_i & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \dots & 0 & \lambda_i \end{pmatrix}.$$

DÉMONSTRATION. On utilise le lemme de décomposition des noyaux sur χ_f et le théorème de Cayley-Hamilton. On regarde ensuite ce qui se passe sur chaque sous-espace caractéristique. On utilise alors la proposition précédente. Pour chaque $i \in [\![1, r]\!]$, on considère une base de trigonalisation $\mathscr{B}_i$ pour l'endomorphisme nilpotent n_{λ_i}. On remarque alors que la matrice de n_{λ_i} dans $\mathscr{B}_i$ est triangulaire supérieure stricte et que la matrice de $\lambda id_{C_{\lambda_i}}$ dans $\mathscr{B}_i$ est $\lambda_i I_{m_i}$. Ainsi la matrice de l'endomorphisme f_{λ_i} dans $\mathscr{B}_i$ est bien de la forme annoncée. Enfin, la base adaptée à la décomposition est la concaténation des bases $\mathscr{B}_i$ en respectant l'ordre.

□

Remarque 1.66 *Si f est diagonalisable, la décomposition en sous-espaces caractéristiques correspond à la décomposition en sous-espaces propres.*

Corollaire 1.67 *Soit un $\mathbb{K}$-espace vectoriel E de dimension finie n et $f \in \mathscr{L}(E)$. f est trigonalisable si et seulement si E est la somme directe des sous-espaces caractéristiques de f.*

Cette décomposition est particulièrement utile quand $\mathbb{K} = \mathbb{C}$ puisqu'alors tout polynôme est scindé. Attention, cette décomposition n'a pas de sens si χ_f n'est pas scindé !

Exemple 1.68 *Considérons la matrice*

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

Après calculs on obtient que $\chi_A(X) = X(X-1)^2$. Donc χ_A est scindé. La valeur propre 0 a pour multiplicité algébrique 1. Donc $E_0 = C_0$. La valeur propre 1 a pour multiplicité algébrique 2 donc C_1 est de dimension 2. Après calculs on obtient que

$$E_0 = \mathrm{Vect}\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \ E_1 = \mathrm{Vect}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \ \textit{et} \ \ker(A - I_3)^2 = \mathrm{Vect}\left\{\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}\right\}.$$

Ainsi

$$C_0 = \mathrm{Vect}\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \ \textit{et} \ C_1 = \mathrm{Vect}\left\{\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}\right\}.$$

Exemple 1.69 *Considérons la matrice*

$$B = \begin{pmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{pmatrix}.$$

Après calculs on obtient que $\chi_B(X) = (X+2)^2(X-1)$. Donc χ_B est scindé. La valeur propre 1 a pour multiplicité algébrique 1. Donc $E_1 = C_1$. La valeur propre -2 a pour multiplicité algébrique 2 donc C_{-2} est de dimension 2. Après calculs on obtient que

$$E_1 = \mathrm{Vect}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \ \textit{et} \ E_{-2} = \mathrm{Vect}\left\{\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}\right\}.$$

Ainsi B est diagonalisable et

$$C_1 = \mathrm{Vect}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \ \textit{et} \ C_{-2} = \mathrm{Vect}\left\{\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}\right\}.$$

1.9 Résumé

On peut à présent résumer les conditions de diagonalisabilité et de trigonalisation établies dans ce chapitre.

Conditions de diagonalisabilité et de trigonalisabilité d'un endomorphisme d'un $\mathbb{K}$-espace vectoriel de dimension finie non nulle

Soit E un espace vectoriel sur $\mathbb{K}$ de dimension finie non nulle égale à n et u un endomorphisme de E de valeurs propres $\lambda_1, \cdots, \lambda_p$. Ci-dessous, pour chaque i entre 1 et p, on a noté m_{λ_i} l'ordre de multiplicité algébrique de la valeur propre λ_i de u (multiplicité dans le polynôme caractéristique) et n_{λ_i} la dimension du sous-espace propre associé $\ker(u - \lambda_i \, id_E)$ (multiplicité géométrique), χ_u le polynôme caractéristique de u et μ_u le polynôme minimal de u.

$$
\begin{array}{rcl}
u \text{ est diagonalisable} & \Longleftrightarrow & \text{il existe une base de } E \text{ dans laquelle la matrice de } u \text{ est diagonale,} \\
& \Longleftrightarrow & \text{il existe une base de } E \text{ constituée de vecteurs propres de } u, \\
& \Longleftrightarrow & E \text{ est la somme directe des sous-espaces propres de } u, \\
& \Longleftrightarrow & n = \displaystyle\sum_{i=1}^{p} n_{\lambda_i}, \\
& \Longleftrightarrow & \chi_u \text{ est scindé sur } \mathbb{K} \text{ et, } \forall i \in \{1, \ldots, p\},\ n_{\lambda_i} = m_{\lambda_i}, \\
& \Longleftrightarrow & \mu_u \text{ est scindé sur } \mathbb{K}, \text{ à racines simples,} \\
& \Longleftrightarrow & \text{il existe un polynôme } P \text{ non nul, scindé sur } \mathbb{K}, \text{ à racines simples et tel que } P(u) = 0_{\mathscr{L}(E)}, \\
& \Longleftarrow & u \text{ possède } n \text{ valeurs propres simples } \Longleftrightarrow \chi_u \text{ est scindé sur } \mathbb{K}, \text{ à racines simples.}
\end{array}
$$

$$
\begin{array}{rcl}
u \text{ est trigonalisable} & \Longleftrightarrow & \text{il existe une base de } E \text{ dans laquelle la matrice de } u \text{ est triangulaire supérieure,} \\
& \Longleftrightarrow & \chi_u \text{ est scindé sur } \mathbb{K}, \\
& \Longleftrightarrow & E \text{ est la somme directe des sous-espaces caractéristiques de } u.
\end{array}
$$

1.10 Application de la réduction au calcul des puissances d'une matrice

L'objectif de cette sous-section est d'utiliser les techniques que nous avons vu dans ce chapitre pour simplifier le calcul des puissances d'une matrice $A \in M_n(\mathbb{K})$.

1.10.1 Grâce au polynôme minimal

Supposons que la matrice A ait un polynôme minimal de degré petit (c'est le cas si la dimension de l'espace ambiant est petite). Dans ce cas, pour $k \in \mathbb{N}$, effectuons la division euclidienne de X^k par μ_A

$$X^k = Q_k \mu_A + R_k$$

avec $\deg R_k < \deg \mu_A$. En évaluant en A, il vient alors que $A^k = R_k(A)$.

Exemple 1.70 *Soit la matrice*

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

Cherchons le polynôme minimal de A. On sait déjà (par le théorème de Cayley-Hamilton) que $\deg \mu_A \leq 3$. *Comme* (I_n, A) *sont libres il vient que* $\deg \mu_A \geq 2$. *On calcule alors* A^2 *et on remarque que*

$$A^2 = \begin{pmatrix} -3 & -8 & 0 \\ 2 & 5 & 0 \\ 2 & 4 & 1 \end{pmatrix} = 2A - I_2.$$

Donc $\mu_A(X) = X^2 - 2X + 1 = (X-1)^2$. *Soit* $k \in \mathbb{N}$, $k \geq 2$. *On effectue la division euclidienne de* X^k *par* μ_A. *On cherche alors* Q_k *et* $R_k(X) = a_k X + b_k$ *tels que*

$$X^k = Q_k(X)(X-1)^2 + a_k X + b_k.$$

Pour cela, on évalue en 1 *et on obtient que* $a_k + b_k = 1$. *De plus, en dérivant l'expression et en évaluant en* 1 *on obtient* $a_k = k$. *On résout facilement et il vient que* $R_k(X) = kX + (1-k)$. *D'où*

$$A^k = kA + (k-1)I_2 = \begin{pmatrix} 1-2k & -4k & 0 \\ k & 1+2k & 0 \\ k & 2k & 1 \end{pmatrix},$$

formule valable en fait pour tout $k \in \mathbb{N}$.

1.10.2 Dans le cas où A est diagonalisable

Supposons que la matrice A soit diagonalisable. Il existe alors une matrice inversible $P \in GL_n(\mathbb{K})$ et une matrice diagonale $D \in M_n(\mathbb{K})$ telles que $A = PDP^{-1}$. Ainsi pour tout $k \in \mathbb{N}$, $A^k = PD^kP^{-1}$ et D^k est simple à calculer. Cette méthode a cependant un écueil, il y a beaucoup de calculs : il faut calculer les sous-espaces propres de la matrice pour déterminer les matrices P et D, calculer ensuite P^{-1}, puis faire le calcul de PD^kP^{-1}.

Exemple 1.71 *Soit la matrice*

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

Soit $k \in \mathbb{N}$, *calculons* B^k. *On calcule d'abord le polynôme caractéristique de* B *et on obtient* $\chi_B(X) = (X+2)(X-3)^2$. *On cherche ensuite les sous-espaces propres de* B. *Après calculs on obtient*

$$E_{-2} = \mathrm{Vect}\begin{pmatrix} -4 \\ 1 \\ 0 \end{pmatrix} \text{ et } E_3 = \mathrm{Vect}\left\{\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right\}.$$

Donc B *est diagonalisable. On définit la matrice inversible* P *et la matrice diagonale associée* D *(en gardant le même ordre entre les valeurs propres et les vecteurs propres)*

$$P = \begin{pmatrix} -4 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ et } D = \begin{pmatrix} -2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

de sorte que $B = PDP^{-1}$. *Il nous reste à calculer* P^{-1} *et* D^k *et après calculs il vient que*

$$P^{-1} = \frac{1}{5}\begin{pmatrix} -1 & 1 & 0 \\ 1 & 4 & 0 \\ 0 & 0 & 5 \end{pmatrix} \text{ et } D^k = \begin{pmatrix} (-2)^k & 0 & 0 \\ 0 & 3^k & 0 \\ 0 & 0 & 3^k \end{pmatrix}.$$

Ainsi

$$B^k = PD^kP^{-1} = \frac{1}{5}\begin{pmatrix} 4\cdot(-2)^k + 3^k & 4\cdot(-2)^{k+1} + 4\cdot 3^k & 0 \\ (-2)^{k+1} + 3^k & (-2)^k + 4\cdot 3^k & 0 \\ 0 & 0 & 5\cdot 3^k \end{pmatrix}.$$

1.10.3 Dans le cas où A est trigonalisable

Dans le cas où A est trigonalisable, il existe une matrice inversible $P \in GL_n(\mathbb{K})$ et une matrice triangulaire supérieure $T \in M_n(\mathbb{K})$ telles que $A = PTP^{-1}$. Ainsi pour tout $k \in \mathbb{N}$, on a $A^k = PT^kP^{-1}$. Toute la difficulté réside dans le fait de trouver la bonne matrice T où les calculs des puissances seront plus faciles. Le recours à une décomposition en sous-espaces caractéristiques s'avère souvent utile. Nous ne donnons pas les détails. Voir les exercices 81, 87 et 88.

1.11 Application de la réduction aux suites récurrentes

1.11.1 Suites récurrentes d'ordre 1 à valeur vectorielle

Soit $n \in \mathbb{N}^*$, une matrice $A \in M_n(\mathbb{K})$ et $U_0 \in \mathbb{K}^n$ (écrit en colonne). Soit la suite $(U_k)_{k\in\mathbb{N}} \in (\mathbb{K}^n)^{\mathbb{N}}$ définie via la relation de récurrence suivante

$$U_{k+1} = AU_k.$$

Une récurrence évidente montre que pour tout $k \in \mathbb{N}$, on a

$$U_k = A^k U_0,$$

de sorte que l'on est ramené au calcul des puissances de A. On peut alors utiliser les techniques de la section précédente.

Exemple 1.72 *Soit $u_0, v_0, w_0 \in \mathbb{K}$. On définit les suites $(u_k)_{k\in\mathbb{N}}$, $(v_k)_{k\in\mathbb{N}}$, $(w_k)_{k\in\mathbb{N}}$ par la relation de récurrence suivante*

$$\forall k \in \mathbb{N}, \begin{cases} u_{k+1} = -u_k - 4v_k \\ v_{k+1} = u_k + 3v_k \\ w_{k+1} = u_k + 2v_k + w_k \end{cases}.$$

On commence par définir

$$U_k = \begin{pmatrix} u_k \\ v_k \\ w_k \end{pmatrix} \text{ et } A = \begin{pmatrix} -1 & -4 & 0 \\ 1 & 3 & 0 \\ 1 & 2 & 1 \end{pmatrix}$$

de sorte que $U_{k+1} = AU_k$. Il nous reste donc à calculer les puissances de A. Cette dernière correspond exactement à la matrice de l'exemple 1.70. Ainsi, pour tout $k \in \mathbb{N}$

$$\begin{cases} u_k = (1-2k)u_0 - 4kv_0 \\ v_k = ku_0 + (1+2k)v_0 \\ w_k = ku_0 + 2kv_0 + w_0 \end{cases}.$$

1.11.2 Suites récurrentes linéaires d'ordre p

Pour simplifier les choses, nous allons nous placer sur $\mathbb{C}$. On se donne un entier non nul p, des nombres complexes $a_0 \cdots, a_{p-1} \in \mathbb{C}$ avec $a_0 \neq 0$. Notons S l'ensemble des suites complexes $(u_n) \in \mathbb{C}^{\mathbb{N}}$ vérifiant $\forall n \in \mathbb{N}$

$$u_{n+p} = a_{p-1}u_{n+p-1} + \cdots + a_1 u_{n+1} + a_0 u_n.$$

On dit que $(u_n)_{n\in\mathbb{N}}$ vérifie une relation de récurrence d'ordre p. Le but de cette sous-section est déterminer l'ensemble solution S. On a le premier résultat suivant.

Proposition 1.73 *L'ensemble S est un sous-espace vectoriel de $\mathbb{C}^{\mathbb{N}}$. De plus, S est de dimension finie et* $\dim S = p$.

DÉMONSTRATION. Il est clair que S est stable par somme et par multiplication par un nombre complexe (la relation de récurrence étant linéaire). On introduit alors l'application

$$\begin{array}{cccc} \Phi & : & S & \to & \mathbb{C}^p \\ & & u & \mapsto & (u_0, u_1, \cdots, u_{p-1)} \end{array}.$$

On constate que Φ est une application linéaire. Au vu de la relation de récurrence satisfaite par un élément de S, Φ est bijective : la connaissance des p premiers termes d'une suite de S est nécessaire et suffisante pour déterminer entièrement la suite. D'où $\dim S = p$. □

Introduisons maintenant l'opérateur de décalage

$$\begin{array}{cccc} d & : & \mathbb{C}^{\mathbb{N}} & \to & \mathbb{C}^{\mathbb{N}} \\ & & (u_n)_{n\in\mathbb{N}} & \mapsto & (u_{n+1})_{n\in\mathbb{N}} \end{array}.$$

On a alors le résultat fondamental suivant.

Proposition 1.74 *Soit $P \in \mathbb{C}[X]$ le polynôme*

$$P(X) = X^p - \sum_{k=0}^{p-1} a_k X^k.$$

Alors $S = \ker P(d)$. De plus, si P se décompose en

$$P(X) = \prod_{k=1}^{r} (X - \lambda_k)^{m_k}$$

où les λ_k sont deux à deux distincts, on a

$$S = \oplus_{k=1}^{r} \ker (d - \lambda_k id)^{m_k}$$

où $\dim \ker (d - \lambda_k id)^{m_k} = m_k$.

Remarque 1.75 *Comme $a_0 \neq 0$, on a $\lambda_k \neq 0$ pour tout k.*

Définition 1.76 *On appelle équation caractéristique de la relation de récurrence d'ordre p, l'équation $P(\lambda) = 0$.*

DÉMONSTRATION. La première partie de la preuve repose sur le fait que pour tout $k \in \mathbb{N}$, $d^k((u_n)_{n\in\mathbb{N}}) = (u_{n+k})_{n\in\mathbb{N}}$. Pour la deuxième partie, on utilise le lemme de décomposition des noyaux et on remarque que $\ker (d - \lambda_k id)^{m_k}$ est le sous-espace caractéristique de l'endomorphisme induit par d sur $\ker P(d)$. □

Soit λ une racine du polynôme P de multiplicité m. On sait que $\lambda \neq 0$ car $a_0 \neq 0$. Déterminons $\ker (d - \lambda id)^m$. On a déjà

$$\ker (d - \lambda id) \subset \ker (d - \lambda id)^m$$

et comme $\ker (d - \lambda id)$ correspond à l'ensemble des suites géométrique de raison λ, nous avons

$$(\lambda^n)_{n\in\mathbb{N}} \in \ker (d - \lambda id)^m.$$

Mais nous pouvons aller plus loin. Commençons par un lemme technique qui correspond au cas où $\lambda = 1$.

Lemme 1.77 *Soit $k \in \mathbb{N}^*$. On a*

$$\ker (d - id)^k = \{(P(n))_{n\in\mathbb{N}} \,,\, P \in \mathbb{C}_{k-1}[X]\}.$$

DÉMONSTRATION. On constate que pour $(u_n)_{n\in\mathbb{N}} \in \mathbb{C}^{\mathbb{N}}$, comme d et id commutent, on a

$$(d - id)^k((u_n)_{n\in\mathbb{N}}) = \sum_{l=0}^{k} \binom{k}{l} (-1)^{n-k} d^k((u_n)_{n\in\mathbb{N}}) = \left(\sum_{l=0}^{k} \binom{k}{l} (-1)^{k-l} u_{n+l} \right)_{n\in\mathbb{N}}.$$

Le sous-espace vectoriel $\ker (d - id)^k$ correspond donc à un ensemble solution d'une suite récurrence d'ordre k, donc par la proposition 1.73, $\dim \ker (d - id)^k = k$. Soit $P \in \mathbb{C}_{k-1}[X]$, $P(X) = \sum_{l=0}^{\deg(P)} b_l X^l$. Considérons la suite $(P(n))_{n\in\mathbb{N}}$. On remarque que

$$(d - id)((P(n))_{n\in\mathbb{N}}) = \left(\sum_{l=0}^{\deg(P)} b_l ((n+1)^l - n^l) \right)_{n\in\mathbb{N}} = \left(\sum_{l=0}^{\deg(P)} \sum_{p=0}^{l-1} b_l \binom{l}{p} n^p \right)_{n\in\mathbb{N}} = (Q(n))_{n\in\mathbb{N}}$$

où Q est un polynôme tel que $\deg Q \leq \deg P - 1$. Ainsi, comme $\deg P < k$, on en déduit que

$$(d - id)^k((P(n))_{n\in\mathbb{N}} = (0)_{n\in\mathbb{N}}.$$

Cela implique que

$$\{(P(n))_{n\in\mathbb{N}} \,,\, P \in \mathbb{C}_{k-1}[X]\} \subset \ker (d - id)^k.$$

L'égalité de ces deux ensembles suit car ils ont la même dimension. □

On peut maintenant traiter le cas général.

Proposition 1.78 *On a*

$$\ker(d-\lambda id)^m = \{(P(n)\lambda^n)_{n\in\mathbb{N}} \in \mathbb{C}^{\mathbb{N}}\ ,\ P \in \mathbb{C}_{m-1}[X]\}.$$

Pour prouver ce résultat nous allons utiliser une méthode qui s'appelle **la variation de la constante**. On se dit qu'une suite qui est dans $\ker(d-\lambda id)^m$ doit avoir une forme similaire à une suite de $\ker(d-\lambda id)$. On cherche donc une solution u de la forme $(\lambda^n)_{n\in\mathbb{N}} \cdot v$ avec v à déterminer.

DÉMONSTRATION. Soit $u \in \ker(d-\lambda id)^m$. Définissons la suite $v = (v_n)_{n\in\mathbb{N}} = \left(\frac{u_n}{\lambda^n}\right)_{n\in\mathbb{N}}$ de sorte que pour tout n on a $u_n = v_n\lambda^n$. On a alors

$$\begin{aligned}
(d-\lambda id)^m u &= \left(\sum_{l=0}^{m}\binom{m}{l}(-\lambda)^{m-l}u_{n+l}\right)_{n\in\mathbb{N}} = \left(\sum_{l=0}^{m}\binom{m}{l}(-\lambda)^{m-l}(\lambda)^{n+l}v_{n+l}\right)_{n\in\mathbb{N}} \\
&= \left(\lambda^{n+m}\sum_{l=0}^{m}\binom{m}{l}(-1)^{m-l}v_{n+l}\right)_{n\in\mathbb{N}} = (\lambda^{m+n})_{n\in\mathbb{N}}\cdot\left(\sum_{l=0}^{m}\binom{m}{l}(-1)^{m-l}v_{n+l}\right)_{n\in\mathbb{N}} \\
&= (\lambda^{m+n})_{n\in\mathbb{N}}\cdot(d-id)^m v,
\end{aligned}$$

de sorte que, comme $(d-\lambda id)^m u = 0$ et que $\lambda \neq 0$, on a $(d-id)^m v = 0$. Par le lemme 1.77, il existe un polynôme $P \in \mathbb{C}_{m-1}[X]$ tel que $v = (P(n))_{n\in\mathbb{N}}$. Ainsi,

$$\ker(d-\lambda id)^m \subset \{(P(n)\lambda^n)_{n\in\mathbb{N}} \in \mathbb{C}^{\mathbb{N}}\ ,\ P \in \mathbb{C}_{m-1}[X]\}$$

et l'égalité de ces deux ensembles suit car ils ont la même dimension. □

Nous pouvons maintenant résumer cette sous-section avec le théorème suivant.

Théorème 1.79 *Soit $p \in \mathbb{N}^*$, $a_0 \cdots, a_{p-1} \in \mathbb{C}$ avec $a_0 \neq 0$. Notons S l'ensemble des suites complexes $(u_n) \in \mathbb{C}^{\mathbb{N}}$ vérifiant la relation de récurrence d'ordre p suivante : $\forall n \in \mathbb{N}$*

$$u_{n+p} = a_{p-1}u_{n+p-1} + \cdots + a_1u_{n+1} + a_0u_n.$$

Soit $P(X) = X^p - \sum_{k=0}^{p-1} a_kX^k$ le polynôme associé à la relation de récurrence de décomposition en facteurs irréductibles

$$P(X) = \prod_{k=1}^{r}(X-\lambda_k)^{m_k}$$

où les λ_k sont deux à deux distincts. Alors

$$S = \left\{\left(\sum_{k=0}^{r}P_k(n)\lambda^k\right)_{n\in\mathbb{N}}\ ,\ P_k \in \mathbb{C}_{m_k-1}[X]\right\}.$$

La principale difficulté pour appliquer ce théorème est d'arriver à obtenir les racines avec multiplicités du polynôme associé. Finissons cette sous-section en adaptant le théorème dans le cas où $p = 2$.

Théorème 1.80 *Soit $a_0, a_1 \in \mathbb{C}$ avec $a_0 \neq 0$. Soit une suite $(u_n) \in \mathbb{C}^{\mathbb{N}}$ vérifiant la relation de récurrence d'ordre 2 suivante : $\forall n \in \mathbb{N}$*

$$u_{n+2} = a_1u_{n+1} + a_0u_n.$$

Soit l'équation caractéristique associée $X^2 - a_1X - a_0 = 0$. On note Δ son discriminant. Plusieurs cas se présentent.

1. *Si $\Delta \neq 0$, l'équation caractéristique admet deux racines distinctes λ et μ et, $\exists a, b \in \mathbb{C}$, $\forall n \in \mathbb{N}$*

$$u_n = a\lambda^n + b\mu^n.$$

2. *Si $\Delta = 0$, l'équation caractéristique admet une racine double λ et $\exists a, b \in \mathbb{C}$, $\forall n \in \mathbb{N}$*

$$u_n = (a+bn)\lambda^n.$$

Remarque 1.81 *Si les coefficients a_0 et a_1 sont réels on peut facilement adapter le résultat précédent. Si $(u_n) \in \mathbb{R}^{\mathbb{N}}$ est une suite réelle vérifiant la relation de récurrence d'ordre 2. On a les alternatives suivantes.*

1. *Si $\Delta > 0$, l'équation caractéristique admet deux racines réelles distinctes λ et μ et, $\exists a, b \in \mathbb{R}$, $\forall n \in \mathbb{N}$*
$$u_n = a\lambda^n + b\mu^n.$$
2. *Si $\Delta = 0$, l'équation caractéristique admet une racine réelle double λ et $\exists a, b \in \mathbb{R}$, $\forall n \in \mathbb{N}$*
$$u_n = (a + bn)\,\lambda^n.$$
3. *Si $a_0, a_1 \in \mathbb{R}$, et $\Delta < 0$, l'équation caractéristique admet deux racines complexes conjugués $re^{i\theta}$ et $re^{-i\theta}$ et $\exists a, b \in \mathbb{R}$, $\forall n \in \mathbb{N}$*
$$u_n = (a\cos(n\theta) + b\sin(n\theta))\,r^n.$$

Pour montrer le dernier point, on remarque que l'ensemble des solutions réelles est un espace vectoriel réelle de dimension 2. Ainsi comme $(r^n e^{in\theta})_{n\in\mathbb{N}}$ est une solution complexe, sa partie réelle et sa partie imaginaire sont des solutions réelles de la relation de récurrence. Il suffit alors de remarquer qu'ils forment une famille libre de l'ensemble des solutions réelles.

Chapitre 2

Formes bilinéaires réelles

Une *application bilinéaire* est l'analogue à deux variables d'une application linéaire. Dans ce chapitre, nous n'étudierons que les formes bilinéaires réelles, mais on peut étendre la notion à des formes bilinéaires complexes.

2.1 Généralités sur les formes bilinéaires réelles

Étant donné un $\mathbb{R}$-espace vectoriel E, une application b de $E \times E$ dans $\mathbb{R}$ est appelée *forme bilinéaire réelle* si

$$\begin{aligned}&\forall (x,x') \in E^2,\ \forall (y,y') \in E^2,\ \forall \lambda \in \mathbb{R},\\ &\qquad b(x+x',y) = b(x,y)+b(x',y),\ b(x,y+y') = b(x,y)+b(x,y'),\ b(\lambda x,y) = b(x,\lambda y) = \lambda\, b(x,y).\end{aligned}$$

On peut exprimer cette définition de manière un peu plus formelle en disant que l'application b est *linéaire par rapport à chacune de ses variables*. Ceci peut être rendu précis de la façon suivante.

Définition 2.1 *Soit E un $\mathbb{R}$-espace vectoriel. Une application b de $E \times E$ dans $\mathbb{R}$ est appelée* ***forme bilinéaire réelle*** *si l'****application partielle à gauche de*** *b, notée $L(b)$, de E dans $\mathscr{L}(E;\mathbb{R})$ définie par*

$$\forall x \in E,\ L(b)(x) = (y \mapsto b(x,y)),$$

*et l'****application partielle à droite de*** *b, notée $R(b)$, de E dans $\mathscr{L}(E;\mathbb{R})$ définie par*

$$\forall y \in E,\ R(b)(y) = (x \mapsto b(x,y))$$

sont toutes deux linéaires.

L'ensemble des formes bilinéaires réelles sur $E \times E$, noté $\mathscr{L}(E,E;\mathbb{R})$, est un $\mathbb{R}$-espace vectoriel. Il est en effet clair que la somme de deux formes bilinéaires réelles est une forme bilinéaire réelle et que le produit d'une forme bilinéaire réelle par un réel est une forme bilinéaire réelle.

Remarque 2.2 *Dans la précédente définition, on a en fait défini une application $L : \mathscr{L}(E,E;\mathbb{R}) \to \mathscr{L}(E;\mathscr{L}(E;\mathbb{R}))$ (resp. $R : \mathscr{L}(E,E;\mathbb{R}) \to \mathscr{L}(E;\mathscr{L}(E;\mathbb{R}))$). Il est facile de vérifier qu'elle est linéaire et que c'est un isomorphisme (dit* ***canonique****).*

On note qu'on a ici linéarité à trois niveaux différents, puisque les applications L, $L(b)$ et, pour tout vecteur x de E, $L(b)(x)$ (resp. R, $R(b)$ et, pour tout vecteur y de E, $R(b)(y)$) sont toutes trois linéaires.

Définition 2.3 (forme bilinéaire réelle symétrique) *Soit E un $\mathbb{R}$-espace vectoriel et b une forme bilinéaire réelle sur $E \times E$. On dit que l'application b est* ***symétrique*** *(resp.* ***antisymétrique****) si et seulement si*

$$\forall (x,y) \in E^2,\ b(x,y) = b(y,x) \text{ (resp. } b(x,y) = -b(y,x)\text{)}.$$

Exemple 2.4 *Presque tout ce qui porte le nom de* produit *est bilinéaire. Ainsi, le produit dans $\mathbb{R}$,*

$$\begin{array}{ccc}\mathbb{R}\times\mathbb{R} & \to & \mathbb{R}\\ (x,y) & \mapsto & xy,\end{array}$$

est une forme bilinéaire réelle symétrique. Ceci résulte de l'associativité et de la commutativité de la multiplication et de la distributivité de la multiplication par rapport à l'addition.

Sur $\mathbb{R}^n$, le "produit scalaire" usuel,

$$\begin{array}{ccc} \mathbb{R}^n \times \mathbb{R}^n & \to & \mathbb{R} \\ (x,y) & \mapsto & x \cdot y, \end{array}$$

défini par

$$x \cdot y = \sum_{i=1}^{n} x_i y_i,$$

où $x = (x_1, \ldots, x_n)$ et $y = (y_1, \ldots, y_n)$, est une forme bilinéaire réelle symétrique.

L'application

$$\begin{array}{ccc} \mathbb{R}^2 \times \mathbb{R}^2 & \to & \mathbb{R} \\ (x,y) & \mapsto & x_1 y_2 + 2x_2 y_1 - x_1 y_1, \end{array}$$

est une forme bilinéaire réelle qui n'est pas symétrique car $b((1,0),(0,1)) = 1$ mais $b((0,1),(1,0)) = 2$.

L'application

$$\begin{array}{ccc} \mathbb{R}^2 \times \mathbb{R}^2 & \to & \mathbb{R} \\ (x,y) & \mapsto & x_1 + y_2 \end{array}$$

n'est une pas forme bilinéaire réelle car l'application qui à $x \in \mathbb{R}^2 \mapsto b(x,(0,1)) = x_1 + 1$ n'est pas linéaire.

2.2 Représentation matricielle d'une forme bilinéaire

Définition 2.5 (matrice d'une forme bilinéaire) *Soit E un espace vectoriel de dimensions finie non nulle n, $\mathscr{B} = \{e_1, \ldots, e_n\}$ une base de E et b une forme bilinéaire réelle sur $E \times E$. La* ***matrice de b relativement à la base $\mathscr{B}$*** *est la matrice dont le coefficient à la i^e ligne et la j^e colonne est égal à $b(e_i, e_j)$.*

Une conséquence de la précédente définition est que, si les matrices colonnes X de $M_{n,1}(\mathbb{R})$ et Y de $M_{n,1}(\mathbb{R})$ contiennent les coordonnées respectives des vecteurs x de E dans la base $\mathscr{B}$ et y de E dans la base $\mathscr{B}$ et que la matrice M de $M_n(\mathbb{R})$ représente la forme bilinéaire b relativement à $\mathscr{B}$, on a

$$X^\top M Y = b(x,y),$$

en identifiant une matrice scalaire au scalaire qu'elle contient. Cette dernière formule montre comment calculer l'image d'un couple de vecteurs par une forme bilinéaire lorsque l'on dispose de la matrice de cette dernière relativement à des bases données. Elle montre en particulier qu'une forme bilinéaire b sur $\mathbb{R}^n \times \mathbb{R}^n$ s'écrit sous la forme

$$\forall (x,y) \in \mathbb{R}^n \times \mathbb{R}^n,\ b(x,y) = \sum_{i=1}^{n} \sum_{j=1}^{n} m_{ij} x_i y_j,$$

c'est-à-dire comme une combinaison linéaire de monômes de degré un en x et y dont les coefficients sont ceux de la matrice de b relativement à la base canonique de $\mathbb{R}^n$. Dans ce cas, la matrice de b relativement à la base canonique est la matrice $M = (m_{ij})$.

Exemple 2.6 *Considérons la forme bilinéaire réelle*

$$\begin{array}{ccc} b : \mathbb{R}^2 \times \mathbb{R}^2 & \to & \mathbb{R} \\ (x,y) & \mapsto & x_1 y_2 + 2x_2 y_1 - x_1 y_1. \end{array}$$

La matrice de b relativement à la base canonique de $\mathbb{R}^2$ est

$$Mat_{\mathscr{C}^2}(b) = \begin{pmatrix} -1 & 1 \\ 2 & 0 \end{pmatrix}.$$

Proposition 2.7 *Soit E un espace vectoriel de dimension finie non nulle, $\mathscr{B}$ une base de E et b une forme bilinéaire réelle sur $E \times E$. La matrice de b relativement à la base $\mathscr{B}$ est symétrique (resp. antisymétrique) si et seulement si la forme b est symétrique (resp. antisymétrique).*

Changement de base

Considérons une forme bilinéaire réelle sur $E\times E$ et des bases $\mathscr{B}$ et $\mathscr{B}'$ de E. On note P la matrice de passage de $\mathscr{B}$ à $\mathscr{B}'$, c'est-à-dire qu'on a $X = PX'$ avec $X = \mathrm{mat}_{\mathscr{B}}(x)$ et $X' = \mathrm{mat}_{\mathscr{B}'}(x)$ pour tout vecteur x de E. Comparons la matrice de b relativement aux bases $\mathscr{B}$ d'une part avec celle relativement aux bases $\mathscr{B}'$ d'autre part. On a

$$(X')^\top M'Y' = b(x,y) = X^\top MY = (PX')^\top M(PY') = (X')^\top P^\top MPY',$$

d'où, par identification,

$$M' = P^\top MP \tag{2.1}$$

Remarque 2.8 *On notera que la matrice transposée $P^\top$ est généralement différente de l'inverse P^{-1}. Par conséquent, les matrices carrées M et M' ne sont généralement pas semblables (on dit qu'elles sont* congrues*). Ceci signifie en particulier que si la matrice d'un endomorphisme de E dans une base donnée est la même que celle d'une forme bilinéaire sur $E\times E$ relativement à cette base, il n'en sera pas nécessairement de même après changement de base.*

Restriction à un sous-espace vectoriel

Soit une forme bilinéaire réelle b sur $E\times E$. Soit F un sous-espace vectoriel de E. On peut définir la restriction de b à F comme étant

$$b_{|F} : \begin{array}{ccc} F\times F & \to & \mathbb{R} \\ (x,y) & \mapsto & b(x,y). \end{array}$$

Ainsi, contrairement à un endomorphisme, on peut restreindre une forme bilinéaire réelle à n'importe quel sous-espace vectoriel.

Exemple 2.9 *Considérons la forme bilinéaire réelle*

$$b : \begin{array}{ccc} \mathbb{R}^2\times\mathbb{R}^2 & \to & \mathbb{R} \\ (x,y) & \mapsto & x_1y_2 + 2x_2y_1 - x_1y_1. \end{array}$$

Considérons la restriction de b au sous-espace vectoriel $F = \mathrm{Vect}\begin{pmatrix}1\\3\end{pmatrix}$. On décompose $x\in F$ et $y\in F$ dans la base $\left\{\begin{pmatrix}1\\3\end{pmatrix}\right\}$ de F. On note $x = u\begin{pmatrix}1\\3\end{pmatrix}$ et $y = v\begin{pmatrix}1\\3\end{pmatrix}$ où $u,v\in\mathbb{R}$. On a alors

$$b(x,y) = u\times 3v + 2\times 3u\times v - u\times v = 8uv.$$

La matrice de $b_{|F}$ relativement à la base $\left\{\begin{pmatrix}1\\3\end{pmatrix}\right\}$ est

$$Mat_{\left\{\begin{pmatrix}1\\3\end{pmatrix}\right\}}(b_{|F}) = \begin{pmatrix}8\end{pmatrix}.$$

2.3 Non dégénérescence d'une forme bilinéaire réelle symétrique

Proposition 2.10 *Soit E est un $\mathbb{R}$-espace vectoriel de dimension finie et b est une forme bilinéaire réelle sur $E\times E$. Soit $\mathscr{B}$ une base quelconque de E. Notons M la matrice de b relativement à la base $\mathscr{B}$. Alors le signe du déterminant de M est indépendant de la base.*

DÉMONSTRATION. Soit $\mathscr{B}'$ une autre base de E et M' la matrice de b relativement à cette base. Si on note P la matrice de passage de $\mathscr{B}$ à $\mathscr{B}'$, on a vu précédemment que $M' = P^\top MP$. En passant au déterminant, on obtient que $\det(M') = \det(P^\top)\det(M)\det(P) = \det(M)\det(P)^2$. D'où le résultat car $\det(P)^2 > 0$. □

Cette proposition motive la définition suivante.

Définition 2.11 (non dégénérescence) *Soit E est un $\mathbb{R}$-espace vectoriel de dimension finie. Une forme bilinéaire réelle symétrique b sur $E\times E$ est dite* ***non dégénérée*** *si et seulement si la matrice associée à b relativement à une base quelconque de E est inversible. Si cette même matrice n'est pas inversible, on dit que b sur $E\times E$ est* ***dégénérée****.*

Remarque 2.12 *Une autre définition de la non dégénérescence est possible. On dit que b sur $E \times E$ est **non dégénérée** si*

$$\{y \in E\ ,\ b(x,y) = 0_E\ \forall x \in E\} = \{0_E\}.$$

Remarque 2.13 *La restriction à un sous-espace vectoriel d'une forme bilinéaire non dégénérée peut être dégénérée. Par exemple la forme définie sur $\mathbb{R}^2 \times \mathbb{R}^2$ par*

$$b(x,y) = x_1y_1 - x_2y_2,$$

de matrice relativement à la base canonique $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, est non dégénérée (car matrice de déterminant non nul). Pourtant, la restriction de b à la droite vectorielle engendrée par $e_1 + e_2$ est dégénérée, car identiquement nulle.

De manière similaire, on peut aussi définir le rang et le noyau d'une forme bilinéaire symétrique.

Définition 2.14 (rang et noyau d'une forme bilinéaire symétrique) *Soit E un $\mathbb{R}$-espace vectoriel de dimension finie et b une forme bilinéaire symétrique sur $E \times E$. On appelle **rang** de b, noté $Rang(b)$, le rang de la matrice associée à b relativement à une base quelconque de E et **noyau** de la forme bilinéaire b, noté* $\ker(b)$*, est*

$$\ker(b) = \{x \in E\ ,\ b(x,y) = 0\ \forall y \in E\}.$$

Concrètement, le rang d'une forme bilinéaire est le rang de la matrice de b relativement à une base quelconque de E. Le noyau lui est relié au noyau d'une matrice de b relativement à une base quelconque de E.

Exemple 2.15 *Considérons la forme bilinéaire réelle*

$$b : \begin{array}{ccc} \mathbb{R}^3 \times \mathbb{R}^3 & \to & \mathbb{R} \\ (x,y) & \mapsto & x_1y_1 - x_2y_1 - y_2x_1 + x_2y_2. \end{array}$$

b est bien symétrique. Sa matrice relativement à la base canonique est

$$Mat_{\mathscr{C}^3}(b) = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Alors $Rang(b) = 1$, $\ker(b) = \mathrm{Vect}\left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$ et b est dégénérée.

Proposition 2.16 *Soit E est un $\mathbb{R}$-espace vectoriel de dimension finie. Soit b une forme bilinéaire réelle symétrique sur $E \times E$. b est non dégénérée si seulement si* $\ker(b) = \{0_E\}$.

Chapitre 3

Formes quadratiques réelles

On aborde dans ce chapitre la notion de *forme quadratique réelle*. Celle-ci intervient dans de nombreux domaines des mathématiques et de la physique. Elle est notamment à la base de la géométrie euclidienne et de sa généralisation dans les espaces de Hilbert. L'étude arithmétique des formes quadratiques a aussi été le point de départ de la théorie des nombres algébriques.

3.1 Généralités

Définition 3.1 (forme quadratique réelle) *Soit E un $\mathbb{R}$-espace vectoriel. Une application q de E dans $\mathbb{R}$ est appelée* ***forme quadratique réelle*** *sur E s'il existe une forme bilinéaire réelle b de $E \times E$ dans $\mathbb{R}$ telle que*

$$\forall x \in E,\ q(x) = b(x, x).$$

Toute forme bilinéaire réelle b sur $E \times E$ donne naissance à une forme quadratique réelle sur E en posant

$$\forall x \in E,\ q(x) = b(x, x).$$

On appelle alors q la *forme quadratique associée à b*. Une même forme quadratique réelle peut ainsi être associée à plusieurs formes bilinéaires réelle. On a toutefois le résultat suivant.

Proposition et définition 3.2 (forme polaire d'une forme quadratique) *Sur un $\mathbb{R}$-espace vectoriel, toute forme quadratique réelle est associée à une unique forme bilinéaire réelle symétrique, cette dernière étant appelée la* ***forme polaire*** *de la forme quadratique.*

DÉMONSTRATION. Soit q une forme quadratique réelle sur un $\mathbb{R}$-espace vectoriel E. Par définition, il existe une forme bilinéaire b, non nécessairement symétrique, telle que

$$\forall x \in E,\ b(x, x) = q(x).$$

Symétrisons la forme b en introduisant la forme bilinéaire c définie par

$$\forall (x, y)x \in E^2,\ c(x, y) = \frac{1}{2}\left(b(x, y) + b(y, x)\right).$$

On a alors

$$\forall x \in E,\ c(x, x) = b(x, x),$$

et q est donc aussi la forme quadratique associée à c. Par ailleurs, la forme c étant symétrique, on a

$$\begin{aligned}\forall (x, y) \in E^2,\ q(x + y) - q(x - y) &= c(x + y, x + y) - c(x - y, x - y)\\ &= c(x, x) + 2\,c(x, y) + c(y, y) - c(x, x) + 2\,c(x, y) - c(y, y)\\ &= 4\,c(x, y),\end{aligned}$$

ce qui donne

$$\forall (x,y)\in E^2,\ c(x,y)=\frac{1}{4}\left(q(x+y)-q(x-y)\right). \tag{3.1}$$

La forme c est donc définie de manière unique. □

La formule (3.1) obtenue dans la démonstration ci-dessus est appelée *identité de polarisation*. On notera que l'on a également

$$\begin{aligned}\forall (x,y)\in E^2,\ q(x+y)-q(x)-q(y)&=c(x+y,x+y)-c(x,x)-c(y,y)\\&=c(x,x)+2\,c(x,y)+c(y,y)-c(x,x)-c(y,y)\\&=2\,c(x,y),\end{aligned}$$

d'où la formule

$$\forall (x,y)\in E^2,\ c(x,y)=\frac{1}{2}\left(q(x+y)-q(x)-q(y)\right),$$

qui constitue une deuxième identité de polarisation. Une troisième identité de polarisation est donnée par

$$\forall (x,y)\in E^2,\ c(x,y)=\frac{1}{2}\left(q(x)+q(y)-q(x-y)\right).$$

En pratique, si l'on exhibe une forme bilinéaire symétrique ayant une forme quadratique q associée, c'est nécessairement la forme polaire de q. On n'a donc pas toujours besoin d'utiliser les identités de polarisation.

Un *polynôme homogène de degré deux* sur $\mathbb{R}$ en les variables $x_1,\dots,x_n$ est une expression de la forme

$$\sum_{i=1}^{n}\alpha_{ii}\,{x_i}^2+\sum_{i=1}^{n}\sum_{j=i+1}^{n}\alpha_{ij}\,x_ix_j,$$

où les coefficients α_{ij}, $1\leq i\leq j\leq n$, appartiennent à $\mathbb{R}$.

Proposition 3.3 *Soit E un $\mathbb{R}$-espace vectoriel de dimension finie non nulle et $\mathscr{B}=\{e_1,\dots,e_n\}$ une base de E. Les formes quadratiques réelles sur E sont les applications q telles que, pour tout vecteur x de E, $q(x)$ est un polynôme homogène de degré deux en les coordonnées de x dans la base $\mathscr{B}$.*

DÉMONSTRATION. Soit b la forme polaire de q. Posons

$$\forall (i,j)\in\{1,\dots,n\}^2,\ a_{ij}=b(e_i,e_j)$$

et

$$\forall x\in E,\ x=\sum_{i=1}^{n}x_ie_i.$$

On a alors

$$\begin{aligned}\forall x\in E,\ q(x)&=b\left(\sum_{i=1}^{n}x_ie_i,\sum_{j=1}^{n}x_je_j\right)\\&=\sum_{i=1}^{n}\sum_{j=1}^{n}x_ix_j\,b(e_i,e_j)\\&=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}\,x_ix_j\\&=\sum_{i=1}^{n}a_{ii}\,{x_i}^2+2\sum_{i=1}^{n}\sum_{j=i+1}^{n}a_{ij}\,x_ix_j,\end{aligned}$$

car

$$\forall (i,j)\in\{1,\dots,n\}^2,\ a_{ij}=a_{ji}.$$

□

Exemple 3.4 *L'application*

$$q : \begin{array}{ccc} \mathbb{R}^3 & \to & \mathbb{R} \\ x & \mapsto & x_1^2 - x_2^2 + 3x_1x_2 \end{array}$$

est une forme quadratique réelle de forme polaire

$$b : \begin{array}{ccc} \mathbb{R}^3 \times \mathbb{R}^3 & \to & \mathbb{R} \\ (x, y) & \mapsto & x_1y_1 - x_2y_2 + \frac{3}{2}x_1y_2 + \frac{3}{2}y_1x_2. \end{array}$$

L'application

$$\tilde{q} : \begin{array}{ccc} \mathbb{R}^2 & \to & \mathbb{R} \\ x & \mapsto & x_1^2 - x_2 \end{array}$$

n'est pas une forme quadratique réelle car q n'est pas un polynôme homogène de degré 2 car on a par exemple $q(0,3) = -3 \neq -3^2 = 3^2 q(0,1)$.

Remarque 3.5 *Une forme quadratique étant donnée par un polynôme homogène de degré deux, sa forme polaire peut être déterminée en* polarisant *chaque monôme de ce polynôme. Ainsi, un monôme de la forme $a\,{x_i}^2$ est polarisé en $a\,x_i y_i$, tandis qu'un monôme de la forme $a\,x_i x_j$ est polarisé en $\frac{a}{2}\left(x_i y_j + x_j y_i\right)$.*

Définitions 3.6 (matrice, noyau et rang d'une forme quadratique) *Soit q une forme quadratique réelle sur un $\mathbb{R}$-espace vectoriel, supposé de dimension finie. La matrice de q est celle de sa forme polaire. Le noyau et le rang de q sont aussi définis comme étant ceux de sa forme polaire.*

Il découle de ces définitions que l'on peut exprimer matriciellement cette dernière en utilisant ce qui a été vu dans le précédent chapitre. En notant M la matrice d'une forme quadratique réelle q de A relativement à une base $\mathscr{B}$ de E et en notant X la matrice colonne représentant un élément x de E dans la base $\mathscr{B}$, on a

$$q(x) = X^T M X.$$

Une autre conséquence de cette définition est que l'espace vectoriel des formes quadratiques réelles sur E, noté $\mathscr{Q}(E)$, est isomorphe à $S_n(\mathbb{R})$, l'ensemble des matrices symétriques à coefficients dans $\mathbb{R}$. En particulier, si l'espace E est de dimension n, alors l'espace $\mathscr{Q}(E)$ est de dimension $\frac{1}{2}n(n+1)$.

Définition 3.7 (forme quadratique non dégénérée) *Soit E un $\mathbb{R}$-espace vectoriel de dimension finie. Une forme quadratique réelle sur E est dite* ***non dégénérée*** *si et seulement si son noyau est réduit au vecteur nul (ou son rang est égal à la dimension de l'espace E).*

Définition 3.8 (formes quadratiques équivalentes) *Deux formes quadratiques q et q' sur un $\mathbb{R}$-espace vectoriel E sont dites* ***équivalentes*** *s'il existe un automorphisme u de E tel que*

$$\forall x \in E,\ q'(x) = q(u(x)).$$

En dimension finie, l'équivalence entre formes quadratiques se traduit matriciellement par le fait que les matrices des formes relativement à une même base sont congrues.

3.2 Orthogonalité

Soit E un $\mathbb{R}$-espace vectoriel, q une forme quadratique réelle sur E et b la forme polaire de q.

Définition 3.9 (cône isotrope) *Un vecteur x de E est dit* ***isotrope*** *pour q si $q(x) = 0$. On appelle* ***cône isotrope*** *de q l'ensemble $\mathscr{C}_q$ des vecteurs isotropes pour q.*

Définition 3.10 (forme quadratique réelle définie) *On dit que q est* ***définie*** *si $\mathscr{C}_q = \{0_E\}$. Autrement dit, si $x \in E$ vérifie $q(x) = 0$, alors $x = 0_E$.*

Définition 3.11 (forme bilinéaire réelle alternée) *Une forme bilinéaire réelle b sur $E \times E$ est dite* alternée *si et seulement si tout vecteur de E est isotrope pour sa forme quadratique associée, c'est-à-dire si*

$$\forall x \in E,\ b(x,x) = 0.$$

Remarque 3.12 *On a toujours l'inclusion* $\ker(q) \subset \mathscr{C}_q$*, mais le cône isotrope* $\mathscr{C}_q$ *n'est généralement pas un sous-espace vectoriel de* E*. Il en découle qu'une forme quadratique définie est non dégénérée, mais que la réciproque est généralement fausse.*

Exemple 3.13 *Prenons* $E = \mathbb{R}^2$*.*

La forme quadratique réelle q_1 *vérifiant, pour tout* $x \in E$*,* $q_1(x) = x_1 x_2$ *satisfait* $\ker(q_1) = \{0_E\}$ *mais* $\mathscr{C}_{q_1} = \{x_1 = 0\} \cup \{x_2 = 0\}$ *n'est pas un sous-espace vectoriel de* E*.*

La forme quadratique réelle q_2 *vérifiant, pour tout* $x \in E$*,* $q_2(x) = {x_1}^2 - {x_2}^2$ *est non dégénérée (*$\ker(q_2) = \{0_E\}$*) mais pas définie (*$\mathscr{C}_{q_2} = \{x_1 = x_2\} \cup \{x_1 = -x_2\}$*).*

La forme quadratique réelle q_3 *vérifiant, pour tout* $x \in E$*,* $q_3(x) = {x_1}^2 + 3{x_2}^2$ *est non dégénérée et définie.*

Définition 3.14 (orthogonalité) *On dit que deux vecteurs* x *et* y *de* E *sont* ***orthogonaux*** *pour la forme* q *(ou pour sa forme polaire* b*), ou simplement* q*-orthogonaux, si et seulement si* $b(x, y) = 0$*, ce que l'on note* $x \perp y$*.*

Définition 3.15 (base q-orthogonale) *Une base* $\mathscr{B}$ *de* E *est dite* ***q-orthogonale*** *si ses éléments sont deux à deux orthogonaux pour* q*.*

Si E est de dimension finie non nulle égale à n et que $\mathscr{B} = \{\varepsilon_1, \dots, \varepsilon_n\}$ est une base q-orthogonale de E, alors on a

$$q\left(\sum_{i=1}^{n} x_i\,\varepsilon_i\right) = \sum_{i=1}^{n} {x_i}^2\, q(\varepsilon_i).$$

Autrement dit, la matrice de q relativement à la base $\mathscr{B}$ est *diagonale*. Avant de montrer l'existence d'une base q-orthogonale, nous avons besoin d'un peu de vocabulaire.

Définition 3.16 (Dual) *Si* E *est un* $\mathbb{R}$*-espace vectoriel de dimension finie, on note* $E^* = \mathscr{L}(E, \mathbb{R})$ *appelé aussi dual de* E*. Un élément de* E^* *est appelé forme linéaire sur* E*.*

Notons que $\dim(E^*) = \dim(E)$ et que si φ est une forme linéaire non nul, $\ker(\varphi)$ est un hyperplan (sous-espace vectoriel de E de dimension $n-1$).

Théorème 3.17 *On suppose que l'espace* E *est de dimension finie non nulle. Alors, il existe une base* q*-orthogonale de* E*.*

DÉMONSTRATION. On va raisonner par récurrence sur la dimension de E, qu'on note n.

Si $n = 1$, il n'y a rien à montrer. On suppose donc que l'entier n est strictement plus grand que 1. Si la forme q est nulle, toute base est q-orthogonale. Sinon, il existe un vecteur ε_1 de E tel que $q(\varepsilon_1) \neq 0$ et l'hypothèse de récurrence suppose alors qu'il existe une base q-orthogonale de tout sous-espace vectoriel de E de dimension $n-1$. Posons

$$H = \{x \in E \,|\, b(\varepsilon_1, x) = 0\}.$$

Le sous-espace H est un hyperplan de E car c'est le noyau de la forme linéaire non nulle $x \in E \mapsto b(\varepsilon_1, x)$ (qui est non nulle car l'image de ε_1 est non nul). En notant $\{\varepsilon_2, \dots, \varepsilon_n\}$ une base q-orthogonale de H, on a que $\{\varepsilon_1, \dots, \varepsilon_n\}$ est une base q-orthogonale de E. □

Ce résultat assure l'existence d'une base q-orthogonale en dimension finie. Soit $\mathscr{B} = \{\varepsilon_1, \dots, \varepsilon_n\}$ une telle base. En posant $\mu_i = q(\varepsilon_i)$ pour tout entier i de $\{1, \dots, n\}$, on a

$$\forall x \in E,\ q(x) = q\left(\sum_{i=1}^{n} x_i\,\varepsilon_i\right) = \sum_{i=1}^{n} \mu_i\, {x_i}^2.$$

On a ainsi écrit q comme une combinaison linéaire de carrés de formes linéaires linéairement indépendantes (les formes $l_i : x \in E \mapsto x_i$ où x_i est la i-ème coordonnée de x dans la base $\mathscr{B}$). En pratique, de telles formes linéaires peuvent être déterminées grâce à un procédé algorithmique de complétion des carrés portant le nom de *réduction de Gauss*.

Réduction de Gauss d'une forme quadratique

Soit E un $\mathbb{R}$-espace vectoriel de dimension n et q une forme quadratique réelle sur E s'écrivant

$$\forall x \in E,\ q(x) = \sum_{i=1}^{n} a_{ii}\, {x_i}^2 + \sum_{i=1}^{n} \sum_{j=i+1}^{n} a_{ij}\, x_i x_j.$$

En procédant itérativement, on va écrire la forme q comme une combinaison linéaire de carrés de formes linéaires sur E linéairement indépendantes. Deux cas se présentent.

1. Il existe au moins un indice i dans $\{1,\dots,n\}$ pour lequel le coefficient a_{ii} est non nul, par exemple $a_{11} = a \neq 0$. On peut alors écrire q sous la forme

$$\forall x \in E,\ q(x) = a\, {x_1}^2 + x_1 \ell(x_2,\dots,x_n) + \tilde{q}(x_2,\dots,x_n),$$

où ℓ et $\tilde{q}$ sont des formes respectivement linéaire et quadratique en les variables $x_2,\dots,x_n$, que l'on réécrit comme

$$\forall x \in E,\ q(x) = a\left(x_1 + \frac{1}{2a}\ell(x_2,\dots,x_n)\right)^2 + \tilde{q}(x_2,\dots,x_n) - \frac{1}{4a}(\ell(x_2,\dots,x_n))^2.$$

La forme q est alors la somme d'une constante multipliée par le carré d'une forme linéaire et d'une forme quadratique en les $n-1$ variables $x_2,\dots,x_n$. On réitère ensuite le procédé sur cette dernière forme jusqu'à arriver à la forme souhaitée.

2. Pour tout indice i dans $\{1,\dots,n\}$, le coefficient a_{ii} est nul. Si la forme q est nulle, on a achevé la réduction. Sinon, il existe au moins un coefficient a_{ij}, pour lequel l'indice i est strictement inférieur à l'indice j, non nul, par exemple $a_{12} = a \neq 0$. On peut alors écrire q sous la forme

$$\forall x \in E,\ q(x) = a\, x_1 x_2 + x_1 \ell_1(x_3,\dots,x_n) + x_2 \ell_2(x_3,\dots,x_n) + \tilde{q}(x_3,\dots,x_n),$$

où ℓ_1 et ℓ_2 sont des formes linéaires et $\tilde{q}$ est une forme quadratique en les variables $x_3,\dots,x_n$, que l'on réécrit comme[a]

$$\begin{aligned}\forall x \in E,\ q(x) =& a\left(x_1 + \frac{1}{a}\ell_2(x_3,\dots,x_n)\right)\left(x_2 + \frac{1}{a}\ell_1(x_3,\dots,x_n)\right) + \tilde{q}(x_3,\dots,x_n)\\ &- \frac{1}{a}\ell_1(x_3,\dots,x_n)\ell_2(x_3,\dots,x_n)\\ =& \frac{a}{4}\left(x_1 + x_2 + \frac{1}{a}(\ell_1(x_3,\dots,x_n) + \ell_2(x_3,\dots,x_n))\right)^2\\ &- \frac{a}{4}\left(x_1 - x_2 + \frac{1}{a}(\ell_2(x_3,\dots,x_n) - \ell_1(x_3,\dots,x_n))\right)^2 + \tilde{q}(x_3,\dots,x_n)\\ &- \frac{1}{a}\ell_1(x_3,\dots,x_n)\ell_2(x_3,\dots,x_n).\end{aligned}$$

Le premier terme dans le membre de droite de la dernière égalité contient deux carrés de formes linéaires linéairement indépendantes et une forme quadratique en les $n-2$ variables $x_3,\dots,x_n$. On réitère ensuite le procédé sur cette dernière forme jusqu'à arriver à la forme souhaitée.

a. On notera qu'on utilise ici l'identité remarquable

$$\forall (a,b) \in \mathbb{R}^2,\ ab = \frac{1}{4}\left[(a+b)^2 - (a-b)^2\right].$$

En appliquant le procédé de réduction de Gauss, on arrive à une *forme réduite* de q,

$$\forall x \in E,\ q(x) = \sum_{i=1}^{r} \alpha_i (\ell_i(x))^2, \tag{3.2}$$

c'est-à-dire l'écriture de q sous la forme d'une combinaison linéaire de r carrés de formes linéaires linéairement indépendantes (on n'a pas prouvé ce dernier point, mais il est facile de le vérifier), l'entier r étant le rang de q.

3.3 Classification des formes quadratiques réelles

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

Définition 3.18 *Une forme quadratique réelle sur E est dite **positive** (resp. **négative**) si et seulement si*

$$\forall x \in E,\ q(x) \geq 0 \text{ (resp. } q(x) \leq 0).$$

*On dit qu'elle est **définie positive** (resp. **définie négative**) si et seulement si*

$$\forall x \in E,\ x \neq 0_E,\ q(x) > 0 \text{ (resp. } q(x) < 0).$$

Par extension, une matrice symétrique est dite définie positive/positive/définie négative/négative si c'est la matrice d'une forme quadratique réelle définie positive/positive/définie négative/négative sur un $\mathbb{R}$-espace vectoriel de dimension finie.

Attention : une matrice symétrique (définie) positive n'est pas nécessairement une matrice à coefficients (strictement) positifs, comme le montre le contre-exemple suivant :

$$M = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 2 \\ 0 & 2 & 3 \end{pmatrix},\ X = \begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix} \text{ on a } X^\top M X = -1.$$

Théorème et définition 3.19 (« loi d'inertie de Sylvester ») *On suppose que l'espace E est de dimension finie non nulle égale à n. Soit q une forme quadratique réelle sur E. Il existe une base $\{\varepsilon_i\}_{i=1,\ldots,n}$ de E et des entiers p et r, avec $0 \leq p \leq r \leq n$, tels que l'on a*

$$q(x) = \sum_{i=1}^{p} {x_i}^2 - \sum_{i=p+1}^{r} {x_i}^2$$

*pour tout vecteur x de E tel que $x = \sum_{i=1}^{n} x_i\, \varepsilon_i$ (dans cette expression, la première (resp. seconde) somme est nulle si $p = 0$ (resp. $p = r$)). De plus, les entiers p et r sont indépendants de la base choisie pour mettre q sous cette forme et r est en particulier le rang de q. Enfin, le couple d'entiers $(p, r-p)$ est appelé la **signature** de q.*

DÉMONSTRATION. Hors programme. □

Le théorème précédent montre que la signature et le rang d'une forme quadratique ne dépendent pas de la base relativement à laquelle sa matrice est diagonale : on dit que ce sont des[1] *invariants* de la forme.

On en déduit le résultat suivant.

Corollaire 3.20 (équivalence des formes quadratiques réelles) *Deux formes quadratiques sur des $\mathbb{R}$-espaces vectoriels de dimension finie sont équivalentes si et seulement si elles ont même signature.*

Remarque 3.21 *Les entiers p et $r-p$ sont respectivement la dimension maximale des sous-espaces de E sur lesquels la forme q est définie positive et la dimension maximale des sous-espaces de E sur lesquels la forme q est définie négative. On ne peut cependant parler du « plus grand » sous-espace vectoriel sur lequel q est définie positive car un tel sous-espace n'existe généralement pas, comme le montre l'exemple suivant.*

Exemple : sur $\mathbb{R}^2$, la forme $q(x) = {x_1}^2 - {x_2}^2$, de forme polaire $b(x,y) = x_1 y_1 - x_2 y_2$, a pour signature $(1,1)$. On remarque que, si $|x_1| > |x_2|$, on a $q(x) > 0$. La restriction de q à toute droite dont la pente a une valeur absolue strictement plus petite que 1 est donc définie positive. De la même manière, sa restriction à toute droite dont la pente a une valeur absolue strictement plus grande que 1 est définie négative. Elle n'est cependant pas définie positive ou définie négative sur $\mathbb{R}^2$ (hormis $\{0_{\mathbb{R}^2}\}$ et $\mathbb{R}^2$ lui-même, les sous-espaces vectoriels de $\mathbb{R}^2$ sont des droites caractérisées par leurs pentes).

Il découle du précédent résultat qu'une forme quadratique est

1. Ce ne sont pas les seuls.

— positive si sa signature est $(r,0)$,
— négative si sa signature est $(0,r)$,
— définie positive si sa signature est $(n,0)$,
— définie négative si sa signature est $(0,n)$,
— non dégénérée si sa signature est $(p,n-p)$, *i.e.* $r=n$.

Remarque 3.22 *Il est important de noter que la matrice d'une forme quadratique définie positive (ou définie négative), relativement à n'importe quelle base de l'espace, est inversible, puisque de rang égal à son ordre.*

Un résultat utile pour montrer qu'une forme quadratique est définie positive est basé sur une propriété caractérisant les matrices réelles symétriques définies positives.

Théorème 3.23 (« critère de Sylvester ») *Soit n un entier naturel non nul. Une matrice réelle symétrique M d'ordre n est définie positive si et seulement si tous ses mineurs principaux dominants (ou primaires) sont strictement positifs, c'est-à-dire*

$$\forall k\in\{1,\dots,n\},\ \det(M_k)>0,$$

où

$$M_k=\begin{pmatrix} m_{11} & \dots & m_{1k}\\ \vdots & & \vdots\\ m_{k1} & \dots & m_{kk}\end{pmatrix}.$$

DÉMONSTRATION. Montrons tout d'abord que la condition est nécessaire. Pour cela, introduisons une forme quadratique q sur un espace vectoriel réel E de dimension égale à n, dont la matrice symétrique M est la matrice relativement à une base donnée de E. Si M est définie positive, alors q également et l'on sait (voir le théorème 3.19) qu'il existe une base $\{\varepsilon_i\}_{i=1,\dots,n}$ relativement à laquelle la matrice de q est la matrice identité, c'est-à-dire qu'il existe une matrice inversible P telle que

$$I_n=P^\top MP.$$

On a ainsi

$$1=\det(P)^2\det(M),$$

dont on déduit que $\det(M)$ est strictement positif. En considérant la restriction de la forme q au sous-espace vectoriel engendré $\mathrm{Vect}(\{\varepsilon_1,\dots,\varepsilon_k\})$ et en procédant de manière similaire pour k allant de 1 à $n-1$, on montre que tout mineur principal dominant de M est strictement positif.

Pour montrer que la condition est suffisante, on raisonne par récurrence sur l'ordre de la matrice, c'est-à-dire sur l'entier naturel non nul n. Pour $n=1$, c'est évident, puisque qu'on peut identifier M_1 à un réel strictement positif. Supposons le résultat vrai pour toute matrice d'ordre $n-1$, avec n un entier supérieur ou égal à 2, et considérons la matrice symétrique M_n, que l'on peut écrire par blocs sous la forme

$$M_n=\begin{pmatrix} M_{n-1} & m\\ m^\top & \alpha\end{pmatrix},$$

avec m appartenant à $M_{n-1,1}(\mathbb{R})$ et α appartenant à $\mathbb{R}$. La matrice M_{n-1} étant symétrique définie positive par hypothèse de récurrence, elle est inversible et ses colonnes forment une base de $M_{n-1,1}(\mathbb{R})$ et l'on peut par conséquent écrire m comme une (unique) combinaison linéaire de ces dernières, dont on note $\alpha_1,\dots,\alpha_{n-1}$ les coefficients. En considérant alors la matrice inversible d'ordre n

$$P=\begin{pmatrix} 1 & 0 & \dots & 0 & -\alpha_1\\ 0 & \ddots & \ddots & \vdots & \vdots\\ \vdots & \ddots & \ddots & 0 & \vdots\\ \vdots & & \ddots & \ddots & -\alpha_{n-1}\\ 0 & \dots & \dots & 0 & 1\end{pmatrix},$$

on vérifie que

$$P^\top M_nP=\begin{pmatrix} M_{n-1} & 0\\ 0^\top & \alpha\end{pmatrix},$$

d'où $\det(P^\top M_n P) = (\det(P))^2 \det(M_n) = \det(M_{n-1})\alpha$. Puisque les mineurs $\det(M_n)$ et $\det(M_{n-1})$ sont strictement positifs, on en déduit que α l'est aussi. Le caractère défini positif de M_n étant équivalent à celui de la matrice $P^\top M_n P$, on peut alors conclure en écrivant que

$$\forall X \in M_{n,1}(\mathbb{R}),\ X^\top P^\top M_n P X = \begin{pmatrix} x_1 \\ \vdots \\ x_{n-1} \end{pmatrix}^\top M_{n-1} \begin{pmatrix} x_1 \\ \vdots \\ x_{n-1} \end{pmatrix} + (x_n)^2\alpha,$$

et en utilisant que la matrice M_{n-1} est définie positive et le réel α est strictement positif. □

Remarque 3.24 *On déduit de ce résultat un critère pour déterminer si une matrice symétrique réelle est définie* négative *en utilisant le fait que son opposée est définie positive.*

Remarque 3.25 *On pourrait penser qu'un critère analogue au critère de Sylvester pour une matrice symétrique positive serait que tous les mineurs dominants principaux soient positifs. Ceci est faux, comme l'illustre l'exemple de la matrice*

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

dont les mineurs principaus dominants sont nuls et qui est symétrique négative. On peut en revanche montrer qu'une condition nécessaire et suffisante est que tous [2] *les mineurs principaux de la matrice soient positifs.*

Terminons cette section en nous focalisant sur les formes quadratiques positives.

Théorème 3.26 (« inégalité de Cauchy–Schwarz ») *Soit q une forme quadratique positive sur E, de forme polaire b. On a*

$$\forall (x,y) \in E^2,\ |b(x,y)| \leq \sqrt{q(x)}\sqrt{q(y)}.$$

Si de plus la forme q est définie, il y a égalité dans l'inégalité si et seulement si les vecteurs x et y sont linéairement dépendants.

DÉMONSTRATION. La forme q étant positive, on a

$$\forall t \in \mathbb{R},\ \forall (x,y) \in E^2,\ q(t\,x+y) = t^2\,q(x) + 2t\,b(x,y) + q(y) \geq 0.$$

Si $q(x) = 0$, l'inégalité ci-dessus devient

$$\forall t \in \mathbb{R},\ 2t\,b(x,y) + q(y) \geq 0,$$

ce qui entraîne que $b(x,y) = 0$. Sinon, le trinôme du second degré en t à gauche de l'inégalité possède un discriminant négatif, ce qui s'écrit encore

$$b(x,y)^2 - q(x)q(y) \leq 0,$$

conduisant à l'inégalité annoncée.

Si q est de plus définie, on suppose que le vecteur x est non nul (l'inégalité étant triviale lorsque ce n'est pas le cas). Alors $q(x)$ est non nul de sorte qu'on a égalité si et seulement le discriminant ci-dessus est nul, c'est-à-dire si et seulement s'il existe un réel t_0 tel que $q(t_0\,x+y) = 0$, ce qui équivaut à $t_0\,x+y = 0_E$. □

Corollaire 3.27 *Soit q une forme quadratique positive. Le cône isotrope de q est égal à son noyau.*

DÉMONSTRATION. On a toujours $\ker(q) \subset \mathscr{C}_q$ et il faut donc juste prouver l'inclusion réciproque. Soit x un vecteur appartenant à $\mathscr{C}_q$ et y un vecteur de E. En vertu de l'inégalité de Cauchy–Schwarz, on a

$$0 \leq |b(x,y)| \leq \sqrt{q(x)}\sqrt{q(y)} = 0,$$

ce qui implique que $b(x,y) = 0$, d'où x appartient à $\ker(q)$. □

Ce résultat vaut également pour une forme quadratique négative. Une forme positive (ou négative) est donc non dégénérée si et seulement si elle est définie.

2. Pour une matrice d'ordre n, ces déterminants sont au nombre de $\sum_{k=1}^{n} \binom{n}{k} = 2^n - 1$.

Corollaire 3.28 (inégalité de Minkowski) *Soit q une forme quadratique positive. On a*

$$\forall (x,y)\in E^2,\ \sqrt{q(x+y)} \leq \sqrt{q(x)}+\sqrt{q(y)}.$$

DÉMONSTRATION. C'est une conséquence directe de l'inégalité de Cauchy–Schwarz. On a en effet

$$\forall (x,y)\in E^2,\ q(x+y) = q(x)+2\,b(x,y)+q(y) \leq q(x)+2\sqrt{q(x)}\sqrt{q(y)}+q(y) = \left(\sqrt{q(x)}+\sqrt{q(y)}\right)^2.$$

□

Chapitre 4

Espaces euclidiens

La notion fondamentale d'*espace euclidien* généralise de manière naturelle la géométrie « classique », introduite par Euclide dans les *Éléments*. Elle est définie par la donnée d'un espace vectoriel sur le corps des réels, de dimension finie, muni d'un *produit scalaire*, et permet, entre autres choses, de « mesurer » des distances et des angles.

4.1 Définitions

Définition 4.1 (produit scalaire) *Soit E un espace vectoriel sur $\mathbb{R}$. Une application $\langle\cdot,\cdot\rangle$ de $E\times E$ dans $\mathbb{R}$ est appelée un* ***produit scalaire*** *sur E si elle est bilinéaire, symétrique, définie positive.*

D'après un corollaire de l'inégalité de Cauchy–Schwarz (voir le corollaire 3.27), la forme polaire d'une forme quadratique q positive sur E est non dégénérée si et seulement si $q$est définie. À ce titre, on peut aussi définir un produit scalaire comme une forme bilinéaire, symétrique, non dégénérée et positive.

Nombre de propriétés et de résultats vus dans ce chapitre ne dépendent pas du fait que l'espace vectoriel E soit de dimension finie. Néanmoins, on appelle *espace préhilbertien réel* tout espace vectoriel réel muni d'un produit scalaire. Si l'espace vectoriel réel est de dimension finie, on parle d'*espace euclidien*.

Exemples :
$E=\mathbb{R}^n$, $\langle x,y\rangle=\sum_{i=1}^N x_iy_i$ est un produit scalaire sur E, qui munit $\mathbb{R}^n$ de sa structure euclidienne dite *canonique*.

$E=\mathscr{C}^0([a,b],\mathbb{R})$, $\langle f,g\rangle=\int_a^b f(t)g(t)\,\mathrm{d}t$ est un produit scalaire sur E

$E=\ell^2=\{(u_k)_{k\in\mathbb{N}}\in\mathbb{R}^{\mathbb{N}}\mid \sum_{k\in\mathbb{N}}(u_k)^2<+\infty\}$, $\langle u,v\rangle=\sum_{k\in\mathbb{N}}u_kv_k$ est un produit scalaire sur E. C'est l'extension naturelle du premier exemple de produit scalaire en dimension infinie.

$E=\mathbb{R}[X]$, $\langle P,Q\rangle=\int_{-1}^1 P(t)Q(t)\,\mathrm{d}t$, $\langle P,Q\rangle=\int_0^{+\infty}P(t)Q(t)e^{-t}\,\mathrm{d}t$ et $\langle P,Q\rangle=\int_{-1}^1\frac{P(t)Q(t)}{\sqrt{1-t^2}}\,\mathrm{d}t$ sont trois produits scalaires sur E

Théorème 4.2 *Soit E un espace préhilbertien réel, de produit scalaire noté $\langle\cdot,\cdot\rangle$. L'application de E dans $\mathbb{R}$ définie par*

$$\forall x\in E,\ \|x\|=\sqrt{\langle x,x\rangle},$$

est une ***norme*** *sur E, c'est-à-dire une application satisfaisant aux propriétés suivantes :*
- *séparation : $\forall x\in E$, $\|x\|=0\implies x=0_E$,*
- *absolue homogénéité : $\forall(\lambda,x)\in\mathbb{R}\times E$, $\|\lambda x\|=|\lambda|\,\|x\|$,*
- *sous-additivité : $\forall(x,y)\in E^2$, $\|x+y\|\leq\|x\|+\|y\|$.*

On l'appelle ***norme associée au*** *(ou* ***dérivant du****)* ***produit scalaire****.*

DÉMONSTRATION. On a tout d'abord

$$\forall x\in E,\ \|x\|=0\iff\sqrt{\langle x,x\rangle}=0\iff\langle x,x\rangle=0\iff x=0_E,$$

puisque la forme quadratique associée au produit scalaire est définie positive.

On a ensuite

$$\forall \lambda \in \mathbb{R},\ \forall x \in E,\ \|\lambda\, x\| = \sqrt{\langle \lambda\, x, \lambda\, x\rangle} = \sqrt{\lambda^2 \langle x, x\rangle} = |\lambda| \sqrt{\langle x, x\rangle} = |\lambda|\, \|x\|.$$

Enfin, la sous-additivité découle de l'inégalité de Minkowski. □

Une norme est toujours positive. En effet, on a

$$\forall x \in E,\ 0 = \|0_E\| = \|x - x\| \leq \|x\| + \|-x\| = \|x\| + \|x\| = 2\,\|x\|.$$

Une application de E dans $\mathbb{R}$ qui ne satisfait que les propriétés d'homogénéité et de sous-additivité énoncées plus haut est appelée une *semi-norme*.

L'inégalité de Cauchy–Schwarz dans un espace préhilbertien réel se réécrit :

$$\forall (x, y) \in E^2,\ |\langle x, y\rangle| \leq \|x\|\|y\|.$$

Un espace vectoriel muni d'une norme est appelé un *espace vectoriel normé*.

Toutes les normes sur un espace vectoriel ne dérivent pas nécessairement d'un produit scalaire. Donner un exemple.

Proposition 4.3 *Soit E un espace préhilbertien réel et $\|\cdot\|$ une norme euclidienne sur E. Pour tous vecteurs x et y de E, on a*

$$\|x + y\| = \|x\| + \|y\|$$

si et seulement si la famille $\{x, y\}$ est positivement liée, i.e. si $x = 0$ ou s'il existe un réel positif λ tel que $y = \lambda\, x$.

DÉMONSTRATION. Si le vecteur x est nul, l'égalité est évidente. Si x est non nul et que $y = \lambda\, x$, avec λ un réel positif, on a

$$\|x + y\| = \|(1 + \lambda)x\| = (1 + \lambda)\|x\| = \|x\| + \lambda\|x\| = \|x\| + \|\lambda x\| = \|x\| + \|y\|.$$

Réciproquement, si on a égalité alors

$$\|x\|^2 + 2\langle x, y\rangle + \|y\|^2 = \|x + y\|^2 = (\|x\| + \|y\|)^2 = \|x\|^2 + 2\,\|x\|\|y\| + \|y\|^2,$$

d'où $\langle x, y\rangle = \|x\|\|y\|$, ce qui correspond au cas d'égalité dans l'inégalité de Cauchy-Schwarz et implique que la famille $\{x, y\}$ est liée. Ainsi, soit x est nul, soit $y = \lambda\, x$ avec λ un réel. On a donc

$$\lambda\|x\|^2 = \lambda\langle x, y\rangle = \|x\|\|y\| \geq 0 \implies \lambda \geq 0.$$

□

Théorème 4.4 (identité du parallélogramme) *Soit E un espace préhilbertien réel. On a*

$$\forall (x, y) \in E^2,\ \|x + y\|^2 + \|x - y\|^2 = 2\left(\|x\|^2 + \|y\|^2\right).$$

DÉMONSTRATION. On a d'une part

$$\forall (x, y) \in E^2,\ \|x + y\|^2 = \|x\|^2 + 2\langle x, y\rangle + \|y\|^2,$$

et d'autre part

$$\forall (x, y) \in E^2,\ \|x - y\|^2 = \|x\|^2 - 2\langle x, y\rangle + \|y\|^2.$$

L'égalité de l'énoncé s'obtient en sommant les deux précédentes. □

Remarques :

Il existe une réciproque à ce théorème. Si $(E, \|\cdot\|)$ est un espace vectoriel normé dans lequel l'identité est vérifiée, alors $\langle x, y\rangle = \frac{1}{4}\left(\|x + y\|^2 - \|x - y\|^2\right)$ est un produit scalaire associé à la norme $\|\cdot\|$ (voir exercice de TD).

L'identité s'interprète géométriquement de la manière suivante : la somme des carrés des longueurs des quatre côtés d'un parallélogramme est égale à la somme des carrés des longueurs de ses deux diagonales. Elle est par ailleurs équivalente au résultat suivant, connu sous le nom du *théorème de la médiane* ou *d'Apollonius* (de Perge) : *soit ABC un triangle du plan et I le milieu du segment $[BC]$, alors*

$$AB^2 + AC^2 = 2\left(BI^2 + AI^2\right).$$

Faire un dessin.

La définition suivante est une conséquence de l'inégalité de Cauchy–Schwarz.

4.2 Orthogonalité dans les espaces euclidiens

4.2.1 Généralités

Dans la suite, on notera $\langle \cdot, \cdot \rangle$ le produit scalaire.

Définitions 4.5 *Soit E un espace préhilbertien réel. Deux vecteurs x et y de E sont dits **orthogonaux**, ce que l'on note $x \perp y$, si et seulement si $\langle x, y \rangle = 0$.*

*Soit A une partie non vide de E. On appelle **orthogonal de** A, et on note $A^{\perp}$, l'ensemble*

$$A^{\perp} = \{x \in E \mid \forall a \in A,\ \langle x, a \rangle = 0\}.$$

Proposition 4.6 *Soit E un espace préhilbertien réel et A une partie non vide de E. Alors $A^{\perp}$ est un sous-espace vectoriel de E.*

DÉMONSTRATION. On constate que $A^{\perp}$ est l'intersection de sous-espace vectoriels :

$$A^{\perp} = \cap_{a \in A} \{x \in E \mid \langle x, a \rangle = 0\} = \cap_{a \in A} \ker(x \in E \mapsto \langle x, a \rangle).$$

□

4.2.2 Bases orthonormées

Définitions 4.7 *Soit E un espace préhilbertien réel. Une famille de vecteurs non nuls $\{e_i\}_{i \in I}$ de E est dite **orthogonale** si elle vérifie*

$$\forall (i, j) \in I^2,\ i \neq j,\ \left\langle e_i, e_j \right\rangle = 0.$$

*Une telle famille est dite **orthonormale** (ou **orthonormée**) si l'on a de plus*

$$\forall i \in I,\ \|e_i\| = 1.$$

Proposition 4.8 (relation de Pythagore) *Soit E un espace euclidien, p un entier naturel non nul et $\{e_i\}_{i=1,\ldots,p}$ une famille orthogonale de vecteurs de E. On a*

$$\|\sum_{i=1}^{p} e_i\|^2 = \sum_{i=1}^{p} \|e_i\|^2.$$

DÉMONSTRATION. laissée en exercice (récurrence) □

Proposition 4.9 *Une famille orthogonale est libre.*

DÉMONSTRATION. laissée en exercice □

Il découle de cette dernière propriété que l'on peut parler de *base orthogonale* (ou *orthonormale*) d'un espace préhilbertien réel.

Proposition 4.10 (coordonnées dans une base orthonormée) *Si E est un espace euclidien et si $\{e_1, \ldots, e_n\}$ est une base orthonormée de E, on a*

$$\forall x \in E,\ x = \sum_{i=1}^{n} x_i e_i,\ \forall y \in E,\ y = \sum_{i=1}^{n} y_i e_i,\ \langle x, y \rangle = \sum_{i=1}^{n} x_i y_i.$$

De plus, les coordonnées d'un vecteur x de E dans cette base vérifient

$$\forall i \in \{1, \ldots, n\},\ x_i = \langle x, e_i \rangle,$$

et l'on a

$$\|x\|^2 = \sum_{i=1}^{n} \langle x, e_i \rangle^2.$$

DÉMONSTRATION. laissée en exercice □

Remarque 4.11 *Noter que la notion de base orthonormée n'a pas de sens pour une forme quadratique quelconque car il n'est pas toujours possible d'assurer la positivité en chaque élément de la base.*

On retiendra de cette proposition que l'évaluation du produit scalaire entre deux vecteurs est aisée lorsque l'on travaille dans une base orthonormée.

Remarque 4.12 *Cette façon de déterminer les coordonnées d'un vecteur s'applique également lorsque l'on veut déterminer la matrice d'un endomorphisme u de E dans une base orthonormée $\mathscr{B} = \{e_1, \dots, e_n\}$. On a en effet*

$$\forall (i,j) \in \{1,\dots,n\}^2,\ (mat_{\mathscr{B}}(u))_{ij} = \left\langle u(e_j), e_i \right\rangle.$$

Une manière de construire une base orthonormée à partir d'une base quelconque de l'espace est donnée par le *procédé d'orthonormalisation de Gram–Schmidt*, décrit dans la preuve du théorème suivant.

Théorème 4.13 *Soit E un espace euclidien de dimension n et $\{f_1, \dots, f_n\}$ une base de E. Alors, il existe une unique base orthonormée $\{e_1, \dots, e_n\}$ de E telle que*

$$\forall j \in \{1,\dots,n\},\ \mathrm{Vect}(\{e_1,\dots,e_j\}) = \mathrm{Vect}(\{f_1,\dots,f_j\}) \text{ et } \left\langle f_j, e_j \right\rangle > 0.$$

DÉMONSTRATION. On va construire la base orthonormée par récurrence en procédant comme suit.

On pose tout d'abord $e_1 = \frac{f_1}{\|f_1\|}$. On suppose ensuite avoir construit une famille $\{e_1, \dots, e_j\}$, avec j un entier de $\{1, \dots, n-1\}$. On définit alors

$$e_{j+1} = \frac{f_{j+1} - \sum_{k=1}^{j} \left\langle f_{j+1}, e_k \right\rangle e_k}{\|f_{j+1} - \sum_{k=1}^{j} \left\langle f_{j+1}, e_k \right\rangle e_k\|}.$$

Vérifions les différentes propriétés énoncées dans le théorème. Il est tout d'abord clair que chaque vecteur e_j, avec j appartenant à $\{1, \dots, n\}$, appartient à $\mathrm{Vect}(\{f_1, \dots, f_j\})$ en utilisant les formules ci-dessus et en raisonnant par récurrence. Réciproquement, chaque vecteur f_j, avec j appartenant à $\{1, \dots, n\}$, appartient à $\mathrm{Vect}(\{e_1, \dots, e_j\})$.

Le fait de pouvoir normaliser le vecteur à chaque étape en divisant $f_{j+1} - \sum_{k=1}^{j} \left\langle f_{j+1}, e_k \right\rangle e_k$ par sa norme provient du fait que $\{f_1, \dots, f_n\}$ est une une famille libre et de la propriété précédente. Ceci permet en particulier d'assurer que le processus itératif qu'on a introduit est bien défini (on n'a pas de division par zéro).

Il faut ensuite montrer que la famille obtenue est orthogonale. Pour cela, il suffit de montrer que le vecteur e_j, avec j appartenant à $\{2, \dots, n\}$, est orthogonal à tout vecteur e_i, avec i appartenant à $\{1, \dots, j-1\}$, en raisonnant par récurrence. Pour $j = 2$, on a

$$\langle e_2, e_1 \rangle = \left\langle \frac{f_2 - \langle f_2, e_1 \rangle e_1}{\|f_2 - \langle f_2, e_1 \rangle e_1\|}, e_1 \right\rangle = \frac{1}{\|f_2 - \langle f_2, e_1 \rangle e_1\|} \left(\langle f_2, e_1 \rangle - \langle f_2, e_1 \rangle \|e_1\|^2 \right) = 0.$$

Pour j supérieur ou égal à 3, on fait l'hypothèse de récurrence

$$\forall (i,k) \in \{1,\dots,j-1\}^2, \langle e_i, e_k \rangle = \delta_{ik}.$$

On a alors

$$\begin{aligned}
\forall i \in \{1,\dots,j-1\},\ \left\langle e_j, e_i \right\rangle &= \frac{1}{\|f_j - \sum_{k=1}^{j-1} \left\langle f_j, e_k \right\rangle e_k\|} \left\langle f_j - \sum_{k=1}^{j-1} \left\langle f_j, e_k \right\rangle e_k, e_i \right\rangle \\
&= \frac{1}{\|f_j - \sum_{k=1}^{j-1} \left\langle f_j, e_k \right\rangle e_k\|} \left(\left\langle f_j, e_i \right\rangle - \sum_{k=1}^{j-1} \left\langle f_j, e_k \right\rangle \langle e_k, e_i \rangle \right) \\
&= \frac{1}{\|f_j - \sum_{k=1}^{j-1} \left\langle f_j, e_k \right\rangle e_k\|} \left(\left\langle f_j, e_i \right\rangle - \sum_{k=1}^{j-1} \left\langle f_j, e_k \right\rangle \delta_{ik} \right) \\
&= \frac{1}{\|f_j - \sum_{k=1}^{j-1} \left\langle f_j, e_k \right\rangle e_k\|} \left(\left\langle f_j, e_i \right\rangle - \left\langle f_j, e_i \right\rangle \right) \\
&= 0.
\end{aligned}$$

Enfin, on a, pour tout entier j de $\{1,\dots,n\}$,

$$\|e_j\|^2 = \left\langle e_j, e_j\right\rangle = \frac{1}{\|f_j - \sum_{k=1}^{j-1}\left\langle f_j, e_k\right\rangle e_k\|}\left\langle f_j - \sum_{k=1}^{j-1}\left\langle f_j, e_k\right\rangle e_k, e_j\right\rangle = \frac{\left\langle f_j, e_j\right\rangle}{\|f_j - \sum_{k=1}^{j-1}\left\langle f_j, e_k\right\rangle e_k\|},$$

d'où $\left\langle f_j, e_j\right\rangle > 0$.

Il nous reste à montrer que la base est unique. Toute base $\{e_1,\dots,e_n\}$ satisfaisant aux propriétés précédentes est telle que le vecteur e_1 appartient à $\mathrm{Vect}(\{f_1\})$, c'est-à-dire que $e_1 = \lambda f_1$. Puisque $\|e_1\| = 1$ et $\langle e_1, f_1\rangle > 0$, on trouve que $\lambda = \frac{1}{\|f_1\|}$.

Supposons ensuite que les vecteurs $e_1,\dots,e_j$ soient de la forme donnée pour j superieur ou égal à 1. On sait que e_{j+1} appartient à $\mathrm{Vect}(\{f_1,\dots,f_{j+1}\}) = \mathrm{Vect}(\{e_1,\dots,e_j,f_{j+1}\})$, d'où

$$e_{j+1} = \sum_{k=1}^{j} \lambda_k\, e_k + \lambda\, f_{j+1}.$$

En considérant successivement les produits scalaires de ce vecteur avec les vecteurs e_i pour i appartenant à $\{1,\dots,j\}$, on trouve que $\lambda_i = -\lambda \left\langle f_{j+1}, e_i\right\rangle$ et on a par conséquent

$$e_{j+1} = \lambda\left(f_{j+1} - \sum_{k=1}^{j}\left\langle f_{j+1}, e_k\right\rangle e_k\right).$$

Les conditions $\|e_{j+1}\| = 1$ et $\left\langle e_{j+1}, f_{j+1}\right\rangle > 0$ permettent alors de déterminer λ, conduisant à la forme donnée. □

Remarque 4.14 *Le procédé de Gram–Schmidt peut aussi s'appliquer en dimension infinie, quand on dispose d'une suite $(u_n)_{n\in\mathbb{N}}$ de vecteurs d'un espace préhilbertien réel E, telle que $\{u_1,\dots,u_n\}$ soit libre pour tout entier naturel non nul n, pour former une suite de vecteurs orthonormés.*

Exemple 4.15 *Considérons $E = \mathbb{R}[X]$ muni du produit scalaire $\langle P, Q\rangle = \int_{-1}^{1} P(t)Q(t)\,\mathrm{d}t$ (le fait que cela nous donne bien un produit scalaire est laissé au lecteur). On peut appliquer le procédé d'orthonormalisation de Gram–Schmidt à la base canonique de $\mathbb{R}[X]$. On obtient une famille de polynômes orthogonaux appelée polynômes de Legendre.*

Dans la construction ci-dessus, on peut observer que la matrice de passage de la base $\{f_1,\dots,f_n\}$ à la base $\{e_1,\dots,e_n\}$ est triangulaire supérieure, à coefficients diagonaux strictement positifs (ce sont les inverses de normes de vecteurs).

Corollaire 4.16 *Un espace euclidien non réduit au vecteur nul possède une base orthonormée.*

Proposition 4.17 *Soit E un espace euclidien et A un sous-espace vectoriel de E. On a*

$$E = A \oplus A^{\perp} \text{ et } (A^{\perp})^{\perp} = A$$

On a aussi $\dim(A^{\perp}) = \dim(E) - \dim(A)$.

DÉMONSTRATION. Supposons que l'espace E est de dimension n et que le sous-espace A est de dimension $p \leq n$. En vertu du théorème de la base incomplète, il est possible de compléter toute base de A en une base de E. En appliquant le procédé d'orthonormalisation de Gram–Schmidt à une telle base, on obtient une base orthonormée de E dont il n'est pas difficile de voir (il suffit pour cela d'utiliser les propriétés énoncées dans le théorème 4.13) que ses p premiers vecteurs forment une base orthonormée de A et les $n-p$ suivants une base orthonormée de $A^{\perp}$. Ceci permet de conclure que $E = A \oplus A^{\perp}$, et donc que $\dim(E) = \dim(A) + \dim(A^{\perp})$.

Pour prouver la seconde assertion, on a $A \subset (A^{\perp})^{\perp}$. En effet si $a \in A$, et si $y \in A^{\perp}$ on a $\langle x, y\rangle = 0$ donc x est orthogonal à tout élément de $A^{\perp}$ donc $x \in (A^{\perp})^{\perp}$. On utilise enfin que $\dim(A) = \dim(E) - \dim(A^{\perp}) = \dim((A^{\perp})^{\perp})$. □

Remarque 4.18 *Cette proposition reste vraie dans un espace préhilbertien réel si le sous-espace A est de dimension* finie.

4.2.3 Projection orthogonale

Proposition et définition 4.19 (projection et symétrie orthogonales) *Soit E un espace euclidien et A un sous-espace vectoriel de E. Tout vecteur x de E s'écrit, de manière unique, comme la somme $x = u + v$ avec u appartenant à A et v appartenant à $A^{\perp}$. On appelle* ***projection orthogonale de x sur*** *A le vecteur u ainsi défini, et on note p_A l'application de E dans E telle que $x \mapsto u = p_A(x)$ ainsi construite. On appelle* ***symétrie orthogonale de x par rapport à*** *A le vecteur $u - v$ et on note s_A l'application de E dans E telle que $x \mapsto u - v = 2\,p_A(x) - x = s_A(x)$ ainsi construite.*

Remarque 4.20 *Rappelons que, étant donné un espace vectoriel E, A un sous-espace vectoriel de E et B un supplémentaire de A, on définit la* projection de E sur A parallèlement à B *comme l'endomorphisme $p_{A,B}$ de E défini par*

$$\forall x \in E,\ p_{A,B}(x) = x_A,$$

où $x = x_A + x_B$ est l'unique décomposition du vecteur x comme somme d'un vecteur de A et d'un vecteur de B. La projection orthogonale p_A est alors la projection sur A parallèlement à $A^{\perp}$. Notons aussi qu'une projection est idempotente, i.e. $p \circ p = p$.

Proposition 4.21 (propriétés d'une projection orthogonale) *Soit A un sous-espace vectoriel d'un espace euclidien E. On a les propriétés suivantes :*

1. *$\forall x \in E,\ x = p_A(x) + p_{A^{\perp}}(x)$, $x - p_A(x) \in A^{\perp}$ et $\|p_A(x)\| \leq \|x\|$.*
2. *Si $\{e_1, \dots, e_p\}$ est une base orthonormée de A, alors*
$$\forall x \in E,\ p_A(x) = \sum_{i=1}^{p} \langle x, e_i \rangle\ e_i.$$
3. *La projection orthogonale sur A se caractérise variationnellement par la propriété de minimalité suivante :*
$$\forall x \in E,\ \|x - p_A(x)\| = \inf_{y \in A} \|x - y\|.$$
4. *Si p est un endomorphisme de E idempotent et si $\operatorname{Im}(p) = \ker(p)^{\perp}$, alors p est la projection orthogonale sur $\operatorname{Im}(p)$.*

DÉMONSTRATION.

1. C'est une conséquence du fait que $E = A \oplus A^{\perp}$ et que $(A^{\perp})^{\perp} = A$, l'inégalité découlant de la relation de Pythagore.
2. Il suffit d'écrire la projection de x comme une combinaison linéaire des vecteurs de la base orthonormée et on en détermine les coefficients en utilisant que $x - p_A(x)$ appartient à $A^{\perp}$, c'est-à-dire en écrivant que
$$\forall x \in E,\ \forall i \in \{1, \dots, p\},\ \langle x - p_A(x), e_i \rangle = 0.$$
3. On écrit que
$$\forall x \in E,\ \forall y \in A,\ \|x - y\|^2 = \|x - p_A(x) + p_A(x) - y\|^2 = \|x - p_A(x)\|^2 + \|p_A(x) - y\|^2,$$
la seconde égalité découlant de l'orthogonalité.
4. On montre que $\ker(p) \oplus \operatorname{Im}(p) = E$ en utilisant le théorème du rang et en servant de l'orthogonalité entre $\ker(p)$ et $\operatorname{Im}(p)$ pour en déduire que $\ker(p) \cap \operatorname{Im}(p) = \{0_E\}$. On conclut par la définition de la projection orthogonale.

□

Remarque : la seconde propriété permet de réécrire la formule donnant les vecteurs de la base orthonormée obtenue par le procédé de Gram-Schmidt sous la forme

$$\forall j \in \{1, \dots, n\},\ e_j = \frac{f_j - p_{\mathrm{Vect}(\{e_1,\dots,e_{j-1}\})}(f_j)}{\|f_j - p_{\mathrm{Vect}(\{e_1,\dots,e_{j-1}\})}(f_j)\|}.$$

Application à la résolution d'un problème de régression

Supposons que l'on dispose de m observations, représentées par les réels y_i, $i = 1, \dots, m$, associées à m conditions, représentées par les réels x_i, $i = 1, \dots, m$. On souhaite approcher le nuage de points (x_i, y_i) du plan par le graphe d'une fonction φ « simple » qui dépend linéairement de n coefficients c_i, $i = 1, \dots, n$. Le *problème de régression* consiste à déterminer « au mieux » les coefficients de la fonction φ, c'est-à-dire de manière à minimiser (en un sens à préciser) la distance entre le vecteur $(y_1, \dots, y_m)$ et le vecteur $(\varphi(x_1), \dots, \varphi(x_m))$.

D'un point de vue algébrique, le problème de régression se pose comme la résolution du système linéaire

$$AC = Y,$$

où $C = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$, $Y = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}$ et A est une matrice de $M_{m,n}(\mathbb{R})$ dépendant des conditions x_i, $i = 1, \dots, m$. Étant donné qu'on a $m > n$ dans la plupart des cas en pratique, ce système est sur-déterminé et il n'admet en général pas de solution. En effet, l'image de A est un sous-espace vectoriel de $M_{m,1}(\mathbb{R})$ de dimension au plus égale à n et le vecteur Y n'en fait généralement pas partie. L'introduction du *problème aux moindres carrés* vise à pallier ce problème en considérant la détermination d'un vecteur C minimisant la distance (euclidienne) entre Y et AC.

Définition 4.22 (problème aux moindres carrés) *Étant donné une matrice A de $M_{m,n}(\mathbb{R})$, avec $m > n$, et un vecteur Y de $M_{m,1}(\mathbb{R})$, le problème aux moindres carrés consiste à trouver un vecteur C de $M_{n,1}(\mathbb{R})$ minimisant la norme euclidienne de $Y - AC$.*

Note : en pratique, la fonction φ est généralement une fonction polynomiale et l'entier $n-1$ est son degré. Si $n = m$, il existe une unique fonction polynomiale de degré $m-1$, $\varphi(x) = \sum_{i=1}^{m} c_i\, x^{i-1}$, telle que $\varphi(x_i) = y_i$ pour $i = 1, \dots, m$. En effet, le système du problème de régression s'écrit dans ce cas

$$\begin{pmatrix} 1 & x_1 & {x_1}^2 & \dots & {x_1}^{m-1} \\ 1 & x_2 & {x_2}^2 & \dots & {x_2}^{m-1} \\ \vdots & \vdots & & \vdots & \\ 1 & x_m & {x_m}^2 & \dots & {x_n}^{m-1} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}$$

et il possède une unique solution dès que les réels x_i sont distincts deux à deux. La fonction polynomiale ainsi obtenue est celle correspondant au *polynôme d'interpolation de Lagrange* associé aux couples (x_i, y_i).

On a cependant $m > n$ en général, comme dans le problème de *régression linéaire*, pour lequel $n = 2$ et $\varphi(x) = c_1 + c_2\, x$.

Cherchons le vecteur C réalisant $\inf_{C \in M_{n,1}(\mathbb{R})} \|Y - AC\|$. Par la caractérisation variationnelle de la projection orthogonale (voir la proposition 4.21), ce vecteur C sera tel que AC est la projection orthogonale de Y sur $\operatorname{Im}(A)$, ce qui signifie encore que le résidu $Y - AC$ doit être orthogonal $\operatorname{Im}(A)$, c'est-à-dire à AZ pour tout choix de vecteur Z de $M_{n,1}(\mathbb{R})$. Dans une base orthonormée, on aura ainsi

$$\forall Z \in M_{n,1}(\mathbb{R}),\ (AZ)^\top (Y - AC) = 0 \iff \forall Z \in M_{n,1}(\mathbb{R}),\ Z^\top A^\top (Y - AC) = 0,$$

ce qui équivaut à avoir

$$A^\top AC = A^\top Y,$$

ce dernier système portant le nom d'*équations normales*. On peut montrer (à titre d'exercice) que la matrice $A^\top A$ est inversible si et seulement si le rang de A est égal à n. Il en découle que le problème aux moindres carrés possède une unique solution si et seulement si le rang de la matrice A est égal au nombre de colonnes de cette dernière. Lorsque c'est le cas, la solution de $AC = Y$ au sens des moindres carrés est donnée par

$$\hat{C} = (A^\top A)^{-1} A^\top Y,$$

la matrice à n lignes et m colonnes $A^+ = (A^\top A)^{-1}A^\top$ étant appelée le *pseudo-inverse* (ou l'*inverse généralisé*) de A.

4.3 Structure du dual d'un espace euclidien

Commençons par un exemple en considérant l'espace $E = \mathbb{R}^3$. Étant donné trois réels α, β et γ, l'application $l_{\alpha,\beta,\gamma}(x) = \alpha x_1 + \beta x_2 + \gamma x_3$ est une forme linéaire sur E, et donc un élément du dual de E. On peut alors remarquer qu'on peut écrire

$$\forall x \in E,\ l_{\alpha,\beta,\gamma}(x) = \langle x, y\rangle$$

en posant $y = (\alpha, \beta, \gamma)$. Ce résultat se généralise à tout espace euclidien.

Théorème 4.23 (théorème de représentation de Riesz) *Soit E un espace euclidien. Pour tout forme linéaire $\ell \in \mathscr{L}(E, \mathbb{R})$, il existe un unique vecteur y de E tel que*

$$\forall x \in E,\ \ell(x) = \langle x, y\rangle\,.$$

DÉMONSTRATION. Soit l'application linéaire i_E de E dans E^* définie par

$$\forall (x, y) \in E^2,\ i_E(y)(x) = \langle x, y\rangle\,.$$

Si $i_E(y) = 0_{E^*}$, on a en particulier que

$$0 = i_E(y)(y) = \langle y, y\rangle\,,$$

d'où $y = 0_E$. L'application est ainsi injective, et donc surjective, puisque $\dim(E) = \dim(E^*)$. □

4.4 Endomorphismes des espaces euclidiens

Dans toute cette section, sauf mention contraire, E désigne un espace euclidien de dimension n non nulle.

4.4.1 Adjoint d'un endomorphisme

Théorème et définition 4.24 (adjoint d'un endomorphisme) *Soit u un endomorphisme de E. Il existe un unique endomorphisme de E, noté u^* et appelé **adjoint de** u, tel que*

$$\forall (x, y) \in E^2,\ \langle u(x), y\rangle = \langle x, u^*(y)\rangle\,.$$

De plus, si $\mathscr{B}$ est une base orthonormée de E, on a

$$mat_{\mathscr{B}}(u^*) = mat_{\mathscr{B}}(u)^\top.$$

DÉMONSTRATION. Soit $\mathscr{B}$ une base orthonormée de E et M la matrice de l'endomorphisme u dans la base $\mathscr{B}$. On cherche un endomorphisme v de E qui soit l'adjoint de u, de matrice N dans la base $\mathscr{B}$. La relation de l'énoncé s'écrit matriciellement

$$\forall (X, Y) \in (M_{n,1}(\mathbb{R}))^2,\ (MX)^\top Y = X^\top N Y \iff \forall (X, Y) \in (M_{n,1}(\mathbb{R}))^2,\ X^\top M^\top Y = X^\top N Y.$$

La matrice N est donc entièrement et uniquement déterminée, et l'endomorphisme v également. □

Remarque 4.25 *On a montré dans la preuve précédente que si $\mathscr{B}$ une base orthonormée de E et M la matrice de l'endomorphisme u dans la base $\mathscr{B}$, alors $M^\top$ est la matrice de l'endomorphisme u^* dans la base $\mathscr{B}$.*

Proposition 4.26 *Soit u un endomorphisme de E. Alors $(u^*)^* = u$ et*

$$\ker(u^*) = (\mathrm{Im}\,(u))^\perp \text{ et } \mathrm{Im}\,(u^*) = (\ker(u))^\perp.$$

DÉMONSTRATION. Il est clair que $(u^*)^* = u$ $((M^\top)^\top = M)$. Montrons que $\ker(u^*) = (\mathrm{Im}\,(u))^\perp$. Soit $x \in \ker(u^*)$ et $y \in \mathrm{Im}\,(u)$. Alors il existe $z \in E$ tel que $y = u(z)$. On alors comme $x \in \ker(u^*)$,

$$\langle x, y\rangle = \langle x, u(z)\rangle = \langle u^*(x), z\rangle = 0.$$

On vient de montrer qu'un élément de $\ker(u^*)$ est orthogonal à tout élément de $\mathrm{Im}\,(u)$. Donc $\ker(u^*) \subset (\mathrm{Im}\,(u))^\perp$. Réciproquement, si $x \in (\mathrm{Im}\,(u))^\perp$, alors pour tout $z \in E$, on a

$$0 = \langle x, u(z)\rangle = \langle u^*(x), z\rangle\,.$$

On vient de montrer que l'image par u^* d'un élément de $(\mathrm{Im}\,(u))^\perp$ est orthogonal à tout élément de E. Comme le produit scalaire est définie, cela oblige $u^*(x) = 0$. Donc $(\mathrm{Im}\,(u))^\perp \subset \ker(u^*)$. Ainsi on a bien $\ker(u^*) = (\mathrm{Im}\,(u))^\perp$.

Enfin, en appliquant le résultat que l'on vient de montrer à u^*, on obtient

$$\mathrm{Im}\,(u^*) = ((\mathrm{Im}\,(u^*))^\perp)^\perp = (\ker((u^*)^*))^\perp = (\ker(u))^\perp.$$

□

4.4.2 Isométries vectorielles

Définition 4.27 (isométrie vectorielle) *Soit E un espace préhilbertien réel et u un endomorphisme de E. On dit que u est une **isométrie vectorielle** (ou encore un **endomorphisme orthogonal**) si*

$$\forall x \in E,\ \|u(x)\| = \|x\|.$$

On note $O(E)$ l'ensemble des isométries vectorielles de E.

Toute application qui conserve la norme n'est pas nécessairement linéaire, et n'est donc pas nécessairement une isométrie vectorielle. Un exemple est donné par l'application $x \mapsto \|x\|\, e$, avec e un vecteur unitaire.

On déduit immédiatement de la définition que les seules valeurs propres réelles possibles pour une isométrie vectorielle sont -1 et 1.

Exemple 4.28 *Les seules homothéties $x \mapsto \lambda\, x$ qui sont des isométries vectorielles sont $-id_E$ et id_E.*

Proposition 4.29 *Soit E un espace préhilbertien réel. Une application u de E dans E est une isométrie vectorielle si et seulement si*

$$\forall (x, y) \in E^2,\ \langle u(x), u(y)\rangle = \langle x, y\rangle\,.$$

DÉMONSTRATION. Supposons que u est une isométrie vectorielle de E. Par définition, c'est un endomorphisme de E vérifiant

$$\forall x \in E,\ \|u(x)\| = \|x\|.$$

On a alors, en vertu de l'identité de polarisation associée au produit scalaire sur E et au carré de la norme associée,

$$\begin{aligned}
\forall (x, y) \in E^2,\ \langle u(x), u(y)\rangle &= \frac{1}{4}\left(\|u(x) + u(y)\|^2 - \|u(x) - u(y)\|^2\right)\\
&= \frac{1}{4}\left(\|u(x + y)\|^2 - \|u(x - y)\|^2\right)\\
&= \frac{1}{4}\left(\|x + y\|^2 - \|x - y\|^2\right)\\
&= \langle x, y\rangle\,.
\end{aligned}$$

Réciproquement, en choisissant $x = y$ dans l'égalité de l'énoncé, on trouve que

$$\forall x \in E,\ \|u(x)\|^2 = \|x\|^2,$$

et l'application u préserve la norme. Il reste à montrer qu'elle est linéaire. On a, pour tout réel λ et tous vecteurs x et y de E,

$$\begin{aligned}
\|u(\lambda x+y)-\lambda u(x)-u(y)\|^2 &= \|u(\lambda x+y)\|^2+\|\lambda u(x)\|^2+\|u(y)\|^2\\
&\quad -2\langle u(\lambda x+y),\lambda u(x)\rangle -2\langle u(\lambda x+y),u(y)\rangle+2\langle \lambda u(x),u(y)\rangle\\
&= \|u(\lambda x+y)\|^2+\lambda^2\|u(x)\|^2+\|u(y)\|^2-2\lambda\langle u(\lambda x+y),u(x)\rangle\\
&\quad -2\langle u(\lambda x+y),u(y)\rangle+2\lambda\langle u(x),u(y)\rangle\\
&= \|\lambda x+y\|^2+\lambda^2\|x\|^2+\|y\|^2-2\lambda\langle \lambda x+y,x\rangle\\
&\quad -2\langle \lambda x+y,y\rangle+2\lambda\langle x,y\rangle\\
&= \|\lambda x+y\|^2+\|\lambda x\|^2+\|y\|^2-2\langle \lambda x+y,\lambda x\rangle\\
&\quad -2\langle \lambda x+y,y\rangle+2\langle \lambda x,y\rangle\\
&= \|\lambda x+y-\lambda x-y\|^2\\
&= 0.
\end{aligned}$$

Par propriété de la norme, il vient alors que

$$\forall\lambda\in\mathbb{R},\ \forall(x,y)\in E^2,\ u(\lambda x+y)-\lambda u(x)-u(y)=0_E,$$

d'où la conclusion. □

Remarque 4.30 *Dans la précédente proposition, on observera que l'application n'est pas supposée linéaire. Comme on peut le voir dans la preuve, la linéarité se déduit de la préservation du produit scalaire.*

Il découle de cette caractérisation qu'une isométrie vectorielle conserve l'orthogonalité, mais toute application conservant l'orthogonalité n'est pas nécessairement une isométrie vectorielle, comme le montre l'exemple d'une homothétie de rapport différent de -1 ou 1.

Théorème 4.31 *Une isométrie vectorielle de E est un automorphisme de E dont l'inverse est donné par son adjoint.*

DÉMONSTRATION. Soit u une isométrie vectorielle de E. Pour tout vecteur x appartenant au noyau de u, on a

$$0=\|u(x)\|=\|x\| \implies x=0_E,$$

et l'endomorphisme est donc injectif. L'espace E étant de dimension finie, ceci équivaut à dire que u est bijectif.

Enfin, on a, d'après la précédente proposition,

$$\forall(x,y)\in E^2,\ \langle x,y\rangle=\langle u(x),u(y)\rangle=\langle x,u^*(u(y))\rangle,$$

d'où $u^*\circ u=id_E$. □

Remarque 4.32 *Quand E est de dimension finie, on peut montrer que $O(E)$ est un sous-groupe de $GL(E)$, le groupe des automorphismes de E (que l'on munit de la composition des applications et dont l'élément neutre est l'application identité). On dit alors que $O(E)$ est le* groupe orthogonal *de E. En revanche, en dimension infinie, une isométrie vectorielle est toujours injective, mais pas nécessairement surjective, et l'ensemble $O(E)$ n'est donc pas nécessairement un groupe.*

Proposition 4.33 *Soit u une isométrie vectorielle de E. Si A est un sous-espace vectoriel de E stable par u, alors son orthogonal $A^\perp$ est aussi stable par u.*

DÉMONSTRATION. L'endomorphisme u étant injectif, on a $\dim(u(A))=\dim(A)$ et, par stabilité de A par u, on peut conclure que $u(A)=A$. On a alors

$$\forall x\in A^\perp,\ \forall y\in A,\ \langle u(x),u(y)\rangle=\langle x,y\rangle=0,$$

d'où $u(x)$ appartient à $(u(A))^\perp=A^\perp$. □

Remarque 4.34 *Ce résultat est encore vrai si A est un sous-espace vectoriel de dimension finie d'un espace préhilbertien réel.*

Proposition 4.35 *Un endomorphisme u de E est une isométrie vectorielle si et seulement si l'image d'une base orthonormale de E par u est une base orthonormale de E.*

DÉMONSTRATION. Soit $\mathscr{B}=\{e_1,\dots,e_n\}$ une base orthonormale de E. Supposons que l'endomorphisme u soit une isométrie vectorielle. On a

$$\forall(i,j)\in\{1,\dots,n\}^2,\ \left\langle u(e_i),u(e_j)\right\rangle=\left\langle e_i,e_j\right\rangle=\delta_{ij}.$$

La famille $\{u(e_1),\dots,u(e_n)\}$ est ainsi une famille libre de n vecteurs, c'est donc une base de E.

Réciproquement, supposons que l'endomorphisme u transforme la base $\mathscr{B}$ en une base orthonormale de E. On a alors

$$\forall x\in E,\ x=\sum_{i=1}^n x_i\,e_i,\ \|u(x)\|^2=\|\sum_{i=1}^n x_i\,u(e_i)\|^2=\sum_{i=1}^n {x_i}^2\|u(e_i)\|^2=\sum_{i=1}^n {x_i}^2=\|x\|^2,$$

et u est une isométrie vectorielle. □

Proposition 4.36 (propriétés matricielles d'une isométrie vectorielle) *Soit $\mathscr{B}$ une base orthonormale de E. Un endomorphisme u de E est une isométrie vectorielle si et seulement si sa matrice M dans la base $\mathscr{B}$ est telle que*

$$M^\top M=MM^\top=I_n.$$

DÉMONSTRATION. Posons $\mathscr{B}=\{e_1,\dots,e_n\}$. Supposons que l'endomorphisme u soit une isométrie vectorielle. On a

$$\forall(i,j)\in\{1,\dots,n\}^2,\ (M^\top M)_{ij}={C_i}^\top C_j,$$

où C_i et C_j sont respectivement les i^{e} et j^{e} colonnes de la matrice M, et par conséquent

$$\forall(i,j)\in\{1,\dots,n\}^2,\ (M^\top M)_{ij}=\left\langle u(e_i),u(e_j)\right\rangle=\left\langle e_i,e_j\right\rangle=\delta_{ij},$$

ce qui signifie encore que $M^\top M=I_n$. La matrice M est donc inversible, d'inverse $M^\top$, et on a donc aussi $MM^\top=I_n$.

Réciproquement, si la matrice M est telle que $M^\top M=MM^\top=I_n$, cela signifie que

$$\forall(i,j)\in\{1,\dots,n\}^2,\ \left\langle u(e_i),u(e_j)\right\rangle=\delta_{ij}.$$

L'image de la base $\mathscr{B}$ par l'endomorphisme u est donc une base orthonormée de E et on en déduit que u est une isométrie vectorielle en utilisant la proposition 4.35. □

Définition 4.37 (matrice orthogonale) *Soit n un entier naturel non nul. On appelle* ***matrice orthogonale*** *une matrice d'ordre n à coefficients réels telle que*

$$M^\top M=MM^\top=I_n.$$

On note $O(n)$ l'ensemble des matrices orthogonales d'ordre n.

La précédente proposition nous dit qu'un endomorphisme de E est une isométrie vectorielle si et seulement si sa matrice dans une base orthonormale quelconque de E est une matrice orthogonale. Ce résultat est *faux* si la base considérée n'est pas orthonormale.

Théorème 4.38 *Soit n un entier naturel non nul. Pour toute matrice M de $O(n)$, on a*

$$\det(M)=\pm1.$$

L'ensemble $O(n)$ est un sous-groupe de $GL(n,\mathbb{R})$, appelé ***groupe orthogonal****.*

DÉMONSTRATION. Par propriété du déterminant, on a

$$\forall M \in M_n(\mathbb{R}),\ \det(M) = \det(M^\top),$$

et donc

$$\forall M \in O(n),\ (\det(M))^2 = \det(MM^\top) = \det(I_n) = 1.$$

Il en résulte que $O(n)$ est un sous-ensemble de $GL(n,\mathbb{R})$. Par ailleurs, la matrice I_n appartient à $O(n)$ et on a

$$\forall (M,n) \in (O(n))^2,\ (M^{-1})^{-1} = (M^\top)^{-1} = (M^{-1})^\top \text{ et } (MN)^{-1} = N^{-1}M^{-1} = N^\top M^\top = (MN)^\top,$$

ce qui permet de montrer que $O(n)$ est bien un sous-groupe de $GL(n,\mathbb{R})$. □

Remarque 4.39 *L'ensemble des matrices orthogonales de déterminant égal à* 1 *forme un sous-groupe du groupe orthogonal, appelé* groupe spécial orthogonal *et noté* $SO(n)$*.*

4.4.3 Endomorphismes auto-adjoints

Définition 4.40 (endomorphisme auto-adjoint) *Un endomorphisme u de E est dit* ***auto-adjoint*** *(ou* ***symétrique****) si et seulement si* $u^* = u$*, ce qui équivaut encore à avoir*

$$\forall (x,y) \in E^2,\ \langle u(x), y\rangle = \langle x, u(y)\rangle\,.$$

Par ailleurs, l'endomorphisme u est dit ***anti-auto-adjoint*** *(ou* ***antisymétrique****) si et seulement si* $u^* = -u$*, c'est-à-dire si et seulement si*

$$\forall (x,y) \in E^2,\ \langle u(x), y\rangle = -\langle x, u(y)\rangle\,.$$

On peut montrer (exercice) que la matrice dans une base orthonormale d'un endomorphisme auto-adjoint est *symétrique*. Réciproquement, si la matrice dans une base orthonormale d'un endomorphisme est symétrique, alors l'endomorphisme est auto-adjoint.

Exemple 4.41 *Toute projection orthogonale est auto-adjoint. En effet si* $x, y \in E$ *et p est une projection orthogonale, on a* $p(x)$ *orthogonal à* $y - p(y)$ *et* $p(y)$ *orthogonal à* $x - p(x)$ *de sorte que*

$$\begin{aligned}
\langle p(x), y\rangle &= \langle p(x), y - p(y) + p(y)\rangle = \langle p(x), y - p(y)\rangle + \langle p(x), p(y)\rangle = \langle p(x), p(y)\rangle \\
\langle x, p(y)\rangle &= \langle p(x) - x + x, p(y)\rangle = \langle p(x) - x, p(y)\rangle + \langle x, p(y)\rangle = \langle p(x), p(y)\rangle\,.
\end{aligned}$$

On peut même montrer qu'une projection est auto-adjoint si et seulement si elle la projection est orthogonale.

On a la proposition suivante, conséquence de la proposition 4.26.

Proposition 4.42 *Soit E un espace euclidien et u un endomorphisme auto-adjoint ou anti-auto-adjoint de E. Alors*

$$\ker(u) = (\mathrm{Im}\,(u))^\perp.$$

Proposition 4.43 *Soit E un espace euclidien et u un endomorphisme auto-adjoint ou anti-auto-adjoint de E. Si A est un sous-espace vectoriel de E stable par u, alors son orthogonal* $A^\perp$ *est aussi stable par u.*

DÉMONSTRATION. On a

$$\forall x \in A^\perp, \forall y \in A,\ \langle u(x), y\rangle = \pm\langle x, u(y)\rangle = 0,$$

selon que l'endomorphisme u est auto-adjoint ou anti-auto-adjoint. Donc l'image par u d'un élément de $A^\perp$ est orthogonal à tout élément de A. Ainsi $u(F^\perp) \subset F^\perp$. □

Théorème 4.44 (diagonalisation des endomorphismes auto-adjoints) *Soit un espace euclidien et un endomorphisme auto-adjoint de cet espace. Alors, l'endomorphisme est diagonalisable et ses sous-espaces propres sont deux à deux orthogonaux. Il existe donc une base orthonormale de l'espace formée de vecteurs propres.*

DÉMONSTRATION. On note E l'espace et u l'endomorphisme auto-adjoint de l'énoncé. On montre tout d'abord que le polynôme caractéristique de l'endomorphisme est scindé sur $\mathbb{R}$. On choisit pour cela une base orthonormée de E et on note M la matrice de l'endomorphisme u dans cette base. La matrice M est réelle symétrique. Il existe une valeur propre λ de M, *a priori* complexe, et une matrice colonne non nulle $Z = (Z_i)$ *a priori* de coordonnées complexes, telles que $MZ = \lambda Z$. Notons pour un vecteur $Y = (Y_i)$, le vecteur $\overline{Y}$ comme $\overline{Y} = (\overline{Y_i})$. On remarque alors que comme M est symétrique réelle

$$(\overline{MZ})^{\perp} Z = \overline{\lambda}\,\overline{Z}^{\perp} Z = \overline{\lambda} \sum_i |Z_i|^2 \text{ et } (\overline{MZ})^{\perp} Z = \overline{Z}^{\perp}\overline{M}^{\perp} Z = \overline{Z}^{\perp} M^{\perp} Z = \overline{Z}^{\perp} M Z = \lambda\,\overline{Z}^{\perp} Z = \lambda \sum_i |Z_i|^2.$$

Comme Z n'est pas nul, on a $\sum_i |Z_i|^2 > 0$ et on en déduit que λ est réel.

On montre ensuite que les sous-espaces propres sont deux à deux orthogonaux. Soit λ et μ deux valeurs propres distinctes de u, respectivement associées à des vecteurs propres x et y. On a alors

$$\lambda \langle x, y \rangle = \langle u(x), y \rangle = \langle x, u(y) \rangle = \mu \langle x, y \rangle,$$

d'où $(\lambda - \mu) \langle x, y \rangle = 0$ et donc $\langle x, y \rangle = 0$.

On montre enfin que u est diagonalisable. Pour cela, on suppose que l'endomorphisme possède p valeurs propres distinctes $\lambda_1, \dots, \lambda_p$ et l'on considère la somme des sous-espaces propres

$$A = E_{\lambda_1} \oplus E_{\lambda_2} \oplus \cdots \oplus E_{\lambda_p}.$$

Il est clair que le sous-espace $u(A)$ est inclus dans A et, par la proposition 4.43, on a de plus que l'orthogonal $A^{\perp}$ est stable par u. La restriction de u à $A^{\perp}$, qui est un endomorphisme autoadjoint, possède alors une valeur propre si $A^{\perp}$ n'est pas réduit au vecteur nul. Ceci est impossible, puisque tous les vecteurs propres de u appartiennent à A par construction. Ainsi, le sous-espace A est égal à E et l'endomorphisme est par conséquent diagonalisable. □

On en déduit immédiatement le résultat suivant.

Corollaire 4.45 (diagonalisation des matrices symétriques réelles) *Soit n un entier naturel non nul. Pour tout matrice M réelle symétrique d'ordre n, il existe une matrice orthogonale O d'ordre n telle que $O^{\top} M O$ est une matrice diagonale réelle.*

On rappelle qu'une matrice M symétrique est dite positive si la forme quadratique réelle qui lui est canoniquement associée est positive. Elle est dite définie positive si la forme quadratique réelle est définie positive. Le dernier corollaire montre par conséquent que M est positive si et seulement si toutes ses valeurs propres sont positives, et définie positive si et seulement si toutes ses valeurs propres sont strictement positives. Plus généralement, la signature d'une forme quadratique réelle est donnée par les nombres de valeurs propres strictement positives et strictement négatives d'une matrice qui la représente.

Corollaire 4.46 *Soit une forme quadratique réelle sur un espace euclidien. Alors, il existe une base orthonormale de l'espace relativement à laquelle la matrice de la forme est diagonale.*

DÉMONSTRATION. On note E l'espace et q la forme quadratique de l'énoncé. Soit $\mathscr{B}$ une base orthonormée de E et M la matrice de la forme q relativement à la base $\mathscr{B}$. Cette dernière est symétrique et, d'après le précédent corollaire, il existe une matrice orthogonale O telle que la matrice $O^{\top} M O$ est diagonale réelle. La matrice O définit alors un changement de base faisant passer de $\mathscr{B}$ à une autre base orthonormée de E, qui est orthogonale pour la forme q (puisque la matrice de cette dernière relativement à cette nouvelle base est diagonale). □

On peut relever une différence entre ce corollaire et le théorème 3.17 assurant l'existence d'une base q-orthogonale. En effet, la base diagonalisant la forme quadratique a ici la propriété additionnelle d'être orthonormée pour le produit scalaire sur l'espace.

Une application importante du résultat de diagonalisation d'un endomorphisme auto-adjoint est donnée par le corollaire suivant.

Corollaire 4.47 (décomposition en valeurs singulières) *Soit m et n des entiers naturels non nuls. Toute matrice M de $M_{m,n}(\mathbb{R})$ admet une* ***décomposition en valeurs singulières****, c'est-à-dire qu'il existe des matrices U de $M_m(\mathbb{R})$ et V de $M_n(\mathbb{R})$, toutes deux orthogonales, et Σ de $M_{m,n}(\mathbb{R})$, « diagonale » à coefficients positifs, telles que*

$$M = U \Sigma V^{\top},$$

les coefficients non nuls de Σ, appelés ***valeurs singulières*** *de M et notés σ_i, $i = 1, \ldots, r$, étant déterminés de manière unique. Si ces valeurs singulières sont distinctes, les colonnes des matrices U et V, respectivement appelées* ***vecteurs singuliers à gauche*** *et* ***à droite*** *de M, sont également déterminées de façon unique à un facteur multiplicatif unitaire près.*
Les valeurs singulières sont les valeurs propres non nulles des matrices $M^\top M$ et $MM^\top$. Les colonnes de U sont les vecteurs propres de $MM^\top$, celles de V les vecteurs propres de $M^\top M$, et l'on a

$$\forall k \in \{1, \ldots, r\},\ {U_k}^\top M = \sigma_k V_k \text{ et } MV_k = \sigma_k U_k.$$