ALGÈBRE BILINÉAIRE

ALGÈBRE BILINÉAIRE

TABLE DES MATIÈRES

La plupart des lois physiques, des processus biologiques sont décrits par des équations non linéaires. Une des façons les plus simples d'étudier de tels phénomènes est de les « linéariser », c'est-à-dire de les approcher par des équations linéaires. La formule de Taylor indique que l'on peut approcher une fonction de $\mathbb{R}$ dans $\mathbb{R}$ qui est régulière, au premier ordre par une fonction linéaire et au second ordre par une fonction « quadratique ». L'interprétation géométrique de ces approximations est la suivante : une courbe est approchée au premier ordre par sa tangente ; au second ordre elle est approchée par une courbe appelée conique. Les notions relevant de l'algèbre linéaire ont été introduites aux S2 et S3. Cette UE est consacrée à l'algèbre bilinéaire. Outre les formes quadratiques, différentes notions relevant de l'algèbre bilinéaire sont présentées, dont celle de produit scalaire. L'algèbre bilinéaire est une porte naturelle vers les notions d'orthogonalité et d'espace euclidien.

Sauf mention du contraire tous les espaces vectoriels considérés sont des espaces vectoriels sur $\mathbb{R}$, autrement dit le corps de base est $\mathbb{R}$.

CHAPITRE 1

FORMES BILINÉAIRES ET FORMES QUADRATIQUES

1.1. Formes bilinéaires

Rappelons qu'une application f entre deux $\mathbb{R}$-espaces vectoriels E et F est *linéaire* lorsque

$$\forall\,\mathbf{x}\in E,\ \forall\,\mathbf{y}\in E,\ \forall\,\lambda\in\mathbb{R},\ \forall\,\mu\in\mathbb{R}\qquad f(\lambda\mathbf{x}+\mu\mathbf{y})=\lambda f(\mathbf{x})+\mu f(\mathbf{y}).$$

De manière « analogue » une application f « à deux variables » de $E\times E$ dans E est bilinéaire si elle est linéaire par rapport à chacune de ces deux variables. Plus précisément

Définition 1.1.1

Soient E et F deux $\mathbb{R}$-espaces vectoriels. L'application

$$\varphi\colon E\times E\to F,\qquad\qquad (\mathbf{x},\mathbf{y})\mapsto\varphi(\mathbf{x},\mathbf{y})$$

est *bilinéaire* si pour tous $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}$ dans E et pour tous λ, μ dans $\mathbb{R}$ nous avons

$$\varphi(\lambda\mathbf{x}+\mu\mathbf{y},\mathbf{z})=\lambda\varphi(\mathbf{x},\mathbf{z})+\mu\varphi(\mathbf{y},\mathbf{z})\qquad\varphi(\mathbf{x},\lambda\mathbf{y}+\mu\mathbf{z})=\lambda\varphi(\mathbf{x},\mathbf{y})+\mu\varphi(\mathbf{x},\mathbf{z}).$$

Si $\mathbf{x}=\displaystyle\sum_{i=1}^{p}\lambda_i x_i$ et $\mathbf{y}=\displaystyle\sum_{j=1}^{p}\mu_j y_j$ avec λ_i, μ_j dans $\mathbb{R}$ et x_i, y_j dans E, alors

(1.1.1)
$$\varphi(\mathbf{x},\mathbf{y})=\varphi\left(\sum_{i=1}^{p}\lambda_i x_i,\mathbf{y}\right)=\sum_{i=1}^{p}\lambda_i\varphi(x_i,\mathbf{y})=\sum_{i=1}^{p}\lambda_i\varphi\left(x_i,\sum_{j=1}^{p}\mu_j y_j\right)=\sum_{i=1}^{p}\sum_{j=1}^{p}\lambda_i\mu_j\varphi(x_i,y_j).$$

De plus, pour tout $\mathbf{x}\in E$, nous avons

$$\varphi(\mathbf{x},\mathbf{0}_E)=\mathbf{0}_F,\qquad\qquad\varphi(\mathbf{0}_E,\mathbf{x})=\mathbf{0}_F.$$

Nous notons $\mathcal{L}_2(E\times E,F)$ l'ensemble des applications bilinéaires de $E\times E$ dans F. L'ensemble des applications de $E\times E$ dans F est un $\mathbb{R}$-espace vectoriel dont l'ensemble des applications bilinéaires de $E \times E$ dans F est un sous-espace vectoriel ; il possède ainsi lui-même une structure de $\mathbb{R}$-espace vectoriel : pour tous f, g dans $\mathcal{L}_2(E \times E, F)$ et pour tous α, β dans $\mathbb{R}$ nous avons

$$\alpha f + \beta g \in \mathcal{L}_2(E \times E, F).$$

Définitions 1.1.1

On appelle *forme* toute application définie sur un espace vectoriel qui est à valeurs dans le corps de base de cet espace vectoriel.
Ainsi, une *forme bilinéaire* sur le $\mathbb{R}$-espace vectoriel E est une application bilinéaire f de $E \times E$ à valeurs dans $\mathbb{R}$.

Remarques 1.1.1. — 1. Rappelons que sauf mention explicite du contraire le corps de base des espaces vectoriels est le corps des nombres réels $\mathbb{R}$. Signalons toutefois qu'une théorie similaire à celle que nous présentons existe pour des espaces vectoriels sur des corps différents, comme par exemple le corps $\mathbb{C}$ des nombres complexes.

Exemples 1.1.1. — 1. L'application $\varphi \colon \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$ définie pour $\mathbf{x} = (x_1, x_2, x_3)$ et $\mathbf{y} = (y_1, y_2, y_3)$ par

$$\varphi(\mathbf{x}, \mathbf{y}) = x_1y_1 + x_2y_2 + x_3y_3$$

est une forme bilinéaire sur $E = \mathbb{R}^3$.

2. Dans l'espace vectoriel $\mathbb{R}^4$ l'application $\varphi \colon \mathbb{R}^4 \times \mathbb{R}^4 \to \mathbb{R}$ définie pour $\mathbf{x} = (x_1, x_2, x_3, x_4)$ et $\mathbf{y} = (y_1, y_2, y_3, y_4)$ par

$$\varphi(\mathbf{x}, \mathbf{y}) = x_1y_1 + x_2y_2 + x_3y_3 - c^2x_4y_4$$

où c désigne un paramètre réel est une forme bilinéaire [1].

3. Sur l'espace vectoriel $\mathcal{C}([0,1], \mathbb{R})$ des applications continues de $[0,1]$ dans $\mathbb{R}$, l'application $\varphi \colon \mathcal{C}([0,1], \mathbb{R}) \times \mathcal{C}([0,1], \mathbb{R}) \to \mathbb{R}$ définie par

$$\varphi(f, g) = \int_0^1 f(t)g(t)\,\mathrm{d}t$$

est une forme bilinéaire. Ce résultat découle des propriétés de l'intégrale.

4. Sur l'espace vectoriel $\mathrm{M}(n, \mathbb{R})$ des matrices carrées de taille n à coefficients réels, l'application φ définie par

$$\varphi \colon \mathrm{M}(n, \mathbb{R}) \times \mathrm{M}(n, \mathbb{R}) \to \mathbb{R}, \qquad (A, B) \mapsto \mathrm{tr}({}^{\mathrm{t}}AB)$$

est une forme bilinéaire. Le résultat découle de la linéarité de l'application trace qui à une matrice de $\mathrm{M}(n, \mathbb{R})$ associe la somme de ses éléments diagonaux.

1. Cette forme bilinéaire intervient dans la théorie de la relativité, elle porte le nom de *forme de Lorentz*.

1.2. Cas d'un espace vectoriel en dimension finie

1.2.1. Rappel sur le lien entre les matrices et les applications linéaires. — Soient E un $\mathbb{R}$-espace vectoriel de dimension p muni d'une base $\mathcal{B}$ et F un $\mathbb{R}$-espace vectoriel de dimension n muni d'une base $\mathcal{B}'$. À une application linéaire f de E dans F on peut associer une matrice relativement aux bases $\mathcal{B}$ et $\mathcal{B}'$. C'est une matrice de taille (n, p) à coefficients réels dont la $j$ième colonne est constituée des coordonnées de l'image du $j$ième vecteur de la base de départ $\mathcal{B}$ par rapport à la base d'arrivée $\mathcal{B}'$. Nous la notons $\mathrm{Mat}_{\mathcal{B},\mathcal{B}'}(f)$ et nous disons que la matrice $\mathrm{Mat}_{\mathcal{B},\mathcal{B}'}(f)$ représente f dans les bases $\mathcal{B}$ et $\mathcal{B}'$.

En pratique, en notant $\mathcal{B} = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_p)$ et $\mathcal{B}' = (\mathbf{e}'_1, \mathbf{e}'_2, \ldots, \mathbf{e}'_n)$ les deux bases nous avons

$$\mathrm{Mat}_{\mathcal{B},\mathcal{B}'}(f) = \begin{pmatrix} a_{1,1} & a_{1,2} & \ldots & a_{1,p} \\ a_{2,1} & a_{2,2} & \ldots & a_{2,p} \\ \vdots & \vdots & & \vdots \\ a_{n,1} & a_{n,2} & \ldots & a_{n,p} \end{pmatrix}$$

où $f(\mathbf{e}_1) = a_{1,1}\mathbf{e}'_1 + a_{2,1}\mathbf{e}'_2 + \ldots + a_{n,1}\mathbf{e}'_n$, $f(\mathbf{e}_2) = a_{1,2}\mathbf{e}'_1 + a_{2,2}\mathbf{e}'_2 + \ldots + a_{n,2}\mathbf{e}'_n$, $\ldots$, $f(\mathbf{e}_p) = a_{1,p}\mathbf{e}'_1 + a_{2,p}\mathbf{e}'_2 + \ldots + a_{n,p}\mathbf{e}'_n$. Cette identification entre les applications linéaires et les matrices permet d'utiliser toute une classe de résultats issue de l'algèbre matricielle.

1.2.2. Formes bilinéaires et matrices. — Comme pour les applications linéaires il est possible de représenter les formes bilinéaires par des matrices.

⚠ Une telle représentation n'est possible que pour des formes bilinéaires et non pour des applications bilinéaires à valeurs dans un espace vectoriel quelconque.

Soit E un $\mathbb{R}$-espace vectoriel de dimension n muni d'une base $\mathcal{B} = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$. Considérons un couple $(\mathbf{x}, \mathbf{y}) \in E^2$ donné dans la base $\mathcal{B}$ par

$$\mathbf{x} = \sum_{i=1}^{n} x_i \mathbf{e}_i, \qquad \mathbf{y} = \sum_{j=1}^{n} y_j \mathbf{e}_j.$$

La bilinéarité de φ entraîne que $\varphi(\mathbf{x}, \mathbf{y})$ peut s'écrire sous la forme

$$\varphi(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i y_j \varphi(\mathbf{e}_i, \mathbf{e}_j). \tag{1.2.1}$$

Autrement dit la connaissance des valeurs de $\varphi(\mathbf{e}_i, \mathbf{e}_j)$ pour tout couple $(i, j) \in \{1, \ldots, n\}^2$ permet d'évaluer $\varphi(\mathbf{x}, \mathbf{y})$ dès que l'on sait décomposer $\mathbf{x}$ et $\mathbf{y}$ dans la base $\mathcal{B}$: une forme bilinéaire est déterminée par les valeurs $a_{ij} = \varphi(\mathbf{e}_i, \mathbf{e}_j)$ qu'elle prend sur une base de E.

Définition 1.2.1

Soient E un $\mathbb{R}$-espace vectoriel de dimension n muni d'une base $\mathcal{B} = (\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n)$ et φ une forme bilinéaire sur E. On appelle *matrice de la forme bilinéaire* φ par rapport à la base $\mathcal{B}$ la matrice de $\mathrm{M}(n, \mathbb{R})$ dont les coefficients sont $\varphi(\mathbf{e}_i, \mathbf{e}_j)$ pour $(i, j) \in \{1, \dots, n\}^2$. On la note

$$\mathrm{Mat}_{\mathcal{B}}(\varphi) = \begin{pmatrix} \varphi(\mathbf{e}_1, \mathbf{e}_1) & \varphi(\mathbf{e}_1, \mathbf{e}_2) & \dots & \varphi(\mathbf{e}_1, \mathbf{e}_n) \\ \varphi(\mathbf{e}_2, \mathbf{e}_1) & \varphi(\mathbf{e}_2, \mathbf{e}_2) & \dots & \varphi(\mathbf{e}_2, \mathbf{e}_n) \\ \vdots & \vdots & & \vdots \\ \varphi(\mathbf{e}_n, \mathbf{e}_1) & \varphi(\mathbf{e}_n, \mathbf{e}_2) & \dots & \varphi(\mathbf{e}_n, \mathbf{e}_n) \end{pmatrix}$$

et on dit que la matrice $\mathrm{Mat}_{\mathcal{B}}(\varphi)$ représente φ dans la base $\mathcal{B}$.

La relation (1.2.1) s'écrit matriciellement

$$\varphi(\mathbf{x}, \mathbf{y}) = {}^{\mathrm{t}}XAY \tag{1.2.2}$$

où $A = \mathrm{Mat}_{\mathcal{B}}(\varphi)$ et où X et Y sont les matrices colonnes représentant $\mathbf{x}$ et $\mathbf{y}$ dans la base $\mathcal{B}$:

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \qquad Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}.$$

Remarques 1.2.1. — 1. Inversement toute matrice de $\mathrm{M}(n, \mathbb{R})$ définit une forme bilinéaire sur E via la formule (1.2.2).

2. L'application qui à une forme bilinéaire φ associe sa matrice est un isomorphisme d'espaces vectoriels entre l'espace vectoriel $\mathcal{L}_2(E \times E, \mathbb{R})$ et l'espace vectoriel $\mathrm{M}(n, \mathbb{R})$. L'existence de cet isomorphisme permet d'affirmer que les dimensions du $\mathbb{R}$-espace vectoriel $\mathcal{L}_2(E \times E, \mathbb{R})$ et du $\mathbb{R}$-espace vectoriel $\mathrm{M}(n, \mathbb{R})$ sont les mêmes. Sachant que l'espace vectoriel $\mathrm{M}(n, \mathbb{R})$ est de dimension n^2, on en déduit que $\dim_{\mathbb{R}} \mathcal{L}_2(E \times E, \mathbb{R}) = n^2$.

Exemples 1.2.1. — 1. Pour déterminer la matrice de la forme bilinéaire définie sur $\mathbb{R}^3$ par

$$\forall \mathbf{x} \in \mathbb{R}^3, \forall \mathbf{y} \in \mathbb{R}^3 \quad \varphi(\mathbf{x}, \mathbf{y}) = x_1y_1 + x_2y_2 + x_3y_3$$

dans la base canonique de $\mathbb{R}^3$, on calcule les valeurs de $\varphi(\mathbf{e}_i, \mathbf{e}_j)$ pour tous les couples $(i, j) \in \{1, 2, 3\}^2$. Nous avons par exemple $\varphi(\mathbf{e}_1, \mathbf{e}_1) = 1$ et $\varphi(\mathbf{e}_1, \mathbf{e}_2) = 0$. Nous obtenons

$$\mathrm{Mat}_{\mathcal{B}_{\mathrm{can}}}(\varphi) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

2. La forme de Lorentz définie sur $\mathbb{R}^4$ par

$$\varphi(\mathbf{x}, \mathbf{y}) = x_1y_1 + x_2y_2 + x_3y_3 - c^2x_4y_4$$

avec $c \in \mathbb{R}$, admet pour matrice dans la base canonique de $\mathbb{R}^4$

$$\mathrm{Mat}_{\mathcal{B}_{\mathrm{can}}}(\varphi) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -c^2 \end{pmatrix}.$$

⚠ Une matrice carrée de taille n à coefficients réels peut représenter aussi bien une application linéaire entre deux espaces vectoriels de dimension n qu'une forme bilinéaire sur un espace de dimension n. Attention à ne pas confondre ces deux notions (applications linéaires de E dans F et formes bilinéaires sur E) même si toutes les deux peuvent être représentées par le même objet mathématique : une matrice.

1.3. Changement de base

Bien entendu, comme pour les applications linéaires, la notion de matrice représentant une forme bilinéaire dépend de la base choisie sur l'espace vectoriel E. Il est donc important de savoir comment est modifiée cette représentation lorsque l'on change de base.

1.3.1. Changement de base pour représenter une application linéaire. — Commençons par effectuer quelques rappels concernant les applications linéaires. Considérons une application linéaire f entre deux $\mathbb{R}$-espaces vectoriels E et F. On munit E des bases $\mathcal{B}$ et $\mathcal{B}'$ et F des bases $\mathcal{C}$ et $\mathcal{C}'$. Les deux matrices $M = \mathrm{Mat}_{\mathcal{B},\mathcal{C}}(f)$ et $M' = \mathrm{Mat}_{\mathcal{B}',\mathcal{C}'}(f)$ vérifient

$$M' = Q^{-1}MP \tag{1.3.1}$$

où P est la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$ et Q est la matrice de passage de $\mathcal{C}$ à $\mathcal{C}'$. Deux matrices M et M' pour lesquelles il existe deux matrices inversibles P et Q satisfaisant l'égalité (1.3.1) sont dites *équivalentes*. Deux matrices sont donc équivalentes si elles représentent la même application linéaire dans des bases différentes.

Considérons maintenant un endomorphisme f de E (c'est-à-dire une application linéaire de E dans E) et deux bases $\mathcal{B}$ et $\mathcal{B}'$ de l'espace vectoriel E. Les deux matrices $M = \mathrm{Mat}_{\mathcal{B}}(f)$ et $M' = \mathrm{Mat}_{\mathcal{B}'}(f)$ vérifient

$$M' = P^{-1}MP \tag{1.3.2}$$

où P désigne la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$. Deux matrices M et M' pour lesquelles il existe une matrice inversible P satisfaisant l'égalité (1.3.2) sont dites *semblables*. Deux matrices sont donc semblables si elles représentent le même endomorphisme dans deux bases différentes.

Rappelons que $\mathrm{GL}(n, \mathbb{R})$ désigne le sous-espace vectoriel de $\mathrm{M}(n, \mathbb{R})$ formé par les matrices inversibles.

1.3.2. Changement de base pour représenter une application bilinéaire. — Intéressons-nous désormais au changement de base dans la représentation matricielle d'une application bilinéaire.

Proposition 1.3.1

Soient $\mathcal{B}$ et $\mathcal{B}'$ deux bases d'un espace vectoriel E de dimension finie et P la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$. Si φ est une forme bilinéaire sur E, de matrice M dans la base $\mathcal{B}$ et de matrice M' dans la base $\mathcal{B}'$, alors

$$M' = {}^{\mathrm{t}}PMP.$$

Démonstration. — Soient $\mathcal{B} = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ et $\mathcal{B}' = (\mathbf{e}'_1, \mathbf{e}'_2, \ldots, \mathbf{e}'_n)$ deux bases d'un espace vectoriel E de dimension finie n. Soit φ une forme bilinéaire sur E. Par définition de la matrice d'une forme bilinéaire dans une base donnée les matrices $M = \mathrm{Mat}_{\mathcal{B}}(\varphi)$ et $M' = \mathrm{Mat}_{\mathcal{B}'}(\varphi)$ ont respectivement pour coefficients $(m_{i,j})_{1\leq i,\, j\leq n}$ et $(m'_{i,j})_{1\leq i,\, j\leq n}$ où pour tous $1 \leq i,\, j \leq n$

$$m_{i,j} = \varphi(\mathbf{e}_i, \mathbf{e}_j), \qquad m'_{i,j} = \varphi(\mathbf{e}'_i, \mathbf{e}'_j).$$

En notant $P = (p_{i,j})_{1\leq i,\, j\leq n}$ la matrice de passage de la base $\mathcal{B}$ à la base $\mathcal{B}'$ nous avons pour tout $1 \leq i \leq n$

$$\mathbf{e}'_i = \sum_{k=1}^{n} p_{ki}\mathbf{e}_k.$$

En utilisant la bilinéarité de l'application φ nous obtenons pour tous $1 \leq i,\, j \leq n$

$$m'_{i,j} = \varphi\left(\sum_{k=1}^{n} p_{ki}\mathbf{e}_k, \sum_{\ell=1}^{n} p_{\ell j}\mathbf{e}_\ell\right) = \sum_{k=1}^{n}\sum_{\ell=1}^{n} p_{ki}p_{\ell j}\varphi(\mathbf{e}_k, \mathbf{e}_\ell).$$

Cette égalité se réécrit en faisant intervenir les coefficients $m_{ij} = \varphi(\mathbf{e}_i, \mathbf{e}_j)$ de la manière suivante :

$$m'_{i,j} = \sum_{k=1}^{n}\sum_{\ell=1}^{n} p_{ki}p_{\ell j}m_{k\ell} = \sum_{k=1}^{n} p_{ki}\left(\sum_{\ell=1}^{n} m_{k\ell}p_{\ell j}\right) = \sum_{k=1}^{n} p_{ki}q_{kj}$$

où q_{kj} désigne le coefficient (k, j) de la matrice MP. Le membre de droite de l'égalité correspond donc à l'expression du terme (i, j) du produit matriciel ${}^{\mathrm{t}}P(MP)$, *i.e.*

$$m'_{i,j} = ({}^{\mathrm{t}}PMP)_{i,j}.$$

Cette relation étant vraie pour tout $1 \leq i,\, j \leq n$ nous avons bien $M' = {}^{\mathrm{t}}PMP$. □

1.3.3. Matrices congrues. —

Définition 1.3.1

Une matrice M de $\mathrm{M}(n, \mathbb{R})$ est dite *congrue* à une matrice M' de $\mathrm{M}(n, \mathbb{R})$ s'il existe un élément P de $\mathrm{GL}(n, \mathbb{R})$ tel que $M' = {}^{\mathrm{t}}PMP$.

Remarque 1.3.1. — La Définition 1.3.1 indique que deux matrices carrées de même taille sont congrues si elles représentent la même forme bilinéaire dans deux bases différentes.

Proposition 1.3.2

La relation de congruence entre matrices de $\mathrm{M}(n, \mathbb{R})$ définit une relation d'équivalence sur $\mathrm{M}(n, \mathbb{R})$.

Démonstration. —

- ◇ Toute matrice M de $\mathrm{M}(n, \mathbb{R})$ est congrue à elle-même (propriété de réflexivité) puisque $M = {}^{\mathrm{t}}\mathrm{Id} M \mathrm{Id}$, $\mathrm{Id} \in \mathrm{GL}(n, \mathbb{R})$.
- ◇ Soient M et M' deux matrices de $\mathrm{M}(n, \mathbb{R})$. Montrons que si M est congrue à M', alors M' est congrue à M (propriété de symétrie). La matrice M est congrue à M' s'il existe P dans $\mathrm{GL}(n, \mathbb{R})$ tel que $M' = {}^{\mathrm{t}}PMP$. En multipliant à gauche par $({}^{\mathrm{t}}P)^{-1}$ et à droite par P^{-1} nous en déduisons que $({}^{\mathrm{t}}P)^{-1} M' P^{-1} = M$, puis [2] $M = {}^{\mathrm{t}}(P^{-1}) M' P^{-1}$ qui permet de conclure que M' est congrue à M.
- ◇ Soient M, M' et M'' trois matrices de $\mathrm{M}(n, \mathbb{R})$. Montrons que si M est congrue à M' et si M' est congrue à M'', alors M est congrue à M'' (propriété de transitivité). D'une part si M est congrue à M', alors il existe $P \in \mathrm{GL}(n, \mathbb{R})$ tel que

$$M' = {}^{\mathrm{t}}PMP; \tag{1.3.3}$$

d'autre part si M' est congrue à M'' alors il existe $Q \in \mathrm{GL}(n, \mathbb{R})$ tel que

$$M'' = {}^{\mathrm{t}}QM'Q. \tag{1.3.4}$$

En combinant les égalités (1.3.3) et (1.3.4) nous obtenons $M'' = {}^{\mathrm{t}}Q{}^{\mathrm{t}}PMPQ$. Comme ${}^{\mathrm{t}}Q{}^{\mathrm{t}}P = {}^{\mathrm{t}}(PQ)$ et que le produit de deux matrices inversibles est une matrice inversible, cette relation s'écrit sous la forme $M'' = {}^{\mathrm{t}}(PQ)M(PQ)$. Ainsi M'' est congrue à M.

□

2. Nous utilisons ici le fait que ${}^{\mathrm{t}}(P^{-1}) = ({}^{\mathrm{t}}P)^{-1}$ pour tout P dans $\mathrm{GL}(n, \mathbb{R})$.

Si deux matrices sont congrues, alors elles sont équivalentes. Par ailleurs, le fait que deux matrices équivalentes aient toujours le même rang justifie la définition suivante.

Définition 1.3.2

Soit φ une forme bilinéaire sur un $\mathbb{R}$-espace vectoriel E de dimension finie. Nous appelons *rang* de la forme bilinéaire φ le rang de la matrice de φ dans une base quelconque de E.

Exemple 1.3.1. — Calculons le rang de la forme de Lorentz dont la matrice dans la base canonique de $\mathbb{R}^4$ s'écrit

$$\mathrm{Mat}_{\mathcal{B}_{\mathrm{can}}}(\varphi) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -c^2 \end{pmatrix}.$$

Les trois premiers vecteurs lignes de cette matrice forment une famille de trois vecteurs libres dans $\mathbb{R}^4$ (ce sont en effet les trois premiers vecteurs de la base canonique de $\mathbb{R}^4$).

- ◇ Si $c = 0$, alors le quatrième vecteur ligne de la matrice $\mathrm{Mat}_{\mathcal{B}_{\mathrm{can}}}(\varphi)$ est le vecteur nul, il est par conséquent lié aux trois premiers et l'espace vectoriel engendré par les quatre vecteurs lignes de la matrice $\mathrm{Mat}_{\mathcal{B}_{\mathrm{can}}}(\varphi)$ est un espace vectoriel de dimension 3.
- ◇ Si c est non nul, alors les quatre vecteurs lignes de la matrice $\mathrm{Mat}_{\mathcal{B}_{\mathrm{can}}}(\varphi)$ forment une famille libre de $\mathbb{R}^4$ et engendrent un espace vectoriel de dimension 4. Dans ce cas, le rang de la forme de Lorentz vaut 4.

1.4. Formes bilinéaires symétriques

Définition 1.4.1

Soit φ une forme bilinéaire sur un $\mathbb{R}$-espace vectoriel E.

- ◇ φ est *symétrique* si pour tous $\mathbf{x}$, $\mathbf{y}$ dans E nous avons $\varphi(\mathbf{x}, \mathbf{y}) = \varphi(\mathbf{y}, \mathbf{x})$;
- ◇ φ est *antisymétrique* si pour tous $\mathbf{x}$, $\mathbf{y}$ dans E nous avons $\varphi(\mathbf{x}, \mathbf{y}) = -\varphi(\mathbf{y}, \mathbf{x})$.

Nous noterons $\mathcal{L}_2^s(E \times E, \mathbb{R})$ l'ensemble des formes bilinéaires symétriques sur E et $\mathcal{L}_2^a(E \times E, \mathbb{R})$ l'ensemble des formes bilinéaires antisymétriques sur E. Ces deux ensembles sont des sous-espaces vectoriels de l'espace vectoriel $\mathcal{L}_2(E \times E, \mathbb{R})$ des formes bilinéaires sur E.

Exemples 1.4.1. — 1. La forme bilinéaire sur $\mathbb{R}^3$ définie par

$$\varphi(\mathbf{x}, \mathbf{y}) = x_1y_1 + x_2y_2 + x_3y_3$$

est une forme bilinéaire symétrique.

2. La forme bilinéaire sur $\mathrm{M}(n,\mathbb{R})$ définie par

$$\mathrm{M}(n,\mathbb{R}) \times \mathrm{M}(n,\mathbb{R}) \to \mathbb{R}, \qquad (A,B) \mapsto \mathrm{tr}({}^{\mathrm{t}}AB)$$

est une forme bilinéaire symétrique. La trace d'une matrice étant égale à la somme des éléments diagonaux et les éléments diagonaux restant inchangés lors de la transposition, nous avons pour tous A, B dans $\mathrm{M}(n,\mathbb{R})$

$$\varphi(A,B) = \mathrm{tr}({}^{\mathrm{t}}AB) = \mathrm{tr}({}^{\mathrm{t}}({}^{\mathrm{t}}AB));$$

puisque ${}^{\mathrm{t}}({}^{\mathrm{t}}AB) = {}^{\mathrm{t}}BA$ nous en déduisons que $\varphi(A,B) = \mathrm{tr}({}^{\mathrm{t}}BA) = \varphi(B,A)$.

L'énoncé suivant résulte de la Définition 1.4.1 ainsi que de la représentation matricielle des formes bilinéaires en dimension finie introduite dans la partie précédente.

Proposition 1.4.1

Soient E un $\mathbb{R}$-espace vectoriel de dimension finie, $\mathcal{B}$ une base de E et φ une forme bilinéaire sur E. La forme bilinéaire φ est symétrique (resp. antisymétrique) si et seulement si la matrice $\mathrm{Mat}_{\mathcal{B}}(\varphi)$ est symétrique (resp. antisymétrique).

Rappelons qu'une matrice A de $\mathrm{M}(n,\mathbb{R})$ est *symétrique* si ${}^{\mathrm{t}}A = A$ et *antisymétrique* si ${}^{\mathrm{t}}A = -A$. L'ensemble des matrices symétriques de taille n à coefficients réels est noté $\mathrm{S}(n,\mathbb{R})$ et l'ensemble des matrices antisymétriques de taille n à coefficients réels est noté $\mathrm{A}(n,\mathbb{R})$. Les deux sous-espaces vectoriels $\mathrm{S}(n,\mathbb{R})$ et $\mathrm{A}(n,\mathbb{R})$ sont en somme directe dans $\mathrm{M}(n,\mathbb{R})$. Plus précisément nous avons l'énoncé ci-dessous basé sur la décomposition de toute matrice M de $\mathrm{M}(n,\mathbb{R})$ comme la somme d'une matrice symétrique S et d'une matrice antisymétrique A

$$M = \underbrace{\frac{1}{2}(M + {}^{\mathrm{t}}M)}_{S} + \underbrace{\frac{1}{2}(M - {}^{\mathrm{t}}M)}_{A}$$

Proposition 1.4.2

Les deux sous-espaces $\mathrm{S}(n,\mathbb{R})$ et $\mathrm{A}(n,\mathbb{R})$ sont supplémentaires dans $\mathrm{M}(n,\mathbb{R})$:

$$\mathrm{S}(n,\mathbb{R}) \oplus \mathrm{A}(n,\mathbb{R}) = \mathrm{M}(n,\mathbb{R}).$$

De plus, les dimensions des $\mathbb{R}$-espaces vectoriels de $\mathrm{S}(n,\mathbb{R})$ et $\mathrm{A}(n,\mathbb{R})$ sont

$$\dim_{\mathbb{R}} \mathrm{S}(n,\mathbb{R}) = \frac{1}{2}n(n+1), \qquad \dim_{\mathbb{R}} \mathrm{A}(n,\mathbb{R}) = \frac{1}{2}n(n-1).$$

La Proposition 1.4.1 assure que toutes les propriétés des matrices symétriques ou antisymétriques réelles pourront être interprétées comme propriétés des formes bilinéaires symétriques ou antisymétriques. L'enoncé suivant en est l'illustration :

Proposition 1.4.3

Soit E un $\mathbb{R}$-espace vectoriel. Les deux sous-espaces $\mathcal{L}_2^s(E \times E, \mathbb{R})$ et $\mathcal{L}_2^a(E \times E, \mathbb{R})$ sont supplémentaires dans $\mathcal{L}_2(E \times E, \mathbb{R})$:

$$\mathcal{L}_2^s(E \times E, \mathbb{R}) \oplus \mathcal{L}_2^a(E \times E, \mathbb{R}) = \mathcal{L}_2(E \times E, \mathbb{R}).$$

Dans le cas où E est un $\mathbb{R}$-espace vectoriel de dimension finie $n \in \mathbb{N}^*$ nous avons

$$\dim_{\mathbb{R}} \mathcal{L}_2^s(E \times E, \mathbb{R}) = \frac{1}{2}n(n+1), \qquad \dim_{\mathbb{R}} \mathcal{L}_2^a(E \times E, \mathbb{R}) = \frac{1}{2}n(n-1).$$

⚠ La Proposition 1.4.3 n'est pas un Corollaire de la Proposition 1.4.2 car elle couvre aussi le cas où E n'est pas de dimension finie. La démonstration reste identique à celle qui est généralement utilisée pour démontrer la Proposition 1.4.2 ; en particulier on montre que toute forme bilinéaire φ sur un $\mathbb{R}$-espace vectoriel E se décompose de la façon suivante : pour tous $\mathbf{x}$, $\mathbf{y}$ dans E nous avons

$$\varphi(\mathbf{x}, \mathbf{y}) = \underbrace{\frac{1}{2}\Big(\varphi(\mathbf{x}, \mathbf{y}) + \varphi(\mathbf{y}, \mathbf{x})\Big)}_{\varphi_s(\mathbf{x}, \mathbf{y})} + \underbrace{\frac{1}{2}\Big(\varphi(\mathbf{x}, \mathbf{y}) - \varphi(\mathbf{y}, \mathbf{x})\Big)}_{\varphi_a(\mathbf{x}, \mathbf{y})}$$

où l'application $\varphi_s \colon (\mathbf{x}, \mathbf{y}) \mapsto \frac{1}{2}\Big(\varphi(\mathbf{x}, \mathbf{y}) + \varphi(\mathbf{y}, \mathbf{x})\Big)$ est une forme bilinéaire symétrique et où $\varphi_a \colon (\mathbf{x}, \mathbf{y}) \mapsto \frac{1}{2}\Big(\varphi(\mathbf{x}, \mathbf{y}) - \varphi(\mathbf{y}, \mathbf{x})\Big)$ est une forme bilinéaire antisymétrique.

Remarque 1.4.1. — Une forme bilinéaire φ_a antisymétrique sur E vérifie : $\varphi_a(\mathbf{x}, \mathbf{x}) = 0$ pour tout $\mathbf{x} \in E$; en effet, pour tout $\mathbf{x} \in E$ nous avons $\varphi_a(\mathbf{x}, \mathbf{x}) = -\varphi_a(\mathbf{x}, \mathbf{x})$.

1.5. Formes quadratiques

Définition 1.5.1

Une application q d'un espace vectoriel E dans $\mathbb{R}$ est une *forme quadratique* s'il existe une forme bilinéaire φ sur E telle que pour tout $\mathbf{x} \in E$

$$q(\mathbf{x}) = \varphi(\mathbf{x}, \mathbf{x}).$$

Remarque 1.5.1. — Soient H un sous-espace vectoriel d'un $\mathbb{R}$-espace vectoriel E et q une forme quadratique sur E. L'application $q_{|H} \colon x \in H \mapsto q(x)$ est une forme quadratique sur H. C'est la restriction de la forme quadratique q au sous-espace vectoriel H.

Par définition, toute forme quadratique q est issue d'une forme bilinéaire φ. Cette forme bilinéaire n'a aucune raison d'être unique ni d'être symétrique. Nous avons toutefois le résultat suivant :

Proposition 1.5.1: (Formule de polarisation)

Soit E un $\mathbb{R}$-espace vectoriel. Pour toute forme quadratique q sur E, il existe une unique forme bilinéaire symétrique φ_0 pour laquelle nous avons

$$\forall \mathbf{x} \in E \qquad q(\mathbf{x}) = \varphi_0(\mathbf{x}, \mathbf{x}).$$

Cette forme bilinéaire symétrique φ_0 s'appelle la *forme polaire* de q. Elle est donnée par la relation

$$\forall \mathbf{x}, \mathbf{y} \in E, \quad \varphi_0(\mathbf{x}, \mathbf{y}) = \frac{1}{2}(q(\mathbf{x}+\mathbf{y}) - q(\mathbf{x}) - q(\mathbf{y})) = \frac{1}{4}(q(\mathbf{x}+\mathbf{y}) - q(\mathbf{x}-\mathbf{y}))$$

appelée *formule de polarisation*.

Démonstration. — Considérons une forme quadratique q définie sur un $\mathbb{R}$-espace vectoriel. Par définition il existe une forme bilinéaire $\varphi\colon E \times E \to \mathbb{R}$ telle que

$$\forall \mathbf{x} \in E \quad q(\mathbf{x}) = \varphi(\mathbf{x}, \mathbf{x}). \tag{1.5.1}$$

Dans un premier temps, montrons que l'on peut trouver une forme bilinéaire symétrique φ_0 satisfaisant (1.5.1). D'après la Proposition 1.4.3 la forme bilinéaire φ se décompose comme la somme d'une forme bilinéaire symétrique φ_s et d'une forme bilinéaire antisymétrique φ_a : $\varphi = \varphi_s + \varphi_a$. La Remarque 1.4.1 assure que pour tout $\mathbf{x} \in E$ nous avons $\varphi_a(\mathbf{x}, \mathbf{x}) = 0$. Ainsi la forme quadratique q vérifie

$$\forall \mathbf{x} \in E \quad q(\mathbf{x}) = \varphi_s(\mathbf{x}, \mathbf{x}).$$

L'application φ_s est une forme bilinéaire symétrique ; nous posons $\varphi_0 = \varphi_s$.

Nous allons maintenant démontrer que toute forme bilinéaire symétrique satisfaisant l'égalité (1.5.1) est donnée par la formule de polarisation. Cela permettra de démontrer l'unicité de φ_0. Considérons une forme bilinéaire symétrique φ satisfaisant la relation (1.5.1). En utilisant la bilinéarité de φ, pour $\mathbf{x}, \mathbf{y} \in E$ nous avons

$$q(\mathbf{x}+\mathbf{y}) = \varphi(\mathbf{x}+\mathbf{y}, \mathbf{x}+\mathbf{y}) = \varphi(\mathbf{x}, \mathbf{x}) + \varphi(\mathbf{x}, \mathbf{y}) + \varphi(\mathbf{y}, \mathbf{x}) + \varphi(\mathbf{y}, \mathbf{y}).$$

Comme φ est supposée symétrique, cette relation s'écrit encore

$$q(\mathbf{x}+\mathbf{y}) = q(\mathbf{x}) + 2\varphi(\mathbf{x}, \mathbf{y}) + q(\mathbf{y}). \tag{1.5.2}$$

Nous en déduisons que $\varphi(\mathbf{x}, \mathbf{y}) = \frac{1}{2}(q(\mathbf{x}+\mathbf{y}) - q(\mathbf{x}) - q(\mathbf{y}))$ ce qui correspond à la première formule de polarisation. Ainsi φ est unique.

Pour établir la seconde formule de polarisation nous utilisons la relation (1.5.2) puis cette même relation en substituant $-\mathbf{y}$ à $\mathbf{y}$. Puisque $q(-\mathbf{y}) = \varphi(-\mathbf{y}, -\mathbf{y}) = \varphi(\mathbf{y}, \mathbf{y}) = q(\mathbf{y})$, nous en déduisons que :

$$q(\mathbf{x}+\mathbf{y}) = q(\mathbf{x}) + 2\varphi(\mathbf{x},\mathbf{y}) + q(\mathbf{y}), \qquad q(\mathbf{x}-\mathbf{y}) = q(\mathbf{x}) - 2\varphi(\mathbf{x},\mathbf{y}) + q(\mathbf{y}).$$

En ajoutant membre à membre ces deux relations, nous avons la seconde formule de polarisation $\varphi(\mathbf{x},\mathbf{y}) = \frac{1}{4}\big(q(\mathbf{x}+\mathbf{y}) - q(\mathbf{x}-\mathbf{y})\big)$. □

La façon la plus immédiate d'obtenir une forme quadratique est d'utiliser une forme bilinéaire φ et de construire l'application $\mathbf{x} \in E \mapsto \varphi(\mathbf{x},\mathbf{x})$.

Exemples 1.5.1. — 1. L'application $q\colon \mathbb{R}^3 \to \mathbb{R}$ définie par

$$q(\mathbf{x}) = x_1^2 + x_2^2 + x_3^2$$

est une forme quadratique sur $E = \mathbb{R}^3$. Elle est construite à partir de la forme bilinéaire sur $\mathbb{R}^3$ définie par

$$\varphi(\mathbf{x},\mathbf{y}) = x_1y_1 + x_2y_2 + x_3y_3.$$

Puisque cette forme bilinéaire est symétrique, nous déduisons de la Proposition 1.5.1 que φ est la forme polaire de q.

2. La forme bilinéaire

$$\mathcal{C}([0,1],\mathbb{R}) \times \mathcal{C}([0,1],\mathbb{R}) \to \mathbb{R}, \qquad (f,g) \mapsto \int_0^1 f(t)g(t)\,\mathrm{d}t$$

permet de construire la forme quadratique sur $\mathcal{C}([0,1],\mathbb{R})$ définie par

$$q(f) = \int_0^1 f(t)^2\,\mathrm{d}t.$$

3. Sur l'espace vectoriel $\mathrm{M}(n,\mathbb{R})$ l'application $q\colon \mathrm{M}(n,\mathbb{R}) \to \mathbb{R}$ définie par

$$q(A) = \mathrm{tr}({}^{\mathrm{t}}AA)$$

est une forme quadratique. Sa forme polaire est l'application bilinéaire

$$\mathrm{M}(n,\mathbb{R}) \times \mathrm{M}(n,\mathbb{R}) \to \mathbb{R}, \qquad (A,B) \mapsto \mathrm{tr}({}^{\mathrm{t}}AB).$$

Pour montrer qu'une application q définie sur un $\mathbb{R}$-espace vectoriel E et à valeurs dans $\mathbb{R}$ est une forme quadratique, on peut utiliser la caractérisation suivante :

Proposition 1.5.2

Une application q d'un espace vectoriel E dans $\mathbb{R}$ est une forme quadratique si et seulement si

- ◇ $\forall\, \mathbf{x} \in E$, $\forall\, \lambda \in \mathbb{R}$, $q(\lambda\mathbf{x}) = \lambda^2 q(\mathbf{x})$,
- ◇ l'application

$$E \times E \to \mathbb{R}, \qquad (\mathbf{x},\mathbf{y}) \mapsto \frac{1}{2}\big(q(\mathbf{x}+\mathbf{y}) - q(\mathbf{x}) - q(\mathbf{y})\big)$$

est une forme bilinéaire.

Démonstration. — Supposons que l'application $q\colon E \to \mathbb{R}$ soit une forme quadratique. Soit φ la forme bilinéaire la définissant. Pour tout $\lambda \in \mathbb{R}$, pour tout $\mathbf{x} \in E$ nous avons

$$q(\lambda\mathbf{x}) = \varphi(\lambda\mathbf{x}, \lambda\mathbf{x}) = \lambda^2\varphi(\mathbf{x},\mathbf{x}) = \lambda^2 q(\mathbf{x}).$$

L'application donnée par

$$E \times E \to \mathbb{R}, \qquad (\mathbf{x},\mathbf{y}) \mapsto \frac{1}{2}\big(q(\mathbf{x}+\mathbf{y}) - q(\mathbf{x}) - q(\mathbf{y})\big)$$

correspond à la forme polaire de la forme quadratique q, c'est donc une forme bilinéaire.

Réciproquement, supposons que l'application q vérifie les deux assertions de l'énoncé et considérons l'application

$$\varphi\colon E \times E \to \mathbb{R}, \qquad (\mathbf{x},\mathbf{y}) \mapsto \frac{1}{2}\big(q(\mathbf{x}+\mathbf{y}) - q(\mathbf{x}) - q(\mathbf{y})\big)$$

La seconde assertion assure que φ est une forme bilinéaire. D'après la première assertion nous avons

$$\forall\, \mathbf{x} \in E \qquad \varphi(\mathbf{x},\mathbf{x}) = \frac{1}{2}\big(q(2\mathbf{x}) - q(\mathbf{x}) - q(\mathbf{x})\big) = \frac{1}{2}\big(4q(\mathbf{x}) - 2q(\mathbf{x})\big) = q(\mathbf{x}).$$

Cela implique, d'après la Définition 1.5.1, que q est une forme quadratique. □

Exemple 1.5.2. — Montrons en utilisant la Proposition 1.5.2 que l'application q définie sur le $\mathbb{R}$-espace vectoriel $\mathbb{R}_2[X]$ des polynômes de degré 2 à coefficients réels par

$$q\colon \mathbb{R}_2[X] \to \mathbb{R}, \qquad P \mapsto P(0)P(1)$$

est une forme quadratique.

D'une part, pour tout $P \in \mathbb{R}_2[X]$ et tout $\lambda \in \mathbb{R}$ nous avons

$$q(\lambda P) = \lambda P(0)\lambda P(1) = \lambda^2 q(P).$$

La première condition de la Proposition 1.5.2 est satisfaite.

Montrons d'autre part que l'application

$$\mathbb{R}_2[x] \times \mathbb{R}_2[x] \to \mathbb{R}, \qquad (P,Q) \mapsto \frac{1}{2}\big(q(P+Q) - q(P) - q(Q)\big)$$

est une application bilinéaire sur $\mathbb{R}_2[X]$. Pour tous P, Q dans $\mathbb{R}_2[X]$ nous avons

$$\varphi(P,Q) = \frac{1}{2}(P(0)Q(1) + P(1)Q(0)).$$

Remarquons tout d'abord que pour tous P, Q dans $\mathbb{R}_2[X]$ nous avons $\varphi(P,Q) = \varphi(Q,P)$, autrement dit, l'application φ est symétrique. Ainsi, pour démontrer que φ est bilinéaire, il suffit de démontrer qu'elle est linéaire par rapport à sa première variable. Soient P, Q et R dans $\mathbb{R}_2[X]$ et soient λ, μ dans $\mathbb{R}$. Nous avons

$$\begin{aligned}
\varphi(\lambda P + \mu Q, R) &= \frac{1}{2}\Big((\lambda P + \mu Q)(0)R(1) + (\lambda P + \mu Q)(1)R(0)\Big) \\
&= \frac{1}{2}\Big(\lambda P(0)R(1) + \mu Q(0)R(1) + \lambda P(1)R(0) + \mu Q(1)R(0)\Big) \\
&= \frac{1}{2}\lambda\Big(P(0)R(1) + P(1)R(0)\Big) + \frac{1}{2}\mu\Big(Q(0)R(1) + Q(1)R(0)\Big) \\
&= \lambda\varphi(P,R) + \mu\varphi(Q,R)
\end{aligned}$$

ce qui montre la linéarité de φ par rapport à la première variable.

Définition 1.5.2

Soient E un $\mathbb{R}$-espace vectoriel de dimension finie, $\mathcal{B}$ une base de E et q une forme quadratique sur E.

- ◇ On appelle *matrice* de q dans la base $\mathcal{B}$ la matrice de la forme polaire de q dans la base $\mathcal{B}$.
- ◇ Le *rang de la forme quadratique* q est le rang de la matrice de q.

En désignant par X la matrice colonne dont les éléments sont les composantes d'un vecteur $\mathbf{x} \in E$ dans une base donnée $\mathcal{B}$ de l'espace vectoriel E, par A la matrice d'une forme quadratique q sur E dans cette même base $\mathcal{B}$ et par φ_0 la forme polaire de q, l'égalité $q(\mathbf{x}) = \varphi_0(\mathbf{x},\mathbf{x})$ s'écrit matriciellement

$$q(\mathbf{x}) = {}^{\mathrm{t}}XAX.$$

Remarques 1.5.2. — 1. Soit E un $\mathbb{R}$-espace vectoriel de dimension n muni d'une base $\mathcal{B} = (\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n)$. Considérons une forme quadratique q sur E dont la forme polaire est notée φ_0 et la matrice associée A. Si $\mathbf{x} \in E$ est donné dans la base $\mathcal{B}$ par $\mathbf{x} = \sum_{i=1}^{n} x_i\mathbf{e}_i$ alors

$$q(\mathbf{x}) = \varphi_0(\mathbf{x},\mathbf{x}) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j.$$

Dans la somme sur l'indice noté j nous pouvons distinguer les cas $j = i$, $j > i$ et $j < i$. On remarque que puisque φ_0 est la forme polaire de q, sa matrice A est symétrique et les coefficients a_{ij} et a_{ji} sont égaux. Nous obtenons

$$q(\mathbf{x}) = \sum_{i=1}^{n} a_{ii}x_i^2 + 2\sum_{i=1}^{n}\sum_{k=i+1}^{n} a_{ij}x_ix_j.$$

2. Réciproquement si A est une matrice symétrique, l'expression

$$q(\mathbf{x}) = \sum_{i=1}^{n} a_{ii}x_i^2 + 2\sum_{i=1}^{n}\sum_{k=i+1}^{n} a_{ij}x_ix_j.$$

définit la forme quadratique $q\colon \mathbf{x} \mapsto q(\mathbf{x})$ associée à la forme bilinéaire symétrique donnée par

$$\varphi_0(\mathbf{x},\mathbf{y}) = \sum_{i=1}^{n} a_{ii}x_iy_i + \sum_{i=1}^{n}\sum_{j=i+1}^{n} a_{ij}(x_ix_j + x_jy_i).$$

Matriciellement cela permet d'associer à toute matrice symétrique A une forme quadratique q sur $\mathbb{R}^n$ telle que $q(\mathbf{x}) = {}^{\mathrm{t}}XAX$, le vecteur X correspondant à la matrice colonne dont les éléments sont les composantes du vecteur $\mathbf{x}$ dans la base canonique de $\mathbb{R}^n$. La forme polaire φ_0 associée à q est déterminée matriciellement par $\varphi_0(\mathbf{x},\mathbf{y}) = {}^{\mathrm{t}}XAX$.

Exemple 1.5.3. — L'application $q\colon \mathbb{R}^3 \to \mathbb{R}$ définie par

$$q(\mathbf{x}) = 3x_1^2 + x_2^2 + 2x_1x_2 - 3x_1x_3$$

est une forme quadratique sur $E = \mathbb{R}^3$. La matrice de q dans la base canonique de $\mathbb{R}^3$ est

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

Définition 1.5.3

Soit E un $\mathbb{R}$-espace vectoriel. Soit q une forme quadratique sur E.
On dit que la forme quadratique q est *non dégénérée* si son rang est égal à la dimension de E ; elle est dite *dégénérée* sinon.

1.6. Orthogonalité

Soient $\mathcal{D}_1$ et $\mathcal{D}_2$ deux droites dans le plan ; on connaît des critères permettant de dire si elles sont ou non perpendiculaires. Par exemple, en considérant les vecteurs directeurs

$$\mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}, \qquad \mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix},$$

des droites $\mathcal{D}_1$ et $\mathcal{D}_2$ il suffit d'évaluer la quantité $u_1v_1 + u_2v_2$ appelée produit scalaire des vecteurs $\mathbf{u}$ et $\mathbf{v}$. Si cette valeur est nulle, alors les droites sont perpendiculaires, sinon elles ne sont pas perpendiculaires. Comme nous allons le voir cette notion familière d'orthogonalité est liée à un choix d'application bilinéaire sur le plan et peut être généralisée à des espaces de dimension quelconque.

Définition 1.6.1

Soient E un $\mathbb{R}$-espace vectoriel, q une forme quadratique sur E et φ_0 la forme polaire de q. Deux vecteurs $\mathbf{x}$ et $\mathbf{y}$ de E sont dits *orthogonaux* selon q (ou selon φ_0) si $\varphi_0(\mathbf{x}, \mathbf{y}) = 0$.

Remarque 1.6.1. — Une forme polaire étant symétrique, si x et y sont orthogonaux selon q, alors y et x sont orthogonaux selon q.

Exemples 1.6.1. —

1. Dans l'espace vectoriel $\mathbb{R}^2$ la forme quadratique $q\colon \mathbf{x} \mapsto x_1^2 + x_2^2$ fournit la notion d'orthogonalité usuelle dans le plan. La forme polaire de q est l'application $\varphi_0\colon (\mathbf{x}, \mathbf{y}) \mapsto x_1y_1 + x_2y_2$ de sorte que deux vecteurs $\mathbf{u}$ et $\mathbf{v}$ de composantes (u_1, u_2) et (v_1, v_2) sont orthogonaux selon cette forme quadratique q si et seulement si $u_1v_1 + u_2v_2 = 0$. Ainsi selon la forme quadratique q les vecteurs
$$\mathbf{u}' = \begin{pmatrix} 4 \\ 0 \end{pmatrix}, \qquad \mathbf{v}' = \begin{pmatrix} 5 \\ 5 \end{pmatrix},$$
ne sont pas orthogonaux car $\varphi_0(\mathbf{u}', \mathbf{v}') = 20 \neq 0$ mais les vecteurs
$$\mathbf{u} = \begin{pmatrix} 4 \\ 0 \end{pmatrix}, \qquad \mathbf{v} = \begin{pmatrix} 0 \\ 5 \end{pmatrix},$$
le sont car $\varphi_0(\mathbf{u}, \mathbf{v}) = 0$.
2. Dans l'espace vectoriel $\mathbb{R}^2$ la forme quadratique $q\colon \mathbf{x} \mapsto x_1^2 - 2x_1x_2$ a pour forme polaire $\varphi_0\colon (\mathbf{x}, \mathbf{y}) \mapsto x_1y_1 - x_1y_2 - x_2y_1$. Selon cette forme quadratique q les vecteurs
$$\mathbf{u}' = \begin{pmatrix} 4 \\ 0 \end{pmatrix}, \qquad \mathbf{v}' = \begin{pmatrix} 5 \\ 5 \end{pmatrix},$$
sont orthogonaux car $\varphi_0(\mathbf{u}', \mathbf{v}') = 0$ mais les vecteurs
$$\mathbf{u} = \begin{pmatrix} 4 \\ 0 \end{pmatrix}, \qquad \mathbf{v} = \begin{pmatrix} 0 \\ 5 \end{pmatrix},$$
ne le sont pas car $\varphi_0(\mathbf{u}, \mathbf{v}) = -20 \neq 0$.

Définitions 1.6.1: (orthogonal)

Soient A et B deux sous-ensembles d'un $\mathbb{R}$-espace vectoriel E, q une forme quadratique sur E et φ_0 la forme polaire associée à q.

- On appelle *orthogonal* de A selon q (ou selon φ_0) l'ensemble
$$A^\perp = \{\mathbf{x} \in E \mid \forall \mathbf{y} \in A \quad \varphi_0(\mathbf{x}, \mathbf{y}) = 0\}.$$
- Les deux ensembles A et B sont dit *orthogonaux* selon q (ou selon φ_0) et on note $A \perp B$ si
$$\forall \mathbf{x} \in A \ \forall \mathbf{y} \in B \ \ \varphi_0(\mathbf{x}, \mathbf{y}) = 0.$$

Notons que $\{\mathbf{0}\}^\perp = E$ car pour tout $\mathbf{x} \in E$ on a $\varphi_0(\mathbf{x}, \mathbf{0}) = 0$. Pour la même raison l'ensemble $\{\mathbf{0}\}$ est orthogonal à tous les autres sous-ensembles de E.

De nombreuses formes quadratiques peuvent être définies sur un espace vectoriel E ; il existe donc plusieurs relations d'orthogonalité sur un même espace. La notation $\perp$ ne précise pas la forme quadratique pour laquelle il y a orthogonalité. En cas d'ambiguïté nous noterons $\perp_{\varphi_0}$ ou $\perp_q$ pour faire référence à la forme quadratique ou forme bilinéaire symétrique considérée.

Proposition 1.6.1

Soient E un $\mathbb{R}$-espace vectoriel, q une forme quadratique sur E et A un sous-ensemble de E.

- L'ensemble $A^\perp$ est un sous-espace vectoriel de E.
- Nous avons $A \subset A^{\perp\perp}$.

Démonstration. — Nous avons l'inclusion $A^\perp \subset E$. D'après la définition de l'orthogonal $A^\perp$, $\mathbf{0}$ appartient à $A^\perp$. Soient $\mathbf{x}$, $\mathbf{y}$ dans $A^\perp$, soient λ, μ dans $\mathbb{R}$ et soit $\mathbf{z}$ dans A ; alors
$$\varphi_0(\lambda\mathbf{x} + \mu\mathbf{y}, \mathbf{z}) = \lambda \underbrace{\varphi_0(\mathbf{x}, \mathbf{z})}_{0 \text{ car } \mathbf{x} \in A^\perp} + \mu \underbrace{\varphi_0(\mathbf{y}, \mathbf{z})}_{0 \text{ car } \mathbf{y} \in A^\perp} = 0$$
autrement dit $\lambda\mathbf{x} + \mu\mathbf{y}$ appartient à $A^\perp$.

Dire qu'un élément $\mathbf{z} \in E$ appartient à $A^{\perp\perp}$ signifie
$$\forall \mathbf{y} \in A^\perp \quad \varphi_0(\mathbf{z}, \mathbf{y}) = 0.$$
Soit $\mathbf{x} \in A$; pour tout $y \in A^\perp$, par définition de $A^\perp$ nous avons $\varphi_0(\mathbf{x}, \mathbf{y}) = 0$. Nous en déduisons que $\mathbf{x}$ appartient à $A^{\perp\perp}$ ce qui montre que $A \subset A^{\perp\perp}$. $\square$

Exemple 1.6.2. — Soient $E = \mathbb{R}_2[X]$ et

$$q\colon \mathbb{R}_2[X] \to \mathbb{R}, \qquad P \mapsto P(0)P(1).$$

Remarquons que q est une forme quadratique sur $\mathbb{R}_2[X]$ dont la forme polaire $\varphi_0\colon \mathbb{R}_2[X] \times \mathbb{R}_2[X] \to \mathbb{R}$ est définie par

$$\varphi_0(P, Q) = \frac{1}{2}\Big(P(0)Q(1) + P(1)Q(0)\Big).$$

Considérons le polynôme $P = X^2 + X + 1 \in \mathbb{R}_2[X]$ et le sous-ensemble $A = \{P\}$ de $\mathbb{R}_2[X]$. Par définition un polynôme $Q \in \mathbb{R}_2[X]$ appartient à $A^\perp$ si et seulement si $\varphi_0(P, Q) = 0$. Nous avons donc

$$Q \in A^\perp \iff \frac{1}{2}\Big(P(0)Q(1) + P(1)Q(0)\Big) = 0.$$

Compte tenu de l'expression du polynôme P nous obtenons

$$Q \in A^\perp \iff Q(1) + 3Q(0) = 0.$$

En écrivant Q dans la base canonique de $\mathbb{R}_2[X]$ sous la forme $Q = aX^2 + bX + c$ avec a, b et c trois nombres réels nous avons

$$Q \in A^\perp \iff a + b + 4c = 0.$$

L'ensemble $A^\perp$ est par conséquent constitué des polynômes Q de la forme $Q = (-b - 4c)X^2 + bX + c$, *i.e.* $Q = b(X - X^2) + c(1 - 4X^2)$, b et c étant deux réels quelconques. Nous en déduisons que

$$A^\perp = \mathrm{Vect}(X - X^2, 1 - 4X^2).$$

De la même manière un polynôme $R = \alpha X^2 + \beta X + \gamma \in \mathbb{R}_2[X]$ appartient à $A^{\perp\perp}$ si et seulement si

$$\begin{cases} \varphi_0(R, X - X^2) = 0 \\ \varphi_0(R, 1 - 4X^2) = 0 \end{cases}$$

si et seulement si

$$\frac{1}{2}\Big(R(0)(1-1) + R(1)(0-0)\Big) = \frac{1}{2}\Big(R(0)(1-4) + R(1)(1-0)\Big) = 0$$

si et seulement si $R(0)(-3) + R(1) = 0$ si et seulement si $\gamma(-3) + \alpha + \beta + \gamma = 0$ si et seulement si $\alpha + \beta - 2\gamma = 0$. Nous en déduisons que R appartient à $A^{\perp\perp}$ si et seulement s'il est de la forme $R = \alpha\left(X^2 + \frac{1}{2}\right) + \beta\left(X + \frac{1}{2}\right)$, α et β étant deux réels quelconques. Nous avons donc

$$A^{\perp\perp} = \mathrm{Vect}\left(X^2 + \frac{1}{2}, X + \frac{1}{2}\right) = \mathrm{Vect}\Big(X^2 - X, X^2 + X + 1\Big)$$

Remarquons que l'inclusion $A \subset A^{\perp\perp}$ est stricte.

Définition 1.6.2

Soit E un $\mathbb{R}$-espace vectoriel. Soient q une forme quadratique sur E et φ_0 la forme polaire de q.
On appelle *noyau de la forme quadratique q* l'ensemble

$$N(q) = E^{\perp} = \{\mathbf{x} \in E \,|\, \forall\, \mathbf{y} \in E \quad \varphi_0(\mathbf{x}, \mathbf{y}) = 0\}.$$

Proposition 1.6.2

Soient E un $\mathbb{R}$-espace vectoriel et q une forme quadratique sur E.
La forme q est non dégénérée si et seulement si $N(q) = \{0\}$ si et seulement si $\det \mathrm{Mat}(q) \neq 0$.

Définitions 1.6.2

Soient E un $\mathbb{R}$-espace vectoriel et q une forme quadratique sur E.
On appelle *cône isotrope* de q l'ensemble

$$C(q) = \{\mathbf{x} \in E \,|\, q(\mathbf{x}) = 0\}.$$

Les éléments du cône isotrope sont appelés des *vecteurs isotropes*.

Remarques 1.6.2. — 1. Nous avons toujours $\mathbf{0} \in N(q)$ et $\mathbf{0} \in C(q)$.

2. ⚠ Il ne faut pas confondre $N(q)$ et $C(q)$. L'ensemble $N(q)$ est l'ensemble des vecteurs qui sont orthogonaux à tous les vecteurs de E alors que $C(q)$ est l'ensemble des vecteurs qui sont orthogonaux à eux-mêmes. Ainsi nous avons l'inclusion $N(q) \subset C(q)$ et cette inclusion est en général stricte.

Exemple 1.6.3. — Considérons la forme quadratique définie sur $\mathbb{R}^2$ par

$$q\colon \mathbb{R}^2 \to \mathbb{R}, \qquad \mathbf{x} \mapsto x_1^2 - x_2^2.$$

La forme polaire de q est l'application

$$\varphi_0\colon \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}, \qquad (\mathbf{x}, \mathbf{y}) \mapsto x_1y_1 - x_2y_2$$

et le noyau $N(q)$ de la forme quadratique correspond à l'ensemble des éléments $\mathbf{x} \in \mathbb{R}^2$ tels que $x_1y_1 - x_2y_2 = 0$ pour tout $\mathbf{y} \in \mathbb{R}^2$. En choisissant successivement $\mathbf{y} = (1, 0)$ et $\mathbf{y} = (0, 1)$ nous montrons que $N(q) = \{\mathbf{0}\}$.

Le cône isotrope de cette forme quadratique q n'est pas égal à $\{\mathbf{0}\}$ puisque l'élément $\mathbf{x} = (1, 1)$ appartient à $C(q)$ (il suffit de vérifier que $q(1, 1) = 0$). Plus précisément le cône isotrope de q est

$$C(q) = \{\mathbf{x} \in \mathbb{R}^2 \,|\, x_1 = x_2 \text{ ou } x_1 = -x_2\}.$$

Intéressons-nous au cas où E est un espace vectoriel de dimension finie n. Étant donnée une base $\mathcal{B} = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ de E, à toute forme quadratique q sur E est associée une matrice $M \in \mathrm{S}(n, \mathbb{R})$. Dire qu'un élément $\mathbf{x}$ de E, représenté par la matrice colonne $X \in \mathbb{R}^n$ dans la base $\mathcal{B}$, appartient au noyau de la forme quadratique q signifie que

$$\forall Y \in \mathbb{R}^n \quad {}^{\mathrm{t}}YMX = 0.$$

En choisissant pour Y le vecteur $Y = MX$ nous en déduisons que $\mathbf{x}$ appartient à $N(q)$ si et seulement si $MX = 0$. Ainsi, en dimension finie, le noyau d'une forme quadratique correspond exactement au noyau de l'endomorphisme ayant la même matrice que la forme quadratique dans une même base donnée. En particulier, le théorème du rang s'écrit dans le cas de l'algèbre bilinéaire sous la forme suivante : *la dimension de E est la somme du rang de q et de la dimension du noyau de q, i.e. en dimension finie*

$$\dim E = \mathrm{rg}(q) + \dim N(q).$$

1.7. Bases orthogonales

Commençons par quelques rappels. Étant donnés un espace vectoriel E et un ensemble I nous appelons *famille de vecteurs indexée par I* toute application

$$i \in I \mapsto \mathbf{v}_i \in E.$$

Nous la notons $(\mathbf{v}_i)_{i \in I}$. L'ensemble de départ I est appelé *ensemble des indices de la famille*. Lorsque I est un ensemble fini (resp. infini), on dit que $(\mathbf{v}_i)_{i \in I}$ est une *famille finie* (resp. *infinie*).

Définitions 1.7.1

Soient $\mathcal{F}$ une famille de vecteurs d'un $\mathbb{R}$-espace vectoriel E et q une forme quadratique sur E de forme polaire associée φ_0.

- ◇ La famille $\mathcal{F}$ est *q-orthogonale* si

$$\forall \mathbf{f} \in \mathcal{F}, \forall \mathbf{f}' \in \mathcal{F} \quad \mathbf{f} \neq \mathbf{f}' \;\Rightarrow\; \varphi_0(\mathbf{f}, \mathbf{f}') = 0.$$

- ◇ La famille $\mathcal{F}$ est *q-orthonormale* si la famille $\mathcal{F}$ est q-orthogonale et $\varphi_0(\mathbf{f}, \mathbf{f}) = 1$ pour tout $\mathbf{f} \in \mathcal{F}$.

Remarque 1.7.1. — Une famille à un seul élément non nul est considérée comme q-orthogonale pour toute forme quadratique q.

Exemples 1.7.1. — 1. Sur l'espace vectoriel $\mathcal{C}([0,1], \mathbb{R})$ des applications continues de $[0,1]$ dans $\mathbb{R}$, l'application $\varphi\colon \mathcal{C}([0,1], \mathbb{R}) \times \mathcal{C}([0,1], \mathbb{R}) \to \mathbb{R}$ définie par

$$\varphi(f, g) = \int_0^1 f(t)g(t)\,\mathrm{d}t$$

est une forme bilinéaire dont la fome quadratique associée est l'application

$$q\colon \mathcal{C}([0,1],\mathbb{R}) \to \mathbb{R}, \qquad\qquad f \mapsto \int_0^1 f(t)^2\,\mathrm{d}t.$$

Considérons la famille $(f_n)_{n\in\mathbb{N}^*}$ de $\mathcal{C}([0,1],\mathbb{R})$ indexée par l'ensemble des entiers naturels non nuls et définie par

$$\forall\, n \in \mathbb{N}^* \qquad f_n\colon t \in [0,1] \mapsto \cos(2\pi n t).$$

Pour tous n, $m \in \mathbb{N}^*$ nous avons

$$\begin{aligned}\varphi(f_n, f_m) &= \int_0^1 \cos(2\pi n t)\cos(2\pi m t)\mathrm{d}t\\ &= \frac{1}{2}\left(\int_0^1 \cos\left(2\pi(n+m)t\right)\mathrm{d}t + \int_0^1 \cos\left(2\pi(n-m)t\right)\mathrm{d}t\right)\end{aligned}$$

Puisque $\int_0^1 \cos(2\pi\lambda t)\mathrm{d}t = 0$ pour tout réel $\lambda \neq 0$, nous en déduisons que lorsque $n \neq m$ nous avons $\varphi(f_n, f_m) = 0$ ce qui montre que la famille $(f_n)_{n\in\mathbb{N}^*}$ est une famille orthogonale.

Cette famille n'est pas orthonormale puisque pour tout $n \in \mathbb{N}^*$ nous avons

$$\varphi(f_n, f_n) = \frac{1}{2}\left(\int_0^1 \cos\left(4\pi n t\right)\mathrm{d}t + \int_0^1 1\,\mathrm{d}t\right) = \frac{1}{2} \neq 1.$$

Par contre, en considérant la famille $(g_n)_{n\in\mathbb{N}^*}$ définie par $g_n = \sqrt{2}f_n$, alors on a d'une part, pour tous $n \in \mathbb{N}^*$, $m \in \mathbb{N}^*$ tels que $n \neq m$ la relation $\varphi(g_n, g_m) = 2\varphi(f_n, f_m) = 0$ et d'autre part pour tout $n \in \mathbb{N}^*$ la relation $\varphi(g_n, g_n) = 2\varphi(f_n, f_n) = 1$. Ainsi la famille $(g_n)_{n\in\mathbb{N}^*}$ est orthonormale.

2. Dans l'espace vectoriel $\mathbb{R}^4$ la forme de Lorentz est la forme bilinéaire définie par

$$\varphi\colon \mathbb{R}^4 \times \mathbb{R}^4 \to \mathbb{R}, \qquad\qquad (\mathbf{x}, \mathbf{y}) \mapsto x_1y_1 + x_2y_2 + x_3y_3 - c^2x_4y_4$$

où c désigne un paramètre réel. La base canonique de $\mathbb{R}^4$, c'est-à-dire la famille finie $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4\}$, est une famille orthogonale pour cette forme bilinéaire puisque pour tous $1 \leq i, j \leq 4$, $i \neq j$ nous avons $\varphi(\mathbf{e}_i, \mathbf{e}_j) = 0$. Par contre, la famille $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4\}$ n'est pas orthonormale pour la forme de Lorentz puisque nous avons $\varphi(\mathbf{e}_4, \mathbf{e}_4) = -c^2 \neq 1$.

Rappelons qu'une famille finie $(\mathbf{v}_i)_{1\leq i\leq p}$ d'un espace vectoriel est dite *liée* s'il existe des réels α_1, α_2, …, α_p dont un au moins est non nul, tels que

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \ldots + \alpha_p\mathbf{v}_p = 0.$$

Si une famille n'est pas liée, on dit qu'elle est *libre* (ou que les vecteurs de la famille sont linéairement indépendants).

Ces définitions de familles liées et libres se généralisent aux familles infinies de la manière suivante : une famille infinie est *libre* si toutes ses sous-familles finies sont libres et une famille infinie est *liée* si elle n'est pas libre.

Concernant les familles q-orthonormales, on a l'énoncé suivant :

Proposition 1.7.1

Soient E un $\mathbb{R}$-espace vectoriel et q une forme quadratique sur E. Toute famille de vecteurs de E qui est q-orthonormale est une famille libre.

Démonstration. — Soit $(\mathbf{v}_i)_{i\in I}$ une famille de vecteurs de E qui est q-orthogonale. Si cette famille est une famille infinie, pour montrer que c'est une famille libre, il faut montrer que toutes ses sous-familles finies sont libres. Ainsi, si on montre que tout famille finie de vecteurs de E qui est q-orthonormale est une famille libre alors l'énoncé sera démontré aussi bien pour les familles finies que pour les familles infinies.

Considérons une famille finie $\mathcal{F} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p\}$ q-orthonormale. Supposons qu'il existe des réels $\lambda_1, \lambda_2, \ldots, \lambda_p$ tels que $\lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 + \ldots + \lambda_p\mathbf{v}_p = 0$. Pour tout $1 \leq k \leq p$ nous avons

$$\varphi_0(\lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 + \ldots + \lambda_p\mathbf{v}_p, \mathbf{v}_k) = 0.$$

La bilinéarité de l'application φ_0 implique que

$$\lambda_1\varphi_0(\mathbf{v}_1, \mathbf{v}_k) + \lambda_2\varphi_0(\mathbf{v}_2, \mathbf{v}_k) + \ldots + \lambda_p\varphi_0(\mathbf{v}_p, \mathbf{v}_k) = 0.$$

Puisque la famille $\mathcal{F}$ est orthonormale tous les termes $\varphi_0(\mathbf{v}_i, \mathbf{v}_k)$ sont nuls excepté celui correspondant à $i = k$ qui vaut 1. Nous en déduisons que $\lambda_k = 0$ pour $1 \leq k \leq p$. Cela montre que tous les λ_k sont nuls : la famille $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p\}$ est libre. □

Une famille peut être q-orthogonale sans pour autant être libre. Par exemple la famille $\{(1, 1), (2, 2)\}$ de $\mathbb{R}^2$ est q-orthogonale pour la forme quadratique $q(\mathbf{x}) = x_1^2 - x_2^2$ mais cette famille n'est manifestement pas libre puisque l'un des vecteurs est le double de l'autre.

Cas de la dimension finie. — Dans le cas où on considère un $\mathbb{R}$-espace vectoriel de dimension finie, certaines familles de vecteurs jouent un rôle particulier : ce sont les bases.

Définition 1.7.1

Soient E un $\mathbb{R}$-espace vectoriel de dimension finie et q une forme quadratique sur E. Une base $\mathcal{B}$ de E est *q-orthogonale* (resp. *q-orthonormale*) si la famille $\mathcal{B}$ est q-orthogonale (resp. q-orthonormale).

Si on dispose d'une base q-orthonormale $\mathcal{B} = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ de E est alors on peut déterminer la matrice de la forme quadratique q dans cette base : par définition cette matrice est celle de la forme polaire φ_0 associée à la forme quadratique q. Les coefficients de cette matrice sont donc les réels $\varphi_0(\mathbf{e}_i, \mathbf{e}_j)$ sont nuls sauf ceux correspondant à $i = j$ qui valent 1. La matrice de la forme quadratique est donc la matrice identité. Nous avons ainsi démontré l'énoncé suivant :

Proposition 1.7.2

Soient E un $\mathbb{R}$-espace vectoriel de dimension finie et q une forme quadratique sur E. La matrice de q dans une base q-orthonormale est la matrice identité.

Remarque 1.7.2. — La matrice de q dans une base q-orthogonale est une matrice diagonale.

En dimension finie si $\mathcal{B} = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ est une base de E alors d'après la relation (1.2.1) nous avons

$$\forall \mathbf{x} \in E \qquad q(\mathbf{x}) = \varphi_0(\mathbf{x}, \mathbf{x}) = \sum_{i=1}^{n}\sum_{j=1}^{n} x_i x_j \varphi_0(\mathbf{e}_i, \mathbf{e}_j)$$

où $x_1, x_2, \ldots, x_n$ désignent les composantes de $\mathbf{x}$ dans la base $\mathcal{B}$. Si la base considérée est q-orthogonale, alors tous les coefficients $\varphi_0(\mathbf{e}_i, \mathbf{e}_j)$ sont nuls sauf lorsque $i = j$ où $\varphi_0(\mathbf{e}_i, \mathbf{e}_j) = q(\mathbf{e}_i)$. Nous avons donc

$$\forall \mathbf{x} \in E \qquad q(x) = \sum_{i=1}^{n} q_i^2 q(\mathbf{e}_i).$$

Lorsque la base considérée est q-orthonormale, les coefficients $q(\mathbf{e}_i)$ sont égaux à 1 et nous avons

$$\forall \mathbf{x} \in E \qquad q(\mathbf{x}) = \sum_{i=1}^{n} q_i^2.$$

On voit donc l'intérêt de disposer d'une base q-orthonormale (ou q-orthogonale) : la donnée des images d'une telle base par q permet de connaître l'image de n'importe quel élément de E.

Étant donnée une forme quadratique on peut se demander s'il existe toujours une base orthogonale pour celle-ci. L'énoncé suivant apporte la réponse :

Théorème 1.7.1

Soient E un $\mathbb{R}$-espace vectoriel de dimension finie et q une forme quadratique sur E. L'espace vectoriel E admet une base q-orthogonale.

Démonstration. — Nous allons démontrer l'énoncé par récurrence sur la dimension de E. Pour tout $n \in \mathbb{N}^*$ nous allons montrer la propriété $\mathcal{P}_n$ suivante :

> Pour tout $\mathbb{R}$-espace vectoriel E de dimension n et toute forme quadratique q sur E il existe une base q-orthogonale de E.

Si $n = 1$, alors une base d'un $\mathbb{R}$-espace vectoriel de dimension $n = 1$ est formé par n'importe quel élément non nul de E. Pour toute forme quadratique, une telle base est q-orthogonale : $\mathcal{P}_1$ est donc vraie.

Supposons que la propriété $\mathcal{P}_n$ soit vraie pour un entier $n \in \mathbb{N}^*$; montrons que cela entraîne que la propriété $\mathcal{P}_{n+1}$ est vraie. Soient E un $\mathbb{R}$-espace vectoriel de dimension $n+1$ et q une forme quadratique sur E. Si q est l'application nulle, alors toute base de E est q-orthogonale. Intéressons-nous au cas où q n'est pas l'application nulle : il existe $\mathbf{v} \in E$ tel que $q(\mathbf{v}) \neq 0$. La forme polaire φ_0 associée à q étant bilinéaire, l'application

$$E \to \mathbb{R}, \qquad \mathbf{x} \mapsto \varphi_0(\mathbf{v}, \mathbf{x})$$

est une forme linéaire. Ce n'est pas l'application nulle car $f(\mathbf{v}) = \varphi_0(\mathbf{v}, \mathbf{v}) = q(\mathbf{v})$ et $q(\mathbf{v}) \neq 0$. Le noyau H de cette forme linéaire est un sous-espace vectoriel de E de dimension n. D'après l'hypothèse de récurrence $\mathcal{P}_n$ il existe une base $(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ de H orthogonale pour la forme quadratique $q_{|H}$. La famille $(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n, \mathbf{v})$ est une base q-orthogonale de E : la propriété $\mathcal{P}_{n+1}$ est vraie. $\square$

Jusqu'à présent nous avons utilisé les propriétés matricielles (produit d'une matrice par un vecteur, changement de bases, etc) pour obtenir des résultats sur les formes linéaires ou quadratiques. Dans l'énconcé suivant c'est le contraire : la formulation du Théorème 1.7.1 sous forme matricielle implique le résultat suivant :

Proposition 1.7.3

Toute matrice symétrique réelle est congrue à une matrice diagonale. En d'autres termes, pour toute matrice $A \in \mathrm{S}(n, \mathbb{R})$ il existe une matrice P telle que ${}^{\mathrm{t}}PAP$ soit diagonale. De plus, la matrice P est inversible.

Démonstration. — À toute matrice $A \in \mathrm{S}(n, \mathbb{R})$ on peut associer la forme quadratique q sur $\mathbb{R}^n$ définie par

$$\forall \mathbf{x} \in \mathbb{R}^n \quad q(\mathbf{x}) = {}^{\mathrm{t}}XAX,$$

le vecteur X correspondant à la matrice colonne dont les éléments sont les composantes du vecteur $\mathbf{x}$ dans la base canonique de $\mathbb{R}^n$. Le Théorème 1.7.1 assure l'existence d'une base $\mathcal{B}$ de $\mathbb{R}^n$ qui est q-orthogonale. La matrice D représentant la forme quadratique q dans cette base q-orthogonale est diagonale. Si P désigne la matrice de passage de la base canonique de $\mathbb{R}^n$ à la base $\mathcal{B}$, alors P est une matrice inversible et nous avons $D = {}^{\mathrm{t}}PAP$. La matrice ${}^{\mathrm{t}}PAP$ est donc diagonale. $\square$

1.8. Décomposition en somme de carrés de formes linéaires

Quelques notions sur l'ensemble des formes linéaires. — L'ensemble $\mathcal{L}(E,\mathbb{R})$ des formes linéaires sur un espace vectoriel E possède lui-même une structure de $\mathbb{R}$-espace vectoriel. Cette espace vectoriel $\mathcal{L}(E,\mathbb{R})$ est généralement noté E^* et est appelé le *dual* algébrique de E.

On s'intéresse au cas où l'espace vectoriel E est un espace vectoriel de dimension finie n. Le dual E^* est un espace vectoriel de dimension finie et nous avons $\dim E = \dim E^*$. Considérons une base $(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n)$ de E et définissons n formes linéaires $\mathbf{e}_j^*$, $1 \leq j \leq n$ par

$$\forall\, 1 \leq j \leq n \qquad \mathbf{e}_j^* \colon E \to \mathbb{R},\ \mathbf{x} \mapsto x_j$$

où les réels x_1, x_2, ..., x_n désignent les composantes de $\mathbf{x}$ dans la base $(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n)$. Les formes linéaires $\mathbf{e}_j^*$ sont caractérisées par les valeurs qu'elles prennent sur une base de E. Nous avons

$$\forall\, 1 \leq i,\, j \leq n \qquad \mathbf{e}_j^*(\mathbf{e}_i) = \delta_{ij}.$$

On peut vérifier que les n formes linéaires $\mathbf{e}_1^*$, $\mathbf{e}_2^*$, ..., $\mathbf{e}_n^*$ sont linéairement indépendantes. Puisque l'espace vectoriel E^* des formes linéaires sur E est un espace vectoriel de dimension n, nous en déduisons que la famille $(\mathbf{e}_1^*, \mathbf{e}_2^*, \dots, \mathbf{e}_n^*)$ est une base de E^*.

Proposition-Définition 1.8.1

Soient E un espace vectoriel de dimension finie et $\mathcal{B} = (\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n)$ une base de E. La famille de formes linéaires $(\mathbf{e}_1^*, \mathbf{e}_2^*, \dots, \mathbf{e}_n^*)$ définie par

$$\forall\, 1 \leq i,\, j \leq n \qquad \mathbf{e}_j^*(\mathbf{e}_i) = \delta_{ij}.$$

est une base de E^*. Cette base est appelée la *base duale* de $\mathcal{B}$ et est notée $\mathcal{B}^*$.

Le principe des méthodes dite « par dualité », est d'obtenir des informations sur un vecteur $\mathbf{x} \in E$ à partir des images de ce vecteur par des formes linéaires sur E. L'énoncé suivant en est une illustration.

Proposition 1.8.1

Soit $\mathbf{x}$ un élément d'un espace vectoriel E de dimension finie. Nous avons l'équivalence

$$(\forall f \in E^*\ f(\mathbf{x}) = 0) \quad \Longleftrightarrow \quad \mathbf{x} = \mathbf{0}.$$

Démonstration. — Considérons une base $(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n)$ de E et désignons par x_1, x_2, ..., x_n les composantes de $\mathbf{x}$ dans cette base. Par définition des formes linéaires $\mathbf{e}_i^*$, nous avons $\mathbf{e}_i^*(\mathbf{x}) = x_i$ pour tout $1 \leq i \leq n$. Si pour toute forme linéaire $f \in E^*$ nous avons $f(\mathbf{x}) = 0$ alors en considérant le cas particulier des formes linéaires $\mathbf{e}_i^*$ nous obtenons $\mathbf{e}_i^*(\mathbf{x}) = 0$ pour tout $1 \leq i \leq n$. Comme $\mathbf{e}_i^*(\mathbf{x}) = x_i$, toutes les composantes de $\mathbf{x}$ sont nulles et par conséquent $\mathbf{x} = \mathbf{0}$.

La réciproque découle du fait que l'image du vecteur nul par une forme linéaire est le réel 0. □

Nous avons construit, pour toute base de E, une base E^* appelée base duale. L'énoncé suivant montre qu'inversement à partir d'une base $\widetilde{\mathcal{B}}$ de E^* il est possible de trouver une base de E dont $\widetilde{\mathcal{B}}$ est la base duale.

Proposition 1.8.2

Soient E un espace vectoriel de dimension finie et $\widetilde{\mathcal{B}} = (f_1, f_2, \ldots, f_n)$ une base de E^*. Il existe une base $\mathcal{B} = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ de E telle que $\widetilde{\mathcal{B}}$ soit la base duale de $\mathcal{B}$, c'est-à-dire telle que

$$\forall\, 1 \leq i \leq n \qquad f_i = \mathbf{e}_i^*.$$

Démonstration. — Puisque $\widetilde{\mathcal{B}}$ est une base de l'espace vectoriel E^* nous pouvons construire en utilisant la Proposition 1.8.1 la base duale $(f_1^*, f_2^*, \ldots, f_n^*)$ de $\widetilde{\mathcal{B}}$ qui est une base de l'espace vectoriel $(E^*)^*$. Montrons que l'application

$$\Phi\colon E \to (E^*)^*, \qquad\qquad \mathbf{x} \mapsto \varphi_{\mathbf{x}}\colon f \in E^* \mapsto f(x) \in \mathbb{R}$$

est un isomorphisme, *i.e.* une application linéaire bijective. Autrement dit montrons que les deux espaces vectoriels E et $(E^*)^*$ sont isomorphes via Φ.

Pour montrer que l'application Φ est un isomorphisme commençons par remarquer que l'application Φ est linéaire : pour tous λ, $\mu \in \mathbb{R}$, pour tous $\mathbf{x}$, $\mathbf{y} \in E$, $\Phi(\lambda\mathbf{x}+\mu\mathbf{y})$ est l'application qui à une forme linéaire $f \in E^*$ associe le réel $f(\lambda\mathbf{x}+\mu\mathbf{y})$. Ainsi pour toute application linéaire f nous avons

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

i.e. $\Phi(\lambda\mathbf{x} + \mu\mathbf{y}) = \lambda\Phi(\mathbf{x}) + \mu\Phi(\mathbf{y})$.

Montrons ensuite que l'application Φ est injective : si $\Phi(\mathbf{x}) = 0$, alors pour toute application linéaire $f \in E^*$ nous avons $f(\mathbf{x}) = 0$. La Proposition 1.8.1 implique que $\mathbf{x} = \mathbf{0}$: l'application Φ est donc bien injective.

Puisque $\dim E = \dim(E^*)^*$ nous obtenons que Φ est une application linéaire bijective, et par conséquent un isomorphisme de E sur $(E^*)^*$.

La famille $(f_1^*, f_2^*, \ldots, f_n^*)$ étant une base de l'espace vectoriel $(E^*)^*$, la famille $\big(\Phi^{-1}(f_1^*),$ $\Phi^{-1}(f_2^*), \ldots, \Phi^{-1}(f_n^*)\big)$ est une base de l'espace vectoriel E. Il ne reste plus qu'à démontrer que nous avons

$$\forall\, 1 \leq i \leq n \qquad \big(\Phi^{-1}(f_i^*)\big)^* = f_i.$$

Soit $1 \leq i \leq n$ un entier fixé. Pour montrer que les deux formes linéaires $\left(\Phi^{-1}(f_i^*)\right)^*$ et f_i sont égales, il suffit de montrer qu'elles agissent de la même façon sur une base donnée de E. Considérant la base $(\Phi^{-1}(f_1^*), \Phi^{-1}(f_2^*), \ldots, \Phi^{-1}(f_n^*))$ de E il suffit donc de montrer que

$$\forall 1 \leq j \leq n \qquad \left(\Phi^{-1}(f_i^*)\right)^* (\Phi^{-1}(f_j^*)) = f_i(\Phi^{-1}(f_j^*)) \tag{1.8.1}$$

Tout d'abord, par définition d'une base duale, nous avons

$$\forall 1 \leq j \leq n \quad \left(\Phi^{-1}(f_i^*)\right)^* (\Phi^{-1}(f_j^*)) = \delta_{ij}. \tag{1.8.2}$$

D'autre part, par définition de l'application Φ, nous avons $\Phi(\mathbf{x})(f) = f(\mathbf{x})$ pour tous $\mathbf{x} \in E$ et $f \in E^*$. En particulier pour $f = f_i$ nous avons $\Phi(\Phi^{-1}(f_j^*))(f_i) = f_i(\Phi^{-1}(f_j^*))$. Nous obtenons donc

$$\forall 1 \leq j \leq n \quad f_i(\Phi^{-1}(f_j^*)) = f_j^*(f_i) = \delta_{ij} \tag{1.8.3}$$

Les égalités (1.8.2) et (1.8.3) impliquent (1.8.1). □

Proposition 1.8.3

Soient E un espace vectoriel de dimension finie $n \in \mathbb{N}^*$ et $(f_1, f_2, \ldots, f_r)$ une famille de r formes linéaires indépendantes. Nous avons

$$\dim \left(\bigcap_{i=1}^{r} \ker f_i \right) = n - r.$$

Démonstration. — Le théorème de la base incomplète assure que la famille libre $(f_1, f_2, \ldots, f_r)$ de l'espace vectoriel E^* peut être complétée en une base $\widetilde{\mathcal{B}} = (f_1, f_2, \ldots, f_n)$ de E^*. D'après la Proposition 1.8.2 il existe une base $\mathcal{B} = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ de E dont la base duale est $\widetilde{\mathcal{B}}$.

Soit $\mathbf{x}$ un élément de E dont la décomposition dans la base $\mathcal{B}$ est $\mathbf{x} = \sum_{j=1}^{n} x_j \mathbf{e}_j$ avec $(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$. Nous avons les équivalences suivantes :

$$\begin{aligned}
\mathbf{x} \in \bigcap_{i=1}^{r} \ker f_i &\iff \forall 1 \leq i \leq r \quad f_i(\mathbf{x}) = 0 \\
&\iff \forall 1 \leq i \leq r \quad f_i \left(\sum_{j=1}^{n} x_j \mathbf{e}_j \right) = 0 \\
&\iff \forall 1 \leq i \leq r \quad \sum_{j=1}^{n} x_j f_i(\mathbf{e}_j) = 0
\end{aligned}$$

Puisque $\widetilde{\mathcal{B}}$ est la base duale de la base $\mathcal{B}$, la Proposition 1.8.1 assure que pour tous $1 \leq i, j \leq n$ nous avons $f_i(\mathbf{e}_j) = \delta_{ji}$ de sorte que

$$\begin{aligned}\mathbf{x} \in \bigcap_{i=1}^{r} \ker f_i &\iff \forall\, 1 \leq i \leq r \quad x_i = 0 \\ &\iff \mathbf{x} \in \mathrm{Vect}(\mathbf{e}_{r+1}, \ldots, \mathbf{e}_n).\end{aligned}$$

Autrement dit nous avons montré que $\bigcap_{i=1}^{r} \ker f_i = \mathrm{Vect}(\mathbf{e}_{r+1}, \ldots, \mathbf{e}_n)$. Les $n-r$ vecteurs $\mathbf{e}_{r+1}$, $\mathbf{e}_{r+2}$, …, $\mathbf{e}_n$ étant des vecteurs linéairement indépendants, l'espace vectoriel $\mathrm{Vect}(\mathbf{e}_{r+1}, \ldots, \mathbf{e}_n)$ est de dimension $n - r$. □

Remarques 1.8.1. — 1. Un sous-espace vectoriel de dimension $n-1$ d'un espace vectoriel E de dimension n est appelé un *hyperplan* de E. Par exemple les hyperplans de $\mathbb{R}^3$ sont les plans vectoriels alors que les hyperplans de $\mathbb{R}^2$ correspondent aux droites vectorielles. La Proposition 1.8.3 assure que le noyau d'une forme linéaire non nulle est un hyperplan (prendre $r = 1$).

2. Il existe un résultat plus général que celui énoncé dans la Proposition 1.8.3 qui est le suivant :

> *tout sous-espace vectoriel de dimension $n - r$ d'un espace vectoriel de dimension n est l'intersection de r noyaux de formes linéaires indépendantes.*

Par exemple tout hyperplan est le noyau d'une forme linéaire non nulle.

Exemples 1.8.1. — 1. L'application

$$\mathbb{R}^2 \to \mathbb{R} \qquad (x_1, x_2) \mapsto a_1x_1 + a_2x_2$$

où $(a_1, a_2) \in \mathbb{R}^2$ est une forme linéaire sur $\mathbb{R}^2$. Elle est non nulle si $(a_1, a_2) \neq (0, 0)$. Dans ce cas son noyau, qui est l'ensemble

$$\{(x_1, x_2) \in \mathbb{R}^2 \mid a_1x_1 + a_2x_2 = 0\}$$

est un espace vectoriel de dimension 1.

2. Étant donnés deux vecteurs $\mathbf{a} = (a_1, a_2, a_3)$ et $\mathbf{b} = (b_1, b_2, b_3)$ dans $\mathbb{R}^3$, les deux applications

$$\mathbb{R}^3 \to \mathbb{R} \qquad (x_1, x_2, x_3) \mapsto a_1x_1 + a_2x_2 + a_3x_3$$

et

$$\mathbb{R}^3 \to \mathbb{R} \qquad (x_1, x_2, x_3) \mapsto b_1x_1 + b_2x_2 + b_3x_3$$

sont deux formes linéaires sur $\mathbb{R}^3$. Elles sont indépendantes si les deux vecteurs $\mathbf{a}$ et $\mathbf{b}$ sont indépendants. D'après la Proposition 1.8.3 l'ensemble

$$\left\{(x_1, x_2, x_3) \in \mathbb{R}^3 \mid \left\{\begin{array}{l} a_1x_1 + a_2x_2 + a_3x_3 = 0 \\ b_1x_1 + b_2x_2 + b_3x_3 = 0 \end{array}\right.\right\}$$

est un espace vectoriel de dimension 1.

La méthode de Gauss et le théorème de Gauss. — Le premier exemple de forme quadratique en dimension finie que nous avons vu est l'application

$$\mathbb{R}^3 \to \mathbb{R}, \qquad \mathbf{x} \mapsto x_1^2 + x_2^2 + x_3^2.$$

Sur $\mathbb{R}^3$ nous avons aussi introduit la forme quadratique de Lorentz définie par

$$\mathbb{R}^4 \to \mathbb{R}, \qquad \mathbf{x} \mapsto x_1^2 + x_2^2 + x_3^2 - c^2 x_4^2.$$

Ces deux formes quadratiques sont des sommes de formes quadratiques « simples » du type

$$\mathbb{R}^n \to \mathbb{R}, \qquad \mathbf{x} \mapsto x_i^2.$$

La forme quadratique sur $\mathbb{R}^2$ définie par

$$q\colon \mathbb{R}^2 \to \mathbb{R}, \qquad \mathbf{x} \mapsto 4x_1x_2$$

n'a pas, semble-t-il, cette structure. Néanmoins, pour tout $(x_1, x_2) \in \mathbb{R}^2$ en écrivant $4x_1x_2 = (x_1 + x_2)^2 - (x_1 - x_2)^2$ nous obtenons que la forme quadratique q peut aussi s'écrire comme somme de « carrés ». Ce résultat se généralise à toutes les formes quadratiques sur un espace vectoriel de dimension finie et la méthode pour déterminer la décomposition est appelée la méthode de Gauss. Nous allons maintenant décrire cette méthode.

Nous avons établi que toute forme quadratique sur un $\mathbb{R}$-espace vectoriel E de dimension n peut s'écrire sous la forme

$$q(\mathbf{x}) = \sum_{i=1}^{n} a_{ii}x_i^2 + \sum_{i=1}^{n}\sum_{j=i+1}^{n} a_{ij}x_ix_j.$$

chaque a_{ij} désignant un nombre réel et les x_i désignant les composantes de $\mathbf{x}$ dans une base fixée de E.

Afin de faire apparaître uniquement des « carrés », nous allons utiliser une méthode connue sous le nom de *méthode de Gauss*. Selon les coefficients a_{ij} deux cas sont à envisager.

◇ Premier cas. Il existe au moins un entier i tel que $a_{ii} \neq 0$; supposons qu'il s'agisse de a_{11}. En notant $\mathbf{x}' = (x_2, x_3, \dots, x_n)$ nous écrivons dans un premier temps q sous la forme

$$q(\mathbf{x}) = a_{11}x_1^2 + x_1B(\mathbf{x}') + C(\mathbf{x}')$$

où B est une forme linéaire sur un espace vectoriel de dimension $n-1$ et C une forme quadratique sur un espace vectoriel de dimension $n-1$. La forme quadratique q peut ensuite s'écrire comme la somme du carré d'une forme linéaire et d'une forme quadratique en $\mathbf{x}'$ de la manière suivante :

$$q(\mathbf{x}) = a_{11}\left(x_1 + \frac{B(\mathbf{x}')}{2a_{11}}\right)^2 + \left(C(\mathbf{x}') - \frac{B(\mathbf{x}')^2}{4a_{11}^2}\right)$$

◇ Deuxième cas. Tous les termes a_{ii} sont nuls et il existe un coefficient a_{ij}, $i \neq j$, non nul. Supposons qu'il s'agisse de a_{12}. En notant $\mathbf{x}'' = (x_3, x_4, \dots, x_n)$ on écrit dans un premier temps q sous la forme

$$q(\mathbf{x}) = a_{12}x_1x_2 + x_1B(\mathbf{x}'') + x_2C(\mathbf{x}'') + D(\mathbf{x}'')$$

où B et C sont des formes linéaires et D une forme quadratique. Nous écrivons ensuite q comme la somme de deux carrés de formes linéaires et d'une forme quadratique en $\mathbf{x}''$ de la façon suivante :

$$\begin{aligned} q(\mathbf{x}) &= a_{12}\left(x_1 + \frac{C(\mathbf{x}'')}{a_{12}}\right)\left(x_2 + \frac{B(\mathbf{x}'')}{a_{12}}\right) + \left(D(\mathbf{x}'') - \frac{B(\mathbf{x}'')C(\mathbf{x}'')}{a_{12}}\right) \\ &= \frac{a_{12}}{4}\left(\left(x_1 + x_2 + \frac{B(\mathbf{x}'') + C(\mathbf{x}'')}{a_{12}}\right)^2 - \left(x_1 - x_2 + \frac{C(\mathbf{x}'') - B(\mathbf{x}'')}{a_{12}}\right)^2\right) + \left(D(\mathbf{x}'') - \frac{B(\mathbf{x}'')C(\mathbf{x}'')}{a_{12}}\right) \end{aligned}$$

On peut alors itérer ce processus pour décomposer une forme quadratique comme la somme de carrés de formes linéaires sur l'espace vectoriel E et d'une forme quadratique sur un espace vectoriel de dimension $m < n$. Au terme du processus nous avons démontré l'énoncé suivant.

Théorème 1.8.1: Théorème de Gauss

Soit q une forme quadratique sur un $\mathbb{R}$-espace vectoriel E de dimension n. Les assertions suivantes sont équivalentes :

- ◇ le rang de q est égal à r ;
- ◇ il existe r formes linéaires indépendantes f_1, f_2, …, f_r sur E et il existe r réels non nuls γ_1, γ_2, …, γ_r tels que
$$q = \sum_{i=1}^{r} \gamma_i f_i^2.$$

Démonstration. — La construction qui a été effectuée précédemment indique que la première assertion implique la seconde.

Montrons la réciproque. Supposons qu'il existe r formes linéaires indépendantes f_1, f_2, …, f_r et qu'il existe r réels non nuls γ_1, γ_2, …, γ_r tels que $q = \sum_{i=1}^{r} \gamma_i f_i^2$. D'après le théorème de la base incomplète la famille libre $(f_1, f_2, \dots, f_r)$ de l'espace vectoriel E^* peut être complétée en une base $\widetilde{\mathcal{B}} = (f_1, f_2, \dots, f_n)$ de E^*. La Proposition 1.8.2 assure l'existence d'une base $\mathcal{B} = (\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n)$ de E dont la base duale est $\widetilde{\mathcal{B}}$. Déterminons la matrice de q dans cette base en évaluant $\varphi_0(\mathbf{e}_i, \mathbf{e}_j)$ pour tous $1 \leq i, j \leq n$ où φ_0 est la forme polaire de q définie par

$$\forall\, \mathbf{x}, \mathbf{y} \in E \quad \varphi_0(\mathbf{x}, \mathbf{y}) = \sum_{k=1}^{r} \gamma_k f_k(\mathbf{x}) f_k(\mathbf{y}).$$

Par définition d'une base duale nous avons $f_i(\mathbf{e}_j) = \delta_{ij}$ pour tous $1 \leq i,\, j \leq n$ et par conséquent

$$\forall\, 1 \leq i,\, j \leq n \qquad \varphi_0(\mathbf{e}_i, \mathbf{e}_j) = \sum_{k=1}^{r} \gamma_k \delta_{ik} \delta_{jk}.$$

Nous avons $\delta_{ik} \neq 0$ si et seulement si $k = i$ et $\delta_{jk} \neq 0$ si et seulement si $k = j$ donc $\varphi_0(\mathbf{e}_i, \mathbf{e}_j) \neq 0$ si et seulement si $i = j$. La matrice de q dans la base $\mathcal{B}$ est donc diagonale. De plus pour tout $1 \leq i \leq r$ nous avons $\varphi_0(\mathbf{e}_i, \mathbf{e}_i) = \gamma_i$ et pour tout $r + 1 \leq i \leq n$ nous avons $\varphi_0(\mathbf{e}_i, \mathbf{e}_i) = 0$. La matrice est donc composée sur la diagonale des réels non nuls γ_1, γ_2, …, γ_r complétée par des 0. Le rang d'une telle matrice est r. □

Exemples 1.8.2. — 1. Décomposons en utilisant la méthode de Gauss la forme quadratique q définie sur $\mathbb{R}^3$ par

$$q(\mathbf{x}) = x_1^2 + 2x_2^2 + 8x_3^2 - 2x_1x_2 + 4x_1x_3.$$

Nous avons

$$\begin{aligned} q(\mathbf{x}) &= x_1^2 + x_1(-2x_2 + 4x_3) + 2x_2^2 + 8x_3^2 \\ &= \left(x_1^2 + 2x_1(-x_2 + 2x_3) + (-x_2 + 2x_3)^2\right) - (-x_2 + 2x_3)^2 + 2x_2^2 + 8x_3^2 \\ &= (x_1 - x_2 + 2x_3)^2 + x_2^2 + 4x_2x_3 + 4x_3^2 \\ &= (x_1 + 2x_3 - x_2)^2 + (x_2 + 2x_3)^2. \end{aligned}$$

Ainsi

$$q(\mathbf{x}) = f_1(\mathbf{x})^2 + f_2(\mathbf{x})^2 \qquad \text{avec} \qquad \begin{cases} f_1(\mathbf{x}) = x_1 + 2x_3 - x_2 \\ f_2(\mathbf{x}) = x_2 + 2x_3 \end{cases}$$

2. Décomposons en utilisant la méthode de Gauss la forme quadratique q définie sur $\mathbb{R}^4$ par

$$q(\mathbf{x}) = x_1x_2 + x_1x_3 + x_1x_4 + x_2x_3 + x_2x_4 + x_3x_4.$$

Aucun carré n'apparaissant nous regroupons tous les termes faisant apparaître x_1 et x_2 de la façon suivante :

$$\begin{aligned} q(\mathbf{x}) &= x_1x_2 + x_1(x_3 + x_4) + x_2(x_3 + x_4) + x_3x_4 \\ &= (x_1 + x_3 + x_4)(x_2 + x_3 + x_4) - (x_3 + x_4)^2 + x_3x_4 \end{aligned}$$

Or d'une part

$$\begin{aligned} q(\mathbf{x}) &= (x_1 + x_3 + x_4)(x_2 + x_3 + x_4) - (x_3 + x_4)^2 + x_3x_4 \\ &= \frac{1}{4}\Big((x_1 + x_3 + x_4 + x_2 + x_3 + x_4)^2 - (x_1 + x_3 + x_4 - x_2 - x_3 - x_4)^2\Big) \\ &= \frac{1}{4}\Big((x_1 + x_2 + 2x_3 + 2x_4)^2 - (x_1 - x_2)^2\Big) \end{aligned}$$

et d'autre part

$$-(x_3+x_4)^2+x_3x_4=-x_3^2-x_4^2-x_3x_4$$

donc

$$q(\mathbf{x})=\frac{1}{4}\Big(x_1+x_2+2(x_3+x_4)\Big)^2-\frac{1}{4}(x_1-x_2)^2-x_3^2-x_4^2-x_3x_4.$$

Il reste ensuite à s'occuper des termes ne faisant pas apparaître x_1 et x_2, *i.e.* de la forme quadratique sur $\mathbb{R}^2$ définie par

$$\widetilde{q}(x_3,x_4)=x_3^2+x_3x_4+x_4^2.$$

Nous regroupons dans un même terme tous les termes où est présent x_3 :

$$\widetilde{q}(x_3,x_4)=\left(x_3+\frac{1}{2}x_4\right)^2-\frac{1}{4}x_4^2+x_4^2=\left(x_3+\frac{1}{2}x_4\right)^2+\frac{3}{4}x_4^2.$$

Nous obtenons finalement

$$q(\mathbf{x})=f_1(\mathbf{x})^2-f_2(\mathbf{x})^2-f_3(\mathbf{x})^2-3f_4(\mathbf{x})^2$$

où les formes linéaires f_i sont données par $f_1(\mathbf{x})=\frac{1}{2}(x_1+x_2+2x_3+2x_4)$, $f_2(\mathbf{x})=\frac{1}{2}(x_1-x_2)$, $f_3(\mathbf{x})=x_3+\frac{1}{2}x_4$ et $f_4(\mathbf{x})=\frac{1}{2}x_4$.

Remarque 1.8.2. — Il existe bien sûr d'autres décompositions d'une forme quadratique comme combinaison linéaire de carrés de formes quadratiques que celle donnée par la méthode de Gauss. Par exemple la forme quadratique q définie sur $\mathbb{R}^2$ par

$$q(\mathbf{x})=(x_1+x_2)^2-x_1^2+(x_1-x_2)^2$$

est une combinaison linéaire de carrés de formes linéaires. On a $q(\mathbf{x})=g_1(\mathbf{x})^2-g_2(\mathbf{x})^2+g_3(\mathbf{x})^2$ avec

$$g_1(\mathbf{x})=x_1+x_2,\qquad g_2(\mathbf{x})=x_1,\qquad g_3(\mathbf{x})=x_1-x_2.$$

Par contre les trois formes linéaires g_1, g_2 et g_3 ne sont pas linéairement indépendantes car $g_1+g_3=2g_2$.

Les combinaisons linéaires obtenues dans l'Exemple 1.8.2 font intervenir des formes linéaires indépendantes.

Un des avantages de la méthode de Gauss est qu'elle assure l'indépendance des formes linéaires obtenues.

Théorème 1.8.2

Soit E un $\mathbb{R}$-espace vectoriel de dimension finie. Soient $\ell_1, \ell_2, \ldots, \ell_r\colon E \to \mathbb{R}$ des formes linéaires linéairement indépendantes et $a_1, a_2, \ldots, a_r$ des réels non nuls tels que

$$q(\mathbf{x}) = a_1\ell_1(\mathbf{x})^2 + a_2\ell_2(\mathbf{x})^2 + \ldots + a_r\ell_r(\mathbf{x})^2$$

quel que soit $u \in E$. Alors le noyau de q est le sous-espace vectoriel des vecteurs $u \in E$ tel que

$$\ell_1(u) = \ell_2(u) = \ldots = \ell_r(u) = 0.$$

Démonstration. — Notons F le sous-espace vectoriel défini par

$$F = \{u \in E \,|\, \ell_1(u) = \ell_2(u) = \ldots = \ell_r(u) = 0\}.$$

Puisque les formes $\ell_1, \ell_2, \ldots, \ell_r$ sont linéairement indépendantes, F est de dimension $n-r$. Soit $\varphi\colon E \times E \to \mathbb{R}$ l'application définie pour tous $u, v \in E$ par

$$\varphi(u,v) = a_1\ell_1(u)\ell_1(v) + a_2\ell_2(u)\ell_2(v) + \ldots + a_r\ell_r(u)\ell_r(v). \tag{1.8.4}$$

Puisque φ est combinaison linéaire de produits de deux formes linéaires, φ est une forme linéaire. De plus, φ est symétrique et $q(u) = \varphi(u,u)$ pour tout $u \in E$. Par suite φ est la forme polaire de q. D'après (1.8.4) nous avons $F \subset \ker q$.

Le théorème de la base incomplète assure que nous pouvons compléter les r formes linéaires linéairement indépendantes $\ell_1, \ell_2, \ldots, \ell_r$ en une base $(\ell_1, \ell_2, \ldots, \ell_n)$ de l'espace dual E^*. Il existe une base $(u_1, u_2, \ldots, u_n)$ de E telle que $u_i^* = \ell_i$ pour tout $1 \leq i \leq n$ (Proposition 1.8.2). Par suite $\ell_i(u_j) = \delta_{ij}$. D'après (1.8.4) il vient $\varphi(u_i, u_i) = a_i$ et $\varphi(u_i, u_j) = 0$ si $i \neq j$. La matrice de q dans la base $(u_1, u_2, \ldots, u_n)$ est donc la matrice diagonale dont les r premiers coefficients diagonaux sont $a_1, a_2, \ldots, a_r$ et les suivants sont nuls. Comme $\operatorname{rg} q = \operatorname{rgMat}_{(u_1,u_2,\ldots,u_n)}(q)$, nous obtenons $\operatorname{rg} q = r$. Par suite $\dim\ker q = n - r = \dim F$. À partir de $\dim\ker q = \dim F$ et de $F \subset \ker q$ nous obtenons $F = \ker q$. $\square$

1.9. Classification des formes quadratiques

Un problème récurrent en mathématiques est de classifier les objets. Peut-on classer les formes quadratiques par type ? et selon quels critères ? Le théorème de Gauss fournit une première classification selon le rang de la forme quadratique : toute forme quadratique de rang r est une somme de r carrés de formes linéaires. On peut être plus précis au niveau de la classification en prenant en compte le signe devant les carrés dans la décomposition.

Soient q une forme quadratique sur un $\mathbb{R}$-espace vectoriel E de dimension n et $\mathcal{B} = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ une base q-orthogonale de E. Posons

$$\alpha(\mathcal{B}) = \#\{i \in \mathbb{N} \,|\, 1 \leq i \leq n \text{ et } q(\mathbf{e}_i) > 0\}$$

$$\beta(\mathcal{B}) = \#\{i \in \mathbb{N} \,|\, 1 \leq i \leq n \text{ et } q(\mathbf{e}_i) < 0\}.$$

Nous avons vu que la matrice de q dans la base q-orthogonale $\mathcal{B}$ est une matrice diagonale, les coefficients diagonaux de cette matrice étant les réels $q(\mathbf{e}_i)$. Les entiers $\alpha(\mathcal{B})$ et $\beta(\mathcal{B})$ correspondent respectivement aux nombres de valeurs propres strictement positives et strictement négatives de la matrice de q dans la base $\mathcal{B}$. Nous en déduisons que le rang de q, que l'on sait être indépendant de la base choisie, vaut $\mathrm{rg}(q) = \alpha(\mathcal{B}) + \beta(\mathcal{B})$.

L'énoncé suivant indique que non seulement $\alpha(\mathcal{B}) + \beta(\mathcal{B})$ est indépendant de la base orthogonale $\mathcal{B}$ mais que les deux entiers $\alpha(\mathcal{B})$ et $\beta(\mathcal{B})$ sont eux-mêmes indépendants de ce choix.

Théorème 1.9.1: Loi d'inertie de Sylvester

Soient q une forme quadratique sur un $\mathbb{R}$-espace vectoriel E de dimension n et $\mathcal{B} = (\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n)$ une base q-orthogonale de E. Posons

$$\alpha(\mathcal{B}) = \#\{i \in \mathbb{N} \,|\, 1 \leq i \leq n \text{ et } q(\mathbf{e}_i) > 0\}$$

$$\beta(\mathcal{B}) = \#\{i \in \mathbb{N} \,|\, 1 \leq i \leq n \text{ et } q(\mathbf{e}_i) < 0\}.$$

Les entiers $\alpha(\mathcal{B})$ et $\beta(\mathcal{B})$ sont indépendants de la base orthogonale $\mathcal{B}$.

Démonstration. — Afin de montrer que les deux entiers $\alpha(\mathcal{B})$ et $\beta(\mathcal{B})$ ne dépendent pas de la base orthogonale choisie, nous allons considérer deux bases q-orthogonales $\mathcal{B} = (\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n)$ et $\mathcal{C} = (\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_n)$ et montrer que $\alpha(\mathcal{B}) = \alpha(\mathcal{C})$ puis que $\beta(\mathcal{B}) = \beta(\mathcal{C})$. Quitte à échanger l'ordre des vecteurs des deux bases $\mathcal{B}$ et $\mathcal{C}$ on peut supposer que

$$\begin{cases} q(\mathbf{b}_i) > 0 \text{ pour } 1 \leq i \leq \alpha(\mathcal{B}) \\ q(\mathbf{b}_i) < 0 \text{ pour } \alpha(\mathcal{B}) + 1 \leq i \leq \alpha(\mathcal{B}) + \beta(\mathcal{B}) = \mathrm{rg}(q) \\ q(\mathbf{c}_i) > 0 \text{ pour } 1 \leq i \leq \alpha(\mathcal{C}) \\ q(\mathbf{c}_i) < 0 \text{ pour } \alpha(\mathcal{C}) + 1 \leq i \leq \alpha(\mathcal{C}) + \beta(\mathcal{C}) = \mathrm{rg}(q) \end{cases}$$

Considérons les sous-espaces vectoriels

$$G_+ = \mathrm{Vect}\big(\mathbf{b}_i \mid 1 \leq i \leq \alpha(\mathcal{B})\big) \qquad H_- = \mathrm{Vect}\big(\mathbf{c}_i \mid \alpha(\mathcal{C}) + 1 \leq i \leq n\big)$$

Montrons dans un premier temps que $G_+ \cap H_- = \{\mathbf{0}\}$. L'inclusion $\{\mathbf{0}\} \subset G_+ \cap H_-$ est immédiate car l'intersection de deux sous-espaces vectoriels est un sous-espace vectoriel et tout sous-espace vectoriel contient $\mathbf{0}$. Soit $\mathbf{x}$ un élément de $G_+ \cap H_-$. Supposons que $\mathbf{x} \neq \mathbf{0}$. Le vecteur $\mathbf{x}$ se décompose de manière unique dans la base $\big(\mathbf{b}_i \mid 1 \leq i \leq \alpha(\mathcal{B})\big)$ de G_+ :

$$\mathbf{x} = \sum_{i=1}^{\alpha(\mathcal{B})} \lambda_i \mathbf{b}_i$$

les λ_i désignant des réels non tous nuls. Puisque $\mathcal{B}$ est q-orthogonale nous en déduisons que $q(\mathbf{x}) = \sum_{i=1}^{\alpha(\mathcal{B})} \lambda_i^2 q(\mathbf{b}_i)$. De plus, $q(\mathbf{b}_i) > 0$ pour tout $1 \leq i \leq \alpha(\mathcal{B})$ ce qui implique l'inégalité $q(\mathbf{x}) > 0$. En raisonnant de manière analogue nous montrons que $x \in H_- \smallsetminus \{\mathbf{0}\}$ entraîne

$q(\mathbf{x}) \leq 0$: contradiction. Ainsi si $\mathbf{x}$ appartient à $G_+ \cap H_-$, alors nécessairement $\mathbf{x} = \mathbf{0}$. Autrement dit $G_+ \cap H_- \subset \{\mathbf{0}\}$. Finalement nous avons $G_+ \cap H_- = \{\mathbf{0}\}$.

Puisque $G_+ \cap H_- = \{\mathbf{0}\}$ nous avons l'égalité

$$\dim G_+ + \dim H_- = \dim(G_+ \oplus H_-).$$

Par ailleurs l'inclusion $G_+ \oplus H_- \subset E$ assure que $\dim(G_+ \oplus H_-) \leq \dim E$. On en déduit que $\dim G_+ + \dim H_- \leq \dim E$, c'est-à-dire que $\alpha(\mathcal{B}) + n - \alpha(\mathcal{C}) \leq n$ qui se réécrit $\alpha(\mathcal{B}) \leq \alpha(\mathcal{C})$.

En effectuant un raisonnement analogue en échangeant le rôle de $\mathcal{B}$ et celui de $\mathcal{C}$, *i.e.* en introduisant les espaces

$$G_- = \mathrm{Vect}\big(\mathbf{b}_i \,|\, \alpha(\mathcal{B}) + 1 \leq i \leq n\big) \qquad\qquad H_- = \mathrm{Vect}\big(\mathbf{c}_i \,|\, 1 \leq i \leq \alpha(\mathcal{C})\big)$$

nous obtenons l'inégalité $\alpha(\mathcal{C}) \leq \alpha(\mathcal{B})$. Finalement $\alpha(\mathcal{B}) = \alpha(\mathcal{C})$.

Puisque le rang de la forme quadratique q ne dépend pas de la base (Définition 1.3.2) nous avons

$$\mathrm{rg}(q) = \alpha(\mathcal{B}) + \beta(\mathcal{B}) = \alpha(\mathcal{C}) + \beta(\mathcal{C}).$$

L'égalité des entiers $\alpha(\mathcal{B})$ et $\alpha(\mathcal{C})$ implique celle des entiers $\beta(\mathcal{B})$ et $\beta(\mathcal{C})$. $\square$

Définition 1.9.1

Soient q une forme quadratique sur un espace vectoriel E de dimension n, $\mathcal{B} = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ une base q-orthogonale de E et α, β les entiers

$$\alpha = \#\{i \in \mathbb{N} \,|\, 1 \leq i \leq n \text{ et } q(\mathbf{e}_i) > 0\}$$
$$\beta = \#\{i \in \mathbb{N} \,|\, 1 \leq i \leq n \text{ et } q(\mathbf{e}_i) < 0\}.$$

Le couple $\varepsilon(q) = (\alpha, \beta)$ est appelé *signature de la forme quadratique q*.

Calculer la signature d'une forme quadratique en utilisant la définition des entiers α et β requiert une base orthogonale pour cette forme quadratique. Le Théorème de Gauss permet de déterminer la signature d'une forme quadratique sans avoir recours à une telle base. Plus précisément nous avons l'énoncé suivant.

Théorème 1.9.2

Soient E un $\mathbb{R}$-espace vectoriel de dimension finie et q une forme quadratique sur E dont une décomposition en carrés de formes linéaires indépendantes est $q = \displaystyle\sum_{i=1}^{r} \gamma_i f_i^2$. La signature de q est (α, β) avec

$$\alpha = \#\{i \in \mathbb{N} \,|\, 1 \leq i \leq r \text{ et } \gamma_i > 0\}$$
$$\beta = \#\{i \in \mathbb{N} \,|\, 1 \leq i \leq r \text{ et } \gamma_i < 0\}.$$

Démonstration. — Le théorème de la base incomplète permet de compléter la famille libre $(f_1, f_2, \ldots, f_r)$ en une base de $\widetilde{\mathcal{B}} = (f_1, f_2, \ldots, f_n)$ de l'espace vectoriel E^*. La Proposition 1.8.2 assure l'existence d'une base $\mathcal{B} = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ de E dont la base duale est $\widetilde{\mathcal{B}}$. Dans cette base $\mathcal{B}$ la matrice de q est une matrice diagonale qui s'écrit

$$\mathrm{Mat}_{\mathcal{B}}(q) = \left[\begin{array}{cccc|c} \gamma_1 & 0 & \ldots & 0 & \\ 0 & \ddots & \ddots & \vdots & 0 \\ \vdots & \ddots & \ddots & 0 & \\ 0 & \ldots & 0 & \gamma_r & \\ \hline & & & & \\ & & 0 & & 0 \\ & & & & \end{array}\right]$$

En particulier, la base $\mathcal{B}$ est orthogonale selon q et nous avons

$$\alpha = \#\{i \in \mathbb{N} \,|\, 1 \leq i \leq r \text{ et } \gamma_i > 0\}$$
$$\beta = \#\{i \in \mathbb{N} \,|\, 1 \leq i \leq r \text{ et } \gamma_i < 0\}.$$

□

Exemples 1.9.1. — 1. La forme quadratique de Lorentz définie sur $\mathbb{R}^4$ par

$$\varphi \colon \mathbb{R}^4 \to \mathbb{R}, \qquad (\mathbf{x}, \mathbf{y}) \mapsto x_1y_1 + x_2y_2 + x_3y_3 - c^2x_4y_4$$

où c désigne un paramètre réel, admet pour forme quadratique associée l'application

$$\mathbf{q} \colon \mathbb{R}^4 \to \mathbb{R}, \qquad \mathbf{x} \mapsto x_1^2 + x_2^2 + x_3^2 - c^2x_4^2.$$

Déterminons la signature de q en utilisant le Théorème 1.9.2. Remarquons que l'expression $q(\mathbf{x}) = x_1^2 + x_2^2 + x_3^2 - c^2x_4^2$ est une décomposition en somme de quatre carrés de formes linéaire ; de plus ces quatre formes linéaires ($\mathbf{x} \mapsto x_1$, $\mathbf{x} \mapsto x_2$, $\mathbf{x} \mapsto x_3$, $\mathbf{x} \mapsto x_4$) sont indépendantes. Selon la valeur du coefficient c nous avons

- ◇ $\varepsilon(q) = (3, 0)$ si $c = 0$;
- ◇ $\varepsilon(q) = (3, 1)$ si $c \neq 0$.

2. La forme quadratique sur $\mathbb{R}^4$ définie par

$$q \colon \mathbb{R}^4 \to \mathbb{R}, \qquad \mathbf{x} \mapsto x_1x_2 + x_1x_3 + x_1x_4 + x_2x_3 + x_2x_4 + x_3x_4$$

a pour décomposition en somme de carrés de formes linéaires indépendantes

$$q(\mathbf{x}) = f_1(\mathbf{x})^2 - f_2(\mathbf{x})^2 - f_3(\mathbf{x})^2 - 3f_4(\mathbf{x})^2$$

avec

$$f_1(\mathbf{x}) = \frac{1}{2}(x_1 + x_2 + 2x_3 + 2x_4), \quad f_2(\mathbf{x}) = \frac{1}{2}(x_1 - x_2), \quad f_3(\mathbf{x}) = x_3 + \frac{1}{2}x_4, \quad f_4(\mathbf{x}) = \frac{1}{2}x_4.$$

D'après le Théorème 1.8.1 le rang de q est 4 et d'après le Théorème 1.9.2 sa signature vaut $(1, 3)$.

Autre critère de classification. — On dit que deux formes quadratiques q et q' sur un $\mathbb{R}$-espace vectoriel E de dimension finie sont *isométriques* s'il existe une application linéaire et bijective u de E dans E telle que

$$\forall\, x \in \mathbf{E} \qquad q(\mathbf{x}) = q'(u(\mathbf{x})).$$

Notons φ (resp. φ') la forme polaire de q (resp. de q'). En utilisant la relation de polarisation donnée à la Proposition 1.5.1 nous établissons que les deux formes quadratiques q et q' sont isométriques si et seulement si

$$\forall\, \mathbf{x}, \mathbf{y} \in E \qquad \varphi(\mathbf{x}, \mathbf{y}) = \varphi'(u(\mathbf{x}), u(\mathbf{y})). \tag{1.9.1}$$

Considérons la matrice A (resp. B) représentant q (resp. q') dans une base $\mathcal{B}$ de E. Matriciellement la relation (1.9.1) se traduit par

$$\forall\, X, Y \in \mathbb{R}^n \qquad {}^{\mathrm{t}}XAY = {}^{\mathrm{t}}(PX)B(PY) = {}^{\mathrm{t}}X\,{}^{\mathrm{t}}PBPY \tag{1.9.2}$$

où P est la matrice de l'application linéaire u dans la base $\mathcal{B}$ et où X et Y sont les vecteurs colonnes représentant les éléments $\mathbf{x}$ et $\mathbf{y}$ de E dans cette base. La relation (1.9.2) étant valable pour tous $X, Y \in \mathbb{R}^n$ nous en déduisons que q et q' sont isométriques si et seulement s'il existe une matrice $P \in \mathrm{GL}(n, \mathbb{R})$ telle que

$$A = {}^{\mathrm{t}}PBP. \tag{1.9.3}$$

La relation d'isométrie est une relation d'équivalence sur l'ensemble des formes quadratiques sur E. Classifier les formes quadratiques pour cette relation revient à trouver les classes d'équivalence pour cette relation d'équivalence. Nous disposons ainsi de deux façons pour classer les formes quadratiques : les ranger selon leur signature ou les ranger par classe d'isométrie. On peut montrer que ces deux classifications sont identiques :

Proposition 1.9.1

Deux formes quadratiques sur un $\mathbb{R}$-espace vectoriel de dimension finie sont isométriques si et seulement si elles ont même signature.

CHAPITRE 2

PRODUIT SCALAIRE SUR UN ESPACE VECTORIEL

2.1. Produit scalaire

Définition 2.1.1

Soient E un $\mathbb{R}$-espace vectoriel et q une forme quadratique sur E.

- ⋄ La forme quadratique q est dite *positive* si $q(\mathbf{x}) \geq 0$ pour tout $\mathbf{x} \in E$.
- ⋄ La forme quadratique q est dite *définie* si $q(\mathbf{x}) \neq 0$ pour tout $\mathbf{x} \in E$.

Remarque 2.1.1. — Pour montrer qu'une forme quadratique q sur un $\mathbb{R}$-espace vectoriel E est définie il faut montrer que si $\mathbf{x} \neq 0$, alors $q(\mathbf{x}) \neq 0$. En pratique, il est souvent plus simple d'utiliser la contraposée : on montre que si $q(\mathbf{x}) = 0$ alors nécessairement $\mathbf{x} = \mathbf{0}$.

Réciproquement pour montrer qu'une forme quadratique q sur un $\mathbb{R}$-espace vectoriel E n'est pas définie, il suffit de trouver un vecteur $\mathbf{x} \neq 0$ pour lequel $q(\mathbf{x}) = 0$.

⚠ Le terme « définie » n'est pas à prendre dans son sens usuel en mathématiques. Pour une forme quadratique q sur un espace vectoriel E, le réel $q(\mathbf{x})$ « existe » pour tout $\mathbf{x} \in E$.

Exemples 2.1.1. — 1. La forme quadratique

$$q\colon \mathbb{R}^n \to \mathbb{R}, \qquad \mathbf{x} \mapsto \sum_{i=1}^{n} x_i^2$$

est une forme quadratique définie et une forme quadratique positive. En effet, cette application est positive car une somme de carrés est toujours positive. Elle est définie car $q(\mathbf{x}) = 0$ implique que $x_i = 0$ pour tout $1 \leq i \leq n$, donc que $\mathbf{x} = \mathbf{0}$ (une somme de réels positifs est nulle si et seulement si tous les termes la composant sont nuls).

2. La forme quadratique
$$q\colon \mathcal{C}([0,1],\mathbb{R}) \to \mathbb{R}, \qquad f \mapsto \int_0^1 f(t)^2\,\mathrm{d}t$$
est aussi une forme quadratique définie et une forme quadratique positive. La positivité provient de la positivité d'une fonction positive. C'est la continuité de f qui permet d'affirmer que $\varphi(f,f) \neq 0$ lorsque $f \neq 0$: l'application q est donc une forme quadratique définie.
3. La forme quadratique associée à la forme de Lorentz donnée par
$$q\colon \mathbb{R}^4 \to \mathbb{R}, \qquad \mathbf{x} \mapsto x_1^2 + x_2^2 + x_3^2 - c^2x_4^2$$
avec $c < 0$ n'est ni positive (puisque $q(0,0,0,1) = -c^2 < 0$), ni définie (puisque $(q(c,0,0,1) = 0)$.
4. La forme quadratique
$$q\colon \mathrm{M}(n,\mathbb{R}) \mapsto \mathbb{R}, \qquad A \mapsto \mathrm{tr}({}^{\mathrm{t}}AA)$$
est une forme quadratique définie et une forme quadratique positive. En effet, si on note $(a_{i,j})_{1\leq i,\,j\leq n}$ les coefficients de la matrice A alors le coefficient (i,j) de la matrice ${}^{\mathrm{t}}AA$ vaut $\displaystyle\sum_{k=1}^{n} a_{ki}a_{kj}$. La trace d'une matrice étant la somme de ces coefficients diagonaux, nous obtenons
$$q(A) = \sum_{i=1}^{n}\sum_{k=1}^{n} a_{ki}^2.$$
Nous nous retrouvons dans le même cas que celui du premier des exemples décrits ici : pour tout $A \in \mathrm{M}(n,\mathbb{R})$, $q(A)$ est positive car c'est une somme de carrés. Elle est définie car $q(A) = 0$ implique que pour tous $1 \leq i,\ j \leq n$ nous avons $a_{ij} = 0$ autrement dit que $A = 0$.

Définitions 2.1.1

Nous dirons qu'une forme quadratique q sur un $\mathbb{R}$-espace vectoriel E est *définie positive* si elle est à la fois définie et positive. Autrement dit une forme quadratique q est définie positive si $q(\mathbf{x}) > 0$ pour tout $\mathbf{x} \in E \smallsetminus \{\mathbf{0}\}$.
Une forme bilinéaire φ sur un $\mathbb{R}$-espace vectoriel E est *définie positive* si la forme quadratique qui lui est associée $E \to \mathbb{R}$, $\mathbf{x} \mapsto \varphi(\mathbf{x},\mathbf{x})$ est une forme quadratique définie positive.

Proposition 2.1.1

Soient E un $\mathbb{R}$-espace vectoriel et q une forme quadratique sur E.

- La forme quadratique q est une forme quadratique définie si et seulement si son cône isotrope $C(q)$ est $\{\mathbf{0}\}$.
- Si la forme quadratique q est une forme quadratique positive, alors son cône isotrope $C(q)$ et son noyau $N(q)$ coïncident.

Démonstration. — D'après la définition du cône isotrope d'une forme quadratique une forme quadratique est définie si et seulement si son cône isotrope est $\{\mathbf{0}\}$.

Considérons une forme quadratique q sur un espace vectoriel E et notons $N(q)$ le noyau de cette forme quadratique et $C(q)$ son cône isotrope. L'inclusion $N(q) \subset C(q)$ étant toujours vraie, il s'agit de prouver que si la forme quadratique q est positive alors $N(q) \subset C(q)$.

Supposons donc que la forme quadratique q est positive et considérons sa décomposition en carrés de formes linéaires indépendantes $q = \sum_{i=1}^{r} \alpha_i f_i^2$ où $r \leq n$ et où chaque α_i est un réel non nul. Nous montrons par l'absurde que tous les coefficients α_i sont strictement positifs : supposons que $\alpha_{i_0} < 0$ pour un certain $1 \leq i_0 \leq r$. La Proposition 1.8.3 assure que l'espace vectoriel $\bigcap_{i=1, i\neq i_0}^{r} \ker f_i$ est un espace vectoriel de dimension $n-(r-1)$ alors que l'espace vectoriel $\bigcap_{i=1}^{r} \ker f_i$ est un espace vectoriel de dimension $n-r$. Il existe donc un élément non nul $\mathbf{x}$ dans $(\bigcap_{i=1, i\neq i_0}^{r} \ker f_i) \smallsetminus \ker f_{i_0}$. Pour ce choix de $\mathbf{x} \neq \mathbf{0}$ nous avons

$$q(\mathbf{x}) = \sum_{i=1}^{r} \alpha_i f_i(\mathbf{x})^2 = \alpha_{i_0} f_{i_0}(\mathbf{x})^2 < 0$$

ce qui contredit le fait que la forme quadratique q est positive.

Ainsi si $\mathbf{x} \in C(q)$, alors nous avons $q(\mathbf{x}) = \sum_{i=1}^{r} \alpha_i f_i(\mathbf{x})^2 = 0$ avec $\alpha_i > 0$ pour tout $1 \leq i \leq r$ ce qui implique que $f_i(\mathbf{x}) = 0$ pour tout $1 \leq i \leq r$. La forme polaire de q qui est

$$\varphi_0 \colon E \times E \to \mathbb{R}, \qquad (\mathbf{z}, \mathbf{y}) \mapsto \sum_{i=1}^{r} \alpha_i f_i(\mathbf{z}) f_i(\mathbf{y})$$

vérifie donc $\varphi_0(\mathbf{x}, \mathbf{y}) = 0$ pour tout $\mathbf{y} \in E$. Cela implique que $\mathbf{x}$ appartient à $N(q)$. Nous avons donc montré l'inclusion $C(q) \subset N(q)$. $\square$

Remarque 2.1.2. — On déduit de la Proposition 2.1.1 que si une forme bilinéaire φ sur un $\mathbb{R}$-espace vectoriel est une forme bilinéaire définie alors on a

$$\forall\, \mathbf{x} \in E \quad \varphi(\mathbf{x}, \mathbf{y}) = 0 \quad \Longleftrightarrow \quad \mathbf{y} = \mathbf{0}.$$

En effet, si $\mathbf{y} = \mathbf{0}$, alors $\varphi(\mathbf{x}, \mathbf{y}) = 0$ pour tout $\mathbf{x} \in E$. Réciproquement si $\varphi(\mathbf{x}, \mathbf{y}) = 0$ pour tout $\mathbf{x} \in E$, alors $\varphi(\mathbf{y}, \mathbf{y}) = 0$, *i.e.* $q(\mathbf{y}) = 0$. Mais φ est définie donc $C(q) = \{\mathbf{0}\}$ d'où $\mathbf{y} = \mathbf{0}$.

Par linéarité nous en déduisons que si φ est une forme bilinéaire définie alors

$$\forall\, \mathbf{x} \in E \quad \varphi(\mathbf{x}, \mathbf{a}) = \varphi(\mathbf{x}, \mathbf{b}) \quad \Longleftrightarrow \quad \mathbf{a} = \mathbf{b}.$$

Définition 2.1.2

Nous appelons *produit scalaire* sur un $\mathbb{R}$-espace vectoriel E une forme bilinéaire symétrique définie positive.

Pour établir qu'une forme bilinéaire φ est un produit scalaire il faut montrer que

◇ l'application φ est symétrique : $\forall\, \mathbf{x},\, \mathbf{y} \in E,\ \varphi(\mathbf{x}, \mathbf{y}) = \varphi(\mathbf{y}, \mathbf{x})$;

◇ la forme quadratique $q\colon E \to \mathbb{R},\ \mathbf{x} \mapsto \varphi(\mathbf{x}, \mathbf{x})$ est définie positive :
  - $\forall\, \mathbf{x} \in E\ q(\mathbf{x}) \geq 0$,
  - $\forall\, \mathbf{x} \in E\ q(\mathbf{x}) = 0 \Rightarrow \mathbf{x} = \mathbf{0}$.

Exemples 2.1.2. — 1. La forme bilinéaire

$$\varphi\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}, \qquad (\mathbf{x}, \mathbf{y}) \mapsto \sum_{i=1}^{n} x_i y_i$$

est un produit scalaire sur $\mathbb{R}^n$ appelé *produit scalaire canonique*. Matriciellement nous avons

$$\varphi(\mathbf{x}, \mathbf{y}) = {}^{\mathrm{t}}XY$$

X et Y étant les matrices colonnes ayant pour coordonnées les composantes de $\mathbf{x}$ et $\mathbf{y}$ dans la base canonique de $\mathbb{R}^n$.

2. La forme bilinéaire

$$q\colon \mathcal{C}([0,1], \mathbb{R}) \times \mathcal{C}([0,1], \mathbb{R}) \to \mathbb{R}, \qquad (f, g) \mapsto \int_0^1 f(t)g(t)\, \mathrm{d}t.$$

est un produit scalaire. Nous avons vu précédemment que cette application est bilinéaire. Elle est symétrique et le fait qu'elle soit définie positive a aussi été établi précédemment.

3. Sur l'espace vectoriel $\mathrm{M}(n, \mathbb{R})$ des matrices carrées de taille n à coefficients réels l'application φ définie par
$$\varphi \colon \mathrm{M}(n, \mathbb{R}) \times \mathrm{M}(n, \mathbb{R}) \to \mathbb{R}, \qquad (A, B) \mapsto \mathrm{tr}({}^{\mathrm{t}}AB).$$
Nous avons établi qu'il s'agit d'une forme bilinéaire et que la forme quadratique qui lui est associée est définie positive.

Proposition 2.1.2

Soient E un $\mathbb{R}$-espace vectoriel de dimension finie égale à n et q une forme quadratique sur E. La forme polaire de q est un produit scalaire si et seulement si sa signature est $(n, 0)$.

Démonstration. — Supposons que q soit un produit scalaire sur E. Puisque la forme quadratique q est positive, sa décomposition en carrés de formes linéaires indépendantes donnée par le Théorème 1.9.2 s'écrit $q = \sum_{i=1}^{r} \alpha_i f_i^2$. De plus, nous pouvons montrer comme dans la démonstration de la Proposition 2.1.1 que si la forme quadratique q est positive alors tous les coefficients α_i sont strictement positifs. La signature de q est donc $(r, 0)$.

Montrons en raisonnant par l'absurde que $r = n$. Supposons que $r < n$. Puisque les r formes linéaires f_1, f_2, …, f_r sont indépendantes la Proposition 1.8.3 assure que l'espace vectoriel $\bigcap_{i=1}^{r} \ker f_i$ est de dimension $n - r > 0$. Par conséquent il existe $\mathbf{x} \in \bigcap_{i=1}^{r} \ker f_i$ tel que $\mathbf{x} \neq \mathbf{0}$. Ce vecteur non nul satisfait $q(\mathbf{x}) = 0$ ce qui est impossible car q est définie positive.

Réciproquement supposons que la signature de q soit $(n, 0)$. La décomposition de q en carrés de formes linéaires indépendantes (Théorème 1.9.2) s'écrit $q = \sum_{i=1}^{n} \alpha_i f_i^2$ où tous les coefficients α_i sont strictement positifs. Il s'ensuit que pour tout $\mathbf{x} \in E$, $q(\mathbf{x}) \geq 0$ (donc que la forme quadratique q est positive) et que si $q(\mathbf{x}) = 0$ alors $f_i(\mathbf{x}) = 0$ pour tout $1 \leq i \leq n$. Puisque la famille $(f_1, f_2, \ldots, f_n)$ est une base de E^* ceci implique que $f(\mathbf{x}) = 0$ pour toute forme linéaire $f \in E^*$. Nous en concluons que $\mathbf{x} = \mathbf{0}$ (Proposition 1.8.1). □

Remarque 2.1.3. — Une forme quadratique permet de définir un produit scalaire lorsque sa forme polaire associée est un produit scalaire.

Exemples 2.1.3. — 1. La forme quadratique q donnée sur $\mathbb{R}^4$ par
$$q(\mathbf{x}) = x_1x_2 + x_1x_3 + x_1x_4 + x_2x_3 + x_2x_4 + x_3x_4$$
admet pour écriture en somme de carrés de formes linéaires indépendantes :
$$q(\mathbf{x}) = f_1(\mathbf{x})^2 - f_2(\mathbf{x})^2 - f_3(\mathbf{x})^2 - 3f_4(\mathbf{x})^2$$

où les formes linéaires f_i sont données par

$$f_1(\mathbf{x}) = \frac{1}{2}(x_1 + x_2 + 2x_3 + 2x_4), \qquad f_2(\mathbf{x}) = \frac{1}{2}(x_1 - x_2), \qquad f_3(\mathbf{x}) = x_3 + \frac{1}{2}x_4, \qquad f_4(\mathbf{x}) = \frac{1}{2}x_4.$$

Nous en déduisons que la signature de q vaut $(1, 3)$. En vertu de la Proposition 2.1.2 cette forme quadratique ne permet pas de définir un produit scalaire sur $\mathbb{R}^4$. À partir de la décomposition de q en somme de carrés de formes linéaires nous pouvons trouver un vecteur non nul nul $\mathbf{x} \in \mathbb{R}^4$ tel que $q(\mathbf{x}) < 0$ ce qui prouve que q n'est pas une forme quadratique positive ; par exemple $q(1, -1, 0, 0) = -1$ (plus généralement n'importe quel vecteur $\mathbf{x}$ non nul dans le noyau de l'application linéaire f_1 est tel que $q(\mathbf{x}) < 0$).

2. Nous en déduisons que la signature de q vaut $(1, 3)$. En vertu de la Proposition 2.1.2 cette forme quadratique ne permet pas de définir un produit scalaire sur $\mathbb{R}^4$. À partir de la décomposition de q en sommes de carrés de formes linéaires nous pouvons trouver un vecteur non nul nul $\mathbf{x} \in \mathbb{R}^4$ tel que $q(\mathbf{x}) < 0$ ce qui prouve que q n'est pas une forme quadratique positive ; par exemple $q(1, -1, 0, 0) = -1$ (plus généralement n'importe quel vecteur $\mathbf{x}$ non nul dans le noyau de l'application linéaire f_1 est tel que $q(\mathbf{x}) < 0$).

3. La forme quadratique q donnée sur $\mathbb{R}^4$ par

$$q(\mathbf{x}) = f_1(\mathbf{x})^2 + 2f_2(\mathbf{x})^2 + 2f_3(\mathbf{x})^2$$

où f_1, f_2 et f_3 sont les trois linéaires de l'exemple précédent a pour signature $(3, 0)$ et ne définit donc pas un produit scalaire sur $\mathbb{R}^4$. Bien entendu, cette forme quadratique est positive mais puisque sa décomposition en somme de carrés de formes linéaires indépendantes ne comporte que trois termes nous pouvons trouver un élément $\mathbf{x} \in \mathbb{R}^4$ non nul tel que $q(\mathbf{x}) = 0$. Par exemple le vecteur $\mathbf{x} = (1, 1, 1, -2)$ annule les trois formes linéaires f_1, f_2 et f_3 et donc annule aussi la forme quadratique q.

Théorème 2.1.1: Inégalité de Cauchy-Schwarz

Pour tout produit scalaire φ sur un $\mathbb{R}$-espace vectoriel E nous avons

$$\forall \mathbf{x}, \mathbf{y} \in E \quad \varphi(\mathbf{x}, \mathbf{y})^2 \leq \varphi(\mathbf{x}, \mathbf{x})\varphi(\mathbf{y}, \mathbf{y}).$$

Démonstration. — Considérons un produit scalaire φ sur E et $\mathbf{x}$, $\mathbf{y}$ dans E. L'application bilinéaire φ étant positive nous avons $\varphi(\mathbf{x} + t\mathbf{y}, \mathbf{x} + t\mathbf{y}) \geq 0$ pour tout $t \in \mathbb{R}$. De plus, en vertu de la bilinéarité de φ nous avons

$$\varphi(\mathbf{x} + t\mathbf{y}, \mathbf{x} + t\mathbf{y}) = t^2\varphi(\mathbf{y}, \mathbf{y}) + 2t\varphi(\mathbf{x}, \mathbf{y}) + \varphi(\mathbf{x}, \mathbf{x}).$$

Nous en déduisons que l'application

$$\mathbb{R} \to \mathbb{R}, \qquad t \mapsto \varphi(\mathbf{x} + t\mathbf{y}, \mathbf{x} + t\mathbf{y})$$

est une application polynomiale de degré 2 qui est à valeurs positives. Elle admet donc au plus une racine réelle et par conséquent, son discriminant Δ est négatif. Ce discriminant vaut

$$\Delta = 4\varphi(\mathbf{x},\mathbf{y})^2 - 4\varphi(\mathbf{x},\mathbf{x})\varphi(\mathbf{y},\mathbf{y})$$

et l'inégalité $\Delta \leq 0$ correspond exactement à l'inégalité de Cauchy-Schwarz. $\square$

Au cours de la démonstration, nous pouvons remarquer les deux choses suivantes :

- ◇ À aucun moment de la démonstration nous avons utilisé le fait que la forme bilinéaire φ est définie. L'inégalité de Cauchy-Schwarz est donc plus généralement vraie pour une forme bilinéaire symétrique et positive.
- ◇ On peut trouver les conditions sur les vecteurs $\mathbf{x}$ et $\mathbf{y}$ pour que l'inégalité de Cauchy-Schwarz devienne l'égalité

$$\varphi(\mathbf{x},\mathbf{y})^2 = \varphi(\mathbf{x},\mathbf{x})\varphi(\mathbf{y},\mathbf{y}). \tag{2.1.1}$$

  En effet, en reprenant les notations de la démonstration l'égalité (2.1.1) a lieu si le discriminant Δ est nul ce qui correspond à l'existence d'une unique racine t_0 pour l'application polynomiale $t \mapsto \varphi(\mathbf{x}+t\mathbf{y},\mathbf{x}+t\mathbf{y})$:

$$\exists\, t_0 \in \mathbb{R} \quad \varphi(\mathbf{x}+t_0\mathbf{y},\mathbf{x}+t_0\mathbf{y}) = 0.$$

  Compte tenu du fait que l'application φ est définie cela équivaut à

$$\exists\, t_0 \in \mathbb{R} \quad \mathbf{x}+t_0\mathbf{y} = 0.$$

  Ainsi l'inégalité de Cauchy-Schwarz devient une égalité si et seulement si les vecteurs $\mathbf{x}$ et $\mathbf{y}$ sont liés.

L'inégalité de Cauchy-Schwarz permet, en effectuant des choix particuliers de produits scalaires, d'obtenir de nombreuses inégalités classiques :

- ◇ Dans $\mathbb{R}^n$ l'inégalité de Cauchy-Schwarz pour le produit scalaire canonique s'écrit

$$\forall\, \mathbf{x}, \in \mathbb{R}^n,\ \mathbf{y} \in \mathbb{R}^n \quad \left(\sum_{i=1}^n x_i y_i\right)^2 \leq \left(\sum_{i=1}^n x_i^2\right)\left(\sum_{i=1}^n y_i^2\right)$$

- ◇ Si le produit scalaire considéré est

$$\mathcal{C}([0,1],\mathbb{R}) \times \mathcal{C}([0,1],\mathbb{R}) \to \mathbb{R}, \qquad (f,g) \mapsto \int_0^1 f(t)g(t)\,\mathrm{d}t$$

  alors l'inégalité de Cauchy-Schwarz se réécrit

$$\left(\int_0^1 f(t)g(t)\,\mathrm{d}t\right)^2 \leq \left(\int_0^1 f(t)^2\,\mathrm{d}t\right)\left(\int_0^1 g(t)^2\,\mathrm{d}t\right).$$

Proposition 2.1.3

Soit φ un produit scalaire sur un $\mathbb{R}$-espace vectoriel E. L'application

$$N : E \to \mathbb{R}, \qquad \mathbf{x} \mapsto \sqrt{\varphi(\mathbf{x}, \mathbf{x})}$$

est une norme sur E, appelée *norme associée au produit scalaire* φ.

Démonstration. — Pour montrer que l'application N est une norme il faut montrer que :

1. $\forall \mathbf{x} \in E \quad N(\mathbf{x}) \geq 0$,
2. $\forall \mathbf{x} \in E \quad N(\mathbf{x}) = 0 \iff \mathbf{x} = 0$,
3. $\forall \mathbf{x} \in E \quad \forall \lambda \in \mathbb{R} \quad N(\lambda \mathbf{x}) = |\lambda| N(\mathbf{x})$,
4. $\forall \mathbf{x} \in E \quad \forall \mathbf{y} \in E \quad N(\mathbf{x} + \mathbf{y}) \leq N(\mathbf{x}) + N(\mathbf{y})$.

La première condition est évidente d'après la définition de N.

La seconde condition résulte du fait que φ est une forme bilinéaire définie.

La troisième condition découle de manière immédiate de la bilinéarité de φ.

Afin de montrer la dernière condition, considérons $\mathbf{x} \in E$, $\mathbf{y} \in E$. Par définition de l'application N nous avons

$$N(\mathbf{x} + \mathbf{y})^2 = \varphi(\mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y}).$$

L'application φ étant bilinéaire et symétrique nous en déduisons que

$$N(\mathbf{x} + \mathbf{y})^2 = N(\mathbf{x})^2 + 2\varphi(\mathbf{x}, \mathbf{y}) + N(\mathbf{y})^2.$$

En utilisant l'inégalité de Cauchy-Schwarz (Théorème 2.1.1) nous avons

$$\varphi(\mathbf{x}, \mathbf{y}) \leq \sqrt{\varphi(\mathbf{x}, \mathbf{x})}\sqrt{\varphi(\mathbf{y}, \mathbf{y})} = N(\mathbf{x})N(\mathbf{y}).$$

Nous en déduisons que

$$N(\mathbf{x} + \mathbf{y})^2 \leq N(\mathbf{x})^2 + 2N(\mathbf{x})N(\mathbf{y}) + N(\mathbf{y})^2 = \left(N(\mathbf{x}) + N(\mathbf{y})\right)^2.$$

□

2.2. Espace euclidien

Définitions 2.2.1

Nous appelons *espace préhilbertien* un $\mathbb{R}$-espace vectoriel muni d'un produit scalaire.
Nous appelons *espace euclidien* un espace préhilbertien de dimension finie.

Lorsqu'il n'y aura pas d'ambiguïté nous noterons $\langle \mathbf{x}, \mathbf{y} \rangle$ le produit scalaire entre deux éléments $\mathbf{x}$ et $\mathbf{y}$ et $||\mathbf{x}||$ la norme associée. Un espace euclidien correspond à la donnée d'un couple

$(E, \langle \cdot, \cdot \rangle)$ formé d'un $\mathbb{R}$-espace vectoriel de dimension finie et d'un produit scalaire sur cet espace.

Avec ces notations l'inégalité de Cauchy-Schwarz (Théorème 2.1.1) se réécrit

$$\forall \, \mathbf{x}, \, \mathbf{y} \in E \quad |\langle \mathbf{x}, \mathbf{y} \rangle| \leq ||\mathbf{x}|| \, ||\mathbf{y}||.$$

Puisqu'à un produit scalaire $\langle \cdot, \cdot \rangle$ nous pouvons toujours associer une norme $|| \cdot ||$ (Proposition 2.1.3), un espace euclidien $(E, \langle \cdot, \cdot \rangle)$ est toujours un espace vectoriel normé $(E, || \cdot ||)$.

Réciproquement, nous pouvons montrer que tout espace vectoriel normé $(E, || \cdot ||)$ où la norme satisfait la relation suivante, appelée *identité du parallélogramme*,

$$\forall \, \mathbf{x}, \, \mathbf{y} \in E \quad ||\mathbf{x} + \mathbf{y}||^2 + ||\mathbf{x} - \mathbf{y}||^2 = 2(||\mathbf{x}||^2 + ||\mathbf{y}||^2)$$

est un espace préhilbertien.

Le produit scalaire étant une forme bilinéaire, dans un espace préhilbertien, il existe une notion d'orthogonalité intrinsèque qui est définie par ce produit scalaire et nous avons l'énoncé suivant :

Théorème 2.2.1

Soient $\mathbf{x}$ et $\mathbf{y}$ deux vecteurs d'un espace préhilbertien E. Les deux assertions suivantes sont équivalentes :

- ◇ $\mathbf{x}$ et $\mathbf{y}$ sont orthogonaux,
- ◇ $||\mathbf{x} + \mathbf{y}||^2 = ||\mathbf{x}||^2 + ||\mathbf{y}||^2$.

Démonstration. — Dans un espace préhilbertien E pour tous $\mathbf{x}$, $\mathbf{y} \in E$ nous avons

$$||\mathbf{x} + \mathbf{y}||^2 = \langle \mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{x} \rangle + 2\langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{y}, \mathbf{y} \rangle = ||\mathbf{x}||^2 + 2\langle \mathbf{x}, \mathbf{y} \rangle + ||\mathbf{y}||^2.$$

Par définition, deux vecteurs $\mathbf{x}$ et $\mathbf{y}$ sont orthogonaux si et seulement si $\langle \mathbf{x}, \mathbf{y} \rangle = 0$. □

Dans le cas où l'espace préhilbertien E est l'espace vectoriel $\mathbb{R}^2$ muni du produit scalaire canonique, le théorème de Pythagore s'interprète de la manière suivante : « le carré de la longueur de l'hypothénuse d'un triangle rectangle est égal à la somme des carrés des deux autres longueurs des côtés du triangle. »

Supposons que F soit un sous-espace vectoriel de dimension finie d'un espace euclidien. Puisque $F^\perp$ est un supplémentaire de F nous pouvons définir la projection sur F parallèlement à $F^\perp$. Il s'agit de l'application linéaire $p \colon E \to E$ définie par $p(x) = u$ si $x = u + v$, $u \in F$, $v \in F^\perp$. Le noyau de p est $F^\perp$. De plus, pour tout vecteur $x \in E$, on a $p(x) = x$ si et seulement si x appartient à F. On en déduit que l'image de p est F et que $F^\perp$ est le sous-espace vecoriel des vecteurs $x - p(x)$, $x \in E$.

Définition 2.2.1

Soit F un sous-espace vectoriel de dimension finie d'un espace préhilbertien réel E. On appelle *projection orthogonale* sur F la projection sur F parallèlement à $F^\perp$.

Dans ce cas on a $(F^\perp)^\perp = F$ et $E = F \oplus F^\perp$. On peut considérer la projection orthogonale sur $F^\perp$ parallèlement à F.

Lemme 2.2.1

Soit E un espace euclidien. Supposons que F soit un sous-espace vectoriel de dimension finie de E. Si p est la projection orthogonale sur F, alors la projection orthogonale sur $F^\perp$ est $\mathrm{id}_E - p$.

Démonstration. — Soit $x \in E$. Nous avons $x = \underbrace{p(x)}_{\in F} + \underbrace{(x - p(x))}_{\in F^\perp}$. La projeté orthogonal de x sur $F^\perp$ est donc $x - p(x) = (\mathrm{id}_E - p)(x)$. □

Proposition 2.2.1

Soit E un espace euclidien. Soit D la droite de E engendrée par le vecteur non nul $\mathbf{v}$. Si p est la projection orthogonale de E sur D et si q est la projection orthogonale sur $D^\perp$, alors pour tout $x \in E$, on

$$p(x) = \frac{\langle x, \mathbf{v}\rangle}{||\mathbf{v}||^2}\mathbf{v}, \qquad q(x) = x - \frac{\langle x, \mathbf{v}\rangle}{||\mathbf{v}||^2}\mathbf{v}.$$

Démonstration. — Soit $x \in E$. Puisque $p(x)$ appartient à D il existe $\lambda \in R$ tel que $p(x) = \lambda\mathbf{v}$. D'autre part le vecteur $x - p(x)$ est orthogonal à D d'où $0 = \langle x - p(x), \mathbf{v}\rangle$; mais

$$\langle x - p(x), \mathbf{v}\rangle = \langle x, \mathbf{v}\rangle - \langle p(x), \mathbf{v}\rangle = \langle x, \mathbf{v}\rangle - \langle \lambda\mathbf{v}, \mathbf{v}\rangle = \langle x, \mathbf{v}\rangle - \lambda||\mathbf{v}||^2.$$

Finalement $\langle x, \mathbf{v}\rangle = \lambda||\mathbf{v}||^2$ et $\lambda = \frac{\langle x, \mathbf{v}\rangle}{||\mathbf{v}||^2}$.

Enfin $q(x) = \mathrm{id}(x) - p(x) = x - p(x)$ conduit à $q(x) = x - \frac{\langle x, \mathbf{v}\rangle}{||\mathbf{v}||^2}\mathbf{v}$. □

Théorème 2.2.2

Soit F un sous-espace vectoriel de dimension finie d'un espace préhilbertien réel E. Désignons par p_F la projection orthogonale sur F.

- ◇ Si $(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_p)$ est une base orthonormale de F, alors pour tout vecteur x de E nous avons : $p_F(x) = \displaystyle\sum_{i=1}^{p} \langle \mathbf{u}_i, x \rangle \mathbf{u}_i$.
- ◇ Pour tout $x \in E$, le vecteur $p_F(x)$ est l'unique vecteur de F réalisant le minimum de la distance de x à F, c'est-à-dire $||x - p_F(x)|| = \min\limits_{y \in F} (||x - y||)$.

Démonstration. — ◇ **à faire**

- ◇ Soit $x \in E$. Alors $x - p_F(x)$ appartient à $F^\perp$ et pour tout $y \in F$ nous avons d'après le Théorème 2.2.1

$$||x - y||^2 = ||(x - p_F(x)) + (p_F(x) - y)||^2 = ||(x - p_F(x))||^2 + ||(p_F(x) - y)||^2 \geq ||x - p_F(x)||^2$$

Ainsi $p_F(x)$ réalise le minimum de la distance de x à F ; de plus, si $y \in F$ réalise également ce minimum l'inégalité précédente doit être une égalité ce qui implique $||x - p_F(x)|| = 0$, c'est-à-dire $x = p_F(x)$.

□

2.3. Procédé d'orthogonalisation de Schmidt

Le Théorème 1.7.1 assure que dans tout $\mathbb{R}$-espace vectoriel E de dimension finie et pour toute forme quadratique q sur E il existe une base q-orthogonale. La démonstration de ce résultat établit l'existence d'une telle base mais ne fournit pas de moyen de construire une telle base. Dans le cas d'un espace préhilbertien, en plus d'une notion d'orthogonalité « naturelle » (celle associée à la forme quadratique issue du produit scalaire) nous disposons de la notion de produit scalaire. La méthode d'orthogonalisation de Schmidt est une méthode permettant, à l'aide du produit scalaire, de construire des bases orthogonales à partir d'une base quelconque.

Proposition 2.3.1: Orthogonalisation de Schmidt

Soient E un espace préhilbertien et $(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_p)$ une famille libre de E. La famille $(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_p)$ définie par récurrence par

$$\left\{\begin{array}{l}\varepsilon_1 = \mathbf{e}_1\\ \varepsilon_k = \mathbf{e}_k - \displaystyle\sum_{i=1}^{k-1} \frac{\langle \varepsilon_i, \mathbf{e}_k\rangle}{||\varepsilon_i||^2}\,\varepsilon_i \qquad \forall\, 2 \leq k \leq p\end{array}\right. \tag{2.3.1}$$

est une famille orthogonale qui engendre le même espace vectoriel que la famille $(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_p)$

Remarque 2.3.1. — Puisque les vecteurs ε_i sont non nuls, nous avons $||\varepsilon_i|| \neq 0$ et nous pouvons diviser chacun des vecteurs de $(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_p)$ par sa norme afin d'obtenir une famille orthonormale.

Démonstration. — Montrons par récurrence sur l'entier $p \in \mathbb{N}^*$ la propriété $(\mathcal{P}_p)$ suivante :

> Pour toute famille libre à p éléments $(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_p)$ la famille $(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_p)$ définie par la relation (2.3.1) est orthogonale et ces deux familles engendrent le même espace vectoriel.

- ◇ La propriété $(\mathcal{P}_1)$ est vraie : en effet, si une famille à un élément (ε_1) est libre, alors cet élément est non nul. La famille (ε_1) est alors une famille orthogonale puisque toute famille à un seul élément non nul est orthogonale.
- ◇ Supposons que la propriété $(\mathcal{P}_q)$ est vraie pour $q \in \mathbb{N}^*$ et montrons que la propriété $(\mathcal{P}_{q+1})$ est vraie. Considérons une famille à $q+1$ éléments $(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_{q+1})$ et construisons la famille $(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_{q+1})$ en utilisant la relation (2.3.1). Puisque le procédé de construction est itératif, les q premiers vecteurs $(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_q)$ sont construits à partir des q premiers vecteurs $(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_q)$. Par hypothèse de récurrence la famille $(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_q)$ est orthogonale et engendre le même espace vectoriel que la famille $(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_q)$.

  Pour montrer que la famille $(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_{q+1})$ est orthogonale, montrons que le vecteur ε_{q+1} est orthogonal aux autres vecteurs ε_k pour $1 \leq k \leq q$. Nous avons

$$\varepsilon_{q+1} = \mathbf{e}_{q+1} - \sum_{i=1}^{q} \frac{\langle \varepsilon_i, \mathbf{e}_{q+1}\rangle}{||\varepsilon_i||^2}\varepsilon_i,$$

  et en effectuant le produit scalaire de ε_{q+1} avec un vecteur ε_k où $1 \leq k \leq q$ nous obtenons

$$\langle \varepsilon_{q+1}, \varepsilon_k\rangle = \langle \mathbf{e}_{q+1}, \varepsilon_k\rangle - \sum_{i=1}^{q} \frac{\langle \varepsilon_i, \mathbf{e}_{q+1}\rangle}{||\varepsilon_i||^2}\langle \varepsilon_i, \varepsilon_k\rangle.$$

Puisque la famille $(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_q)$ est orthogonale, les termes $\langle \varepsilon_i, \varepsilon_k \rangle$ apparaissant dans la dernière somme sont nuls excepté le terme correspondant à la valeur $i = k$. Par suite nous obtenons

$$\begin{aligned}
\langle \varepsilon_{q+1}, \varepsilon_k \rangle &= \langle \mathbf{e}_{q+1}, \varepsilon_k \rangle - \frac{\langle \varepsilon_k, \mathbf{e}_{q+1} \rangle}{||\varepsilon_k||^2} \langle \varepsilon_k, \varepsilon_k \rangle \\
&= \langle \mathbf{e}_{q+1}, \varepsilon_k \rangle - \frac{\langle \varepsilon_k, \mathbf{e}_{q+1} \rangle}{||\varepsilon_k||^2} ||\varepsilon_k||^2 \\
&= \langle \mathbf{e}_{q+1}, \varepsilon_k \rangle - \langle \varepsilon_k, \mathbf{e}_{q+1} \rangle \\
&= \langle \mathbf{e}_{q+1}, \varepsilon_k \rangle - \langle \mathbf{e}_{q+1}, \varepsilon_k \rangle \\
&= 0;
\end{aligned}$$

la famille $(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_{q+1})$ est donc orthogonale.

Il reste à montrer que la famille $(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_{q+1})$ engendre le même sous-espace vectoriel que la famille $(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_{q+1})$. Par hypothèse de récurrence les deux familles $(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_q)$ et $(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_q)$ engendrent le même espace. Puisque $\mathbf{e}_{q+1}$ s'exprime comme combinaison linéaire de ε_{q+1} et des autres ε_k nous en déduisons que l'espace engendré par $(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_{q+1})$ est inclus dans celui engendré par $(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_{q+1})$. De même ε_{q+1} est défini comme combinaison linéaire de $\mathbf{e}_{q+1}$ et des ε_k, $1 \leq k \leq q$; l'espace engendré par $(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_{q+1})$ est donc inclus dans celui engendré par $(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_{q+1})$. Finalement les espaces vectoriels engendrés par les deux familles $(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_{q+1})$ et $(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_{q+1})$ sont égaux.

□

Exemple 2.3.1. — Considérons le $\mathbb{R}$-espace vectoriel $\mathbb{R}_2[X]$ des polynômes à coefficients réels de degré inférieur ou égal à 2. Il s'agit d'un espace vectoriel de dimension 3 dont la base canonique est $(1, X, X^2)$. En tant que sous-espace vectoriel de l'espace vectoriel $\mathcal{C}^0([0,1], \mathbb{R})$ des fonctions continues sur $[0,1]$ l'espace $\mathbb{R}_2[X]$ peut être muni du produit scalaire suivant :

$$\mathbb{R}_2[X] \times \mathbb{R}_2[X] \to \mathbb{R}, \qquad (P, Q) \mapsto \int_0^1 P(t)Q(t)\,\mathrm{d}t.$$

À partir de la base canonique de l'espace vectoriel $\mathbb{R}_2[X]$ construisons une base orthonormée de l'espace euclidien $(\mathbb{R}_2[X], \varphi)$ en utilisant le procédé d'orthogonalisation de Schmidt. Nous avons

$$\varepsilon_1 = 1, \qquad \varepsilon_2 = X - \frac{\varphi(\varepsilon_1, X)}{\varphi(\varepsilon_1, \varepsilon_1)}\varepsilon_1, \qquad \varepsilon_3 = X^2 - \frac{\varphi(\varepsilon_1, X^2)}{\varphi(\varepsilon_1, \varepsilon_1)}\varepsilon_1 - \frac{\varphi(\varepsilon_2, X^2)}{\varphi(\varepsilon_2, \varepsilon_2)}\varepsilon_2.$$

Calculons la matrice associée au produit scalaire φ dans la base que l'on souhaite orthogonaliser : cette matrice contient toutes les valeurs $\varphi(X^i, X^j)$ utile pour le calcul de chaque ε_k. Nous avons

$$\mathrm{Mat}_{(1,X,X^2)}(\varphi) = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \end{pmatrix}.$$

Nous obtenons

$$\varepsilon_1 = 1, \qquad \varepsilon_2 = X - \frac{1}{2}, \qquad \varepsilon_3 = X^2 - \frac{1}{3} - \left(X - \frac{1}{2}\right) = X^2 - X + \frac{1}{3}.$$

Si nous souhaitons obtenir une base orthonormée il suffit de diviser chaque vecteur de la base orthogonale $(\varepsilon_1, \varepsilon_2, \varepsilon_3)$ par sa norme pour le produit scalaire utilisé

$$\mathbf{f}_1 = \frac{\varepsilon_1}{\varphi(\varepsilon_1, \varepsilon_1)}, \qquad \mathbf{f}_2 = \frac{\varepsilon_2}{\varphi(\varepsilon_2, \varepsilon_2)}, \qquad \mathbf{f}_3 = \frac{\varepsilon_3}{\varphi(\varepsilon_3, \varepsilon_3)}.$$

Nous obtenons ainsi la base orthonormale $(\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3)$ où

$$\mathbf{f}_1 = 1, \qquad \mathbf{f}_2 = 2\sqrt{3}\left(X - \frac{1}{2}\right), \qquad \mathbf{f}_3 = 3\sqrt{\frac{10}{23}}\left(X^2 - X + \frac{1}{3}\right).$$

Remarque 2.3.2. — Considérons une base orthonormale $(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)$ d'un espace vectoriel E de dimension n. Pour tous $\mathbf{x}$, $\mathbf{y} \in E$ de composantes $(x_1, x_2, \ldots, x_n)$ et $(y_1, y_2, \ldots, y_n)$ dans la base $(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_p)$ nous avons

$$\langle \mathbf{x}, \mathbf{y} \rangle = \langle \sum_{i=1}^{n} x_i \varepsilon_i, \sum_{j=1}^{n} x_j \varepsilon_j \rangle = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i y_j \underbrace{\langle \varepsilon_i, \varepsilon_j \rangle}_{=\delta_{ij}} = \sum_{i=1}^{n} x_i y_i$$

et $||\mathbf{x}||^2 = \sum_{i=1}^{n} x_i^2$. En terme matriciel nous avons $\langle \mathbf{x}, \mathbf{y} \rangle = {}^{\mathrm{t}}XY$ et $||\mathbf{x}||^2 = {}^{\mathrm{t}}XX$, X et Y étant les matrices colonnes ayant pour coefficients les composantes de $\mathbf{x}$ et $\mathbf{y}$ dans une base orthonormale de E.

Ainsi dans un espace euclidien de dimension n muni d'une base orthonormale tout se passe comme dans l'espace euclidien $\mathbb{R}^n$ muni du produit scalaire canonique.

2.4. Réduction simultanée de deux formes quadratiques

Proposition 2.4.1

Soient q et q' deux formes quadratiques sur un $\mathbb{R}$-espace vectoriel E. Si q est définie positive, alors il existe une base $\mathcal{B}$ de E dans laquelle

$$\mathrm{Mat}_{\mathcal{B}}(q) = \mathrm{Id} \qquad \text{et} \qquad \mathrm{Mat}_{\mathcal{B}}(q') \text{ est diagonale}$$

Démonstration. — Soit $\mathcal{B}_0 = (e_i)$ une base de E. opsons

$$M = \mathrm{Mat}_{\mathcal{B}_0}(q), \qquad N = \mathrm{Mat}_{\mathcal{B}_0}(q')$$

Soit $\mathcal{B}_1 = (v_i)$ une base orthonormée pour q et P la matrice de passage de $\mathcal{B}_0$ à $\mathcal{B}_1$. Alors

$$M_1 = {}^{\mathrm{t}}PMP = \mathrm{Id}, \qquad N_1 = {}^{\mathrm{t}}PNP.$$

La matrice N_1 étant symétrique il existe une base de vecteurs propres $\mathcal{B}_2 = (w_i)$ **orthogonaux pour** q d'où

$$M_2 = {}^{\mathrm{t}}P_1 M_1 P_1 = {}^{\mathrm{t}}P_1 \mathrm{Id} P_1 = \begin{bmatrix} ||w_1||^2 & 0 & \dots & 0 \\ 0 & ||w_2||^2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & ||w_n||^2 \end{bmatrix}$$

et

$$N_1 = {}^{\mathrm{t}}P_1 N_1 P_1 = \begin{bmatrix} \lambda_1||w_1||^2 & 0 & \dots & 0 \\ 0 & \lambda_2||w_2||^2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n||w_n||^2 \end{bmatrix}.$$

où P_1 est la matrice de passage de $\mathcal{B}_1$ à $\mathcal{B}_2$ et λ_1, λ_2, …, λ_n sont les valeurs de N_1. Considérons finalement la base $\mathcal{B}_3 = \left(\frac{w_i}{||w_i||^2}\right)$. Notons P_2 la matrice de passage de $\mathcal{B}_2$ à $\mathcal{B}_3$. Nous avons

$$M_3 = {}^{\mathrm{t}}P_2 M_1 P_2 = \mathrm{Id}, \qquad N_3 = {}^{\mathrm{t}}P_2 N_1 P_2 = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{bmatrix}.$$

$\square$

CHAPITRE 3

ISOMÉTRIES ET MATRICES ORTHOGONALES

3.1. Isométries

Dans un espace euclidien E, certains endomorphismes ont la propriété de « conserver » le produit scalaire de l'espace euclidien E, ces endomorphismes sont appelés isométries :

Définition 3.1.1

Soient E un espace euclidien et u un endomorphisme de E. On dit que u est une *isométrie* si

$$\forall\, \mathbf{x},\, \mathbf{y} \in E \quad \langle u(\mathbf{x}),\, u(\mathbf{y})\rangle = \langle \mathbf{x},\, \mathbf{y}\rangle.$$

En choisissant $\mathbf{x} = \mathbf{y}$ on établit aisément qu'une isométrie u est une application qui conserve la norme, *i.e.* que $||u(\mathbf{x})|| = ||\mathbf{x}||$ pour tout $\mathbf{x} \in E$. Nous avons plus précisément le résultat suivant.

Proposition 3.1.1

Soient E un espace euclidien et u un endomorphisme de E. L'endomorphisme u est une isométrie si et seulement si

$$\forall\, \mathbf{x} \in E \quad ||u(\mathbf{x})|| = ||\mathbf{x}||.$$

Remarque 3.1.1. — L'ensemble des isométries d'un espace euclidien E est un groupe, plus précisément c'est un sous-groupe du groupe des isomorphismes de E muni de la loi de composition des endomorphismes $(\mathrm{GL}(E), \circ)$. Le groupe des isométries est noté $\mathrm{O}(E)$.

Démonstration. — Nous avons montré précédemment que toute isométrie conserve la norme.

Montrons la réciproque. Supposons que l'endomorphisme u vérifie

$$||u(\mathbf{x})|| = ||\mathbf{x}|| \tag{3.1.1}$$

et considérons deux vecteurs $\mathbf{x}$ et $\mathbf{y}$ de E. En utilisant la formule de polarisation, nous obtenons

$$\langle u(\mathbf{x}),\, u(\mathbf{y})\rangle = \frac{1}{4}\Big(||u(\mathbf{x}) + u(\mathbf{y})||^2 - ||u(\mathbf{x}) - u(\mathbf{y})||^2\Big).$$

La linéarité de l'application u et l'assertion (3.3.1) indiquent que

$$||u(\mathbf{x}) + u(\mathbf{y})||^2 = ||u(\mathbf{x}+\mathbf{y})||^2 = ||\mathbf{x}+\mathbf{y}||^2, \quad ||u(\mathbf{x}) - u(\mathbf{y})||^2 = ||u(\mathbf{x}-\mathbf{y})||^2 = ||\mathbf{x}-\mathbf{y}||^2.$$

Nous obtenons ainsi

$$\langle u(\mathbf{x}),\, u(\mathbf{y})\rangle = \frac{1}{4}\Big(||\mathbf{x}+\mathbf{y}||^2 - ||\mathbf{x}-\mathbf{y}||^2\Big).$$

La formule de polarisation permet de conclure que $\langle u(\mathbf{x}),\, u(\mathbf{y})\rangle = \langle \mathbf{x},\, \mathbf{y}\rangle$. □

En dimension finie une autre manière de caractériser les isométries est de considérer leur action sur une base orthonormale. Nous avons l'énoncé suivant :

Proposition 3.1.2

Soient E un espace euclidien de dimension finie n et $(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ une base orthonormale de E. Un endomorphisme u de E est une isométrie si et seulement si $(u(\mathbf{e}_1), u(\mathbf{e}_2), \ldots, u(\mathbf{e}_n))$ est une base orthonormale de E.

Démonstration. — Supposons que l'endomorphisme u soit une isométrie et considérons une base orthonormale $(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ et E. D'après la définition pour tous $1 \leq i, j \leq n$, $i \neq j$, nous avons $\langle u(\mathbf{e}_i), u(\mathbf{e}_j)\rangle = \langle \mathbf{e}_i, \mathbf{e}_j\rangle = 0$ et pour tout $1 \leq i \leq n$ nous avons $||u(\mathbf{e}_i)|| = ||\mathbf{e}_i|| = 1$. Nous en déduisons $(u(\mathbf{e}_1), u(\mathbf{e}_2), \ldots, u(\mathbf{e}_n))$ est orthonormale. Puisque toute famille orthonormale est libre, la famille $(u(\mathbf{e}_1), u(\mathbf{e}_2), \ldots, u(\mathbf{e}_n))$ est libre. Étant donné que le cardinal de cette famille est égal à la dimension de E, nous en déduisons que cette famille est une base de E.

Réciproquement supposons que les familles $(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ et $(u(\mathbf{e}_1), u(\mathbf{e}_2), \ldots, u(\mathbf{e}_n))$ soient deux bases orthonormales. Pour montrer que l'endomorphisme u est une isométrie il suffit de montrer que u préserve la norme (Proposition 3.1.1). Considérons un vecteur $\mathbf{x}$ de E dont la décomposition dans la base $(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ est $\mathbf{x} = \sum_{i=1}^{n} x_i\mathbf{e}_i$. L'application u étant linéaire, nous avons

$$u(\mathbf{x}) = \sum_{i=1}^{n} u(x_i\mathbf{e}_i) = \sum_{i=1}^{n} x_i u(\mathbf{e}_i).$$

Cette dernière égalité donne la décomposition du vecteur $u(\mathbf{x})$ dans la base $(u(\mathbf{e}_1),\, u(\mathbf{e}_2),\, \dots,\, u(\mathbf{e}_n))$. Les bases considérées étant toutes les deux orthonormées, nous avons $||\mathbf{x}|| = \sum_{i=1}^{n} x_i^2$ et $||u(\mathbf{x})|| = \sum_{i=1}^{n} x_i^2$ ce qui permet de conclure que $||u(\mathbf{x})|| = ||\mathbf{x}||$. □

3.2. Matrices orthogonales

Rappels sur le lien entre les applications linéaires et les matrices. — Rappelons qu'une application linéaire est entièrement déterminée par les images des vecteurs d'une base. Nous pouvons ainsi caractériser une application linéaire par une matrice.

Définition 3.2.1

Soient E un $\mathbb{R}$-espace vectoriel de dimension n muni d'une base $\mathcal{B}$ et f une application linéaire de E dans E.
Nous appelons *matrice associée à f relativement à la base $\mathcal{B}$* la matrice de $\mathrm{M}(n, \mathbb{R})$ dont la $j$ième colonne est constituée par les composantes de l'image du $j$ème vecteur de la base $\mathcal{B}$ exprimé dans la base $\mathcal{B}$.

Cette définition permet d'affirmer que tout endomorphisme d'un $\mathbb{R}$-espace vectoriel E de dimension finie peut être représenté, relativement à une base, sous forme matricielle. Réciproquement une matrice carrée est toujours la représentation d'une application linéaire. Plus précisément nous avons l'énoncé suivant :

Proposition 3.2.1

Soit A une matrice carrée de taille n à coefficients réels. Il existe une unique application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}^n$ admettant A pour matrice associée relativement à la base canonique. Une telle application est *canoniquement associée* à la matrice A.

La Définition 3.2.1 et la Proposition 3.2.1 permettent d'établir un isomorphisme d'espace vectoriel entre l'ensemble $\mathcal{L}_{\mathbb{R}}(E)$ des endomorphismes d'un $\mathbb{R}$-espace vectoriel E de dimension n et l'ensemble $\mathrm{M}(n, \mathbb{R})$ des matrices carrées de taille n. En dimension finie les résultats concernant les applications linéaires peuvent ainsi toujours se traduire en terme matriciel.

Considérons par exemple un endomorphisme $f \in \mathcal{L}_{\mathbb{R}}(E)$ et un couple de vecteurs $(\mathbf{x}, \mathbf{y}) \in E^2$ tel que $\mathbf{y} = f(\mathbf{x})$. Matriciellement si on désigne par $A \in \mathrm{M}(n, \mathbb{R})$ la matrice de f dans une base

$\mathcal{B}$, X et Y les matrices colonnes composées des coordonnées de $\mathbf{x}$ et $\mathbf{y}$ dans la base $\mathcal{B}$, alors l'égalité $\mathbf{y} = f(\mathbf{x})$ s'écrit $Y = AX$.

Les isométries étant des endomorphismes possédant des propriétés particulières, les matrices associées auront elles aussi des propriétés spécifiques.

Définition 3.2.2

Nous appelons *matrice orthogonale* la matrice d'une isométrie dans une base orthonormale d'un espace euclidien.

Comme pour l'ensemble $\mathrm{O}(E)$ des isométries d'un espace euclidien E, l'ensemble des matrices orthogonales de taille n, noté $\mathrm{O}(n, \mathbb{R})$, possède une structure de groupe. Plus précisément, $\mathrm{O}(n, \mathbb{R})$ est un sous-groupe du groupe des matrices inversibles $(\mathrm{GL}(n, \mathbb{R}), \cdot)$. Il est appelé le *groupe orthogonal*.

Considérons une isométrie u d'un espace euclidien $(E, \langle \cdot, \cdot \rangle)$ de dimension n. Par définition, pour tous $\mathbf{x}$, $\mathbf{y} \in E$ nous avons

$$\langle u(\mathbf{x}), u(\mathbf{y}) \rangle = \langle \mathbf{x}, \mathbf{y} \rangle.$$

Matriciellement cette relation se traduit par

$$\langle UX, UY \rangle_{\mathbb{R}^n} = \langle X, Y \rangle_{\mathbb{R}^n}$$

où les matrices colonnes X et Y correspondent aux coordonnées des vecteurs $\mathbf{x}$ et $\mathbf{y}$ dans une base orthonormale de E, U est la matrice de l'endomorphisme u dans cette même base et $\langle \cdot, \cdot \rangle_{\mathbb{R}^n}$ désigne le produit scalaire canonique sur $\mathbb{R}^n$.

Proposition 3.2.2

Pour tous vecteurs X et Y de $\mathbb{R}^n$ et pour toute matrice $M \in \mathrm{M}(n, \mathbb{R})$ nous avons

$$\langle MX, Y \rangle_{\mathbb{R}^n} = \langle X, {}^{\mathrm{t}}MY \rangle_{\mathbb{R}^n}.$$

Démonstration. — Considérons deux vecteurs X et Y de $\mathbb{R}^n$ et une matrice $M \in \mathrm{M}(n, \mathbb{R})$. Nous savons que $\langle X, Y \rangle_{\mathbb{R}^n} = {}^{\mathrm{t}}XY$ donc

- d'une part $\langle MX, Y \rangle_{\mathbb{R}^n} = {}^{\mathrm{t}}(MX)Y = {}^{\mathrm{t}}X{}^{\mathrm{t}}MY$;
- d'autre part $\langle X, {}^{\mathrm{t}}MY \rangle_{\mathbb{R}^n} = {}^{\mathrm{t}}X{}^{\mathrm{t}}MY$.

Nous en déduisons que $\langle MX, Y \rangle_{\mathbb{R}^n} = \langle X, {}^{\mathrm{t}}MY \rangle_{\mathbb{R}^n}$. $\square$

En pratique étant donnée une matrice $U \in \mathrm{M}(n, \mathbb{R})$ il n'est pas simple, à l'aide de la Définition 3.2.2, de déterminer si cette matrice est orthogonale ou non. Nous allons voir que les matrices orthogonales vérifient certaines propriétés les caractérisant.

Étant donnée une matrice orthogonale U la Proposition 3.2.2 permet d'établir que pour tous X, $Y \in \mathbb{R}^n$ nous avons

$$\langle X, {}^{\mathrm{t}}UUY\rangle_{\mathbb{R}^n} = \langle UX, UY\rangle_{\mathbb{R}^n} = \langle X, Y\rangle_{\mathbb{R}^n}.$$

Le produit scalaire étant une forme bilinéaire définie, nous en déduisons que

$$\forall\, Y \in \mathbb{R}^n \quad {}^{\mathrm{t}}UUY = Y.$$

En choisissant pour Y le $i$ème vecteur $\mathbf{e}_i$ de la base canonique de $\mathbb{R}^n$, nous établissons que la ième colonne de la matrice carrée ${}^{\mathrm{t}}UU$ est égale au vecteur $\mathbf{e}_i$. Autrement dit ${}^{\mathrm{t}}UU$ est égale à la matrice identité

$${}^{\mathrm{t}}UU = \mathrm{Id}.$$

Nous venons ainsi de montrer que toute matrice orthogonale U est une matrice inversible et que sa matrice inverse U^{-1} n'est autre que sa matrice transposée ${}^{\mathrm{t}}U$. L'énoncé suivant montre que cette propriété caractérise les matrices orthogonales.

Proposition 3.2.3

Soit $U \in \mathrm{M}(n, \mathbb{R})$. Les assertions suivantes sont équivalentes :

1. la matrice U est une matrice orthogonale ;
2. la matrice U est inversible et $U^{-1} = {}^{\mathrm{t}}U$;
3. les colonnes (resp. les lignes) de la matrice U forment une base orthonormée pour le produit scalaire canonique de $\mathbb{R}^n$.

Démonstration. — L'implication 1. $\Rightarrow$ 2. a été justifiée précédemment. Nous allons montrer successivement que 2. implique 3. puisque 3. implique 1. ce qui établira les équivalences.

Nous considérons sur $\mathbb{R}^n$ le produit scalaire canonique que nous notons $\langle\cdot,\cdot\rangle$ et nous désignons par $(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ la base canonique de $\mathbb{R}^n$. Cette base est orthonormée pour le produit scalaire canonique, *i.e.* $\langle \mathbf{e}_i, \mathbf{e}_j\rangle = 0$ si $i \neq j$ et $\langle \mathbf{e}_i, \mathbf{e}_i\rangle = 1$.

Supposons que l'assertion 2. soit vraie, *i.e.* supposons que la matrice U soit inversible et $U^{-1} = {}^{\mathrm{t}}U$. En particulier ${}^{\mathrm{t}}UU = \mathrm{Id}$. Les colonnes de la matrice U sont les vecteurs $U\mathbf{e}_i$ où les $\mathbf{e}_i$ correspondent aux vecteurs de la base canonique de $\mathbb{R}^n$. Grâce à la Proposition 3.2.2 nous obtenons pour tout $1 \leq i,\, j \leq n$

$$\langle U\mathbf{e}_i, U\mathbf{e}_j\rangle = \langle \mathbf{e}_i, {}^{\mathrm{t}}UU\mathbf{e}_j\rangle = \langle \mathbf{e}_i, \mathbf{e}_j\rangle = \delta_{ij}.$$

Nous en déduisons que la famille $(U\mathbf{e}_1,\, U\mathbf{e}_2, \ldots, U\mathbf{e}_n)$ est aussi une famille orthonormée ; par conséquent $(U\mathbf{e}_1,\, U\mathbf{e}_2, \ldots, U\mathbf{e}_n)$ est une famille libre. Puisque son cardinal est égal à la dimension de $\mathbb{R}^n$ la famille $(U\mathbf{e}_1,\, U\mathbf{e}_2, \ldots, U\mathbf{e}_n)$ est une base de l'espace vectoriel $\mathbb{R}^n$. Ainsi les colonnes de la matrice U forment une base orthonormale de $\mathbb{R}^n$. Le même raisonnement avec la matrice ${}^t U$ à la place de U montre que les lignes de U forment aussi une base orthonormale de $\mathbb{R}^n$. Nous avons donc établi que 2. entraîne 3.

Supposons que l'assertion 3. soit vraie. D'après la Proposition 3.1.1 montrer que la matrice U est une matrice orthogonale est équivalent à montrer que pour tout vecteur $X \in \mathbb{R}^n$ nous avons $||UX||^2 = ||X||^2$. La décomposition d'un vecteur X de $\mathbb{R}^n$ dans la base canonique $(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ sous la forme $X = \sum_{i=1}^{n} x_i \mathbf{e}_i$ implique par linéarité que

$$UX = U\left(\sum_{i=1}^{n} x_i \mathbf{e}_i\right) = \sum_{i=1}^{n} x_i U(\mathbf{e}_i).$$

Les familles $(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ et $(U\mathbf{e}_1, U\mathbf{e}_2, \ldots, U\mathbf{e}_n)$ étant toutes les deux orthonormales nous en déduisons que $||X||^2 = \sum_{i=1}^{n} x_i^2$ et que $||UX||^2 = \sum_{i=1}^{n} x_i^2$. Nous en concluons que $||X||^2 = ||UX||^2$. □

Inverser une matrice n'est en général pas très « simple », par contre, calculer la transposée d'une matrice est « simple ». Un des avantages des matrices orthogonales est que leurs inverses sont leurs transposées (Proposition 3.2.3).

Exemple 3.2.1. — Considérons la matrice

$$U = \begin{pmatrix} \frac{2}{\sqrt{6}} & 0 & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \end{pmatrix}.$$

Il s'agit d'une matrice orthogonale car ses colonnes C_1, C_2 et C_3 forment une base orthonormée de $\mathbb{R}^3$ pour le produit scalaire canonique puisque

$$\langle C_1, C_2\rangle = 0, \quad \langle C_2, C_3\rangle = 0, \quad \langle C_3, C_1\rangle = 0, \quad ||C_1|| = 1, \quad ||C_2|| = 1, \quad ||C_3|| = 1.$$

La Proposition 3.2.3 assure que la matrice U est une matrice inversible et que son inverse est donné par sa transposée :

$$U^{-1} = \begin{pmatrix} \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{6}} \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{pmatrix}.$$

Nous avons indiqué qu'il y avait une correspondance (un isomorphisme) entre l'espace vectoriel $\mathrm{M}(n, \mathbb{R})$ des matrices carrées de taille n à coefficients réels et l'epace vectoriel $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ des applications linéaires de $\mathbb{R}^n$ dans $\mathbb{R}^n$. Il est donc possible de traduire en terme d'applications linéaires les propriétés liées aux matrices orthogonales.

Définition 3.2.3

Soient u une isométrie d'un espace euclidien E et U la matrice de u dans une base orthonormale de E.

- ◇ Nous appelons *endomorphisme adjoint* de u et nous notons ${}^{\mathrm{t}}u$, l'endomorphisme associé à la matrice ${}^{\mathrm{t}}U$.
- ◇ Un sous-ensemble F de E est *stable par l'endomorphisme* u (ou par la matrice U) si pour tout $\mathbf{x} \in F$ nous avons $u(\mathbf{x}) \in F$.

Remarque 3.2.1. — Nous savons que pour une matrice $A \in \mathrm{M}(n, \mathbb{R})$ nous avons ${}^{\mathrm{t}}({}^{\mathrm{t}}A) = A$. D'après la définition précédente nous en déduisons que pour une isométrie u d'un espace euclidien E, nous avons ${}^{\mathrm{t}}({}^{\mathrm{t}}u) = u$.

Remarque 3.2.2. — La Proposition 3.2.3 peut se traduire en terme d'applications linéaires de la façon suivante : u est une isométrie si et seulement si u est inversible et $u^{-1} = {}^{\mathrm{t}}u$.

De plus, un endomorphisme d'un espace euclidien E est symétrique si ${}^{\mathrm{t}}u = u$. En particulier, un projecteur orthogonal est symétrique. En effet, soient F un sous-espace vectoriel de E et p_F le projecteur orthogonal sur F. Pour tout x dans E nous avons $p_F(x) \in F$ et $(x - p_F(x)) \in F^{\perp}$. Pour tous x et y dans E nous avons

$$\begin{aligned}
\langle p_F(x), y\rangle &= \langle p_F(x), p_F(y) + (y - p_F(y))\rangle \\
&= \langle p_F(x), p_F(y)\rangle \\
&= \langle p_F(x) + (x - p_F(x)), p_F(y)\rangle \\
&= \langle x, p_F(y)\rangle.
\end{aligned}$$

Une projection orthogonale distincte de id_E n'est pas un endomorphisme orthogonal, néanmoins

Proposition 3.2.4

Soient F un sous-espace vectoriel d'un espace euclidien E et p_F la projection orthogonale sur F. Alors $s_F = 2p_F - \mathrm{id}_E$ est un endomorphisme orthogonal symétrique.

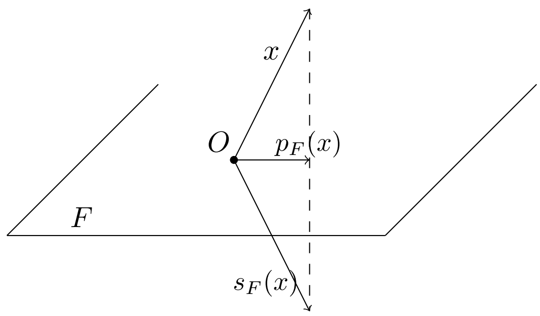
Démonstration. — Nous avons ${}^{\mathrm{t}}s_F = 2{}^{\mathrm{t}}p_F - \mathrm{id}_E = 2p_F - \mathrm{id}_E = s_F$. Par suite s_F est un endomorphisme symétrique.

L'endomorphisme s_F est orthogonal car : ${}^{\mathrm{t}}s_F \circ s_F = (2p_F - \mathrm{id}_E)^2 = 4p_F - 2p_F - 2p_F + \mathrm{id}_E = \mathrm{id}_E$. □

Remarquons que si $x = y + z$ avec $y \in F$ et $z \in F^{\perp}$, alors $s_F(x) = y - z$ d'où la définition suivante :

Définition 3.2.4

Soit F un sous-espace vectoriel d'un espace euclidien E. Soit p_F la projection orthogonale sur F. On appelle *symétrie orthogonale* par rapport à F l'application s_F définie par $s_F = 2p_F - \mathrm{id}_E$.
Si F est un hyperplan, on dit que s_F est une *réflexion* d'hyperplan F.



Nous avons le lien suivant entre les sous-espaces stables par un endomorphisme et ceux stables par l'endomorphisme adjoint.

Proposition 3.2.5

Soient u une isométrie d'un espace euclidien E et F un sous-ensemble de E. Si F est stable par l'endomorphisme u, alors $F^\perp$ est stable par l'endomorphisme adjoint ${}^{\mathrm{t}}u$.

Démonstration. — Considérons une isométrie u de E et un sous-ensemble F de E stable par u. Pour montrer que $F^\perp$ est stable par ${}^{\mathrm{t}}u$ il faut montrer que pour tout $\mathbf{y} \in E$

$$\mathbf{y} \in F^\perp \Rightarrow {}^{\mathrm{t}}u(\mathbf{y}) \in F^\perp.$$

Soit $\mathbf{y}$ un élément de $F^\perp$. Soit $\mathbf{z} \in F$. Matriciellement en notant U la matrice de l'endomorphisme u dans une base donnée de E, Y et Z les vecteurs colonnes représentant $\mathbf{y}$ et $\mathbf{z}$ dans la même base, nous avons

$$\langle {}^{\mathrm{t}}u(\mathbf{y}), \mathbf{z}\rangle = \langle {}^{\mathrm{t}}UY, Z\rangle_{\mathbb{R}^n}$$

En vertu des propriétés du produit scalaire vis-à-vis de la transposition (Proposition 3.2.2), nous en déduisons que

$$\langle {}^{\mathrm{t}}u(\mathbf{y}), \mathbf{z}\rangle = \langle Y, UZ\rangle_{\mathbb{R}^n} = \langle \mathbf{y}, u(\mathbf{z})\rangle.$$

Puisque F est stable par u et que $\mathbf{z}$ appartient à F, nous avons $u(\mathbf{z})$ appartient à F. De plus, comme $\mathbf{y}$ appartient à $F^{\perp}$, nous en déduisons que

$$\langle {}^{\mathrm{t}}u(\mathbf{y}), \mathbf{z}\rangle = \langle \mathbf{y}, u(\mathbf{z})\rangle = 0.$$

ce qui prouve que ${}^{\mathrm{t}}u(\mathbf{y})$ appartient à $F^{\perp}$. □

Proposition 3.2.6

Le déterminant d'une matrice orthogonale vaut 1 ou -1.

Démonstration. — Considérons une matrice orthogonale U. Puisque ${}^{\mathrm{t}}UU = \mathrm{Id}$ nous avons $\det({}^{\mathrm{t}}UU) = \det({}^{\mathrm{t}}U)\det(U) = 1$. Comme le déterminant d'une matrice est égale au déterminant de sa transposée nous obtenons $(\det U)^2 = 1$. la matrice U étant réelle son déterminant est un nombre réel ; nous en déduisons que $\det U = 1$ ou $\det U = -1$. □

Remarque 3.2.3. — L'ensemble des matrices orthogonales de déterminant 1 est noté $\mathrm{SO}(n,\mathbb{R})$: c'est un sous-groupe de $(\mathrm{O}(n,\mathbb{R}),\cdot)$ qui est appelé *groupe spécial orthogonal*.

Cas de la dimension 2. — La troisième assertion de la Proposition 3.2.3 indique que toutes les isométries du plan $\mathbb{R}^2$ s'écrivent matriciellement sous la forme

$$U = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \text{ avec } \begin{cases} a^2 + b^2 = 1 \\ c^2 + d^2 = 1 \\ ab + cd = 0 \end{cases}$$

Les deux premières égalités indiquent qu'il existe ϑ et ϕ dans $[0, 2\pi[$ tels que $a = \cos\vartheta$ et $b = \sin\vartheta$, $c = \cos\phi$ et $d = \sin\phi$. En utilisant la dernière égalité nous avons $\phi \equiv_{2\pi} \vartheta \pm \frac{\pi}{2}$. La matrice U est donc de l'une des deux formes suivantes

$$R_\vartheta = \begin{pmatrix} \cos\vartheta & -\sin\vartheta \\ \sin\vartheta & \cos\vartheta \end{pmatrix} \qquad S_\vartheta = \begin{pmatrix} \cos\vartheta & \sin\vartheta \\ \sin\vartheta & -\cos\vartheta \end{pmatrix}$$

Dans le premier cas nous avons $\det R_\vartheta = 1$ alors que dans le second cas nous avons $\det S_\vartheta = -1$. Notons que les matrices S_ϑ possèdent toujours deux valeurs propres réelles distinctes 1 et -1. La matrice S_ϑ est donc diagonalisable et s'écrit dans une base de vecteurs propres

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

3.3. Description géométrique des isométries

Nous disposons désormais de deux points de vue concernant les isométries : le point de vue des applications linéaires où une isométrie est une application linéaire préservant le produit scalaire et le point de vue matriciel où la matrice d'une isométrie est une matrice orthogonale. Nous allons nous intéresser au point de vue géométrique, tout d'abord dans le cas du plan $\mathbb{R}^2$ puis dans le cas plus général d'un espace vectoriel de dimension finie.

Du point de vue des applications linéaires la Proposition 3.2.6 permet de distinguer les isométries dont le déterminant est 1 de celles dont le déterminant est -1.

Définition 3.3.1

Soit u une isométrie d'un espace euclidien.

- Une isométrie d'un espace euclidien est une *rotation* (ou une *isométrie directe*) si son déterminant vaut 1.
- Une isométrie est dite *indirecte* si son déterminant vaut -1.

Contrairement à l'ensemble des isométries directes, l'ensemble des isométries indirectes ne possède pas de structure de groupe. En effet, si deux matrices ont pour déterminant -1 alors leur produit aura pour déterminant 1 : le produit de deux isométries indirectes n'est pas une isométrie indirecte.

Cas de la dimension $n = 2$. — On déduit de ce qui précède :

Théorème 3.3.1

Dans une base bien choisie, les rotations du plan ont pour matrice

$$R_\vartheta = \begin{pmatrix} \cos\vartheta & -\sin\vartheta \\ \sin\vartheta & \cos\vartheta \end{pmatrix}, \quad \vartheta \in [0, 2\pi[$$

et les isométries indirectes du plan ont pour matrice

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

Géométriquement les isométries indirectes correspondent à des symétries par rapport à une droite vectorielle. Cette droite peut être déterminée en trouvant l'espace propre associé à la valeur propre 1.

Exemple 3.3.1. — Considérons l'endomorphisme u de l'espace vectoriel $\mathbb{R}^2$ dont la matrice dans la base canonique $(\mathbf{e}_1, \mathbf{e}_2)$ de $\mathbb{R}^2$ est

$$A = \mathrm{Mat}_{(\mathbf{e}_1,\mathbf{e}_2)}(u) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

Les deux colonnes de A forment une base orthonormale de $\mathbb{R}^2$, ainsi A est orthogonale. Par conséquent l'endomorphisme u est une isométrie. Puisque le déterminant de A vaut -1 il s'agit d'une isométrie indirecte.

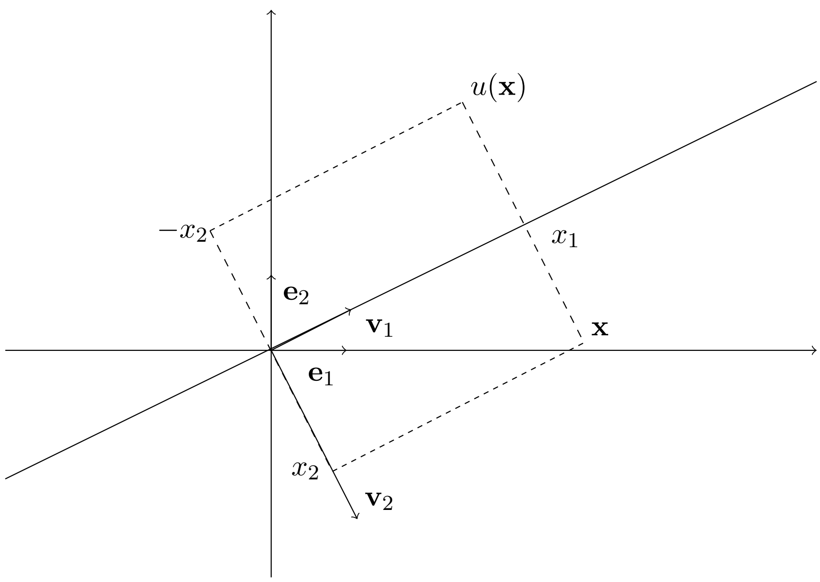
Cherchons une base de $\mathbb{R}^2$ dans laquelle la matrice de u est égale à la matrice $S = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. Pour cela nous diagonalisons la matrice A. Un vecteur propre associé à la valeur propre 1 est le vecteur $\mathrm{v}_1 = (1, \sqrt{2} - 1)$ et un vecteur propre associé à la valeur propre -1 est le vecteur $\mathrm{v}_2 = (1, -\sqrt{2} - 1)$. Ainsi dans la base $(\mathrm{v}_1, \mathrm{v}_2)$ la matrice de l'endomorphisme u vaut S :

$$\mathrm{Mat}_{(\mathrm{v}_1,\mathrm{v}_2)}(u) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Si nous désignons par x_1 et x_2 les deux composantes d'un vecteur $\mathbf{x} \in \mathbb{R}^2$ dans la base $(\mathrm{v}_1, \mathrm{v}_2)$, alors nous avons

$$u(\mathbf{x}) = u(x_1\mathrm{v}_1 + x_2\mathrm{v}_2) = x_1 u(\mathrm{v}_1) + x_2 u(\mathrm{v}_2) = x_1\mathrm{v}_1 - x_2\mathrm{v}_2.$$

L'application u est la symétrie par rapport à la droite vectorielle dirigée par le vecteur v_1.



Symétrie par rapport à une droite vectorielle

Cas de la dimension $n > 2$. — Dans le cas d'un espace euclidien de dimension $n > 2$ nous avons l'énoncé suivant :

Théorème 3.3.2

Soit u une isométrie d'un espace euclidien E. Il existe une base orthonormée de E dans laquelle la matrice u est de la forme

$$U = \begin{pmatrix} \varepsilon_1 & 0 & 0 & \dots & \dots & \dots & \dots & 0 \\ 0 & \varepsilon_2 & 0 & \dots & \dots & \dots & \dots & 0 \\ 0 & 0 & \ddots & \ddots & & & & \vdots \\ 0 & 0 & \ddots & \varepsilon_r & \ddots & & & \vdots \\ \vdots & 0 & \ddots & \ddots & R_{\vartheta_1} & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \ddots & R_{\vartheta_2} & \ddots & \vdots \\ \vdots & \vdots & & & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & \dots & \dots & \dots & 0 & 0 & R_{\vartheta_s} \end{pmatrix}$$

avec $r \in \mathbb{N}$, $s \in \mathbb{N}$ et où pour tout $1 \leq i \leq r$, $\varepsilon_i \in \{-1, 1\}$, et pour tout $1 \leq j \leq s$, $\vartheta_j \in [0, 2\pi[$.

Remarque 3.3.1. — Le Théorème 3.3.2 n'exclut pas d'avoir $r = 0$ ou $s = 0$. Autrement dit, une isométrie d'un espace euclidien peut avoir pour matrice

$$\begin{pmatrix} \varepsilon_1 & & & \\ & \varepsilon_2 & & \\ & & \ddots & \\ & & & \varepsilon_r \end{pmatrix}, \qquad \begin{pmatrix} R_{\vartheta_1} & & & \\ & R_{\vartheta_2} & & \\ & & \ddots & \\ & & & R_{\vartheta_s} \end{pmatrix}.$$

Notons que le second cas ne peut se produire que si la dimension de E est paire.

Démonstration. — La démonstration se fait par récurrence sur la dimension n de l'espace euclidien E.

- ◇ Dans le cas $n = 1$ le résultat est immédiat puisque les seules isométries de $\mathbb{R}$ sont les applications

  $$\mathbb{R} \to \mathbb{R},\ x \mapsto x, \qquad \mathbb{R} \to \mathbb{R},\ x \mapsto -x$$

  dont les matrices dans une base de $\mathbb{R}$ sont respectivement (1) et (-1).
- ◇ Supposons que l'énoncé soit vrai pour tout espace euclidien de dimension inférieure strictement à n. Considérons une isométrie u d'un espace euclidien de dimension n. On doit distinguer deux cas.

- Si u possède au moins une valeur propre réelle alors nous notons $\mathbf{v}$ un vecteur propre associé à cette valeur propre. Nous remarquons que la droite $\mathbb{R}\mathbf{v}$ est stable par u. La Proposition 3.2.5 assure que le sous-espace $H = (\mathbb{R}\mathbf{v})^\perp$ est stable par ${}^{\mathrm{t}}u$. La dimension de H étant strictement inférieure à n nous appliquons l'hypothèse de récurrence à l'endomorphisme ${}^{\mathrm{t}}u_{|H} \colon H \to H$.
- Si u ne possède pas de valeur propre réelle, alors nous exhibons un plan vectoriel P stable par u en considérant l'endomorphisme $f = u + u^{-1}$. Puisque u est une isométrie, ${}^{\mathrm{t}}u = u^{-1}$ et la matrice associée à l'endomorphisme f est une matrice symétrique. Tout endomorphisme symétrique admet une valeur propre réelle (Proposition 4.0.1). Soit $\mathbf{v}$ un vecteur propre de f associé à cette valeur propre λ. Nous avons

$$(u + u^{-1})(\mathbf{v}) = u(\mathbf{v}) + u^{-1}(\mathbf{v}) = \lambda\mathbf{v}. \tag{3.3.1}$$

  Posons $\mathbf{w} = u(\mathbf{v})$. En appliquant l'endomorphisme u à (3.3.1) nous obtenons $u(\mathbf{w}) + \mathbf{v} = \lambda\mathbf{w}$. Nous avons donc $u(\mathbf{v}) = \mathbf{w}$ et $u(\mathbf{w}) = \lambda\mathbf{w} - \mathbf{v}$. Nous en déduisons que le plan vectoriel P engendré par les vecteurs $\mathbf{v}$ et $\mathbf{w}$ est stable par u. L'endomorphisme $u_{|P} \colon P \to P$ restreint au plan P, est une isométrie. D'après l'étude des isométries en dimension 2, l'endomorphisme $u_{|P}$ est une rotation. D'après la Proposition 3.2.5 l'endomorphisme ${}^{\mathrm{t}}u$ est stable par $P^\perp$ et la dimension de $P^\perp$ est strictement inférieure à n. Nous pouvons donc appliquer l'hypothèse de récurrence à l'endomorphisme ${}^{\mathrm{t}}u \colon P^\perp \to P^\perp$.

□

Cas de la dimension $n = 3$. — Si E est un espace de dimension 3, le Théorème 3.3.2 prend la forme suivante :

Théorème 3.3.3

Soit u une isométrie d'un espace euclidien E de dimension 3. Il existe une base orthonormée de E dans laquelle la matrice de u a l'une des formes suivantes :

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\vartheta & -\sin\vartheta \\ 0 & \sin\vartheta & \cos\vartheta \end{pmatrix}, \qquad \begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos\vartheta & -\sin\vartheta \\ 0 & \sin\vartheta & \cos\vartheta \end{pmatrix}$$

avec $\vartheta \in [0, 2\pi[$.

Le premier cas correspond aux rotations (le déterminant de la matrice vaut 1). Le second cas correspond aux *anti-rotations* ; nous distinguons en particulier les deux sous-cas suivants

◇ $\vartheta = 0$, c'est-à-dire $\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, correspondant aux symétries orthogonales par rapport à un plan,

◇ et $\vartheta = \pi$, c'est-à-dire $\begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$, correspondant à la symétrie centrale.

En pratique, lorsque l'on dispose d'une rotation en dimension 3, *i.e.* d'une isométrie de $\mathbb{R}^3$ de déterminant 1, on ne dispose pas toujours d'une base dans laquelle cette rotation a l'expression donnée par le Théorème 3.3.3. Néanmoins nous pouvons déterminer l'angle de cette rotation (c'est-à-dire la valeur de ϑ du Théorème 3.3.3) en calculant la trace de l'isométrie : nous avons $\mathrm{tr}(u) = 2\cos\vartheta + 1$ ce qui permet d'obtenir la valeur de $\cos\vartheta$. Nous déterminons l'angle ϑ à l'aide du signe de $\sin\vartheta$, celui-ci étant donné par le signe du déterminant de la famille de trois vecteurs $(\mathbf{v}, \mathbf{x}, u(\mathbf{x}))$ où $\mathbf{v}$ est le vecteur propre de u associé à la valeur propre 1 et $\mathbf{x}$ un vecteur non colinéaire à l'axe de la rotation.

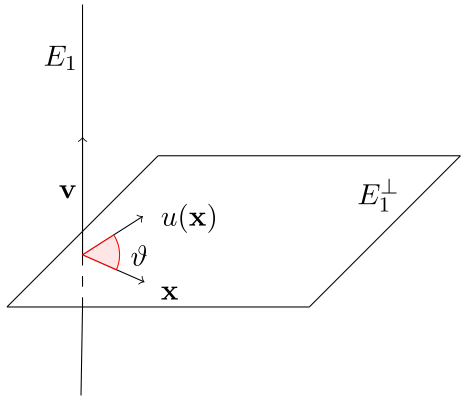
Exemple 3.3.2. — Considérons la matrice

$$U = \begin{pmatrix} \frac{2}{\sqrt{6}} & 0 & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \end{pmatrix}.$$

La matrice U est orthogonale (Exemple 3.2.1). De plus $\det U = 1$. Ainsi l'endomorphisme f_U de $\mathbb{R}^3$, dont la matrice dans la base canonique de $\mathbb{R}^3$ est U, est une rotation.

L'axe de cette rotation est l'ensemble des vecteurs invariants par cette rotation, autrement dit l'axe de la rotation correspond à l'espace propre associé à la valeur propre 1 (un vecteur propre $\mathbf{v}$ associé à la valeur propre 1 vérifie $U\mathbf{v} = \mathbf{v}$). En résolvant le système linéaire $U\mathbf{y} = \mathbf{y}$, nous obtenons que ce vecteur propre est

$$\mathbf{v} = \begin{pmatrix} 1 \\ (\sqrt{2}+1)(\sqrt{3}-\sqrt{2}) \\ \sqrt{3}-\sqrt{2} \end{pmatrix}.$$



Pour déterminer l'angle ϑ de la rotation nous utilisons la trace de la matrice U. Nous avons

$$\operatorname{tr}(U) = \frac{1+\sqrt{2}+\sqrt{3}}{\sqrt{6}} = 2\cos\vartheta + 1$$

d'où $\cos\vartheta = \frac{2+\sqrt{2}+\sqrt{3}-\sqrt{6}}{2\sqrt{6}}$. Puisque le vecteur $\mathbf{x} = (1,0,0)$ n'est pas colinéaire à $\mathbf{v}$, le signe de $\sin\vartheta$ est celui du déterminant de la famille $(\mathbf{v}, \mathbf{x}, u(\mathbf{x}))$. Or

$$\det(\mathbf{v}, \mathbf{x}, u(\mathbf{x})) = \begin{vmatrix} 1 & 1 & \frac{2}{\sqrt{6}} \\ (\sqrt{2}+1)(\sqrt{3}-\sqrt{2}) & 0 & \frac{1}{\sqrt{6}} \\ \sqrt{3}-\sqrt{2} & 0 & -\frac{1}{\sqrt{6}} \end{vmatrix} = \frac{1+\sqrt{2}}{\sqrt{3}}(\sqrt{3}-\sqrt{2}).$$

Ce déterminant étant positif, nous avons l'inégalité $\sin\vartheta > 0$; par suite ϑ appartient à $]0,\pi[$ et la relation $\cos\vartheta = \frac{2+\sqrt{2}+\sqrt{3}-\sqrt{6}}{2\sqrt{6}}$ implique $\vartheta = \arccos\left(\frac{2+\sqrt{2}+\sqrt{3}-\sqrt{6}}{2\sqrt{6}}\right)$. Finalement f_U est la rotation de $\mathbb{R}^3$ d'axe orienté dirigé par $\mathbf{v} = \left(1, (\sqrt{2}+1)(\sqrt{3}-\sqrt{2}), \sqrt{3}-\sqrt{2}\right)$ et d'angle $\arccos\left(\frac{2+\sqrt{2}+\sqrt{3}-\sqrt{6}}{2\sqrt{6}}\right)$.

Exemple 3.3.3. — Considérons la matrice $B = \frac{1}{3}\begin{pmatrix} 2 & 2 & 1 \\ -2 & 1 & 2 \\ -1 & 2 & -2 \end{pmatrix}$.

Les trois colonnes de B forment une base orthonormée de $\mathbb{R}^3$. De plus $\det B = -1$. Par suite l'endomorphisme f_B, dont la matrice dans la base canonique de $\mathbb{R}^3$ est B, est une anti-rotation de $\mathbb{R}^3$.

Pour déterminer l'axe de cette anti-rotation on calcule le sous-espace propre $E_{-1} = \ker(B + \mathrm{Id})$ associé à la valeur propre -1. Nous obtenons $E_{-1} = \operatorname{Vect}(\mathbf{v})$ où $\mathbf{v} = \frac{1}{\sqrt{5}}(0,1,-2)$.

L'angle ϑ de cette anti-rotation satisfait $-1+2\cos\vartheta = \mathrm{tr}(B) = \frac{1}{3}$. Par suite $\cos\vartheta = \frac{2}{3}$ et $\vartheta \equiv_{2\pi} \pm\arccos\left(\frac{2}{3}\right)$. Attention les vecteurs de l'axe ne sont désormais pas fixes mais envoyés sur leur opposé. Néanmoins une anti-rotation α est la composée d'une rotation ρ avec la réflexion s dans le plan de rotation : $\alpha = s\circ\rho$ (on peut aussi écrire $\rho\circ s$ car ρ et s commutent comme on s'en assure sur E_1 et $E_1^{\perp}$). Étant donné que ρ et α ont même angle, nous sommes ramenés, pour l'estimation du signe de $\sin\vartheta$ à l'étude de la rotation ρ. Nous obtenons $\sin\vartheta > 0$ et $\vartheta \equiv_{2\pi} \arccos\frac{2}{3}$. Il en résulte que f_B est la composée de la rotation d'angle $\vartheta = \arccos\frac{2}{3}$ autour de l'axe engendré par $\mathbf{v} = \frac{1}{\sqrt{5}}(0,1,-2)$ et de la symétrie orthogonale par rapport au plan vectoriel orthogonal à $\mathbf{v}$ qui est $\{(x,y,z)\in\mathbb{R}^3 \mid y-2z=0\}$.

Exemple 3.3.4. — Considérons la matrice $C = \frac{1}{3}\begin{pmatrix} 2 & 2 & 1 \\ 2 & -1 & -2 \\ 1 & -2 & 2 \end{pmatrix}$.

Les trois colonnes de C forment une base orthonormée de $\mathbb{R}^3$. De plus $\det C = -1$. Par conséquent l'endomorphisme f_C, dont la matrice dans la base canonique de $\mathbb{R}^3$ est C, est une isométrie indirecte de $\mathbb{R}^3$. Étant donné que $\mathrm{tr}(C) = -1+2\cos\vartheta = 1$ nous obtenons $\cos\vartheta = 1$ et $\vartheta \equiv_{2\pi} 0$. Ainsi f_C est une symétrie orthogonale.

Pour trouver l'axe de cette symétrie, qui est l'orthogonal du plan de la symétrie, nous déterminons $E_{-1} = \ker(C+\mathrm{Id})$. Un calcul montre que $E_{-1} = \mathrm{Vect}(\mathbf{v})$ avec $\mathbf{v} = \frac{1}{\sqrt{6}}(-1,2,1)$. Finalement f_C est la symétrie orthogonale par rapport au plan

$$\left(\mathrm{Vect}(\mathbf{v})\right)^{\perp} = \{(x,y,z)\in\mathbb{R}^3 \mid -x+2y+z=0\}.$$

CHAPITRE 4

DIAGONALISATION DES MATRICES SYMÉTRIQUES RÉELLES

Rappel sur la diagonalisation. —

De manière générale, diagonaliser une matrice $A \in \mathrm{M}(n, \mathbb{R})$, c'est trouver une matrice semblable à A qui soit diagonale. Plus précisément, une matrice A est *diagonalisable* s'il existe une matrice inversible $P \in \mathrm{GL}(n, \mathbb{R})$ et une matrice diagonale $D \in \mathrm{M}(n, \mathbb{R})$ telles que $D = P^{-1}AP$. *Diagonaliser* A c'est trouver D et P.

En terme d'application linéaire un endomorphisme f d'un $\mathbb{R}$-espace vectoriel E de dimension finie est *diagonalisable* sur $\mathbb{R}$ s'il existe une base de E formée de vecteurs propres de f. *Diagonaliser* f c'est trouver une telle base.

Tous les endomorphismes ne sont pas diagonalisables. En effet, un endomorphisme d'un $\mathbb{R}$-espace vectoriel E est diagonalisable si et seulement si les deux conditions suivantes sont satisfaites :

- ◇ le nombre de ses valeurs propres (comptées avec multiplicité) est égal à la dimension de E ;
- ◇ la dimension de chacun des sous-espaces propres est égale à la multiplicité de la valeur propre correspondante.

Il existe des endomorphismes pour lesquels l'une ou l'autre de ces deux conditions (ou les deux à la fois) n'est pas vérifiée. En particulier, on sait que le nombre de valeurs propres correspond au nombre de racines réelles du polynôme caractéristique. Ce polynôme étant de degré égal à n (la dimension de E), le nombre de racines ne peut excéder n mais rien n'affirme que ce polynôme a exactement n racines réelles.

Nous allons montrer dans ce Chapitre que toutes les matrices symétriques à coefficients réels sont diagonalisables. De plus, nous établirons que les matrices de passage P qui permettent de diagonaliser les matrices symétriques réelles peuvent être choisies orthogonales. Commençons par étudier quelques propriétés des matrices symétriques réelles.

Proposition 4.0.1

Toutes les valeurs propres d'une matrice symétrique à coefficients réels sont réelles.

Démonstration. — Toute matrice à coefficients réels admet au moins une valeur propre dans $\mathbb{C}$. Considérons $\lambda \in \mathbb{C}$ une valeur propre complexe d'une matrice symétrique réelle $M \in \mathrm{M}(n, \mathbb{R})$. Soit $\mathbf{v} \in \mathbb{C}^n \setminus \{\mathbf{0}\}$ un vecteur propre associé à la valeur propre λ. Nous avons donc

$$M\mathbf{v} = \lambda\mathbf{v}. \tag{4.0.1}$$

En multipliant (4.0.1) par ${}^{\mathrm{t}}\overline{\mathbf{v}}$ nous obtenons

$${}^{\mathrm{t}}\overline{\mathbf{v}}M\mathbf{v} = \lambda{}^{\mathrm{t}}\overline{\mathbf{v}}\mathbf{v}. \tag{4.0.2}$$

D'autre part en multipliant l'égalité (4.0.1) nous avons ${}^{\mathrm{t}}\mathbf{v}{}^{\mathrm{t}}M = \lambda{}^{\mathrm{t}}\mathbf{v}$. Puisque la matrice M est symétrique, nous en déduisons que ${}^{\mathrm{t}}\mathbf{v}M = \lambda{}^{\mathrm{t}}\mathbf{v}$. En prenant le conjugué de cette égalité nous avons ${}^{\mathrm{t}}\overline{\mathbf{v}}\overline{M} = \overline{\lambda}{}^{\mathrm{t}}\overline{\mathbf{v}}$. La matrice M étant réelle, nous obtenons ${}^{\mathrm{t}}\overline{\mathbf{v}}M = \overline{\lambda}{}^{\mathrm{t}}\overline{\mathbf{v}}$. Enfin, en multipliant cette égalité par $\mathbf{v}$ à droite, nous avons

$${}^{\mathrm{t}}\overline{\mathbf{v}}M\mathbf{v} = \overline{\lambda}{}^{\mathrm{t}}\overline{\mathbf{v}}\mathbf{v}. \tag{4.0.3}$$

À partir de (4.0.2) et (4.0.3) nous obtenons $\lambda{}^{\mathrm{t}}\overline{\mathbf{v}}\mathbf{v} = \overline{\lambda}{}^{\mathrm{t}}\overline{\mathbf{v}}\mathbf{v}$. Puisque $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ est non nul, nous avons ${}^{\mathrm{t}}\overline{\mathbf{v}}\mathbf{v} = |v_1|^2 + |v_2|^2 + \ldots + |v_n|^2 \neq 0$. Nous en déduisons que $\lambda = \overline{\lambda}$, autrement dit que λ est réel. □

Proposition 4.0.2

Les sous-espaces propres d'une matrice symétrique à coefficients réels sont deux à deux orthogonaux.

Démonstration. — Soient λ et μ deux valeurs propres distinctes de la matrice M. La Proposition 4.0.1 assure que ces valeurs propres sont réelles. Considérons deux sous-espaces propres E_λ et E_μ associés respectivement aux deux valeurs λ et μ. Pour montrer que E_λ est orthogonal à E_μ il faut montrer que pour tous $X \in E_\lambda$ et $Y \in E_\mu$ nous avons $\langle X, Y\rangle_{\mathbb{R}^n} = 0$, c'est-à-dire ${}^{\mathrm{t}}YX = 0$.

La matrice M étant symétrique, nous avons ${}^{\mathrm{t}}YMX = {}^{\mathrm{t}}({}^{\mathrm{t}}XMY)$. Or, par définition d'un sous-espace propre, nous avons $MX = \lambda X$ et $MY = \mu Y$. Cela implique que $\lambda{}^{\mathrm{t}}YX = \mu{}^{\mathrm{t}}({}^{\mathrm{t}}XY)$, c'est-à-dire $(\lambda - \mu){}^{\mathrm{t}}YX = 0$. Comme $\lambda \neq \mu$, nous obtenons que ${}^{\mathrm{t}}YX = 0$. □

Théorème 4.0.1

Toute matrice symétrique réelle est diagonalisable dans une base orthonormale. Autrement dit, si M appartient à $\mathrm{S}(n, \mathbb{R})$, alors il existe une matrice orthogonale $P \in \mathrm{O}(n, \mathbb{R})$ telle que $P^{-1}MP$ soit une matrice diagonale réelle.

⚠ Ne pas confondre ce résultat avec la Proposition 1.7.3 qui assure que toute matrice symétrique réelle est congrue à une matrice diagonale. Ici le résultat est plus fort : le Théorème 4.0.1 indique que la matrice de passage est orthogonale.

Démonstration. — On procède par récurrence sur la taille n de la matrice M.

Si $n = 1$, alors M est diagonale et le résultat est vérifié.

Supposons que l'énoncé est démontré pour toutes les matrices symétriques réelles de taille $n-1$ et montrons qu'il est vrai pour toutes les matrices symétriques réelles de taille n. Soit M une matrice symétrique réelle de taille n. La Proposition 4.0.1 assure que M possède au moins une valeur propre réelle. Notons $\lambda_1 \in \mathbb{R}$ cette valeur propre et $\mathbf{v} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ un vecteur propre associé. Le vecteur $\mathbf{e}_1 = \frac{\mathbf{v}}{||\mathbf{v}||}$ est aussi un vecteur propre associé à λ_1 ; il est de norme 1 et la droite $\mathbb{R}\mathbf{e}_1$ est stable par M. D'après la Proposition 3.2.5 nous en déduisons que $H = (\mathbb{R}\mathbf{e}_1)^\perp$ est stable par ${}^{\mathrm{t}}M$. La matrice M étant symétrique, il s'ensuit que l'espace vectoriel H est stable par M.

Désignons par u l'endomorphisme associé à la matrice M. Nous avons montré que u est un endomorphisme de H sur H. Puisque $\dim H = n-1$, nous pouvons appliquer l'hypothèse de récurrence à la matrice de l'endomorphisme $u \colon H \to H$. Il existe une base orthonormale $(\mathbf{e}_2, \mathbf{e}_3, \ldots, \mathbf{e}_n)$ de H dans laquelle $u_{|H}$ est diagonale ; autrement dit pour tout $2 \leq i \leq n$ nous avons $u(\mathbf{e}_i) = \lambda_i \mathbf{e}_i$, $\lambda_i \in \mathbb{R}$. En complétant cette base avec le vecteur $\mathbf{e}_1$ nous obtenons une base de $\mathbb{R}^n$ qui convient puisque d'une part pour tout $1 \leq i \leq n$ il existe $\lambda_i \in \mathbb{R}$ tel que $u(\mathbf{e}_i) = \lambda_i \mathbf{e}_i$ ce qui exprime le fait que dans cette base $(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ la matrice de u est diagonale, et d'autre part la base $(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ est une base orthonormale puisque la famille $(\mathbf{e}_2, \mathbf{e}_3, \ldots, \mathbf{e}_n)$ est orthogonale et que le vecteur $\mathbf{e}_1$ de norme 1 et orthogonale à H, donc à la famille $(\mathbf{e}_2, \mathbf{e}_3, \ldots, \mathbf{e}_n)$. □

Exemple 4.0.1. — La matrice $A = \begin{pmatrix} 16 & -12 \\ -12 & 9 \end{pmatrix}$ est une matrice symétrique réelle, elle est par conséquent diagonalisable dans une base orthonormale.

Pour diagonaliser la matrice A, commençons par chercher ses valeurs propres.

Le polynôme caractéristique P de A est donné par

$$P(\lambda) = \det(A - \lambda \mathrm{Id}) = \begin{vmatrix} 16-\lambda & -12 \\ -12 & 9-\lambda \end{vmatrix} = \lambda(\lambda - 25).$$

Les valeurs propres de A sont $\lambda_1 = 0$ et $\lambda_2 = 25$. Pour déterminer la matrice de passage orthogonale calculons les espaces propres associés à λ_1 et λ_2. L'espace propre associé à λ_1 est l'ensemble des vecteurs $\mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix}$ tels que $A\mathbf{x} = \lambda_1 \mathbf{x}$, *i.e.* tels que $16x - 12y = 0$. L'espace propre associé à la valeur propre λ_1 est donc engendré par $\begin{pmatrix} 3 \\ 4 \end{pmatrix}$. Afin d'obtenir une matrice de passage orthogonale considérons un vecteur colinéaire à $\mathbf{v}_1$ de norme 1 :

$$\mathbf{e}_1 = \frac{\mathbf{v}_1}{||\mathbf{v}_1||} = \begin{pmatrix} \frac{3}{5} \\ \frac{4}{5} \end{pmatrix}$$

Pour calculer le second vecteur propre, nous pouvons :

- ◇ soit procéder comme pour le premier vecteur propre,
- ◇ soit utiliser le fait que la matrice de passage (c'est-à-dire la matrice formée par les vecteurs propres associés aux deux valeurs propres) doit être orthogonale.

En utilisant cette seconde méthode nous obtenons que

$$\mathbf{e}_2 = \begin{pmatrix} -\frac{4}{5} \\ \frac{3}{5} \end{pmatrix}$$

est un vecteur orthogonal à $\mathbf{v}_1$ et de norme 1. La matrice A s'écrit $A = PD{}^{\mathrm{t}}P$ où

$$P = \begin{pmatrix} \frac{3}{5} & -\frac{4}{5} \\ \frac{4}{5} & \frac{3}{5} \end{pmatrix}, \qquad D = \begin{pmatrix} 0 & 0 \\ 0 & 25 \end{pmatrix}$$

INDEX