Algèbre linéaire 3

Guillaume Legendre

(version préliminaire du 24 août 2023)

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

DAUPHINE
UNIVERSITÉ PARIS

Table des matières

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 Sous-espaces stables par un endomorphisme

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

Définition 1.1 (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*

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

Le résultat suivant est immédiat.

Théorème 1.2 *Soit E un $\mathbb{K}$-espace vectoriel et u un endomorphisme de E. Si A est un sous-espace vectoriel stable par u, alors la restriction de u à A, notée $u_{|_A}$, définie un endomorphisme de A.*

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

Théorème 1.3 *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. $\square$

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

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

1.2 Éléments propres d'un endomorphisme

Définitions 1.4 (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 1.5 (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* $\mathrm{Sp}(u)$.

---

[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.

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 1.6 *Un endomorphisme peut ne pas avoir de valeur propre, comme le montrent les exemples suivants. A ECRIRE rotation d'angle $\frac{\pi}{2}$ dans $\mathbb{R}^2$ n'admet pas de valeur propre réelle.*

Proposition et définition 1.7 (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

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

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

Proposition 1.8 *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 1.9 *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.*

1.3 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 1.10 (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 1.11 *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 1.12 *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 1.13 *Les valeurs propres d'une matrice triangulaire sont ses coefficients diagonaux.*

Définition 1.14 (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.*

EXEMPLE du polynôme caractéristique associé à la matrice de rotation dans $\mathbb{R}^2$

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 1.15 (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 1.16 *Pour une matrice M d'ordre deux, 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 1.17 (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. □

1.4 Diagonalisation

Proposition 1.18 *Soit E un $\mathbb{K}$-espace vectoriel de dimension finie non nulle et u un endomorphisme de E. Si λ est une valeur propre de u d'ordre de multiplicité algébrique m_λ, alors on a*

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

DÉMONSTRATION. Soit $(e_1, \dots, e_{d_\lambda})$ une base du sous-espace propre E_λ, l'entier d_λ étant la dimension de E_λ, que l'on complète, si nécessaire, en une base $\mathscr{B}$ de E. La matrice représentative M de u dans la base $\mathscr{B}$ s'écrit

$$M = \begin{pmatrix} \lambda I_{d_\lambda} & M_{12} \\ 0 & M_{22} \end{pmatrix},$$

où la matrice M_{22} est d'ordre $n - d_\lambda$, avec n la dimension de E. En utilisant cette écriture par blocs, on trouve que

$$\chi_u(X) = \det(X I_n - M) = \det\left(\begin{pmatrix} (X-\lambda) I_{d_\lambda} & -M_{12} \\ 0 & X I_{n-d_\lambda} - M_{22} \end{pmatrix}\right) = (X-\lambda)^{d_\lambda} \chi_{M_{22}}(X).$$

Par conséquent, le polynôme $(X - \lambda)^{d_\lambda}$ divise $\chi_u(X)$ et l'ordre de multiplicité de λ est donc supérieur ou égal à d_λ. □

Remarque 1.19 *On a en particulier que le sous-espace propre E_λ est de dimension égale à un si λ est racine simple du polynôme caractéristique.*

On nomme parfois *ordre de multiplicité géométrique* de la valeur propre λ l'entier $\dim(E_\lambda)$, en contraste avec l'ordre de multiplicité algébrique m_λ.

Définition 1.20 (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 1.21 *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

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

Théorème 1.22 (condition nécessaire et suffisante de diagonalisabilité) *Soit un endomorphisme d'un $\mathbb{K}$-espace vectoriel de dimension finie non nulle. 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. Soit E un $\mathbb{K}$-espace vectoriel de dimension finie non nulle égale à n et u un endomorphisme de E.

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 1.8, il existe donc une base de E formée de vecteurs propres de u et la proposition 1.21 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

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

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

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

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 1.23 (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.*

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

$$M = P \operatorname{diag}(\lambda_1, \dots, \lambda_n) P^{-1}.$$

1.5 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 1.24 (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 1.25 (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.

La preuve du résultat suivant est laissée au lecteur.

Corollaire 1.26 *Soit E un $\mathbb{C}$-espace vectoriel de dimension finie non nulle et u un endomorphisme de E pour lequel le polynôme caractéristique peut s'écrire*

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

On a alors

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

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.

1.6 Polynômes annulateurs

Soit $P(X)=a_0+a_1X+a_2X^2+\dots$ 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 $u^k=\underbrace{u\circ\cdots\circ u}_{k\text{ termes}}$.

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

Proposition 1.27 *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.

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

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

Définition 1.28 (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 : 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 1.29 *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.

Théorème 1.30 (« théorème de Cayley–Hamilton ») *Soit u un endomorphisme d'un $\mathbb{K}$-espace vectoriel de dimension finie non nulle. On a*

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

---

[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.

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 1.31 *Dans la démonstration ci-dessus, on a cherché le plus petit sous-espace vectoriel stable par u contenant le vecteur 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 *(voir* (1.2)*). L'endomorphisme induit par u sur le sous-espace est alors dit* cyclique.

Définition 1.32 (polynôme minimal) *Un **polynôme minimal** d'un endomorphisme est un polynôme annulateur de cet endomorphisme, unitaire et de degré minimal.*

Proposition 1.33 *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, et celui-ci divise tout polynôme annulateur de u*

DÉMONSTRATION. Notons E l'espace vectoriel de l'énoncé et n sa dimension. Considérons 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 $\mathbb{N}$, car elle contient l'entier n en vertu du théorème de Cayley–Hamilton. 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 et le polynôme P est donc multiple de Π.

Enfin, si 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. □

Proposition 1.34 *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 1.29, 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 une donc valeur propre de u. □

Le résultat suivant caractérise le polynôme minimal d'un endomorphisme dont le polynôme caractéristique est scindé.

Théorème 1.35 *Soit E un $\mathbb{K}$-espace vectoriel de dimension finie non nulle et u un endomorphisme de E dont le polynôme caractéristique est scindé :*

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

où les entiers naturels m_{λ_i}, $i=1,\dots,p$, sont non nuls et les valeurs propres λ_i, $i=1,\dots,p$, sont deux à deux distinctes. Alors on a

$$\mu_u(X)=\prod_{i=1}^{p}(X-\lambda_i)^{l_{\lambda_i}},$$

où les entiers naturels l_{λ_i}, $i=1,\dots,p$, sont respectivement compris entre 1 et m_{λ_i}.

DÉMONSTRATION. On a précédemment montré que toute racine du polynôme minimal μ_u est une valeur propre de u. On a donc nécessairement que l_{λ_i} est supérieur ou égal à 1 pour tout entier i de $\{1, \dots, p\}$. On sait par ailleurs que le polynôme minimal μ_u divise le polynôme caractéristique χ_u, d'où l_{λ_i} est inférieur ou égal à m_{λ_i} pour tout entier i de $\{1, \dots, p\}$. □

exemple : différence entre polynômes minimal et caractéristique pour une homothétie de rapport λ, $u = \lambda\, id_E$. La matrice de u dans n'importe quelle base de E est diagonale, de coefficients diagonaux tous égaux à λ, d'où $\chi_u(X) = (X - \lambda)^n$ alors que $\mu_u(X) = X - \lambda$.

L'exemple précédent 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 la restriction de l'endomorphisme est une homothétie. Cette observation conduit à une nouvelle caractérisation de la diagonalisation.

Théorème 1.36 *Un endomorphisme d'un $\mathbb{K}$-espace vectoriel de dimension finie non nulle dont le polynôme caractéristique est scindé sur $\mathbb{K}$ est diagonalisable si et seulement si son polynôme minimal est scindé à racines simples.*

DÉMONSTRATION. Notons E l'espace vectoriel et u l'endomorphisme de l'énoncé. Considérons le polynôme $P(X) = \prod_{i=1}^{p}(X - \lambda_i)$ dont les racines sont les valeurs propres de u. Soit x un vecteur propre de u associé à une valeur propre λ_j. L'image de x par $P(u)$ est $P(u)(x) = \prod_{i=1}^{p}(\lambda_j - \lambda_i)x = 0_E$.

Supposons que u est diagonalisable. Il existe alors une base E formée de vecteurs propres de u et $P(u)$ s'annule sur chacun des vecteurs composant cette base. Le polynôme P est donc un polynôme annulateur de u, multiple du polynôme minimal μ_u. Or, par le théorème 1.35, on a que P divise μ_u, d'où $P = \mu_u$.

Réciproquement, supposons que le polynôme minimal de u soit

$$\mu_u(X) = \prod_{i=1}^{p}(X - \lambda_i).$$

Les sous-espaces propres étant en somme directe, on doit simplement montrer que tout vecteur de E s'écrit comme une combinaison linéaire de vecteurs propres. Par une décomposition en éléments simples, on a

$$\frac{1}{\prod_{k=1}^{p}(X - \lambda_k)} = \sum_{i=1}^{p}\left(\frac{1}{\prod_{j=1, j\neq i}^{p}(\lambda_i - \lambda_j)}\right)\frac{1}{X - \lambda_i}.$$

En multipliant cette égalité par le dénominateur du membre de gauche, il vient[3]

$$1 = \sum_{i=1}^{p}\left(\frac{1}{\prod_{j=1, j\neq i}^{p}(\lambda_i - \lambda_j)}\right)\prod_{k=1, k\neq i}^{p}(X - \lambda_k).$$

Posons alors $\alpha_i = \frac{1}{\prod_{j=1, j\neq i}^{p}(\lambda_i - \lambda_j)}$, $P_i(X) = \prod_{k=1, k\neq i}^{p}(X - \lambda_k)$. On a que

$$id_E = \sum_{i=1}^{p}\alpha_i\, P_i(u),$$

d'où

$$\forall x \in E,\ x = \sum_{i=1}^{p}\alpha_i\, P_i(u)(x).$$

Enfin, en utilisant que

$$\forall i \in \{1, \dots, p\},\ P_i(X)(X - \lambda_i) = \prod_{j=1}^{p}(X - \lambda_j) = \mu_u(X),$$

on trouve que

$$0_E = \mu_u(u)(x) = (u - \lambda_i\, id_E)(P_i(u)(x)),$$

---

[3] C'est l'écriture du polynôme constant et égal à 1 dans la base des polynômes de Lagrange associés aux valeurs propres de u.

d'où $P_i(u)(x)$ appartient à E_{λ_i}. □

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 1.37 (« 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, toujours en utilisant l'identité, $x = 0_E + 0_E = 0_E$, d'où $\ker(P(u)) \cap \ker(Q(u)) = \{0_E\}$. □

Il est possible de généraliser ce résultat par récurrence.

Corollaire 1.38 (« 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*

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

Une ultime caractérisation de la diagonalisabilité découle de ce dernier résultat.

Théorème 1.39 *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.

L'implication directe du théorème découle du théorème 1.36. 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 1.38 permet d'écrire que

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

Une base de E adaptée[4] à cette décomposition en somme directe est une base de vecteurs propres de u et l'endomorphisme est donc diagonalisable. □

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. Ci-dessous, on a noté m_{λ_i} l'ordre de multiplicité algébrique de la valeur propre λ_i de u et n_{λ_i} la dimension du sous-espace propre associé $\ker(u - \lambda_i \, id_E)$, $i = 1, \dots, p$, χ_u le polynôme caractéristique de u et μ_u le polynôme minimal de u.

u est diagonalisable $\iff$ il existe une base de E dans laquelle la matrice de u est diagonale,
$\iff$ il existe une base de E constituée de vecteurs propres de u,
$\iff$ E est la somme directe des sous-espaces propres de u,
$\iff$ $n = \sum_{i=1}^{p} n_{\lambda_i}$,
$\iff$ χ_u est scindé sur $\mathbb{K}$ et, $\forall i \in \{1, \dots, p\}$, $n_{\lambda_i} = m_{\lambda_i}$,
$\iff$ μ_u est scindé sur $\mathbb{K}$, à racines simples,
$\iff$ il existe un polynôme P non nul, scindé sur $\mathbb{K}$, à racines simples et tel que $P(u) = 0_{\mathscr{L}(E)}$,
$\impliedby$ u possède n valeurs propres simples $\iff$ χ_u est scindé sur $\mathbb{K}$, à racines simples.

u est trigonalisable $\iff$ il existe une base de E dans laquelle la matrice de u est triangulaire supérieure,
$\iff$ χ_u est scindé sur $\mathbb{K}$,
$\iff$ E est la somme directe des sous-espaces caractéristiques de u.

1.7 Réduction de Jordan (hors programme)

On peut aller plus loin que la simple trigonalisation en construisnat une base de l'espace 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.

1.7.1 Sous-espaces caractéristiques

Définition 1.40 (sous-espace caractéristique) *Soit E un $\mathbb{K}$-espace vectoriel, u un endomorphisme de E et λ une valeur propre de u. On appelle* ***sous-espace caractéristique*** *(ou* ***sous-espace spectral****, ou encore* ***sous-espace propre généralisé****) de u associé à λ le sous-espace vectoriel*

$$F_\lambda = \ker\left((u - \lambda \, id_E)^{l_\lambda}\right),$$

où l'entier l_λ est l'ordre de multiplicité de λ en tant que racine du polynôme minimal de u. Un vecteur non nul x de E appartenant à N_λ est un ***vecteur propre généralisé*** *de u associé à λ.*

Proposition 1.41 *On a les propriétés suivantes.*

- N_λ est stable par u,

- $\dim(F_\lambda) = m_\lambda$, où m_λ est l'ordre de multiplicité algébrique de λ (c'est-à-dire en tant que racine du polynôme caractéristique),

- $E = \bigoplus_{i=1}^{p} F_{\lambda_i}$.

---

[4] 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.

DÉMONSTRATION. A ECRIRE □

Remarques : (à voir si c'est clair dans la preuve) l'exposant l_λ est celui pour lequel le noyau apparaissant dans la définition atteint sa dimension maximale. On a ainsi la chaîne suivante :

$$\ker(u-\lambda\, id_E) \subset \ker\big((u-\lambda\, id_E)^2\big) \subset \dots \subset \ker\big((u-\lambda\, id_E)^{l_\lambda}\big) = \ker\big((u-\lambda\, id_E)^{l_\lambda+1}\big) = \dots$$

Le dernier résultat de la proposition montre qu'un endomorphisme d'un espace vectoriel est trigonalisable si et seulement si l'espace est la somme (directe) des sous-espaces caractéristiques de l'endomorphisme, c'est-à-dire si et seulement s'il existe une base de l'espace formée de vecteurs propres généralisés de l'endomorphisme. Cette caractérisation, déjà donnée dans le tableau récapitulatif en fin de section précédente, rejoint celle basée sur le polynôme caractéristique, qui doit être scindé pour que l'endomorphisme soit trigonalisable.

1.7.2 Décomposition de Jordan–Chevalley

Théorème 1.42 (« décomposition de Jordan–Chevalley ») *Soit u un endomorphisme dont le polynôme caractéristique est scindé sur $\mathbb{K}$. Alors, il existe un unique couple (d, n) d'endomorphismes de E, avec d diagonalisable et n nilpotent, tels que $u = d + n$, $n \circ d = d \circ n$. De plus, il existe des polynômes P_d et P_n tels que et $d = P_d(u)$ et $n = P_n(u)$.*

DÉMONSTRATION. A REPRENDRE
Observons tout d'abord que si $u = d + n$ alors $n = u - d$ et il suffit donc de montrer que l'endomorphisme d est entièrement déterminé par u, c'est-à-dire, puisque d est diagonalisable, de déterminer ses valeurs propres λ_i et les sous-espaces propres associés E_{λ_i} en fonction de u.

Puisque l'endomorphisme n commute avec d, les sous-espaces propres E_{λ_i} sont stables par n et donc par u. Notons u_i la restriction de u à E_{λ_i}, n_i la dimension de E_{λ_i} et ν_i l'indice de nilpotence de la restriction de n à E_{λ_i}, c'est-à-dire $u_i - \lambda_i\, id_{E_{\lambda_i}}$ (on a $\nu_i \leq n_i$).

Le polynôme caractéristique, le polynôme minimal et le spectre de $u_i - \lambda_i\, id_{E_{\lambda_i}}$ sont respectivement X^{n_i}, X^{ν_i} et $\{0\}$, ceux de u_i sont respectivement $(X-\lambda_i)^{n_i}$, $(X-\lambda_i)^{\nu_i}$ et $\{\lambda_i\}$ et ceux de u sont $\prod_{k=1}^{p}(X-\lambda_k)^{n_k}$, le plus petit commun diviseur ? ? ? et $\{\lambda_k\}_{k=1,\dots,p}$. Les valeurs propres de d sont ainsi celle de u. Les sous-espaces propres de d se déduisent aussi de u car ce sont les sous-espaces caractéristiques $\ker((u-\lambda_i\, id_E)^{\nu_i})$. Ils sont stables par u et le lemme de décomposition des noyaux montre que E est somme directe de ces sous-espaces vectoriels. Ceci permet de définir d comme l'endomorphisme de E dont la restriction à chaque E_{λ_i} est l'homothétie de rapport λ_i. □

Remarque : on donne parfois à ce résultat le nom de décomposition de Dunford.

Remarque : application au calcul de l'exponentielle d'un endomorphisme $\exp(u) = \exp(d+n) = \exp(d)\exp(n)$ car d et n commutent. Le calcul de $\exp(d)$ est simple dans une base de diagonalisation et $\exp(n) = \sum_{k=0}^{q-1} \frac{n^k}{k!}$ où q est l'indice de nilpotence de n.

1.7.3 Forme de Jordan

Définitions 1.43 (bloc et forme de Jordan) *Soit l un entier naturel non nul. On a appelle **bloc de Jordan de taille** l une matrice de $M_l(\mathbb{K})$ de la forme*

$$J_l(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \dots & \dots & 0 & \lambda \end{pmatrix},$$

avec λ un scalaire.
*Une matrice est dite sous **forme de Jordan** si elle est diagonale par blocs, chaque bloc diagonal étant un bloc de Jordan.*

Note : une matrice sous forme de Jordan est diagonale si chaque bloc de Jordan est de taille 1.

Forme réduite des endomorphismes nilpotents (lien avec la réduction de Frobenius ?)

Proposition 1.44 (réduction des endomorphismes nilpotents) *Soit u un endomorphisme nilpotent de E. Il existe une base dans laquelle la matrice de u a la forme*

$$\begin{pmatrix} 0 & a_1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & & & \ddots & a_{n-1} \\ 0 & \dots & \dots & \dots & 0 \end{pmatrix},$$

avec $a_i \in \{0,1\}$.

DÉMONSTRATION. A ECRIRE ☐

La décomposition de Jordan–Chevalley alliée à la forme réduite des endomorphismes nilpotents conduit à la réduction de Jordan des endomorphismes.

Théorème 1.45 (réduction de Jordan) *Soit u un endomorphisme de E dont le polynôme caractéristique est scindé sur $\mathbb{K}$. Il existe une base de l'espace dans laquelle la matrice de u est sous forme de Jordan.*

AJOUTER somme des tailles des blocs vaut n les scalaires des blocs sont les valeurs propres, non nécessairement distinctes, de u

DÉMONSTRATION. Notons n la dimension de E et $\lambda_1, \dots, \lambda_p$ les valeurs propres, deux à deux distinctes, de u. Le polynôme caractérisque de u étant scindé sur $\mathbb{K}$, il existe des entiers naturels non nuls $m_1, \dots, m_p$ tels que

$$\chi_u(X) = (X - \lambda_1)^{m_1} \dots (X - \lambda_p)^{m_p}.$$

Il découle alors du lemme de décomposition des noyaux généralisé (voir le corollaire 1.38) que

$$E = \bigoplus_{i=1}^{p} \ker\left((u - \lambda_i\, id_E)^{m_i}\right).$$

Pour tout entier i dans $\{1, \dots, p\}$, notons F_i le sous-espace caractéristique associé à λ_i, $F_i = \ker\left((u - \lambda_i\, id_E)^{m_i}\right)$. Ce sous-espace vectoriel est stable par u et la restriction de u à F_i peut s'écrire $u_{|_{F_i}} = \lambda_i id_{F_i} + N_i$, où $N_i = (u - \lambda_i id_E)_{|_{F_i}}$ est un endomorphisme nilpotent. D'après le proposition 1.44, il existe une base $\mathscr{B}_i$ de F_i dans laquelle la matrice de N_j est diagonale par blocs de la forme $J(0)$ et la matrice $u_{|_{F_i}}$ dans cette base est donc diagonale par blocs de la forme $J(\lambda_i)$. En concaténant alors les bases ainsi obtenues pour chacun des sous-espaces caractéristiques de U, on construit une base de E dans laquelle la matrice de l'endomorphisme est sous la forme de Jordan voulue. ☐

1.8 Applications de la réduction

1.8.1 Calcul des puissances d'une matrice

Étant donné une matrice A de $M_n(\mathbb{K})$, supposons que l'on souhaite calculer les puissances positives de A. Si $A = PBP^{-1}$, alors, pour tout entier naturel k, on a $A = PB^kP^{-1}$ et, si le calcul des puissances de B est plus simple que celui des puissances de A, ce qui est par exemple le cas si B est une matrice diagonale, on a intérêt à utiliser une telle réduction. Lorsque A est inversible, notons que cette approche permet de calculer l'inverse de A, ou plus généralement toute puissance négative de A.

Chapitre 2

Formes bilinéaires

Une *application bilinéaire* est l'analogue à deux variables d'une application linéaire.

2.1 Généralités sur les applications bilinéaires

Étant donné trois $\mathbb{K}$-espaces vectoriels E, F et G, une application b de $E \times F$ dans G est dite *bilinéaire* si et seulement si

$$\forall (x,x') \in E^2,\ \forall (y,y') \in F^2,\ \forall \lambda \in \mathbb{K},$$
$$b(x+x',y) = b(x,y) + b(x',y),\ b(x,y+y') = b(x,y) + b(x,y'),\ b(\lambda\, x,y) = b(x,\lambda\, y) = \lambda\, b(x,y).$$

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, F et G trois $\mathbb{K}$-espaces vectoriels. Une application b de $E \times F$ dans G est dite* ***bilinéaire*** *si et seulement si l'****application partielle à gauche de*** *b, notée $L(b)$, de E dans $\mathscr{L}(F;G)$ 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 F dans $\mathscr{L}(E;G)$ définie par*

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

sont toutes deux linéaires.

Lorsque $G = \mathbb{K}$, on parle de *forme bilinéaire*.

L'ensemble des applications bilinéaires de $E \times F$ dans G, noté $\mathscr{L}(E,F;G)$, est un espace vectoriel sur $\mathbb{K}$. Il est en effet clair que la somme de deux applications bilinéaires est une application bilinéaire et que le produit d'une application bilinéaire par un scalaire est une application bilinéaire.

Remarque 2.2 *Dans la précédente définition, on a en fait défini une application $L : \mathscr{L}(E,F;G) \to \mathscr{L}(E;\mathscr{L}(F;G))$ (resp. $R : \mathscr{L}(E,F;G) \to \mathscr{L}(F;\mathscr{L}(E;G))$). 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 F, $R(b)(y)$) sont toutes trois linéaires.

Définition 2.3 (application bilinéaire symétrique) *Soit E et G deux $\mathbb{K}$-espaces vectoriels et b une application bilinéaire de $E \times E$ dans G. On dit que l'application b est* ***symétrique*** *(resp.* ***antisymétrique****) si et seulement si*

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

exemples : 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 bilinéaire 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.

Soit m, n et p des entiers naturels non nuls. Le produit matriciel de $M_{m,n}(\mathbb{R})\times M_{n,p}(\mathbb{R})$ dans $M_{m,p}(\mathbb{R})$ est une application bilinéaire.

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

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

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

Le *crochet de dualité* $\langle\cdot,\cdot\rangle_{E^*,E}$, défini par

$$\forall x\in E,\ \forall\varphi\in E^*,\ \langle\varphi,x\rangle_{E^*,E}=\varphi(x),$$

est une forme bilinéaire sur $E^*\times E$.

2.2 Représentation matricielle d'une forme bilinéaire

Nous nous limitons à partir de maintenant aux *formes* bilinéaires, c'est-à-dire qu'on aura $G=\mathbb{K}$ dans la suite. Dans ce cas particulier, considérant une forme bilinéaire b de $E\times F$ dans $\mathbb{K}$, l'application linéaire à gauche de b, $L(b)$, envoie E dans $\mathscr{L}(F;\mathbb{K})$, c'est-à-dire dans F^*, le dual de F. De la même manière, l'application linéaire à droite de b, $R(b)$, envoie F dans E^*.

On suppose que $\mathscr{B}=\{e_1,\dots,e_n\}$ soit une base de E et que $\mathscr{C}=\{f_1,\dots,f_m\}$ soit une base de F. On note alors $\mathscr{B}^*$ la base de E^* et $\mathscr{C}^*$ la base de F^* qui sont les bases duales respectives de $\mathscr{B}$ et de $\mathscr{C}$.

On considère la matrice N de l'application $L(b)$ dans les bases $\mathscr{B}$ et $\mathscr{C}^*$. De la même manière, la matrice M représente l'application $R(b)$ dans les bases $\mathscr{C}$ et $\mathscr{B}^*$. On a le résultat suivant.

Lemme 2.4 *La matrice N est la transposée de la matrice M, i.e.*

$$\forall i\in\{1,\dots,m\},\ \forall j\in\{1,\dots,n\},\ n_{ij}=m_{ji}.$$

DÉMONSTRATION. On remarque tout d'abord que la matrice N appartient à $M_{m,n}(\mathbb{K})$ puisque E est de dimension n et F^* de dimension m. De la même façon, la matrice M appartient à $M_{n,m}(\mathbb{K})$.

Ensuite, le coefficient n_{ij} est la i^e coordonnée de la forme $L(b)(e_j)$ dans la base duale $\mathscr{C}^*$. Par propriété de la base duale, il vaut par conséquent $L(b)(e_j)(f_i)$, soit encore $b(e_j,f_i)$. De la même façon, le coefficient m_{ji} est la j^e coordonnée de la forme $R(b)(f_i)$ dans la base duale $\mathscr{B}^*$ et vaut donc $R(b)(f_i)(e_j)=b(e_j,f_i)$. □

Définition 2.5 (matrice d'une forme bilinéaire) *Soit E et F deux espaces vectoriels de dimensions finies non nulles, $\mathscr{B}$ une base de E, $\mathscr{C}$ une base de F et b une forme bilinéaire sur $E\times F$. La* ***matrice de** b **relativement aux bases** $\mathscr{B}$ **et** $\mathscr{C}$ *est celle de l'application $R(b)$ dans les bases $\mathscr{C}$ et $\mathscr{B}^*$.*

Remarque 2.6 *L'application R étant bijective, il y a une correspondance bijective entre les formes bilinéaires sur $E\times F$ et les matrices à n lignes et m colonnes quand les espaces E et F sont respectivement de dimension n et m. Cet isomorphisme* dépend *du choix des bases sur E et F, et n'est donc pas canonique.*

Lemme 2.7 (expression des coefficients de la matrice d'une forme bilinéaire) *Soit E et F deux espaces vectoriels de dimensions finies non nulles, $\mathscr{B}=\{e_1,\dots,e_n\}$ une base de E et $\mathscr{C}=\{f_1,\dots,f_m\}$ une base de F. Le coefficient à la i^e ligne et la j^e colonne de la matrice de la forme bilinéaire b relativement aux bases $\mathscr{B}$ et $\mathscr{C}$ est égal à $b(e_i,f_j)$.*

DÉMONSTRATION. Le coefficient en question a précédemment été noté m_{ij} et l'on a vu dans la preuve du dernier lemme qu'il était égal à $b(e_i, f_j)$. ☐

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

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

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 $\mathbb{R}^n$.

Proposition 2.8 *Soit E un espace vectoriel de dimension finie non nulle, $\mathscr{B}$ une base de E et b une forme bilinaire 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 bases

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

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

d'où, par identification,

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

Remarque 2.9 *On notera que la matrice transposée $P^\top$ est généralement différente de l'inverse P^{-1}. Par conséquent, même lorsque $F = E$ et $\mathscr{C} = \mathscr{B}$ (et donc $Q = P$), 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.*

2.3 Non dégénérescence

Définition 2.10 (non dégénérescence) *Une forme bilinéaire sur $E \times F$ est dite* ***non dégénérée à gauche*** *(resp.* ***à droite****) si et seulement si l'application linéaire associée à gauche $L(b)$ (resp. à droite $R(b)$) est injective. Elle est dite* ***non dégénérée*** *si et seulement si elle est non dégénérée à gauche et à droite.*

Par conséquent, une forme bilinéaire b est non dégénérée à gauche (resp. à droite) si et seulement si $\ker(L(b)) = \{x \in E \mid \forall y \in F,\ b(x,y) = 0\} = \{0_E\}$ (resp. $\ker(R(b)) = \{y \in F \mid \forall x \in E,\ b(x,y) = 0\} = \{0_F\}$), ce sous-espace étant appelé le *noyau à gauche* (resp. *à droite*) de la forme b.

Lemme 2.11 *Si $E = F$ et que la forme bilinéaire est symétrique, la non dégénérescence à gauche est équivalente à la non dégénérescence à droite et les noyaux à gauche et à droite coïncident.*

DÉMONSTRATION. Dans ce cas, on a $L(b) = R(b)$. ☐

Proposition 2.12 *Si les espaces vectoriels E et F sont de dimensions finies et égales et b est une forme bilinéaire sur $E \times F$, les conditions suivantes sont équivalentes :*
- — *La forme b est non dégénérée à droite.*
- — *La forme b est non dégénérée à gauche.*
- — *L'application $L(b)$ est injective.*
- — *L'application $R(b)$ est injective.*
- — *L'application $L(b)$ est surjective.*
- — *L'application $R(b)$ est surjective.*

DÉMONSTRATION. La proposition découle du fait que la matrice de l'application $R(b)$, qui est carrée dans ce cas particulier, est la transposée de la matrice de l'application $L(b)$ et du théorème du rang. ☐

Corollaire 2.13 *Si les espaces vectoriels E et F sont de dimensions finies et égales, une forme bilinéaire sur $E \times F$ est non dégénérée si et seulement si le déterminant de sa matrice relativement à des bases quelconques de E et de F est non nul.*

Remarque 2.14 *La restriction à un sous-espace vectoriel d'une forme bilinéaire non dégénérée peut être dégénérée.*

exemple : la forme définie sur $\mathbb{R}^2 \times \mathbb{R}^2$ par

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

de matrice $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ relativement à la base canonique, est non dégénérée. En effet, l'application $R(b)$ envoie e_1 sur ${e_1}^*$ et e_2 sur $-{e_2}^*$, elle est donc surjective. Pourtant, sa restriction à la droite vectorielle engendrée par $e_1 + e_2$ est dégénérée, car identiquement nulle.

Définition 2.15 (rang d'une application bilinéaire) *Soit E et F des espaces vectoriels de dimensions finies et égales et b une forme bilinéaire sur $E \times F$. Le* ***rang*** *de la forme b est le rang de l'application linéaire associée à droite de b.*

Remarque 2.16 *On aurait tout aussi bien pu prendre pour définition du rang de b le rang de l'application linéaire associée à* gauche*, puisque $L(b)$ et $R(b)$ ont même rang d'après la proposition précédente.*

Par définition de la matrice représentative d'une forme bilinéaire, le rang d'une forme bilinéaire est égal au rang de sa matrice. Ce dernier est bien entendu indépendant du choix des bases, les matrices de changement de base étant des matrices inversibles.

2.4 Formes sesquilinéaires

Lorsque les espaces vectoriels considérés sont construit sur le corps $\mathbb{C}$, on fait couramment appel à la notion de *forme sesquilinéaire*, qui constitue l'équivalent[1] complexe des formes bilinéaires à valeurs réelles.

Définition 2.17 (forme sesquilinéaire) *Soit E et F deux $\mathbb{C}$-espaces vectoriels. Une application s de $E \times F$ dans $\mathbb{C}$ est une* ***forme sesquilinéaire à gauche*** *si et seulement si*
- — *elle est linéaire à droite,*

$$\forall x \in E,\ \forall y \in F,\ \forall z \in F,\ \forall \lambda \in \mathbb{C},\ s(x, y + \lambda z) = s(x,y) + \lambda s(x,z),$$

- — *elle est semi-linéaire à gauche,*

$$\forall x \in E,\ \forall z \in E,\ \forall y \in F,\ \forall \lambda \in \mathbb{C},\ s(x + \lambda z, y) = s(x,y) + \overline{\lambda} s(z,y),$$

De la même manière, on peut définir une forme sesquilinéaire ***à droite****.*

Cette définition justifie l'appellation utilisée, le préfixe *sesqui-* indiquant que le mot préfixé est dans un rapport de un et demi (à gauche, la forme est linéaire par rapport à la somme de vecteurs, mais pas par rapport à la multiplication par un scalaire).

---

[1] Ceci apparaîtra plus clairement dans le prochain chapitre.

Définition 2.18 (symétrie hermitienne) *Une forme sesquilinéaire s sur $E \times E$ vérifie la propriété de* ***symétrie hermitienne*** *si et seulement si*

$$\forall (x, y) \in E \times E,\ s(y, x) = \overline{s(x, y)}.$$

Remarque 2.19 *On peut observer qu'on a en particulier*

$$\forall x \in E,\ s(x, x) = \overline{s(x, x)},$$

ces quantités sont donc réelles.

Représentation matricielle

$$s(x, y) = X^* M Y = \overline{X}^\top M Y.$$

Chapitre 3

Formes quadratiques

On aborde dans ce chapitre la notion de *forme quadratique*. 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) *Soit E un $\mathbb{K}$-espace vectoriel. Une application q de E dans $\mathbb{K}$ est appelée* ***forme quadratique*** *sur E s'il existe une forme bilinéaire b de $E \times E$ dans $\mathbb{K}$ telle que*

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

Toute forme bilinéaire b sur $E \times E$ donne naissance à une forme quadratique 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 peut ainsi être associée à plusieurs formes bilinéaires. On a toutefois le résultat suivant.

Proposition et définition 3.2 (forme polaire d'une forme quadratique) *Sur un espace vectoriel dont le corps est de caractéristique*[1] *différente de 2 (comme $\mathbb{R}$ ou $\mathbb{C}$), toute forme quadratique est associée à une unique forme bilinéaire symétrique, cette dernière étant appelée la* ***forme polaire*** *de la forme quadratique.*

DÉMONSTRATION. Soit q une forme quadratique sur une espace vectoriel E dont le corps est de caractéristique différente de 2. 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

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

---

[1] La *caractéristique* d'un anneau (unitaire) est par définition l'ordre pour la loi additive de l'élément neutre de la loi multiplicative si cet ordre est fini (s'il est infini, la caractéristique de l'anneau est nulle). Soit $(A,+,\times)$ un anneau unitaire d'élément neutre pour la loi additive « + » noté 0_A et d'élément neutre pour la loi multiplicative « × » noté 1_A. La caractéristique de A est alors le plus petit entier naturel n non nul tel que

$$n \times 1_A = \underbrace{1_A + 1_A + \cdots + 1_A}_{n \text{ termes}} = 0_A$$

si un tel entier existe. Dans le cas contraire (autrement dit si 1_A est d'ordre infini), la caractéristique de A est nulle.

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 $\mathbb{K}$ 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 $\mathbb{K}$.

Proposition 3.3 *Soit E un $\mathbb{K}$-espace vectoriel de dimension finie non nulle et $\mathscr{B}=\{e_1,\dots,e_n\}$ une base de E. Les formes quadratiques 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

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

□

Remarque 3.4 *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 3.5 (matrice, noyau et rang d'une forme quadratique) *Soit q une forme quadratique sur un* $\mathbb{K}$*-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 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 sur E, noté $\mathscr{Q}(E)$, est isomorphe à $S_n(\mathbb{K})$, l'ensemble des matrices symétriques à coefficients dans $\mathbb{K}$. 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 3.6 (forme quadratique non dégénérée) *Une forme quadratique est dite* ***non dégénérée*** *si et seulement si son noyau est réduit au vecteur nul.*

Définition 3.7 (formes quadratiques équivalentes) *Deux formes quadratiques q et q′ sur un* $\mathbb{K}$*-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

Dans toute cette section, E désigne un $\mathbb{K}$-espace vectoriel, q une forme quadratique sur E et b la forme polaire de q.

Définition 3.8 (cône isotrope) *Un vecteur x de E est dit* ***isotrope*** *pour q si et seulement si* $q(x) = 0$*. On appelle* ***cône isotrope*** *de q l'ensemble* $\mathscr{C}_q$ *des vecteurs isotropes pour q. On dit enfin que q est* ***définie*** *si et seulement si* $\mathscr{C}_q = \{0_E\}$*.*

Une forme bilinéaire 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 3.9 *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 : $E = \mathbb{R}^2$, $q(x) = x_1 x_2$, $\ker(q) = \{0_E\}$ et $\mathscr{C}_q = \{x_1 = 0\} \cup \{x_2 = 0\}$. $q(x) = {x_1}^2 - {x_2}^2$ est non dégénérée, mais pas définie.

Définition 3.10 (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* $x \perp y$*.*

Si A est une partie de E, on appelle ***orthogonal*** *de A pour la forme q (ou pour sa forme polaire b), et on note* $A^\perp$*, la partie de E dont les éléments sont orthogonaux à ceux de A,*

$$A^\perp = \{y \in E \mid \forall x \in A,\ b(x,y) = 0\}.$$

Enfin, deux sous-ensembles A et B de E sont dits orthogonaux pour la forme q (ou pour sa forme polaire b) si et seulement si

$$\forall x \in A,\ \forall y \in B,\ b(x,y) = 0,$$

ce que l'on note encore $A \perp B$.

Il est immédiat que $A^\perp$ est un sous-espace vectoriel de E. Par ailleurs, pour toute partie A de E, l'orthogonal de A est égal à celui du sous-espace engendré par A, c'est-à-dire que

$$A^\perp = \mathrm{Vect}(A)^\perp.$$

Par conséquent, on considèrera dans la suite uniquement des orthogonaux de parties qui sont des sous-espaces vectoriels de E.

On remarquera par ailleurs qu'il découle des précédentes définitions que

$$\ker(q) = E^\perp,$$

et que tout vecteur isotrope est orthogonal à lui-même.

Remarque 3.11 *La notion d'orthogonalité existe également pour les formes bilinéaires non symétriques et plus généralement définies sur $E \times F$, avec $E \neq F$. Par exemple, lorsque $F = E^*$ et que la forme correspond au crochet de dualité, l'orthogonal d'un sous-espace vectoriel A de E n'est autre que l'*annulateur *de A (c'est-à-dire l'ensemble des formes linéaires s'annulant sur A), i.e. $A^\perp = A^0$.*

Proposition 3.12 *Soit A et B des parties de E. On a*

(i) $A \subset (A^\perp)^\perp$,

(ii) $A \subset B \implies B^\perp \subset A^\perp$.

DÉMONSTRATION.

(i) Soit x un vecteur de A. On a

$$\forall y \in A^\perp,\ b(x,y) = 0,$$

d'où x est orthogonal à tout vecteur de $A^\perp$ et appartient donc à $(A^\perp)^\perp$.

(ii) Soit y un vecteur de $B^\perp$. On a

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

En particulier, cette égalité est vraie pour tout vecteur x de A, puisque A est inclus dans B. Par conséquent, le vecteur y appartient à $A^\perp$.

□

Proposition 3.13 *On suppose que l'espace E est de dimension finie. Pour tout sous-espace vectoriel A de E, on a*

(i) $\dim(A) + \dim(A^\perp) = \dim(E) + \dim(A \cap \ker(q))$,

(ii) $(A^\perp)^\perp = A + \ker(q)$.

DÉMONSTRATION.

(i) Considérons la restriction de $R(b)$ au sous-espace A, que l'on note φ. Par le théorème du rang, on a

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

On a $\ker(\varphi) = A \cap \ker(b)$ d'une part et $\mathrm{Im}(\varphi)^0 = A^\perp$ d'autre part. Par ailleurs, on sait que $\dim(E) = \dim(\mathrm{Im}(\varphi)) + \dim(\mathrm{Im}(\varphi)^0)$. On a donc

$$\begin{aligned}\dim(A^\perp) &= \dim(E) - \dim(\mathrm{Im}(\varphi))\\ &= \dim(E) - (\dim(A) - \dim(\ker(\varphi)))\\ &= \dim(E) - \dim(A) + \dim(A \cap \ker(q)).\end{aligned}$$

(ii) On sait d'après la précédente proposition que $A \subset (A^\perp)^\perp$ et on a par ailleurs que $\ker(q) \subset (A^\perp)^\perp$. On a par conséquent que $A + \ker(\varphi) \subset (A^\perp)^\perp$. En appliquant à $A^\perp$ l'égalité obtenue dans la première partie de la proposition et en utilisant que $\ker(q) \subset A^\perp$, il vient alors

$$\dim(A^\perp) + \dim((A^\perp)^\perp) = \dim(E) + \dim(A^\perp \cap \ker(q)) = \dim(E) + \dim(\ker(q)).$$

En retranchant à cette égalité celle obtenue dans la première partie de la proposition, on trouve

$$\dim((A^\perp)^\perp) - \dim(A) = \dim(\ker(q)) - \dim(A \cap \ker(q)),$$

d'où, par la formule de Grassmann,

$$\dim((A^\perp)^\perp) = \dim(A + \ker(q)),$$

ce qui permet de conclure.

□

Si E est de dimension finie et que la forme q est non dégénérée, on a obtenu que, pour tout sous-espace vectoriel A de E,

$$\dim(A^\perp) = \dim(E) - \dim(A) \text{ et } (A^\perp)^\perp = A.$$

Bien que les dimensions respectives de A et $A^\perp$ soient complémentaires, cela ne veut pas nécessairement dire que ces sous-espaces sont supplémentaires.

contre-exemple : on a précédemment vu que la forme quadratique sur $\mathbb{R}^2$ définie par $q(x) = {x_1}^2 - {x_2}^2$, de forme polaire associée $b(x, y) = x_1 y_1 - x_2 y_2$, est non dégénérée. Pour cette forme, le sous-espace vectoriel engendré par le vecteur $(1, 1)$ est son propre orthogonal.

Définition 3.14 (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*.

Théorème 3.15 *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

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

Le sous-espace H est un hyperplan de E (c'est le noyau d'une forme linéaire non nulle). 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

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

où $\{\varepsilon_i^*\}_{i=1,\dots,n}$ est la base duale de $\mathscr{B}$. On a ainsi écrit q comme une combinaison linéaire de carrés de formes linéaires linéairement indépendantes. 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 espace vectoriel de dimension n sur un corps de caractéristique différente de 2 et q une forme quadratique 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]

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

   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. Si l'entier r est strictement inférieur à n, on peut compléter la famille libre de formes linéaires trouvée en une base de E^* avec des formes linéaires $\ell_{r+1}, \dots, \ell_n$ bien choisies. Il existe alors une unique base $\{\varepsilon_i\}_{i=1,\dots,n}$ de E dont $\{\ell_i\}_{i=1,\dots,n}$ est la base duale (on parle dans ce cas de base *antéduale*), *i.e.*

$$\forall i \in \{1, \dots, n\},\ \varepsilon_i^* = \ell_i.$$

Cette base est bien telle que voulu puisque l'on a

$$\forall i \in \{1, \dots, n\},\ q(\varepsilon_i) = \begin{cases} \alpha_i & \text{si } i \in \{1, \dots, r\} \\ 0 & \text{sinon} \end{cases},$$

ce qui fournit les éléments diagonaux de la matrice de q. La forme polaire de q s'écrivant

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

on a par ailleurs

$$\forall (i,j) \in \{1, \dots, n\}^2,\ i \neq j,\ b(\varepsilon_i, \varepsilon_j) = 0.$$

Remarque 3.16 *Il existe d'autres moyens que la réduction de Gauss pour arriver à l'écriture d'une forme quadratique sous forme réduite. Il n'y a d'ailleurs pas unicité de cette écriture. Néanmoins, le procédé de Gauss possède l'avantage d'être systématique et de toujours conduire à des carrés de formes linéaires linéairement indépendantes.*

Revenons sur le calcul effectif d'une base q-orthogonale de l'espace E à partir d'une forme réduite. Connaissant n formes linéaires, linéairement indépendantes, $\ell_1, \dots, \ell_n$, il s'agit de chercher explicitement une base $\{\varepsilon_1, \dots, \varepsilon_n\}$ de E vérifiant les conditions

$$\forall (i,j) \in \{1, \dots, n\}^2,\ \ell_i(\varepsilon_j) = \delta_{ij}.$$

Étant donnée une base $\{e_1, \dots, e_n\}$ de E, dans laquelle tout vecteur x de E a pour coordonnées $(x_1, \dots, x_n)$, et $\{e_1^*, \dots, e_n^*\}$ sa base duale, on connaît la matrice de passage R de la base $\{e_1^*, \dots, e_n^*\}$ à la base $\{\ell_1, \dots, \ell_n\}$. Pour déterminer les vecteurs $\varepsilon_1, \dots, \varepsilon_n$, il suffit alors de déterminer l'expression de chacun d'eux dans la base $\{e_1, \dots, e_n\}$, c'est-à-dire de déterminer la matrice de passage P de la base $\{e_1, \dots, e_n\}$ à la base $\{\varepsilon_1, \dots, \varepsilon_n\}$.

La matrice P est la matrice de l'application identique de E en choisissant $\{\varepsilon_1, \dots, \varepsilon_n\}$ comme base de E en tant qu'espace de départ et $\{e_1, \dots, e_n\}$ comme base de E en tant qu'espace d'arrivée. On peut alors vérifier que la matrice transposée de P est la matrice de l'application identique de E^* en choisissant $\{e_1^*, \dots, e_n^*\}$ comme base de E^* en tant qu'espace de départ et $\{\varepsilon_1^*, \dots, \varepsilon_n^*\}$, c'est-à-dire $\{\ell_1, \dots, \ell_n\}$, comme base de E^* en tant qu'espace d'arrivée. On en déduit alors que

$$P^\top = R^{-1},$$

ce qui équivaut à

$$P = (R^\top)^{-1}.$$

Proposition 3.17 *On suppose que l'espace E est de dimension finie non nulle. Le noyau d'une forme quadratique q sur E est engendré par les vecteurs isotropes d'une base q-orthogonale de E.*

DÉMONSTRATION. Soit n la dimension de E, r le rang de q (supposé strictement inférieur à n) et $\mathscr{B} = \{\varepsilon_1, \dots, \varepsilon_n\}$ une base q-orthogonale de E pour laquelle $q(\varepsilon_i) \neq 0$ pour i appartenant à $\{1, \dots, r\}$ et $q(\varepsilon_i) = 0$ pour i appartenant à $\{r+1, \dots, n\}$. Pour tout vecteur x de E, on peut écrire $x = \sum_{i=1}^{n} x_i\, \varepsilon_i$, et le vecteur x appartient au noyau de q si et seulement si $b(x,y) = 0$ pour tout vecteur y de E, c'est-à-dire si et seulement si

$$\forall j \in \{1, \dots, n\},\ b(x, \varepsilon_j) = \sum_{i=1}^{n} x_i\, b(\varepsilon_i, \varepsilon_j) = x_j\, q(\varepsilon_j) = 0,$$

soit encore

$$\forall j \in \{1, \ldots, r\},\ x_j = 0.$$

On en déduit que $\ker(q) = \text{Vect}(\{\varepsilon_{r+1}, \ldots, \varepsilon_n\})$. □

On déduit que cette proposition que le rang d'une forme quadratique est donné par le nombre de vecteurs non isotropes d'une base q-orthogonale.

Corollaire 3.18 *Le noyau d'une forme quadratique q est l'intersection des noyaux des formes linéaires linéairement indépendantes apparaissant dans une forme réduite de q.*

DÉMONSTRATION. D'après la précédente proposition, le noyau de q est engendré par les vecteurs isotropes d'une base q-orthogonale de l'espace. Considérons la base antéduale associée à la base de E^* obtenue, éventuellement par complétion, à partir des formes linéaires ℓ_i, $i = 1, \ldots, r$, apparaissant dans une forme réduite de q. Cette base est q-orthogonale et, par la propriété définissant la base antéduale, ses vecteurs isotropes ε_j, $j = r+1, \ldots, n$, sont tels que

$$\forall i \in \{1, \ldots, r\},\ \ell_i(\varepsilon_j) = 0.$$

□

3.3 Classification des formes quadratiques complexes

Dans cette section, on suppose que E est un $\mathbb{C}$-espace vectoriel. On a le résultat suivant.

Théorème 3.19 *On suppose que l'espace E est de dimension finie égale à n. Soit q une forme quadratique sur E. Alors, il existe une base $\{\varepsilon_1, \ldots, \varepsilon_n\}$ de E telle que*

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

pour tout vecteur x de E tel que $x = \sum_{i=1}^{n} x_i\, \varepsilon_i$, l'entier r étant le rang de q.

DÉMONSTRATION. Il suffit de poser, à partir de la base de E^* obtenue en complétant éventuellement la famille de formes linéaires apparaissant dans une forme réduite (3.2) de q,

$$\tilde{\ell}_i = \begin{cases} \sqrt{\alpha_i}\,\ell_i & \text{si } i \in \{1, \ldots, r\} \\ \ell_i & \text{si } i \in \{r+1, \ldots, n\} \end{cases}$$

et de considérer la base antéduale associée à cette base. Relativement à cette dernière, la matrice de q s'écrit par blocs

$$\begin{pmatrix} I_r & 0_{r,n-r} \\ 0_{n-r,r} & 0_{n-r} \end{pmatrix}$$

et est de rang r. □

Les formes quadratiques sur des $\mathbb{C}$-espaces vectoriels de dimension finie sont ainsi complètement classifiées par leur rang.

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

3.4 Classification des formes quadratiques réelles

Dans cette section, on suppose que E est un $\mathbb{R}$-espace vectoriel.

Définition 3.21 *Une forme quadratique sur E est dite* ***positive*** *(resp.* ***négative****) si et seulement si*

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

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

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

Par extension, une matrice symétrique est dite définie positive/positive/définie négative/négative si c'est la matrice d'une forme quadratique 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 3.22 (« 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 sur E. Il existe une base $\{\varepsilon_i\}_{i=1,\dots,n}$ 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. Dans une forme réduite (3.2) de q, réordonnons les termes de manière à ce que

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

avec $\alpha_i > 0$ si i appartient à $\{1,\dots,p\}$ et $\alpha_i < 0$ si i appartient à $\{p+1,\dots,r\}$. En complétant au besoin la famille libre de formes linéaires apparaissant dans cette écriture de q par des formes linéaires $\ell_{r+1},\dots,\ell_n$ de manière à obtenir une base de E^*, on pose alors

$$\tilde{\ell}_i = \begin{cases} \sqrt{\alpha_i}\,\ell_i & \text{si } i \in \{1,\dots,p\} \\ \sqrt{-\alpha_i}\,\ell_i & \text{si } i \in \{p+1,\dots,r\} \\ \ell_i & \text{si } i \in \{r+1,\dots,n\} \end{cases}$$

et on arrive à l'écriture voulue en considérant pour base $\{\varepsilon_i\}_{i=1,\dots,n}$ la base antéduale de $\{\tilde{\ell}_i\}_{i=1,\dots,n}$.

Montrons ensuite l'unicité des entiers p et r. Soit $\{\varepsilon_i\}_{i=1,\dots,n}$ et $\{\varepsilon'_i\}_{i=1,\dots,n}$ deux bases de E. On a

$$\forall x \in E,\ x = \sum_{i=1}^{n} x_i\, \varepsilon_i = \sum_{i=1}^{n} x'_i\, \varepsilon'_i,\ q(x) = \sum_{i=1}^{p} {x_i}^2 - \sum_{i=p+1}^{r} {x_i}^2 = \sum_{i=1}^{p'} x_i'^2 - \sum_{i=p'+1}^{r'} x_i'^2.$$

Posons $F = \text{Vect}(\{\varepsilon_1,\dots,\varepsilon_p\})$, $G = \text{Vect}(\{\varepsilon_{p+1},\dots,\varepsilon_r\})$, $H = \text{Vect}(\{\varepsilon_{r+1},\dots,\varepsilon_n\})$, $F' = \text{Vect}(\{\varepsilon'_1,\dots,\varepsilon'_{p'}\})$, $G' = \text{Vect}(\{\varepsilon'_{p'+1},\dots,\varepsilon'_{r'}\})$ et $H' = \text{Vect}(\{\varepsilon'_{r'+1},\dots,\varepsilon'_n\})$. On a $F \cap G' = \{0_E\}$. En effet, s'il existe un vecteur x non nul appartenant à $F \cap G'$, on a $q(x) > 0$ et $q(x) < 0$, ce qui est impossible. On a aussi $F \cap H' = \{0_E\}$, car sinon, il existerait un vecteur x non nul tel que $q(x) > 0$ et $q(x) = 0$, ce qui est impossible. Il vient ainsi que F, G' et H' sont en somme directe et on a par conséquent $\dim(F) + \dim(G') + \dim(H') \leq \dim(E)$, soit encore

$$p + r' - p' + (n - r') \leq n \implies p \leq p'.$$

De la même manière, on a $F' \cap G = \{0_E\}$ et $F' \cap H = \{0_E\}$, ce qui permet de montrer que $p' \leq p$ et donc que $p = p'$.

En utilisant le même raisonnement avec G et $F' \cup H'$ pour trouver que

$$(r - p) + (n - (r' - p)) \leq n \implies r \leq r',$$

puis avec G' et $F \cup H$ pour obtenir que

$$(r' - p) + (n - (r - p)) \leq n \implies r' \leq r,$$

on montre que $r = r'$.

Enfin, la matrice de la forme q relativement à la base $\{\varepsilon_i\}_{i=1,\dots,n}$ introduite plus haut s'écrit par blocs

$$\begin{pmatrix} I_p & 0_{p,r-p} & 0_{p,n-r} \\ 0_{r-p,p} & -I_{r-p} & 0_{r-p,n-r} \\ 0_{n-r,p} & 0_{n-r,r-p} & 0_{n-r} \end{pmatrix}$$

et est donc de rang r. □

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[2] *invariants* de la forme.

On en déduit le résultat suivant.

Corollaire 3.23 (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 3.24 *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 $b(x, y) = x_1y_1 - x_2y_2$, 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

— 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 3.25 *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 3.26 (« 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 3.22) 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

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

---

[2] Ce ne sont pas les seuls.

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

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

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 3.27 *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 3.28 *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 principaux dominants sont nuls et qui est symétrique négative. On peut en revanche montrer qu'une condition nécessaire et suffisante est que tous[3] *les mineurs principaux de la matrice soient positifs.*

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

---

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

Théorème 3.29 (« 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 3.30 *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.

Corollaire 3.31 (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

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

□

3.5 Formes hermitiennes

En raison de leurs propriétés, les formes quadratiques définies positives sur un $\mathbb{R}$-espace vectoriel jouent un rôle important, notamment en physique mathématique. On peut par exemple les utiliser pour définir des normes. Sur un $\mathbb{C}$-espace vectoriel, il ne peut y avoir de forme quadratique positive, car si $q(x)$ est une quantité positive alors $q(\mathrm{i}x)=\mathrm{i}^2q(x)=-q(x)$ est une quantité négative. Néanmoins, on peut construire une théorie semblable en utilisant à la place des formes quadratiques des formes *hermitiennes*.

Définition 3.32 (forme hermitienne) *Soit E un $\mathbb{C}$-espace vectoriel. Une application h de E dans $\mathbb{R}$ est une* ***forme hermitienne*** *sur E si et seulement s'il existe une forme s sesquilinéaire à gauche et à symétrie hermitienne sur E telle que*

$$\forall x\in E,\ h(x)=s(x,x).$$

Remarque 3.33 *La forme s de cette définition est complètement déterminée par h. En effet, comme c'était le cas pour les formes quadratiques, on dispose d'identités de polarisation associées :*

$$\begin{aligned}\forall (x,y)\in E^2,\ s(x,y) &= \frac{1}{4}\left(h(x+y)-h(x-y)-\mathrm{i}\,h(x+\mathrm{i}\,y)+\mathrm{i}\,h(x-\mathrm{i}\,y)\right),\\ &= \frac{1}{2}\left(h(x+y)-\mathrm{i}\,h(x+\mathrm{i}\,y)-(1-\mathrm{i})\left(h(x)+h(y)\right)\right),\\ &= \frac{1}{2}\left((1-\mathrm{i})\left(h(x)+h(y)\right)-h(x-y)+\mathrm{i}\,h(x-\mathrm{i}\,y)\right).\end{aligned}$$

Enfin, on a pour les formes hermitiennes les mêmes définitions de matrices représentatives, rang, noyau, etc. que pour les formes quadratiques.

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 et seulement si elle est bilinéaire, symétrique, non dégénérée et positive.*

Si E est un espace vectoriel sur $\mathbb{C}$, on appelle ***produit scalaire hermitien*** *une application $\langle \cdot, \cdot \rangle$ de $E \times E$ dans $\mathbb{C}$ qui est sesquilinéaire, à symétrie hermitienne, non dégénérée et positive.*

D'après un corollaire de l'inégalité de Cauchy–Schwarz (voir le corollaire 3.30), la forme polaire d'une forme quadratique q (resp. d'une forme hermitienne h) positive sur E est non dégénérée si et seulement si q (resp. h) est définie. À ce titre, on peut aussi définir un produit scalaire (resp. un produit scalaire hermitien) comme une forme bilinéaire (resp. sesquilinéaire), symétrique (resp. à symétrie hermitienne), définie 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* tout espace vectoriel muni d'un produit scalaire. Si l'espace vectoriel est de dimension finie, on parle d'*espace euclidien* si l'espace est réel ou d'*espace hermitien* s'il est complexe.

Note (à déplacer ?) : Si E est un espace euclidien et A est un sous-espace vectoriel de E, alors la restriction du produit scalaire de E à A est un produit scalaire sur A.

Exemples :

$E = \mathbb{R}^n$, $\langle x, y \rangle = \sum_{i=1}^{N} x_i y_i$ 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

$E = \ell^2 = \{(u_k)_{k \in \mathbb{N}} \in \mathbb{R}^{\mathbb{N}} \mid \sum_{k \in \mathbb{N}} (u_k)^2 < +\infty\}$, $\langle u, v \rangle = \sum_{k \in \mathbb{N}} u_k v_k$ 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. □

Remarques :
On peut de la même façon définir une norme sur un $\mathbb{C}$-espace vectoriel associée à un produit scalaire hermitien.
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 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 (ou hermitien). Donner un exemple.

Proposition 4.3 *Soit E un espace préhilbertien sur le corps $\mathbb{R}$ 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

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

□

Définition 4.4 (espace de Hilbert) *Un espace préhilbertien qui est complet pour la distance associée à sa norme (définie par $(x, y) \mapsto \|x - y\|$) est appelé un* ***espace de Hilbert****.*

EXEMPLES d'espace de Hilbert
$\mathbb{R}^n$ muni du produit scalaire usuel
$\mathbb{C}^n$ muni du produit hermitien usuel
$L^2([a, b])$, fonctions de carré intégrable sur l'intervalle $[a, b]$, muni de $\langle f, g\rangle = \int_a^b f(t)g(t)\,\mathrm{d}t$.

Théorème 4.5 (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.

Définition 4.6 (écart angulaire) *Soit E un espace préhilbertien réel. Pour tous vecteurs x et y de E, il existe un unique réel θ appartenant à l'intervalle $[0,\pi]$ tel que*

$$\cos(\theta)=\frac{\langle x,y\rangle}{\|x\|\|y\|}.$$

*Ce réel est appelé **écart angulaire entre** x **et** y et correspond à une mesure d'angle non orienté.*

4.2 Orthogonalité dans les espaces euclidiens

La notion d'orthogonalité telle qu'on l'a introduite pour une forme quadratique subsiste bien évidemment avec un produit scalaire ou hermitien, avec l'avantage qu'une telle forme est définie positive.

4.2.1 Généralités

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

Définitions 4.7 *Soit E un espace préhilbertien. 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*

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

Remarques

Le produit scalaire étant une forme symétrique, on a que deux vecteurs x et y sont orthogonaux si et seulement si $\langle y,x\rangle=0$.

On a vu précédemment que $A^\perp$ était un sous-espace vectoriel.

4.2.2 Bases orthonormées

Définitions 4.8 *Soit E un espace préhilbertien. 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 4.9 (relation de Pythagore) *Soit E un espace euclidien, p un entier naturel non nul et $\{e_i\}_{i=1,\dots,p}$ 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 4.10 *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.

Proposition 4.11 (coordonnées dans une base orthonormée) *Si E est un espace euclidien (resp. hermitien) et si $\{e_1,\dots,e_n\}$ est une base orthonormée de E, on a*

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

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

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

et l'on a

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

DÉMONSTRATION. laissée en exercice □

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*

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

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 $\text{Vect}(\{f_1,\dots,f_j\})$ 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 $\text{Vect}(\{e_1,\dots,e_j\})$.

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 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 Vect($\{f_1,\dots,f_{j+1}\}$) = 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. □

Remarques

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 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 : polynômes orthogonaux

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

Proposition 4.15 *Soit E un espace euclidien et A un sous-espace vectoriel de E. On a*

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

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 sait déjà [1] que $A \subset (A^{\perp})^{\perp}$ et on utilise que $\dim(A) = \dim(E) - \dim(A^{\perp}) = \dim((A^{\perp})^{\perp})$. ☐

Remarque 4.16 *Cette proposition reste vraie dans un espace préhilbertien si le sous-espace A est de dimension* finie. *Elle est également vraie si A est un sous-espace vectoriel* fermé *et l'espace E est* complet, *mais elle est fausse en général.*

4.2.3 Projection orthogonale

Proposition et définition 4.17 (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.*

Rappel : 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 $A^{\perp}$.

Note : une projection est idempotente, i.e. $p \circ p = p$.

Proposition 4.18 (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.

---

[1] C'est un résultat de la proposition 3.12.

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
$$\forall x \in E,\ \forall i \in \{1, \ldots, p\},\ \langle x - p_A(x), e_i \rangle = 0.$$
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 $\ker(p) \oplus \mathrm{Im}(p) = E$ en utilisant le théorème du rang et en servant de l'orthogonalité entre $\ker(p)$ et $\mathrm{Im}(p)$ pour en déduire que $\ker(p) \cap \mathrm{Im}(p) = \{0_E\}$. 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
$$\forall j \in \{1, \ldots, n\},\ e_j = \frac{f_j - p_{\mathrm{Vect}(\{e_1, \ldots, e_{j-1}\})}(f_j)}{\|f_j - p_{\mathrm{Vect}(\{e_1, \ldots, e_{j-1}\})}(f_j)\|}.$$

Application à la résolution d'un problème de régression

Supposons que l'on dispose de m observations, représentées par les réels y_i, $i = 1, \ldots, m$, associées à m conditions, représentées par les réels x_i, $i = 1, \ldots, m$. 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, $i = 1, \ldots, n$. 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 $(y_1, \ldots, y_m)$ et le vecteur $(\varphi(x_1), \ldots, \varphi(x_m))$.

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, $i = 1, \ldots, m$. 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 4.19 (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 $i = 1, \ldots, m$. En effet, le système du problème de régression s'écrit dans ce cas
$$\begin{pmatrix} 1 & x_1 & {x_1}^2 & \ldots & {x_1}^{m-1} \\ 1 & x_2 & {x_2}^2 & \ldots & {x_2}^{m-1} \\ \vdots & \vdots & & \vdots & \\ 1 & x_m & {x_m}^2 & \ldots & {x_n}^{m-1} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}$$

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 4.18), 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 E 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 4.20 (théorème de représentation de Riesz) *Soit E un espace euclidien. Pour tout forme linéaire ℓ sur E, 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 4.21 (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

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

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

Remarque

La notion d'adjoint vaut aussi dans le cas d'un espace hermitien et on a alors

$$\mathrm{mat}_{\mathscr{B}}(u^*) = \mathrm{mat}_{\mathscr{B}}(u)^*.$$

4.4.2 Isométries vectorielles

Définition 4.22 (isométrie vectorielle) *Soit E un espace préhilbertien 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.

Si l'espace est hermitien, on parle d'*endomorphisme unitaire*.

Remarques

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 : les seules homothéties $x\mapsto \lambda\,x$ qui sont des isométries vectorielles sont $-id_E$ et id_E.

Proposition 4.23 *Soit E un espace préhilbertien. 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,

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

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. □

Remarques

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 4.24 *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 4.25 *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 (et l'ensemble des endomorphismes unitaires d'un espace hermitien E forme le* groupe unitaire *de E, noté $U(E)$). 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 4.26 *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

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

Proposition 4.27 *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

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

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 4.28 (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

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

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

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

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 4.27. □

Définition 4.29 (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 4.30 *Soit n un entier naturel non nul. Pour toute matrice M de $O(n)$, on a*

$$\det(M)=\pm 1.$$

*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

L'ensemble des matrices orthogonales de déterminant égal à 1 forme un sous-groupe du groupe orthogonal, appelé *groupe spécial orthogonal* et noté $SO(n)$.

Pour un endomorphisme unitaire dans un espace hermitien, de matrice M dans une base orthonormée, on a $M^*M = MM^* = I_n$. On appelle ainsi *matrice unitaire* une matrice carrée U à coefficients complexes telle que $U^*U = UU^* = I_n$. L'ensemble des matrices unitaire d'ordre n, noté $U(n)$, est un sous-groupe de $GL(n,\mathbb{C})$.

Note : les matrices de Pauli forment une base de $SU(2)$.

4.4.3 Endomorphismes auto-adjoints

Définition 4.31 (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.

EXEMPLES : projection orthogonale

Dans un espace hermitien, un endomorphisme est auto-adjoint si et seulement si sa matrice dans une base orthonormale est *hermitienne*, c'est-à-dire qu'elle est égale à sa transposée conjuguée.

Proposition 4.32 *Soit E un espace euclidien ou hermitien 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. Il suffit de reprendre la preuve de la proposition 4.26, en utilisant que

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

selon que l'endomorphisme u est auto-adjoint ou anti-auto-adjoint. □

Théorème 4.33 (diagonalisation des endomorphismes auto-adjoints) *Soit un espace euclidien (ou hermitien) 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. De plus, les valeurs propres de l'endomorphisme sont réelles.*

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 (ou complexe hermitienne). Il existe une valeur propre λ de M, *a priori* complexe, et une matrice colonne non nulle Z, telles que $MZ = \lambda Z$. On a alors [2]

$$(MZ)^*Z = \overline{\lambda}\, Z^*Z = \overline{\lambda}\, \|Z\|^2 \text{ et } (MZ)^*Z = Z^*M^*Z = Z^*MZ = \lambda\, Z^*Z = \lambda\, \|Z\|^2,$$

---

[2] Lorsque E est un espace euclidien, et par conséquent réel, on doit se placer sur le *complexifié* de E, c'est-à-dire l'espace vectoriel sur $\mathbb{C}$ correspondant à $E \times E$ muni de l'addition usuelle et d'une multiplication externe par les nombre complexes, pour effectuer ces calculs.

d'où λ est réelle.

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 4.32, 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 4.34 (diagonalisation des matrices réelles symétriques ou complexes hermitiennes) *Soit n un entier naturel non nul. Pour tout matrice M réelle symétrique (resp. complexe hermitienne) d'ordre n, il existe une matrice orthogonale O (resp. matrice unitaire) d'ordre n telle que $O^\top MO$ (resp. U^*MU) est une matrice diagonale réelle.*

On rappelle qu'une matrice M symétrique (resp. hermitienne) est dite positive si la forme quadratique (resp. hermitienne) qui lui est canoniquement associée est positive. Elle est dite définie positive si la forme quadratique 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 est donnée par les nombres de valeurs propres strictement positives et strictement négatives d'une matrice qui la représente.

Corollaire 4.35 *Soit une forme quadratique (resp. hermitienne) sur un espace euclidien (resp. hermitien). 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 (resp. hermitienne) et, d'après le précédent corollaire, il existe une matrice orthogonale O (resp. unitaire U) telle que la matrice $O^\top MO$ (resp. U^*MU) est diagonale réelle. La matrice O (resp. U) 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 3.15 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 4.36 (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, $i = 1, \dots, r$, 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

$$\forall k \in \{1, \dots, r\},\ {U_k}^\top M = \sigma_k V_k \text{ et } MV_k = \sigma_k U_k.$$

Chapitre 5

Compléments sur les isométries vectorielles et les matrices orthogonales

5.1 Groupe des isométries vectorielles

Rappels

Un groupe $(G, \circ)$ est un couple formé d'un ensemble G et d'une opération $\circ$ sur cet ensemble (parfois appelée loi de composition), qui à deux éléments a et$b$ de G associe un élément $a \circ b$, satisfaisant aux quatre axiomes suivants :

— $\forall (a, b) \in G^2$, $a \circ b \in G$ (loi de compostion interne).

— $\forall (a, b, c) \in G^3$, $(a \circ b) \circ c = a \circ (b \circ c)$ (associativité).

— $\exists e \in G, \forall a \in E$, $e \circ a = a \circ e = a$, e étant appelé l'élément neutre du groupe.

— $\forall a \in G$, $\exists b \in G$, $a \circ b = b \circ a = e$, où e est l'élément neutre du groupe, b étant appelé le symétrique de a.
Un groupe pour lequel on a de plus la propriété

$$\forall (a, b) \in G^2, \ a \circ b = b \circ a,$$

est dit commutatif ou abélien.

exemples : $(\mathbb{Z}, +)$, $(\mathbb{Q} \setminus \{0\}, \times)$, $(GL(E), \circ)$, où $\circ$ est la composition d'applications, sont des groupes.

Remarque sur le vocabulaire utilisé

Lorsque la loi de composition est notée additivement (on écrit $a + b$ pour $a \circ b$), le symétrique est appelé opposé et l'opposé de a est noté $-a$ et l'élément neutre est appelé zéro et noté 0. Lorsque la loi est notée multiplicativement, (on écrit généralement $a \times b$ ou ab pour $a \circ b$), le symétrique est appelé l'inverse et le symétrique de a est noté a^{-1}, l'élément neutre est appelé unité et est noté 1.

Théorème 5.1 *Soit E une espace euclidien de dimension finie non nulle égale à n. L'ensemble des isométries vectorielles $O(E)$, muni de la composition d'applications $\circ$ est un sous-groupe du groupe linéaire $(GL(E), \circ)$.*

DÉMONSTRATION. Il suffit de montrer que $O(E)$ est une partie non vide de $(GL(E), \circ)$, stable par composition et symétrisation.

Tout d'abord, id_E appartient à $O(E)$ car $\forall x \in E$, $id_E(x) = x$ et $(id_E)^{-1} = id_E$.

Considérons ensuite des vecteurs u et v appartenant à $O(E)$. $\forall x \in E$, $\|u(x)\| = \|x\| = \|v(x)\|$, d'où $\|u \circ v(x)\| = \|u(v(x))\| = \|v(x)\| = \|x\|$ (idem pour $v \circ u$).

Enfin, soit u appartenant à $O(E)$ et ayant pour symétrique u^*. On a $u \circ u^* = u^* \circ u = id_E$, d'où $\forall x \in E$, $\|u \circ u^*(x)\| = \|x\|$ et donc $\|x\| = \|u(u^*(x))\| = \|u^*(x)\|$. □

Ce résultat signifie en particulier que la composée de deux isométries vectorielles et la réciproque d'une isométrie vectorielle sont des isométries vectorielles.

A DEPLACER ? exemples d'isométries vectorielles : les symétries orthogonales

Rappels : soit F et G deux sous-espaces vectoriels d'un espace vectoriel E tels que $E = F \oplus G$. $\forall x \in E$, il existe une unique décomposition $x = x_F + x_G$, avec x_F appartenant à F et x_G appartenant à G. La symétrie par rapport à F et parallèlement à G est alors l'application s de $\mathscr{L}(E)$ telle sue $s(x) = X_F - x_G$. On a que $F = \ker(s - id_E)$, $G = \ker(s + id_E)$ et $s \circ s = id_E$. Par ailleurs, $s = 2p - id_E$ où p est la projection sur F parallèlement à G.

Considérons à présent un espace préhilbertien E et un sous-espace A de dimension finie de E. On a $E = A \oplus A^{\perp}$. La symétrie par rapport à A et parallèlement à $A^{\perp}$ est appelée symétrie orthogonale par rapport à A. Cherchons à en déterminer l'expression.

Lorsque $A = \mathrm{Vect}(\{a\})$ est une droite vectorielle dirigée par le vecteur a, on a

$$\forall x \in E,\ s(x) = 2\langle x, a\rangle \frac{a}{\|a\|^2} - x.$$

Lorsque $A = \{b\}^{\perp}$ est un hyperplan, on utilise encore la formule $s(x) = 2p(x) - x$, avec cette fois $p(x) = x - \langle x, b\rangle \frac{b}{\|b\|^2}$. En dimension finie, on parle de réflexion.

Lorsque l'on n'est dans aucun de ces deux cas, on se sert du fait que

$$y = s(x) \iff x + y \in A \text{ et } x - y \in A^{\perp},$$

ou bien, si l'on dispose d'une base orthonormée $\{a_1, \ldots, a_p\}$ de A, on applique la formule

$$s(x) = 2\sum_{i=1}^{p} \langle x, a_i\rangle\, a_i - x.$$

Proposition 5.2 *Soit E un espace euclidien et s un endomorphisme de E. Alors, s est une symétrie orthogonale si et seulement si $s \circ s = id_E$ et $\ker(s + id_E) \subset (\ker(s - id_E))^{\perp}$.*

DÉMONSTRATION. A VOIR □

5.2 Matrices orthogonales

Rappels

Une matrice M de $M_n(\mathbb{R})$ est dite orthogonale si et seulement si $M^{\top}M = MM^{\top} = I_n$. Une matrice orthogonale M est inversible et $M^{-1} = M^{\top}$. Son déterminant vaut ±1.

On note $O_n(\mathbb{R})$ ou simplement $O(n)$ l'ensemble des matrices orthogonales à coefficients réels. C'est un sous-groupe de $GL(n, \mathbb{R})$, appelé groupe orthogonal.

La matrice d'une isométrie vectorielle dans une base orthonormée de l'espace est une matrice orthogonale.

Théorème 5.3 (caractérisation pratique des matrices orthogonales) *La famille des colonnes (resp. des lignes) d'une matrice orthogonale d'ordre n est une base orthonormale de l'espace $M_{n,1}(\mathbb{R})$ (resp. $M_{1,n}(\mathbb{R})$) muni de sa structure euclidienne canonique.*

DÉMONSTRATION. Soit M une matrice orthogonale d'ordre n dont on note $C_1, \ldots, C_n$ les colonnes. L'égalité $M^{\top}M = I_n$ écrite coefficient par coefficient est

$$\forall (i,j) \in \{1, \ldots, n\}^2,\ C_i^{\top} C_j = \delta_{ij},$$

ce que l'on peut encore écrire

$$\forall (i,j) \in \{1, \ldots, n\}^2,\ \left\langle C_i, C_j \right\rangle = \delta_{ij},$$

en faisant appel au produit scalaire de $M_{n,1}(\mathbb{R})$, d'où la conclusion.

On obtient le même résultat pour les lignes de la matrice en transposant l'égalité matricielle utilisée. □

Remarque

Les coefficients d'une matrice orthogonale sont inférieurs ou égal à 1 en valeur absolue.

Théorème 5.4 (caractérisation des matrices orthogonales) *La matrice de passage d'une base orthonormale à une autre base orthonormale est une matrice orthogonale.*

DÉMONSTRATION. Il suffit de considérer l'endomorphisme défini en posant que l'image d'un vecteur d'une base orthonormale par celui-ci est un vecteur de l'autre base orthonormée et d'utiliser le fait que c'est une isométrie vectorielle ayant pour matrice dans la base de départ la matrice de passage en question. □

En conclusion, on retiendra qu'une matrice orthogonale peut s'interpréter comme :

— la matrice d'une isométrie vectorielle dans une base orthonormale,

— la matrice de passage d'une base orthonormale à une autre base orthonormale.

A VOIR : Une isométrie vectorielle est une symétrie orthogonale$\Leftrightarrow$ Sa matrice dans une base orthonormale est symétrique.

5.3 Orientation et produit mixte

Dans cette section, E désigne un espace euclidien de dimension n non nulle.

Définition 5.5 (orientation) *Soit $\mathscr{B}$ et $\mathscr{B}'$ deux bases de E et P la matrice de passage de la première à la seconde de ces bases. On dit que $\mathscr{B}$ et $\mathscr{B}'$ ont la même* ***orientation*** *si et seulement si le déterminant de P est strictement positif.*

Orienter l'espace consiste donc à choisir arbitrairement une base de référence. Les bases de même orientation que cette base seront dites *directes* et les autres *indirectes*.

Remarque

Se placer dans « l'espace $\mathbb{R}^n$ muni de sa structure euclidienne usuelle et orienté » signifie que l'on a muni $\mathbb{R}^n$ du produit scalaire usuel (i.e., canonique) et considéré que la base canonique (qui est une base orthonormale pour le produit scalaire usuel) est directe.

Proposition et définition 5.6 (produit mixte) *On suppose l'espace vectoriel E orienté et que $\mathscr{B}$ est base orthonormale directe de E. Le* ***produit mixte*** *de la famille $\{u_1,\dots,u_n\}$ de vecteurs de E, noté $[u_1,\dots,u_n]$, est le scalaire*

$$[u_1,\dots,u_n]=\det_{\mathscr{B}}(u_1,\dots,u_n).$$

Celui-ci ne dépend pas de la base orthonormale directe choisie.

DÉMONSTRATION. Soit $\mathscr{B}$ et $\mathscr{B}'$ deux bases orthonormales directes de E. La matrice de passage P de $\mathscr{B}$ à $\mathscr{B}'$ est une matrice orthogonale, dont le déterminant vaut 1 puisque les bases ont la même orientation. On a par conséquent

$$\det_{\mathscr{B}}(u_1,\dots,u_n)=\det(P)\det_{\mathscr{B}'}(u_1,\dots,u_n)=\det_{\mathscr{B}'}(u_1,\dots,u_n).$$

□

En démontrant cette proposition, on a au passage prouvé que que le produit mixte d'une base orthonormale directe est égal à 1. Plus généralement, le produit mixte d'une base directe est strictement positif.

remarques

Dans $\mathbb{R}^2$ muni de sa structure euclidienne canonique, le produit mixte $[x,y]$ de deux vecteurs x et y représente l'aire algébrique du parallélogramme défini par x et y.

Dans $\mathbb{R}^3$ muni de sa structure euclidienne canonique, le produit mixte $[x,y,z]$ de trois vecteurs x, y et z représente le volume algébrique du parallélépipède défini par x, y et z.

Dans un esapce orienté, le choix d'un vecteur normel permet d'orienter un hyperplan affine. en effet, si E est orienté et si v est un vecteur normal à un hyperplan affine $\mathscr{H}$, alors on dira qu'une base $\{h_1,\dots,h_{n-1}\}$ de $\mathscr{H}$ est directe si et seulement si $[h_1,\dots,h_{n-1},v]$ est strictement positif. En choisissant $-v$ comme vecteur normal, la même base est indirecte.

Par la formule de transformation des volumes, on a

$$\forall f\in\mathscr{L}(E),\ [f(u_1),\dots,f(u_n)]=\det(f)[u_1,\dots,u_n].\tag{5.1}$$

5.4 Étude des isométries vectorielles

On a vu que l'on pouvait définir le sous-ensemble

$$SO(n) = \{M \in O(n) \,|\, \det(M) = 1\},$$

qui est un sous-groupe de $O(n)$ appelé groupe spécial orthogonal, ce qui permet par suite d'introduire le sous-ensemble

$$O^{-}(n) = O(n) \setminus SO(n) = \{M \in O(n) \,|\, \det(M) = -1\}.$$

Aux matrices de $SO(n)$ et $O^{-}(n)$ sont associés les isométries vectorielles respectivement positives (appartenant à l'ensemble $SO(E)$) et négatives (appartenant à l'ensemble $O^{-}(E)$) de E. D'après la formules (5.1) rappelée dans la section précédente, on observe que les isométries de $SO(E)$ préservent l'orientation.

5.4.1 Le groupe $O(2)$

Intéressons-nous aux isométries vectorielles en dimension deux en considérant une matrice $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ appartenant au groupe $O(2)$. On a $M^{\top}M = I_2$, d'où le système satisfait par les coefficients de la matrice

$$\begin{cases} a^2 + b^2 = 1 \\ c^2 + d^2 = 1 \\ ac + bd = 0 \end{cases}.$$

Ainsi, il existe des réels θ et φ tels que $a = \cos(\theta)$, $b = \sin(\theta)$, $c = \cos(\varphi)$, $d = \sin(\varphi)$ et $\cos(\theta)\cos(\varphi) + \sin(\theta)\sin(\varphi) = 0$, soit encore $\cos(\theta - \varphi) = 0$. Ceci signfie qu'il existe ε valant ± 1 tel que $\varphi = \theta + \varepsilon\frac{\pi}{2}$ modulo 2π. Or, on a

$$\cos\left(\theta + \varepsilon\frac{\pi}{2}\right) = -\varepsilon\sin(\theta) \text{ et } \sin\left(\theta + \varepsilon\frac{\pi}{2}\right) = \varepsilon\cos(\theta),$$

d'où il existe une réel θ et un réel ε prenant les valeurs ± 1 tels que

$$M = \begin{pmatrix} \cos(\theta) & -\varepsilon\sin(\theta) \\ \sin(\theta) & \varepsilon\cos(\theta) \end{pmatrix}.$$

On a enfin $\det(M) = \varepsilon\left((\cos(\theta))^2 + (\sin(\theta))^2\right) = \varepsilon$.

On vient de démontrer le résultat suivant.

Théorème 5.7 *Toute matrice de $SO(2)$ est de la forme*

$$\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$$

pour un certain réel θ. Toute matrice de $O^{-}(2)$ est de la forme

$$\begin{pmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{pmatrix}$$

pour un certain réel θ.

Lemme 5.8 *Le groupe $SO(2)$ est commutatif.*

DÉMONSTRATION. Soit M et N des matrices de $SO(2)$. Il existe des réels α et β tels que $M = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}$ et $N = \begin{pmatrix} \cos(\beta) & -\sin(\beta) \\ \sin(\beta) & \cos(\beta) \end{pmatrix}$. En se servant des formules trigonométriques d'addition, il vient

$$MN = \begin{pmatrix} \cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta) & -\cos(\alpha)\sin(\beta) - \sin(\alpha)\cos(\beta) \\ \sin(\alpha)\cos(\beta) + \cos(\alpha)\sin(\beta) & \cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta) \end{pmatrix} = \begin{pmatrix} \cos(\alpha + \beta) & -\sin(\alpha + \beta) \\ \sin(\alpha + \beta) & \cos(\alpha + \beta) \end{pmatrix} = NM.$$

□

Théorème 5.9 (classification des isométries vectorielles en dimension deux) *Soit E un espace euclidien orienté de dimension égale à 2 et u une isométrie vectorielle de E. Si u appartient à $SO(E)$, la matrice de u ne dépend pas du choix de la base, pourvu que ce soit une base orthonormale directe, et est de la forme*

$$\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$$

pour un certain réel θ, unique modulo 2π. On dit alors que u est une rotation vectorielle d'angle orienté de mesure θ.

Si u appartient à $O^-(E)$, u est une réflexion, c'est-à-dire une symétrie orthogonale par rapport à une droite.

Démonstration. COMPLETER Soit u une isométrie de $SO(E)$, $\mathscr{B}$ et $\mathscr{B}'$ deux bases orthonormales directes de E. Les matrices de u dans ces bases sont des éléments de $SO(2)$. C'est aussi le cas de la matrice de passage P de $\mathscr{B}$ à $\mathscr{B}'$. Par la formule changement de base, on a, en utilisant que le groupe $SO(2)$ est commutatif,

$$\mathrm{mat}_{\mathscr{B}'}(u) = P^{-1}\mathrm{mat}_{\mathscr{B}}(u)P = P^{-1}P\mathrm{mat}_{\mathscr{B}}(u) = \mathrm{mat}_{\mathscr{B}}(u).$$

Enfin, si $\mathscr{B} = \{e_1, e_2\}$, DESSIN

Soit à présent u une isométrie de $O^-(E)$ et $\mathscr{B} = \{e_1, e_2\}$ une base orthonormale directe de E. On observe sur le DESSIN suivant que les vecteurs e_1 et $u(e_1)$ (resp. e_2 et $u(e_2)$) sont symétriques par rapport à la droite engendrée par le vecteur unitaire u_1.

Posons $u_1 = \cos\left(\frac{\theta}{2}\right) e_1 + \sin\left(\frac{\theta}{2}\right) e_2$ et $u_2 = -\sin\left(\frac{\theta}{2}\right) e_1 + \cos\left(\frac{\theta}{2}\right) e_2$. On va montrer que u est la réflexion orthogonale par rapport à le droite engendrée par u_1. Il est clair que $\{u_1, u_2\}$ est une base orthonormale de E. De plus, un calcul facile montre que $u(u_1) = u_1$ et $u(u_2) = -u_2$, de sorte que la matrice de u dans cette base est

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

L'isométrie est donc bien la réflexion par rapport à $\mathrm{Vect}(\{u_1\})$. ☐

remarques

La commutativité du groupe $SO(2)$ signifie géométriquement que deux rotations vectorielles commutent entre elles.

L'indépendance de la matrice d'une rotation vectorielle par rapport au choix de la base orthonormale directe signifie géométriquement que cette rotation agit uniformément sur tous les vecteurs du plan, i.e. de la même manière dans toutes les directions.

On a parlé de rotation « d'angle orienté de mesure θ » sans avoir défini la notion d'angle orienté. Celle-ci est liée à la notion d'écart angulaire introduite dans la définition 4.6 et peut se résumer de la manière suivante.

Théorème 5.10 *Soit E un espace euclidien orienté de dimension égale à 2 et x et y deux vecteurs non nuls de E. Il existe une unique rotation vectorielle u, d'angle de mesure θ, telle que*

$$\frac{y}{\|y\|} = u\left(\frac{x}{\|x\|}\right).$$

On alors $\cos(\theta) = \dfrac{\langle x, y\rangle}{\|x\|\|y\|}$.

Démonstration. Posons $e_1 = \frac{x}{\|x\|}$ et complétons $\{e_1\}$ en une base orthonormale directe $\{e_1, e_2\}$. On note (a_1, a_2) les coordonnées du vecteur $\frac{y}{\|y\|}$ dans cette base. On a en particulier que ${a_1}^2 + {a_2}^2 = 1$. Si u est la rotation vectorielle de matrice

$$\begin{pmatrix} a_1 & -a_2 \\ a_2 & a_1 \end{pmatrix}$$

dans la base $\{e_1, e_2\}$, on a le résultat voulu. On a par ailleurs unicité car la forme d'une matrice de rotation dans une base orthonormale directe est telle qu'on a seuleent besoin de connaître sa première colonne pour la connaître entièrement. Enfin, par définition de r ety de θ, il vient

$$\frac{y}{\|y\|} = \cos(\theta) e_1 + \sin(\theta) e_2,$$

d'où

$$\langle x, y\rangle = \|x\|\|y\|\left\langle \frac{x}{\|x\|}, \frac{y}{\|y\|}\right\rangle = \|x\|\|y\|\,\langle e_1, \cos(\theta) e_1 + \sin(\theta) e_2\rangle = \|x\|\|y\|\cos(\theta).$$

☐

5.4.2 Le groupe $O(3)$

Pour l'analyse de ce groupe, on aura besoin d'une opération propre aux espaces euclidiens orientés de dimension 3 : le *produit vectoriel.*

Proposition et définition 5.11 (produit vectoriel) *Soit E un espace euclidien orienté de dimension égale à* 3 *et x et y deux vecteurs de E. Il existe un unique vecteur de E, noté $x \wedge y$ et appelé produit vectoriel de x et y, pour lequel*

$$\forall z \in E,\ [x, y, z] = \langle x \wedge y, z\rangle .$$

DÉMONSTRATION. On a déjà vu (voir le théorème 4.20) que l'application i_E de E dans E^* définie par

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

était un isomorphisme. À présent, pour tous vecteurs x et y de E, l'application $z \mapsto [x, y, z]$ est une forme linéaire sur E, qui possède par conséquent un unique antécédent par i_E que l'on note $x \wedge y$. Par définition, le vecteur $x \wedge y$ est donc le seul vecteur de E tel que

$$\forall z \in E,\ [x, y, z] = i_E(x \wedge y)(z) = \langle z, x \wedge y\rangle .$$

☐

Remarque : le symbole ∧ utilisé ici pour noter le produit vectoriel est généralement employé en France. Son inconvénient est d'entrer en conflit avec la notation du produit extérieur. Dans la littérature anglophone et allemande (ainsi qu'au Canada francophone, en Suisse, et parfois en Belgique), le produit vectoriel est noté avec le symbole ×. Cette notation est due à Gibbs, à qui l'on doit l'introduction du produit vectoriel. Son inconvénient est d'induire une éventuelle confusion avec le produit des réels et le produit cartésien, mais ces derniers ne portent pas sur des objets de même nature.

Proposition 5.12 (propriétés du produit vectoriel) *Soit E un espace euclidien orienté de dimension égale à* 3 *et x, y et z vecteurs de E. On a les propriétés suivantes.*

1. *Le vecteur $x \wedge y$ est orthogonal à x et y.*
2. *Les vecteurs x et y sont colinéaires si et seulement si $x \wedge y$ est nul.*
3. *Le produit vectoriel est une application bilinéaire alternée de $E \times E$ dans E.*
4. *Si les vecteurs x et y ne sont pas colinéaires alors $\{x, y, x \wedge y\}$ est une base directe de E.*
5. *On suppose la famille $\{x, y\}$ orthonormale. Alors, la famille $\{x, y, z\}$ est une base orthonormale directe de E si et seulement si $z = x \wedge y$.*
6. *Soit $\mathscr{B}$ une base orthonormale directe de E dans laquelle les vecteurs x et y ont pour coordonnées respectives (x_1, x_2, x_3) et (y_1, y_2, y_3). Alors, le vecteur $x \wedge y$ a pour coordonnées dans la base $\mathscr{B}$ $(x_2y_3 - x_3y_2, x_3y_1 - x_1y_3, x_1y_2 - x_2y_3)$.*

DÉMONSTRATION. Avant tout chose, rappelons que tout produit mixte est un déterminant, qui est donc nul pour tout famille liée.

1. Les vecteurs $x \wedge y$ et x sont orthogonaux, car on a

$$\langle x \wedge y, x\rangle = [x, y, x] = 0.$$

   Il en va de même pour $x \wedge y$ et y.

2. Si x et y sont colinéaires, toute famille contenant ces vecteurs est liée et l'on a donc

$$\|x \wedge y\|^2 = \langle x \wedge y, x \wedge y\rangle = [x, y, x \wedge y] = 0.$$

   Réciproquement, supposons que le produit $x \wedge y$ soit nul. La dimension du sous-espace vectoriel engendré par les vecteurs x et y étant inférieure ou égale à 2, il existe un vecteur z appartenant à $E \setminus \text{Vect}(\{x, y\})$. La famille $\{x, y, z\}$ est alors liée, puisque $[x, y, z] = \langle x \wedge y, z\rangle = \langle 0_E, z\rangle = 0$. Le vecteur z n'étant pas combinaison linéaire de x et y, il en découle que la famille $\{x, y\}$ est liée.

3. D'après le point précédent, si le produit vectoriel est bilinéaire, il est aussi alterné. Montrons sa linéarité par rapport à la seconde variable. Pour tous vecteurs x, y et z de E et tout réel λ, on a
$$\begin{aligned}\forall x'\in E,\ \left\langle x\wedge(\lambda y+z)-\lambda(x\wedge y)-x\wedge z,x'\right\rangle &= \left\langle x\wedge(\lambda y+z),x'\right\rangle-\lambda\left\langle x\wedge y,x'\right\rangle-\left\langle x\wedge z,x'\right\rangle\\ &=[x,\lambda y+z,x']-\lambda[x,y,x']-[x,z,x']\\ &=0.\end{aligned}$$
Le vecteur $x\wedge(\lambda y+z)-\lambda(x\wedge y)-x\wedge z$ est ainsi orthogonal à tout vecteur de E et donc nul, d'où
$$\forall(x,y,z)\in E^3,\ \forall\lambda\in\mathbb{R},\ x\wedge(\lambda y+z)=\lambda(x\wedge y)+x\wedge z.$$
4. Supposons que les vecteurs x et y soient non colinéaires. Montrer que la famille $\{x,y,x\wedge y\}$ est une base directe de E revient à montrer que son déterminant dans une base directe de E est strictement positif, ce qui est vérifié puisque l'on a
$$[x,y,x\wedge y]=\langle x\wedge y,x\wedge y\rangle=\|x\wedge y\|^2>0.$$
5. Supposons que la famille $\{x,y\}$ soit orthonormale. La famille $\{x,y,\frac{x\wedge y}{\|x\wedge y\|}\}$ est alors une base orthonormale directe d'après les points 1. et 4. de la proposition. Par ailleurs, on a
$$\|x\wedge y\|=\left\langle x\wedge y,\frac{x\wedge y}{\|x\wedge y\|}\right\rangle=\left[x,y,\frac{x\wedge y}{\|x\wedge y\|}\right]=1.$$
Réciproquement, supposons que la famille $\{x,y,z\}$ soit une base orthonormale directe de E. Le sous-espace vectoriel engendré par les vecteurs x et y étant de dimension égale à 2 et son orthogonal est une droite vectorielle contenant à la fois z et $x\wedge y$, d'où $x\wedge y=\lambda z$ pour un certain réel λ. On a alors
$$\lambda=\langle\lambda z,z\rangle=\langle x\wedge y,z\rangle=[x,y,z]=1.$$
6. Notons e_1, e_2 et E_3 les vecteurs de la base $\mathscr{B}$, les coordonnées du produit $x\wedge y$ dans $\mathscr{B}$ étant alors données par $(\langle x\wedge y,e_1\rangle,\langle x\wedge y,e_2\rangle,\langle x\wedge y,e_3\rangle)$. On a donc
$$\langle x\wedge y,e_1\rangle=[x,y,e_1]=\begin{vmatrix}x_1&y_1&1\\x_2&y_2&0\\x_3&y_3&0\end{vmatrix}=x_2y_3-x_3y_2,$$
et il en va de même pour les autres coordonnées.

□

Remarque

Revenons sur l'interprétation géométrique de l'égalité définissant le produit vectoriel, en supposant que $x\wedge y$ est non nul. On sait déjà que le produit mixte $[x,y,z]$ dans le membre de gauche correspond au volume algébrique du parallélépipède engendré par les vecteurs x, y et z. Dans le membre de droite, on voit, en écrivant $\langle x\wedge y,z\rangle=\|x\wedge y\|\left\langle\frac{x\wedge y}{\|x\wedge y\|},z\right\rangle$, que $\left\langle\frac{x\wedge y}{\|x\wedge y\|},z\right\rangle$ n'est autre que la hauteur (algébrique) de parallélépipède au-dessus du parallélogramme engendré par x et y. Ainsi, $\|x\wedge y\|$ correspond à la valeur de l'aire parallélogramme engendré par x et y.

FAIRE UN DESSIN du parallépipède

Nous sommes à présent en mesure de réaliser l'étude des isométries vectorielles en dimension trois.

Théorème 5.13 (classification des isométries vectorielles en dimension trois) *Soit E un espace euclidien orienté de dimension égale à 3 et u une isométrie vectorielle de E. Il existe une base orthonormale directe $\mathscr{B}$ de E et un réel θ tels que*
$$mat_{\mathscr{B}}(u)=\begin{pmatrix}\pm1&0&0\\0&\cos(\theta)&-\sin(\theta)\\0&\sin(\theta)&\cos(\theta)\end{pmatrix}.$$
Par conséquent, une isométrie vectorielle de E est :

— *une rotation si son déterminant vaut* 1,

— *une* ***antirotation***[1]*, c'est-à-dire la composée d'une rotation et de la symétrie orthogonale (ou réflexion) par rapport au plan orthogonal à l'axe de rotation, si son déterminant vaut* -1. *La mesure de l'angle orienté de rotation est fixée par l'orientation du plan orthogonal à l'axe de rotation, celle-ci étant donnée par l'orientation de l'espace et le choix d'un vecteur directeur de l'axe.*

DÉMONSTRATION. Supposons tout d'abord que l'isométrie vectorielle u appartienne à $SO(E)$ et considérons l'application φ de $\mathbb{R}$ dans lui-même définie par $\varphi(\lambda) = \det(u - \lambda\, id_E)$. Quelle que soit la base choisie pour le calcul de ce déterminant, l'application est une fonction polynomiale de degré 3 de coefficient de terme de plus haut degré égal à -1, valant $\det(u) = 1$ en $\lambda = 0$. Cette fonction étant continue et tendant vers $\pm\infty$ en $\mp\infty$, le théorème des valeurs intermédiaires assure l'existence d'un réel λ_0 strictement positif tel que $\varphi(\lambda_0) = 0$. L'application $u - \lambda_0\, id_E$ n'est donc pas injective et il existe un vecteur x_0 non nul de E tel que $u(x_0) = \lambda_0\, x_0$. L'application u étant une isométrie vectorielle, on a nécessairement $|\lambda_0| = 1$ et, par suite, $\lambda_0 = 1$. Par ailleurs, le sous-espace vectoriel engendré par le vecteur x_0 étant stable par u, son orthogonal $H = \{x_0\}^{\perp}$ l'est également et la restriction u à H, notée v, est un endomorphisme orthogonal de H.

Soit $\{e_2, e_3\}$ une base orthonormale de H et soit

$$R = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

la matrice de v dans cette base ; on sait que R appartient à $O(2)$. En posant $e_1 = \frac{x_0}{\|x_0\|}$ et $\mathscr{B} = \{e_1, e_2, e_3\}$, il vient alors

$$\operatorname{mat}_{\mathscr{B}}(u) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & a & b \\ 0 & c & d \end{pmatrix}.$$

Puisque $\det(u) = 1$, on a $\det(R) = 1$ et donc R appartient à $SO(2)$, ce qui impose sa forme. Si $\mathscr{B}$ n'est pas une base orthonormale directe, il suffit de remplacer le vecteur e_1 par son opposé, ce qui n'affecte pas pas la matrice de u dans la base $\mathscr{B}$.

Supposons à présent que u appartienne à $O^{-}(E)$. Par des arguments similaires à ceux employés ci-dessus, on peut montrer que -1 est valeur propre de l'application u. Notons x_0 un vecteur propre associé. Pour des raisons de signe du déterminant, la restriction v de u à $H = \{x_0\}^{\perp}$ appartient alors à $SO(E)$. En se plaçant dans une base orthonormale directe $\mathscr{B} = \{e_1, e_2, e_3\}$ de E telle que $e_1 = \frac{x_0}{\|x_0\|}$, on a

$$\operatorname{mat}_{\mathscr{B}}(u) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos(\theta) & -\sin(\theta) \\ 0 & \sin(\theta) & \cos(\theta) \end{pmatrix},$$

pour un certain réel θ dans $[0, \pi]$.
Si $\theta = 0$, l'application u est simplement une réflexion par rapport au plan $H = \{x_0\}^{\perp}$. Si $\theta = \pi$, alors $u = id_E$. Enfin, si θ appartient à $]0, \pi[$, on a

$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos(\theta) & -\sin(\theta) \\ 0 & \sin(\theta) & \cos(\theta) \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & -\sin(\theta) \\ 0 & \sin(\theta) & \cos(\theta) \end{pmatrix},$$

l'isométrie vectorielle étant alors la composée de la réflexion par rapport au plan $H = \{x_0\}^{\perp}$ et de la rotation d'axe dirigé et orienté par le vecteur x_0 et d'angle θ. □

Détermination pratique d'une isométrie vectorielle en dimension trois

— Étant donné une matrice M de $M_3(\mathbb{R})$, pour savoir si elle appartient à $O(3)$, on vérifie que ses colonnes C_1, C_2 et C_3 forment une base orthonormale de $M_{3,1}(\mathbb{R})$.

— Si c'est le cas, on ne calcule pas son déterminant pour savoir si elle appartient à $SO(3)$ ou $O^{-}(3)$. Il suffit en effet de calculer le produit vectoriel $C_1 \wedge C_2$, qui vaudra C_3 si M appartient à $SO(3)$ et $-C_3$ si M appartient à $O^{-}(3)$ (le signe de la première composante de $C_1 \wedge C_2$ permet d'ailleurs de conclure

---

[1] On parle aussi de *rotation impropre*, de *roto-réflexion* ou de *rotation-réflexion*.

si elle est non nulle).

- Si M appartient à $SO(3)$, c'est la matrice d'une rotation vectorielle. On détermine son axe en déterminant $\ker(M-I_3)$ et une mesure θ de son angle non orienté à l'aide de sa trace (puisque $\operatorname{tr}(M)=1+2\cos(\theta)$. On oriente ensuite $\ker(M-I_3)$ en faisant le choix d'un vecteur U_1.

  Pour déterminer le signe de θ (à valeurs dans $]-\pi,\pi[$), on fait le choix d'un vecteur unitaire U_2 orthogonal à U_1 et l'on regarde le signe du produit scalaire $\langle MU_2, U_1\wedge U_2\rangle$.
- Si M appartient à $O^-(3)$, on reconnaît facilement le car $M=-I_3$ et M sera la matrice d'une réflexion si et seulement si elle est symétrique. Dans ce cas, on détermine l'hyperplan de la réflexion en déterminant $\ker(M-I_3)$. Sinon, on détermine l'axe de rotation (et donc le plan de réflexion) en déterminant $\ker(M+I_3)$ et l'on obtient la mesure de l'angle orienté en procédant comme précédemment.

5.4.3 Groupe orthogonal en dimension quelconque

Théorème 5.14 *Soit E un espace euclidien de dimension n non nulle et u une isométrie vectorielle de E. Il existe une base orthonormale de E dans laquelle la matrice de u a la forme*

$$\begin{pmatrix} I_p & & & & \\ & -I_q & & & \\ & & R_1 & & \\ & & & \ddots & \\ & & & & R_r \end{pmatrix}$$

où $p+q+2r=n$ (si l'un des entiers est nul, les blocs correspondants n'existent pas) et où $R_i=\begin{pmatrix}\cos(\theta_i) & -\sin(\theta_i)\\ \sin(\theta_i) & \cos(\theta_i)\end{pmatrix}$.

DÉMONSTRATION. La preuve se fait par récurrence sur la dimension n de l'espace. En dimension 1, le résultat est immédiat et il a déjà été démontré en dimensions 2 (voir le théorème 5.9) et 3 (voir le théorème 5.13).

Supposons que n soit supérieur ou égal à 4 et que le résultat soit vrai pour les isométries vectorielles définies sur les espaces euclidiens de dimension inférieure ou égal à $n-1$.

Si l'isométrie vectorielle u admet 1 ou -1 comme valeur propre, alors, pour tout vecteur propre unitaire e_1 associé à cette valeur propre, le sous-espace vectoriel $H=\{e_1\}^\perp$ est stable par u et il existe alors une base orthonormée $\mathscr{B}_H$ de H dans laquelle la matrice de la restriction de u à H est de la forme voulue. Dans la base orthonormale $\{e_1\}\cup\mathscr{B}_H$, la matrice de u est alors de la forme

$$\begin{pmatrix}\pm 1 & 0\\ 0 & \operatorname{mat}_{\mathscr{B}_H}(u_{|_H})\end{pmatrix},$$

qui, en échangeant au besoin le vecteur e_1 avec l'un de ceux de $\mathscr{B}_H$, est bien celle souhaitée.

Si toutes les valeurs propres de u sont complexes non réelles, on a la décomposition $E=\bigoplus_{i=1}^r V_i$, où les sous-espaces vectoriels V_i, $i=1,\dots,r$, sont de dimension 2, deux à deux orthogonaux et stables par u. L'étude du cas $n=2$ montre alors que, pour chaque entier i de $\{1,\dots,r\}$, il existe une base orthonormale de $\mathscr{B}_i$ de V_i dans laquelle la matrice de la restriction de u à V_i est de la forme

$$R_i\begin{pmatrix}\cos(\theta_i) & -\sin(\theta_i)\\ \sin(\theta_i) & \cos(\theta_i)\end{pmatrix},$$

avec θ_i un réel. En concaténant l'ensemble de ces bases, on obtient une base orthonormée de E dans laquelle la matrice de u est

$$\begin{pmatrix} R_1 & 0 & \dots & 0\\ 0 & \ddots & \ddots & \vdots\\ \vdots & \ddots & \ddots & 0\\ 0 & \dots & 0 & R_r\end{pmatrix}.$$

□

Remarque

On a $p=\dim(\ker(u-id_E))$ et $q=\dim(\ker(u+id_E))$. De plus, on a que u appartient à $SO(E)$ (resp. à $O^-(E)$) si et seulement si q est pair (resp. impair).