Licence de Sciences et Technologies

Mathématiques

Unité SC L4 MT 02

ALGEBRE III

Jean-François Havet

Université d'Orléans, Département de Mathématiques
B.P. 6759, 45067 ORLEANS Cedex 2, France

Janvier 2008

Table des matières

1 Dual d'un espace vectoriel

Dans toute cette partie K désignera un corps commutatif.

1.1. Rappels et notations

- Soient F et G des sous-espaces vectoriels d'un K-espace vectoriel E. On dit que G est un *supplémentaire* de F dans E si $E = F \oplus G$.

- *Dans un espace vectoriel (de dimension finie), tout sous-espace vectoriel admet au moins un supplémentaire.*

- Soit E un espace vectoriel de dimension finie n et $\mathcal{B} = (e_1, \ldots, e_n)$ une base de E. Tout vecteur $x \in E$ se décompose de manière unique sur la base $\mathcal{B}$, en $x = \sum_{i=1}^{n} \lambda_i e_i$ et on peut considérer la matrice-colonne des coordonnées de x dans la base $\mathcal{B}$: on note $\operatorname{Mat}_{\mathcal{B}}(x) = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} \in \boldsymbol{M}_{n,1}(K)$.

- Soit $n \in \mathbb{N}^*$ **fixé**. Nous désignerons par $(\varepsilon_1, \ldots, \varepsilon_n)$, la *base canonique* de K^n. Rappelons que la famille $(\varepsilon_j)_{j=1\cdots n}$ est définie par :
  $\varepsilon_1 = (1, 0, \ldots, 0)$, $\varepsilon_2 = (0, 1, 0, \ldots, 0)$, $\varepsilon_j = (0, \ldots, 0, 1, 0, \ldots, 0)$ le 1 étant à la $j^{\text{ème}}$ place, $\varepsilon_n = (0, \ldots, 0, 1)$.
  On a donc $(x_1, x_2, \ldots, x_n) = x_1\varepsilon_1 + x_2\varepsilon_2 + \cdots + x_n\varepsilon_n = \sum_{i=1}^{n} x_i\varepsilon_i$.
  Le lecteur remarquera que cette notation peut devenir ambiguë si on ne précise pas clairement l'entier n, c'est à dire l'espace vectoriel dans lequel on travaille : par exemple ε_1 représente $(1, 0)$ dans K^2 et représente $(1, 0, 0, 0)$ dans K^4.

- Soient m et n **fixés** dans $\mathbb{N}^*$. Nous désignerons par $(E_{ij})_{(i,j)\in[1,m]_{\mathbb{N}}\times[1,n]_{\mathbb{N}}}$, la *base canonique* de $\boldsymbol{M}_{m,n}(K)$, espace vectoriel des matrices à m lignes et n colonnes à coefficients dans K. Rappelons que E_{ij} est la matrice de $\boldsymbol{M}_{m,n}(K)$ dont tous les coefficients sont nuls, sauf celui situé à la ligne i et la colonne j qui vaut 1.

  On a donc pour $A \in \boldsymbol{M}_{m,n}(K)$, $\boxed{A = (a_{ij}) = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}E_{ij}\,.}$

  Cette notation présente la même ambiguïté que la précédente.
  Si on se place dans $\boldsymbol{M}_n(K) = \boldsymbol{M}_{nn}(K)$, on a la règle de multiplication suivante :

$$\boxed{E_{ij}\,E_{k\ell} = \delta_{jk}\,E_{i\ell}\,, \text{ où } \delta_{jk} = \begin{cases} 1 & \text{si } j = k, \\ 0 & \text{sinon} \end{cases} \quad \text{est le } \textit{symbole de Kronecker}.}$$

- Soit $n \in \mathbb{N}^*$. Nous notons $\boldsymbol{D}_n(K)$ le sous-espace vectoriel de $\boldsymbol{M}_n(K)$ constitué des matrices diagonales et $\boldsymbol{S}_n(K)$ le sous-espace vectoriel de $\boldsymbol{M}_n(K)$ constitué des matrices symétriques :

$$\boldsymbol{D}_n(K) = \mathrm{Vect}(\{ E_{i\,i} \;;\; i \in [1, n]_{\mathbb{N}} \}) \quad \text{et} \quad \boldsymbol{S}_n(K) = \{ A \in \boldsymbol{M}_n(k) \;;\; {}^tA = A \} \;.$$

Nous notons $\boldsymbol{GL}_n(K)$ le groupe des matrices inversibles de $\boldsymbol{M}_n(K)$:

$$\boldsymbol{GL}_n(K) = \{ A \in \boldsymbol{M}_n(K) \;;\; \det(A) \neq 0 \} \;.$$

- Si E et F sont des K-espaces vectoriels, alors $\mathcal{L}(E, F)$ est un espace vectoriel de dimension finie et $\dim(\mathcal{L}(E, F)) = \dim(E) \times \dim(F)$.
Soient $\mathcal{B} = (e_1, \ldots, e_m)$, une base de E et $(f_1, \ldots, f_n)$ une base de F. A toute application linéaire $T \in \mathcal{L}(E, F)$ est associée sa matrice $\mathrm{Mat}_{\mathcal{B},\mathcal{B}'}(T) \in \boldsymbol{M}_{n,m}(K)$ relativement aux bases $\mathcal{B}$ et $\mathcal{B}'$. Par définition :

$$\boxed{A = \mathrm{Mat}_{\mathcal{B},\mathcal{B}'}(T) \;\Leftrightarrow\; \forall j \in [1, m]_{\mathbb{N}} \quad T(e_j) = \sum_{i=1}^{n} a_{i\,j} f_i \;.}$$

1.2. Espace dual

1.2.1. Définitions . Soit E un K-espace vectoriel.
On appelle *forme linéaire* sur E toute application linéaire de E dans K.
On appelle *espace dual* de E, noté E^*, l'espace vectoriel des formes linéaires sur E.
On a donc $E^* = \mathcal{L}(E, K)$ et $\varphi \in E^*$ signifie que φ est une application de E dans K telle que : $\forall (x, y) \in E^2 \quad \varphi(x + y) = \varphi(x) + \varphi(y) \quad \text{et} \quad \forall x \in E \quad \forall \lambda \in K \quad \varphi(\lambda x) = \lambda \varphi(x)$.

1.2.2. Exemples .

a) L'application de E dans K, qui à tout vecteur $x \in E$ associe le scalaire $0 \in K$ est une forme linéaire, appelée *forme nulle* sur E.

b) Si $E = \mathcal{C}([a, b], \mathbb{R})$, l'application $f \longmapsto \displaystyle\int_a^b f(t)dt$ est une forme linéaire sur E.

c) Si E est l'espace vectoriel des suites de réels convergentes, alors $u \longmapsto \displaystyle\lim_{n \longrightarrow +\infty} u_n$ est une forme linéaire sur E.

d) Si $E = K[X]$, pour tout $a \in K$, l'application $P \longmapsto P(a)$ est une forme linéaire sur E.

e) Si $E = \boldsymbol{M}_n(K)$, la trace Tr, qui à $A = (a_{i\,j}) \in E$ associe $\mathrm{Tr}(A) = \displaystyle\sum_{i=1}^{n} a_{i\,i}$, est une forme linéaire sur E.

f) Soient E un espace vectoriel de dimension finie et $\mathcal{B} = (e_i)_{i=1\cdots n}$ une base de E. Tout vecteur $x \in E$ se décompose de manière unique sur la base $\mathcal{B}$, en $x = \sum_{i=1}^{n} \lambda_i e_i$. Pour tout $j \in [1, n]_{\mathbb{N}}$, l'application $e_j^* : x \longmapsto \lambda_j$ est une forme linéaire sur E, appelée $j^{\text{ème}}$-*forme coordonnée* relative à la base $\mathcal{B}$.

g) Si $E = \mathbb{R}^3$, l'application qui à $(x, y, z) \in E$ associe $2x + 3y - 5z$ est une forme linéaire sur E.

Plus généralement on a la proposition suivante :

1.2.3. Proposition. *Soit $n \in \mathbb{N}^*$*

(*i*) *Fixons $(a_1, \ldots, a_n) \in K^n$ et considérons l'application φ de K^n dans K qui à tout $x = (x_1, \ldots, x_n) \in K^n$, associe le scalaire $\varphi(x) = a_1x_1 + a_2x_2 + \cdots + a_nx_n$. Alors φ est une forme linéaire sur K^n .*

(*ii*) *Réciproquement, pour toute forme linéaire ψ sur K^n, il existe un unique n-uplet $(a_1, \ldots, a_n) \in K^n$ tel que pour tout $x = (x_1, \ldots, x_n) \in K^n$, on ait $\psi(x) = \sum_{i=1}^{n} a_i x_i$.*

Preuve. (i) Pour tous x et y dans K^n, on a $x + y = (x_1 + y_1, \ldots, x_n + y_n)$ et donc

$$\varphi(x + y) = \sum_{i=1}^{n} a_i(x_i + y_i) = \sum_{i=1}^{n} (a_i x_i + a_i y_i) = \sum_{i=1}^{n} a_i x_i + \sum_{i=1}^{n} a_i y_i = \varphi(x) + \varphi(y) \ .$$

On démontre de même que pour tout $\lambda \in K$, on a $\varphi(\lambda x) = \lambda\varphi(x)$.
(ii) Soit $(\varepsilon_1, \ldots, \varepsilon_n)$ la base canonique de K^n.
Unicité du n-uplet $(a_1, \ldots, a_n) \in K^n$: Supposons que pour tout $x = (x_1, \ldots, x_n) \in K^n$, on ait
$\psi(x) = \sum_{i=1}^{n} a_i x_i$. Alors pour tout $j \in [1, n]_{\mathbb{N}}$, on a $\psi(\varepsilon_j) = a_j$.
Existence : Pour tout $j \in [1, n]_{\mathbb{N}}$, posons $a_j = \psi(\varepsilon_j)$. Pour tout $x = (x_1, \ldots, x_n) \in K^n$ on a $x = \sum_{i=1}^{n} x_i \varepsilon_i$, d'où $\psi(x) = \sum_{i=1}^{n} x_i \psi(\varepsilon_i) = \sum_{i=1}^{n} x_i a_i = \sum_{i=1}^{n} a_i x_i$. ■

1.2.4. Proposition. *Soit E un espace vectoriel de dimension finie. Alors son dual E^* est de dimension finie et $\dim(E^*) = \dim(E)$.*

Preuve. On a : $\dim(E^*) = \dim(\mathcal{L}(E, K)) = \dim(E) \times \dim(K) = \dim(E)$. ■

1.2.5. Proposition. *Soit E un K-espace vectoriel. Toute forme linéaire φ sur E, autre que la forme nulle, est surjective.*

Preuve. Comme φ est non nulle, il existe $x_0 \in E$ tel que $\varphi(x_0) \neq 0$. Soit $\lambda \in K$. Posons $x = \dfrac{\lambda}{\varphi(x_0)} x_0$. Alors $\varphi(x) = \dfrac{\lambda}{\varphi(x_0)} \varphi(x_0) = \lambda$. ■

Intéressons nous maintenant au noyau d'une forme linéaire.

1.3. Hyperplan

1.3.1. Définition. Soit E un K-espace vectoriel. On appelle *hyperplan* de E, le noyau de toute forme linéaire sur E autre que la forme nulle.
Autrement dit, une partie H de E est un hyperplan de E, s'il existe $\varphi \in E^*\backslash\{0\}$ tel que $H = \mathrm{Ker}\,(\varphi)$. On dit alors que la relation $\varphi(x) = 0$ est une *équation* de l'hyperplan H.

1.3.2. Proposition. *Soient E un K-espace vectoriel et H un sous-espace vectoriel de E. Les conditions suivantes sont équivalentes :*

(*i*) *H est un hyperplan de E.*

(*ii*) *$H \neq E$ et pour tout $v \notin H$ on a $E = H \oplus Kv$, où Kv est la droite engendrée par v.*

(*iii*) *H admet une droite pour supplémentaire dans E.*

Si E est de dimension finie, les conditions précédentes sont équivalentes à

(*iv*) $\dim(H) = \dim(E) - 1$.

Preuve. $(i) \Rightarrow (ii)$ Par hypothèse, il existe $\varphi \in E^*\backslash\{0\}$ tel que $H = \mathrm{Ker}\,(\varphi)$. Comme φ est non nulle, H est distinct de E. Soit v est un vecteur n'appartenant pas à H. On a donc $\varphi(v) \neq 0$. Montrons que tout vecteur x de E se décompose de manière unique en $x = h + \lambda v$ avec $h \in H$ et $\lambda \in K$.
Unicité : Si $x = h+\lambda v$ alors $\varphi(x) = \varphi(h)+\lambda\varphi(v) = \lambda\varphi(v)$. D'où $\lambda = \dfrac{\varphi(x)}{\varphi(v)}$ et $h = x-\lambda v$.
Existence : Posons $\lambda = \dfrac{\varphi(x)}{\varphi(v)}$ et $h = x - \lambda v$. On a $\varphi(h) = \varphi(x) - \lambda\varphi(v) = 0$, donc $h \in H$ et $x = h + \lambda v$.
L'implication $(ii) \Rightarrow (iii)$ est immédiate.
$(iii) \Rightarrow (i)$ Soient D une droite telle que $E = H \oplus D$ et v une base de D. Pour tout $x \in E$, il existe un unique couple $(h, \alpha) \in H \times K$ tel que $x = h + \alpha v$; posons alors $\varphi(x) = \alpha$. Autrement dit, $\varphi(x)$ est l'unique scalaire tel que $x - \varphi(x)v$ appartienne à H. Montrons que l'application φ est linéaire. Pour tous x et y dans E, tous λ et μ dans K le vecteur $(\lambda x + \mu y) - (\lambda\varphi(x) + \mu\varphi(y))v = \lambda(x - \varphi(x)v) + \mu(y - \varphi(y)v)$ appartient à H et par unicité $\varphi(\lambda x + \mu y) = \lambda\varphi(x) + \mu\varphi(y)$. D'où φ est une forme linéaire sur E.
De plus $x \in \mathrm{Ker}\,(\varphi)$ si et seulement si $x \in H$.
Si E est de dimension finie, il est clair que les assertions (iii) et (iv) sont équivalentes. ∎

1.3.3. Corollaire. *Deux formes linéaires non nulles sur un espace vectoriel E sont proportionnelles si et seulement si elles ont même noyau.*

Preuve. Soient φ et ψ deux formes linéaires non nulles. Supposons que φ et ψ ont même noyau H. Soit v un vecteur n'appartenant pas à H. Posons $\alpha = \dfrac{\psi(v)}{\varphi(v)}$ et montrons que

$\psi = \alpha\varphi$.
D'après la proposition précédente $E = H \oplus Kv$ et tout vecteur $x \in E$ peut s'écrire $x = h + \lambda v$ avec $h \in H$ et $\lambda \in K$.
D'où $\psi(x) = \psi(h) + \lambda\psi(v) = \lambda\psi(v)$ et $\alpha\varphi(x) = \alpha\varphi(h) + \alpha\lambda\varphi(v) = \lambda\psi(v)$. Donc φ et ψ sont proportionnelles.
La réciproque est immédiate. ■

1.3.4. Remarque. *Soit E un espace vectoriel de dimension finie n. Relativement à une base $\mathcal{B} = (e_i)_{i=1\cdots n}$, un hyperplan H de E admet une équation unique, à un scalaire multiplicatif non nul près, de la forme :* $a_1x_1 + a_2x_2 + \cdots + a_nx_n = 0$.

1.4. Base duale

1.4.1. Proposition. *Soit E un espace vectoriel de dimension finie. Si $\mathcal{B} = (e_i)_{i=1\cdots n}$ est une base de E, la famille des formes coordonnées $B^* = (e_i^*)_{i=1\cdots n}$ est une base du dual E^*. De plus $e_i^*(e_j) = \delta_{i\,j}$ (symbole de Kronecker) pour tous i et j dans $[1, n]_{\mathbb{N}}$.*

Preuve. D'après la définition des e_i^*, on a $e_i^*(e_j) = \delta_{i\,j}$ pour tous i et j dans $[1, n]_{\mathbb{N}}$.
Soit $\psi \in E^*$. Considérons la forme linéaire $\varphi = \sum_{i=1}^{n} \psi(e_i)e_i^*$. Pour tout j on a
$\varphi(e_j) = \sum_{i=1}^{n} \psi(e_i)e_i^*(e_j) = \sum_{i=1}^{n} \psi(e_i)\delta_{i\,j} = \psi(e_j)$.
Les formes linéaires φ et ψ coïncidant sur une base de E sont égales. Par conséquent la famille $(e_i^*)_{i=1\cdots n}$ est génératrice dans E^* ; comme elle comporte n vecteurs dans l'espace vectoriel dual E^* de dimension n, c'est une base de E^*. ■

1.4.2. Définition. La base $\mathcal{B}^*$ de E^* définie dans la proposition précédente est appelée *base duale* de la base $\mathcal{B}$ de E.

1.4.3. Remarque. Soient $\mathcal{B} = (e_i)_{i=1\cdots n}$ une base de E et $B^* = (e_i^*)_{i=1\cdots n}$ sa base duale.

Alors
$$\boxed{\begin{array}{c} \forall x \in E \quad x = \sum_{i=1}^{n} e_i^*(x)e_i \quad \text{et} \quad \forall \varphi \in E^* \quad \varphi = \sum_{i=1}^{n} \varphi(e_i)e_i^* \; . \\ \forall T \in \mathcal{L}(E) \quad a_{i\,j} = e_i^*(T(e_j)) \quad \text{avec} \quad A = \mathrm{Mat}_{\mathcal{B}}(T) \; . \end{array}}$$

1.4.4. Proposition. *Soient $\mathcal{B}$ et $\mathcal{B}'$ des bases de E espace vectoriel de dimension finie et P la matrice de passage de la base $\mathcal{B}$ à la base $\mathcal{B}'$. Alors la matrice de passage de la base $\mathcal{B}^*$ à la base $\mathcal{B}'^*$ est ${}^tP^{-1}$.*

Preuve. Soient $\mathcal{B} = (e_1, \ldots, e_n)$ et $\mathcal{B}' = (f_1, \ldots, f_n)$. Par définition de P on a, pour tout $k \in [1, n]_{\mathbb{N}}$, $f_k = \sum_{\ell=1}^{n} p_{\ell k} e_\ell$. Désignons par Q la matrice de passage de la base $\mathcal{B}^*$ à la base $\mathcal{B}'^*$. On a alors pour tout $j \in [1, n]_{\mathbb{N}}$, $f_j^* = \sum_{i=1}^{n} q_{i\,j} e_i^*$. Or pour tout $(j, k) \in [1, n]_{\mathbb{N}}^2$,

$$\delta_{j\,k} = f_j^*(f_k) = \left(\sum_{i=1}^{n} q_{i\,j} e_i^*\right)\left(\sum_{\ell=1}^{n} p_{\ell k} e_\ell\right) = \sum_{i=1}^{n}\sum_{\ell=1}^{n} q_{i\,j}\, p_{\ell\,k}\, \delta_{i\,\ell} = \sum_{i=1}^{n} q_{i\,j}\, p_{i\,k} = ({}^tQ\,P)_{j\,k} \ .$$

Donc ${}^tQ\,P = I_n$ et par suite $Q = {}^tP^{-1}$. ■

1.4.5. Corollaire. *Soit E un espace vectoriel de dimension finie. Toute base du dual est une base duale : Pour toute base $\mathcal{B}'$ de E^*, il existe une base $\mathcal{B}$ de E telle que $\mathcal{B}'$ soit la base duale de $\mathcal{B}$.*

Preuve. Soit $\mathcal{F}$ une base de E fixée et $\mathcal{F}^*$ sa base duale. Désignons par Q la matrice de passage de la base $\mathcal{F}^*$ à la base $\mathcal{B}'$ et posons $P = {}^tQ^{-1}$. Considérons la base $\mathcal{B}$ de E telle que la matrice de passage de la base $\mathcal{F}$ à $\mathcal{B}$ soit P. D'après la proposition précédente, la matrice de passage de la base $\mathcal{F}^*$ à la base $\mathcal{B}^*$ est ${}^tP^{-1} = Q$. Par suite $\mathcal{B}^* = \mathcal{B}'$. ■

1.4.6. Corollaire. *Soit E un espace vectoriel de dimension finie n.*

(*i*) *Soit $p \in [1, n]_{\mathbb{N}}$. Tout sous-espace vectoriel de E de dimension $n - p$ peut être considéré comme l'intersection de p hyperplans de E.*

(*ii*) *Considérons $(\varphi_i)_{i=1\cdots p}$, une famille de p formes linéaires non nulles de E^* et posons* $G = \bigcap_{i=1}^{p} \mathrm{Ker}\,(\varphi_i)$.
Alors $\dim(G) = n - \dim(\mathrm{Vect}\{\, \varphi_1, \ldots, \varphi_p \,\})$.
En particulier G est un sous-espace vectoriel de dimension $n - p$ si et seulement si la famille $(\varphi_i)_{i=1\cdots p}$ est libre dans E^*.

Preuve. (*i*) Soit F un sous-espace vectoriel de E de dimension $n-p$. Soit $(e_i)_{i=1\cdots n-p}$ une base de F que l'on complète pour obtenir $\mathcal{B} = (e_i)_{i=1\cdots n}$, base de E. Soit $\mathcal{B}^* = (e_i^*)_{i=1\cdots n}$ la base duale de $\mathcal{B}$. Alors $F = \bigcap_{i=n-p+1}^{n} \mathrm{Ker}\,(e_i^*)$.
(*ii*) Supposons tout d'abord que la famille $(\varphi_i)_{i=1\cdots p}$ est libre. Nous pouvons la compléter pour obtenir une base $\mathcal{B}' = (\varphi_i)_{i=1\cdots n}$ de E^*. Soit $(e_i)_{i=1\cdots n}$ la base de E dont $\mathcal{B}'$ est la base duale. On a alors :

$$G = \bigcap_{i=1}^{p} \mathrm{Ker}\,(\varphi_i) = \bigcap_{i=1}^{p} \mathrm{Ker}\,(e_i^*) = \mathrm{Vect}\{\, e_{p+1}, \ldots, e_n \,\} \ .$$

Donc $\dim(G) = n - p$.
Supposons maintenant que la famille $(\varphi_i)_{i=1\cdots p}$ est liée. Posons $\Phi = \mathrm{Vect}\{\varphi_1, \dots, \varphi_p\}$ et $d = \dim(\Phi)$. De la famille $(\varphi_i)_{i=1\cdots p}$, génératrice de Φ, nous pouvons extraire une base. Quitte à renuméroter, nous pouvons supposer que $(\varphi_i)_{i=1\cdots d}$ est une base de Φ.
Soit $j \in [d+1, p]_{\mathbb{N}}$. Il existe $\lambda_1, \dots, \lambda_d$ dans K tels que $\varphi_j = \sum_{i=1}^{d} \lambda_i \varphi_i$. Donc, si x appartient à $\bigcap_{i=1}^{d} \mathrm{Ker}(\varphi_i)$ on a $\varphi_j(x) = 0$. Il en résulte que $\bigcap_{i=1}^{p} \mathrm{Ker}(\varphi_i) = \bigcap_{i=1}^{d} \mathrm{Ker}(\varphi_i)$.
En appliquant à la famille libre $(\varphi_i)_{i=1\cdots d}$ le résultat obtenu précédemment, on a donc $\dim(G) = n - d$. ∎

1.4.7. Exemple. Dans un espace vectoriel de dimension 3, toute droite vectorielle D est l'intersection de deux plans vectoriels distincts. Si (e_1, e_2, e_3) est une base de E, alors D admet dans cette base une équation de la forme : $\begin{cases} a_1x_1 + a_2x_2 + a_3x_3 &= 0 \\ b_1x_1 + b_2x_2 + b_3x_3 &= 0 \end{cases}$ avec (a_1, a_2, a_3) et (b_1, b_2, b_3) non colinéaires dans K^3.

1.4.8. Théorème d'interpolation de Lagrange. *Soient $(a_i)_{i=1\cdots n}$ et $(\lambda_i)_{i=1\cdots n}$ deux familles de n scalaires de K. On suppose les a_i deux à deux distincts. Alors il existe un unique polynôme $P \in K[X]$ avec $d^\circ(P) \leq n-1$ tel que $P(a_i) = \lambda_i$.*

Preuve. Soit Θ l'application qui à $P \in K_{n-1}[X]$ associe $(P(a_i))_{i=1\cdots n} \in K^n$. Cette application est linéaire ; montrons qu'elle est injective. Soit $Q \in \mathrm{Ker}(\Theta)$. On a alors $Q(a_i) = 0$ pour tout $i \in [1, n]_{\mathbb{N}}$ et Q admet n racines distinctes ; comme il est de degré au plus $n-1$, c'est donc le polynôme nul.
De plus comme $\dim(K_{n-1}[X]) = n = \dim(K^n)$, l'application Θ est un isomorphisme. D'où, pour toute famille $(\lambda_i)_{i=1\cdots n} \in K^n$, l'existence et l'unicité d'un polynôme $P \in K_{n-1}[X]$ tel que $P(a_i) = \lambda_i$, à savoir $P = \Theta^{-1}((\lambda_i))$. ∎

1.4.9. Remarque. Reprenons les données du théorème précédent. Nous pouvons donner une expression explicite du polynôme P. Pour tout i, le polynôme $L_i = \prod_{j \neq i} \frac{(X - a_j)}{(a_i - a_j)}$ est de degré inférieur ou égal à $n-1$ et satisfait à $L_i(a_j) = \delta_{i,j}$ pour tout $j \in [1, n]_{\mathbb{N}}$. Il en résulte que $P = \sum_{i=1}^{n} \lambda_i L_i$.

1.4.10. Corollaire. *Soient E un K-espace vectoriel de dimension finie et $f \in \mathcal{L}(E)$. Alors f est diagonalisable si et seulement s'il existe $P \in K[X]$ scindé sur K et à racines simples, annulateur pour f.*

Preuve. Supposons que f est diagonalisable. Soient $\lambda_1, \dots, \lambda_r$ les valeurs propres distinctes de f et soit $P = \prod_{i=1}^{r} (X - \lambda_i)$. Il est scindé sur k et à racines simples. Pour

$x \in E_{\lambda_j}(f)$, espace propre de f relatif à λ_j, on a :

$$P(f)(x) = \prod_{i=1}^{r} (f - \lambda_i \mathrm{Id}_E)(x) = \prod_{i \neq j}^{r} (f - \lambda_i \mathrm{Id}_E) \circ (f - \lambda_j \mathrm{Id}_E)(x) = 0 \ .$$

Soit $y \in E$. Puisque f est diagonalisable, E est somme directe des sous-espaces propres de f et le vecteur y se décompose en $y = \sum_{j=1}^{r} x_j$ avec $x_j \in E_{\lambda_j}(f)$.

Donc $P(f)(y) = \sum_{j=1}^{r} P(f)(x_j) = 0$ et P est annulateur pour f.

Réciproquement, soient $\lambda_1, \ldots, \lambda_p$ les racines de P. Pour tout $j \in [1,p]_{\mathbb{N}}$, considérons le polynôme d'interpolation de Lagrange $L_j \in K_{p-1}[X]$ tel que $L_j(\lambda_i) = \delta_{i\,j}$, pour $i \in [1,p]_{\mathbb{N}}$. Ainsi défini, le polynôme L_j admet tous les λ_i pour $i \neq j$, comme racines. Il en résulte que $(X - \lambda_j)L_j$ admet tous les λ_i pour racines et est donc un multiple de P. Par conséquent $(f - \lambda_j \mathrm{Id}_E) \circ L_j(f) = 0$ $(*)$.

Par ailleurs $(1 - \sum_{j=1}^{p} L_j)(\lambda_i) = 1 - \sum_{j=1}^{p} \delta_{i\,j} = 0$, pour tout i. Le polynôme $1 - \sum_{j=1}^{p} L_j$ admet tous les λ_i pour $i = 1, \ldots, p$ pour racines et est de degré au plus $p-1$; il est donc nul. On obtient alors $\mathrm{Id}_E = \sum_{j=1}^{p} L_j(f)$ $(**)$.

Pour $x \in E$, posons $x_j = L_j(f)(x)$; d'après $(*)$, x_j appartient à $E_{\lambda_j}(f)$ et d'après $(**)$ on a $x = \sum_{j=1}^{p} x_j$. La somme des sous-espaces propres de f est donc égale à E, autrement dit f est diagonalisable. ■

1.5. Transposition

1.5.1. Définition. Soient E et F des K-espaces vectoriels et $f \in \mathcal{L}(E,F)$. On appelle *transposée* de f, l'application notée ${}^t f$ de F^* dans E^* définie par :

$$\forall \psi \in F^* \quad {}^t f(\psi) = \psi \circ f \ .$$

1.5.2. Proposition. *Soient E et F des K-espaces vectoriels. Pour tout $f \in \mathcal{L}(E,F)$, l'application ${}^t f$ appartient à $\mathcal{L}(F^*, E^*)$.*
Si de plus E et F sont de dimension finie et munis respectivement d'une base $\mathcal{B}$ et $\mathcal{C}$, on a :

$$\mathrm{Mat}_{\mathcal{C}^*, \mathcal{B}^*}\left({}^t f\right) = {}^t\left(\mathrm{Mat}_{\mathcal{B},\mathcal{C}}\,(f)\right) \ .$$

(Cette dernière propriété justifie le nom de transposée).

Preuve. Pour tous λ et λ' de K, ψ et ψ' de F^* on a :

$${}^t f(\lambda\psi + \lambda'\psi') = (\lambda\psi + \lambda'\psi') \circ f = \lambda\psi \circ f + \lambda'\psi' \circ f = \lambda\, {}^t f(\psi) + \lambda'\, {}^t f(\psi') \ .$$

Donc ${}^t f$ est linéaire de F^* dans E^*.
Posons $\mathcal{B} = (e_1, \ldots, e_m)$, $\mathcal{C} = (f_1, \ldots, f_n)$ $A = \mathrm{Mat}_{\mathcal{B},\mathcal{C}}(f)$ et $B = \mathrm{Mat}_{\mathcal{C}^*,\mathcal{B}^*}({}^t f)$.
Pour tout $j \in [1, m]_{\mathbb{N}}$ on a : $f(e_j) = \sum\limits_{k=1}^{n} a_{k\,j} f_k$.
Pour tout $i \in [1, n]_{\mathbb{N}}$ et tout $j \in [1, m]_{\mathbb{N}}$, le coefficient $b_{j\,i}$ est la $j^{\text{ème}}$ coordonnée de ${}^t f(f_i^*)$ dans sa décomposition dans la base $\mathcal{B}^*$. Or d'après 1.4.3, on a ${}^t f(f_i^*) = \sum\limits_{j=1}^{m} \left({}^t f(f_i^*)\right)(e_j)\, e_j^*$.
D'où

$$b_{j\,i} = \left({}^t f(f_i^*)\right)(e_j) = f_i^* \circ f(e_j) = f_i^* \left(\sum_{k=1}^{n} a_{k\,j} f_k\right) = \sum_{k=1}^{n} a_{k\,j} \delta_{i\,k} = a_{i\,j}\ . \qquad \blacksquare$$

1.5.3. Proposition. *Soient E, F et G des K-espaces vectoriels.*

(*i*) $\forall (f, f') \in \mathcal{L}(E, F)^2 \quad \forall (\lambda, \lambda') \in K^2 \quad {}^t(\lambda f + \lambda' f') = \lambda\,{}^t f + \lambda'\,{}^t f'$.

(*ii*) $\forall f \in \mathcal{L}(E, F) \quad \forall g \in \mathcal{L}(F, G) \quad {}^t(g \circ f) = {}^t f \circ {}^t g$.

(*iii*) ${}^t(\mathrm{Id}_E) = \mathrm{Id}_{E^*}$.

Preuve. (i) Pour tout $\psi \in F^*$, on a :
${}^t(\lambda f + \lambda' f')(\psi) = \psi \circ (\lambda f + \lambda' f') = \lambda \psi \circ f + \lambda' \psi \circ f' = \lambda\,{}^t f(\psi) + \lambda'\,{}^t f'(\psi) = (\lambda\,{}^t f + \lambda'\,{}^t f')(\psi)$.
(ii) Pour tout $\Psi \in G^*$, on a ${}^t(g \circ f)(\Psi) = \Psi \circ (g \circ f) = (\Psi \circ g) \circ f = {}^t f(\Psi \circ g) = {}^t f \circ {}^t g(\Psi)$.
(iii) Pour tout $\varphi \in E^*$, on a ${}^t(\mathrm{Id}_E)(\varphi) = \varphi \circ \mathrm{Id}_E = \varphi$. $\blacksquare$

1.5.4. Corollaire. *Soient E et F des K-espaces vectoriels de dimension finie.*
L'application $t : \begin{array}{ccc} \mathcal{L}(E, F) & \longrightarrow & \mathcal{L}(F^*, E^*) \\ f & \longmapsto & {}^t f \end{array}$ *est un isomorphisme de K-espaces vectoriels.*

Preuve. D'après la proposition précédente, l'application t est linéaire. Montrons qu'elle est injective ; soit $f \in \mathrm{Ker}\,(t)$ on a ${}^t f = 0$. Choisissons des bases $\mathcal{B}$ et $\mathcal{C}$ respectivement de E et F et soit $A = \mathrm{Mat}_{\mathcal{B},\mathcal{C}}(f)$. Alors ${}^t A = \mathrm{Mat}_{\mathcal{C}^*,\mathcal{B}^*}({}^t f) = 0$. Donc $A = 0$ et par suite $f = 0$.
Par ailleurs $\dim(\mathcal{L}(F^*, E^*)) = \dim(F^*) \times \dim(E^*) = \dim(F) \times \dim(E) = \dim(\mathcal{L}(E, F))$.
L'application t étant linéaire injective entre deux espaces vectoriels de même dimension finie est un isomorphisme. $\blacksquare$

1.5.5. Proposition. *Soient E un K-espace vectoriel de dimension finie, $f \in \mathcal{L}(E)$ et φ une forme linéaire non nulle sur E. Alors l'hyperplan $\mathrm{Ker}\,(\varphi)$ est stable par f si et seulement si φ est vecteur propre pour ${}^t f$.*

Preuve. Supposons que $\mathrm{Ker}\,(\varphi)$ soit stable par f. Si $x \in \mathrm{Ker}\,(\varphi)$ alors $f(x)$ appartient à $\mathrm{Ker}\,(\varphi)$, d'où $\mathrm{Ker}\,(\varphi) \subset \mathrm{Ker}\,(\varphi \circ f)$. Si $\varphi \circ f = 0$ alors ${}^t f(\varphi) = 0 \cdot \varphi$, sinon $\mathrm{Ker}\,(\varphi \circ f)$ est un hyperplan et on a $\mathrm{Ker}\,(\varphi) = \mathrm{Ker}\,(\varphi \circ f)$; d'après le corollaire 1.3.3 la forme linéaire $\varphi \circ f$ est proportionnelle à φ. Dans les 2 cas il existe $\lambda \in K$ tel que ${}^t f(\varphi) = \varphi \circ f = \lambda f$.
Réciproquement, si φ est vecteur propre pour ${}^t f$ relativement à la valeur propre λ, alors $\varphi \circ f = \lambda \varphi$. Pour tout $h \in \mathrm{Ker}\,(\varphi)$ on a $\varphi \circ f(h) = \lambda \varphi(h) = 0$, i.e. $f(h) \in \mathrm{Ker}\,(\varphi)$. $\blacksquare$

2 Algèbre bilinéaire

Dans toute cette partie K désignera un corps commutatif tel que $1_K + 1_K \neq 0_K$, c'est à dire tel que l'on puisse diviser par 2. Le lecteur pourra cette année se contenter de supposer que $K = \mathbb{R}$ ou $K = \mathbb{C}$.
Si X est un ensemble, on notera $\mathcal{F}(X, K)$ l'ensemble des applications de X dans K.

2.1. Formes bilinéaires

2.1.1. Définitions. Soient E, F et G des K-espaces vectoriels.
Une application f de $E \times F$ dans G est dite *bilinéaire* si pour tout $y \in F$ l'application $x \longmapsto f(x, y)$ est linéaire de E dans G et pour tout $x \in E$ l'application $y \longmapsto f(x, y)$ est linéaire de F dans G, c'est à dire :

- $\forall y \in F \quad \forall (x, x') \in E^2 \qquad f(x + x', y) = f(x, y) + f(x', y)$
- $\forall x \in E \quad \forall (y, y') \in F^2 \quad f(x, y + y') = f(x, y) + f(x, y')$
- $\forall x \in E \quad \forall y \in F \quad \forall \lambda \in K \quad f(\lambda x, y) = \lambda f(x, y) = f(x, \lambda y)$.

On appelle *forme bilinéaire* sur E toute application bilinéaire de $E \times E$ dans K.
On note $Bil(E, K)$ l'ensemble des formes bilinéaires sur E.
Une application f de $E \times E$ dans K est dite *symétrique* si pour tout $(x, y) \in E^2$ on a $f(x, y) = f(y, x)$.
On note $\mathcal{S}(E, K)$ l'ensemble des formes bilinéaires symétriques sur E.

2.1.2. Exemples.

a) Le produit de matrices de $\boldsymbol{M}_{mn}(K) \times \boldsymbol{M}_{np}(K)$ dans $\boldsymbol{M}_{mp}(K)$ qui à (A, B) associe $A\,B$ est bilinéaire.

b) L'application $(\varphi, x) \longmapsto \varphi(x)$ est bilinéaire de $E^* \times E$ dans K.

c) Le produit vectoriel $(x, y) \longmapsto x \wedge y$ est une application bilinéaire de $\mathbb{R}^3 \times \mathbb{R}^3$ dans $\mathbb{R}^3$.

d) L'application nulle de $E \times E$ dans K est forme bilinéaire symétrique.

e) Le produit scalaire $(x, y) \longmapsto x \cdot y$ est une forme bilinéaire symétrique sur $\mathbb{R}^3$.

f) L'application $(f, g) \longmapsto \displaystyle\int_0^1 f(t)g(t)\,dt$ est une forme bilinéaire symétrique sur $\mathcal{C}([0, 1], \mathbb{R})$.

g) Soient φ et ψ dans E^* ; l'application $(x, y) \longmapsto \varphi(x)\psi(y)$ est une forme bilinéaire sur E ; si $\varphi = \psi$ alors elle est symétrique.

2.1.3. Remarques.

a) Une application f de $E \times E$ dans K est une forme bilinéaire symétrique si et seulement si f est symétrique et pour tout $y \in E$, l'application $x \longmapsto f(x, y)$ est linéaire de E dans K

b) Soient f une forme bilinéaire sur E, $\lambda_1, \ldots, \lambda_n, \mu_1, \ldots, \mu_p$ des éléments de K et $u_1, \ldots, u_n, v_1, \ldots, v_p$ des vecteurs de E. Alors

$$f\left(\sum_{i=1}^{n} \lambda_i u_i\,,\ \sum_{j=1}^{p} \mu_j v_j\right) = \sum_{i=1}^{n} \lambda_i\, f\left(u_i\,,\ \sum_{j=1}^{p} \mu_j v_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{p} \lambda_i\, \mu_j\, f(u_i, v_j)\ .$$

c) Soient E un espace vectoriel de dimension finie n et $(e_1, \ldots, e_n)$ une base de E. Une forme bilinéaire f sur E est déterminée dès que l'on connaît les scalaires $f(e_i, e_j)$ pour tout $(i, j) \in [1, n]_{\mathbb{N}}^2$:

$$\boxed{\text{Si } \ x = \sum_{i=1}^{n} x_i e_i \ \text{ et } \ y = \sum_{i=1}^{n} y_i e_i \ \text{ alors } \ f(x, y) = \sum_{i=1}^{n}\sum_{j=1}^{n} x_i\, y_j\, f(e_i, e_j).}$$

Ce qui nous incite à poser la définition suivante.

2.1.4. Définitions. Soient E un espace vectoriel de dimension finie, $\mathcal{B} = (e_1, \ldots, e_n)$ une base de E et f une forme bilinéaire sur E. On appelle *matrice de f relativement à $\mathcal{B}$*, que l'on note $\mathrm{Mat}_{\mathcal{B}}(f)$ la matrice $A \in \boldsymbol{M}_n(K)$ telle que $\forall (i, j) \in [1, n]_{\mathbb{N}}^2 \quad a_{i\,j} = f(e_i, e_j)$.

2.1.5. Proposition. *Soient E un espace vectoriel de dimension finie, $\mathcal{B}$ une base de E et f une forme bilinéaire sur E. Posons $A = \mathrm{Mat}_{\mathcal{B}}(f)$.*

(*i*) *Soient x et y des vecteurs de E. Posons $X = \mathrm{Mat}_{\mathcal{B}}(x)$ et $Y = \mathrm{Mat}_{\mathcal{B}}(y)$. Alors*

$$f(x, y) = {}^t X\, A\, Y\ .$$

(*ii*) *f est symétrique si et seulement si A est une matrice symétrique (${}^t A = A$).*

Preuve. Soit $\mathcal{B} = (e_1, \ldots, e_n)$.

(*i*) Si $x = \sum_{i=1}^{n} x_i e_i$ alors $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ et ${}^t X = (x_1, \ldots, x_n)$. De même pour Y.

D'après la remarque précédente, on a : $f(x, y) = \sum_{i=1}^{n}\sum_{j=1}^{n} x_i\, y_j\, f(e_i, e_j) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{i\,j}\, x_i\, y_j$.

Par ailleurs, $A\,Y$ produit de la matrice $A \in \boldsymbol{M}_n(K)$ par la matrice $Y \in \boldsymbol{M}_{n\,1}(K)$ est un élément de $\boldsymbol{M}_{n\,1}(K)$ et pour tout $i \in [1,n]_{\mathbb{N}}$ on a $(AY)_i = \sum_{j=1}^{n} a_{i\,j}\, y_j$. D'où ${}^tX\,A\,Y$, produit de la matrice ${}^tX \in \boldsymbol{M}_{1\,n}(K)$ par la matrice $AY \in \boldsymbol{M}_{n\,1}(K)$ est un élément de $\boldsymbol{M}_{1\,1}(K)$ c'est à dire un scalaire de K. (Nous identifions une matrice à 1 ligne et 1 colonne avec son unique coefficient) et on a :

$$ {}^tX\,A\,Y = \sum_{i=1}^{n} x_i\,(AY)_i = \sum_{i=1}^{n} x_i \sum_{j=1}^{n} a_{i\,j}\, y_j = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{i\,j}\, x_i\, y_j = f(x,y) \ . $$

(ii) Si f est symétrique, on a, pour tout $(i,j) \in [1,n]_{\mathbb{N}}^2$ $\quad a_{j\,i} = f(e_j, e_i) = f(e_i, e_j) = a_{i\,j}$. Donc ${}^tA = A$.
Réciproquement supposons que A soit symétrique. Pour tout couple $(x,y) \in E^2$ on a $f(y,x) = {}^tY\,A\,X$. Or la matrice ${}^tY\,A\,X \in \boldsymbol{M}_{1\,1}(K)$ est égale à sa transposée (elle ne possède qu'un seul coefficient !), d'où

$$ f(y,x) = {}^tY\,A\,X = {}^t\left({}^tY\,A\,X\right) = {}^tX\,{}^tA\,{}^t\left({}^tY\right) = {}^tX\,A\,Y = f(x,y) \ . $$ ∎

2.1.6. Corollaire. *Soient E un espace vectoriel de dimension finie n.*

(*i*) *$Bil(E,K)$ est un sous-espace vectoriel de l'espace vectoriel $\mathcal{F}(E \times E, K)$.*

(*ii*) *Soit $\mathcal{B}$ une base de E. L'application $\Theta : f \longmapsto \mathrm{Mat}_{\mathcal{B}}(f)$ est un isomorphisme d'espaces vectoriels de $Bil(E,K)$ dans $\boldsymbol{M}_n(K)$.*

(*iii*) $\dim(Bil(E,K)) = n^2$.

Preuve. (i) On a déjà remarqué que $Bil(E,K)$ est non vide (2.1.2.(d)) et on vérifie facilement que si f et g sont dans $Bil(E,K)$, λ et μ dans K alors $\lambda f + \mu g$ est bilininéaire.
(ii) De plus : $\mathrm{Mat}_{\mathcal{B}}(\lambda f + \mu g) = \Big((\lambda f + \mu g)(e_i, e_j)\Big)_{i,j} = \lambda \Big(f(e_i, e_j)\Big)_{i,j} + \mu \Big(g(e_i, e_j)\Big)_{i,j}$.
D'où $\Theta(\lambda f + \mu g) = \lambda\Theta(f) + \mu\Theta(g)$.
Si $\Theta(f) = 0$ alors f est nulle compte tenu de la formule (2.1.3.(c)). Donc Θ est injectif.
On vérifie aisément que pour tout $A \in \boldsymbol{M}_n(K)$, l'application f définie sur $E \times E$ par $f(x,y) = {}^tX\,A\,Y$, où $X = \mathrm{Mat}_{\mathcal{B}}(x)$ et $Y = \mathrm{Mat}_{\mathcal{B}}(y)$, appartient à $Bil(E,K)$ et que $\mathrm{Mat}_{\mathcal{B}}(f) = A$. Il en résulte que Θ est surjectif.
L'assertion (iii) résulte immédiatement de (ii). ∎

2.1.7. Corollaire. *Soient E un espace vectoriel de dimension finie n.*

(*i*) *$\mathcal{S}(E,K)$ est un sous-espace vectoriel de $Bil(E,K)$.*

(*ii*) *Soit $\mathcal{B}$ une base de E. L'application $\Theta' : f \longmapsto \mathrm{Mat}_{\mathcal{B}}(f)$ est un isomorphisme d'espaces vectoriels de $\mathcal{S}(E,K)$ dans $\boldsymbol{S}_n(K)$.*

(*iii*) $\dim(S(E,K)) = \dfrac{n(n+1)}{2}$.

Preuve. La démonstration est laissée au lecteur à titre d'exercice. ∎

2.1.8. Théorème. *Soient E un espace vectoriel de dimension finie n, $\mathcal{B}$ et $\mathcal{B}'$ des bases de E et f une forme bilinéaire sur E.*
Si P est la matrice de passage de la base $\mathcal{B}$ à la base $\mathcal{B}'$, alors

$$\mathrm{Mat}_{\mathcal{B}'}(f) = {}^t P\, \mathrm{Mat}_{\mathcal{B}}(f)\ P\ .$$

Preuve. Soit $\mathcal{B}' = (e'_1, \ldots, e'_n)$. Rappelons que par définition, la $k^{\text{ème}}$ colonne de P est formée des coordonnées de e'_k dans la base $\mathcal{B}$. Si $X' = \mathrm{Mat}_{\mathcal{B}'}(x)$, on sait que $X = P\,X'$; de même $Y = P\,Y'$ avec $Y' = \mathrm{Mat}_{\mathcal{B}'}(y)$. D'où

$$f(x,y) = {}^t X\,A\,Y = {}^t(P\,X')\,A\,(P\,Y') = {}^t(X')\,{}^tP\,A\,P\,Y' = {}^t(X')\ \left({}^tP\,A\,P\right)\ Y'\ .$$

Par ailleurs, on a aussi $f(x,y) = {}^t(X')\,\mathrm{Mat}_{\mathcal{B}'}(f)\ Y'$, d'où

$${}^t(X')\,\mathrm{Mat}_{\mathcal{B}'}(f)\ Y' = {}^t(X')\ \left({}^tP\,A\,P\right)\ Y'\ .$$

Cette dernière égalité étant valable pour toutes matrices colonnes X' et Y', on a donc $\mathrm{Mat}_{\mathcal{B}'}(f) = {}^tP\,A\,P$. ∎

2.1.9. Remarque. Le lecteur remarquera la différence fondamentale entre cette formule de changement de bases pour une forme bilinéaire et la formule de changement de bases pour une application linéaire $T \in \mathcal{L}(E)$. (Rappelons que pour une application linéaire $\mathrm{Mat}_{\mathcal{B}'}(T) = P^{-1}\,\mathrm{Mat}_{\mathcal{B}}(T)\ P$).

Pour une forme bilinéaire symétrique f sur E, intéressons nous maintenant à l'application $x \longmapsto f(x,x)$ de E dans K.

2.2. Formes quadratiques

2.2.1. Lemme. *Soit f une forme bilinéaire symétrique sur un K-espace vectoriel E. Pour tout $x \in E$, posons $q(x) = f(x,x)$. Alors*

(i) $\forall x \in E \quad \forall \lambda \in K \quad q(\lambda x) = \lambda^2 q(x)$.

(ii) $\forall (x,y) \in E^2 \quad f(x,y) = \dfrac{1}{2}\Big(q(x+y) - q(x) - q(y)\Big) = \dfrac{1}{4}\Big(q(x+y) - q(x-y)\Big)$.
(Formules de polarisation).

Preuve. (*i*) $\qquad q(\lambda x) = f(\lambda x, \lambda x) = \lambda f(x, \lambda x) = \lambda^2 f(x,x) = \lambda^2 q(x)$.
(*ii*) On a :
$q(x+y) = f(x+y, x+y) = f(x,x) + f(x,y) + f(y,x) + f(y,y) = q(x) + q(y) + 2f(x,y)$.
En remplaçant y par $-y$ on a aussi :
$q(x-y) = q(x) + q(-y) + 2f(x,\,-y) = q(x) + q(y) - 2f(x,y)$.
D'où le résultat. ∎

2.2.2. Définitions. Soit f une forme bilinéaire symétrique sur un K-espace vectoriel E. On appelle *forme quadratique associée à* f l'application de E dans K, notée q_f, définie par $q_f(x) = f(x,x)$, pour tout $x \in E$.
Une application q de E dans K est une *forme quadratique* s'il existe f, forme bilinéaire symétrique sur E telle $q = q_f$. D'après le lemme précédent, la forme bilinéaire symétrique f admettant q pour forme quadratique associée, est unique ; on appelle f la *forme polaire* de q.
On note $\mathcal{Q}(E)$ l'ensemble des formes quadratiques sur E.
Si E est de dimension finie et muni d'une base $\mathcal{B}$, on appelle *matrice de* $q \in \mathcal{Q}(E)$ *relativement à* $\mathcal{B}$ la matrice de sa forme polaire f : $\mathrm{Mat}_{\mathcal{B}}(q) = \mathrm{Mat}_{\mathcal{B}}(f)$.

2.2.3. Remarques.

a) Une forme quadratique q n'est pas linéaire, car on a $q(\lambda x) = \lambda^2 q(x)$, ce qui justifie l'épithète *quadratique*.

b) Si E est muni d'une base finie $\mathcal{B}$, pour tout $x \in E$, on a :

$$\boxed{q(x) = {}^tX\, \mathrm{Mat}_{\mathcal{B}}(q)\ X \ \text{ avec } X = \mathrm{Mat}_{\mathcal{B}}(x)\,.}$$

2.2.4. Exemples.

a) Soient E un espace vectoriel de dimension 3, $\mathcal{B}$ une base de E et f une forme bilinéaire symétrique sur E telle que $\mathrm{Mat}_{\mathcal{B}}(f) = A = \begin{pmatrix} 2 & 1 & 3 \\ 1 & 0 & 0 \\ 3 & 0 & -1 \end{pmatrix}$. La forme quadratique q_f associée a pour expression :

$$q_f(x) = (x_1, x_2, x_3) \begin{pmatrix} 2 & 1 & 3 \\ 1 & 0 & 0 \\ 3 & 0 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 2x_1^2 - x_3^2 + 2x_1x_2 + 6x_1x_3\ .$$

On remarque qu'en fonction des coordonnées, la forme quadratique est une fonction $Q(x_1, x_2, x_3)$ polynômiale homogène de degré 2, c'est à dire combinaison linéaire des monômes x_i^2 ou x_ix_j. De plus

$$\begin{aligned} \frac{\partial Q}{\partial x_1}(x_1,x_2,x_3) &= 4x_1 + 2x_2 + 6x_3 \\ \frac{1}{2}\frac{\partial Q}{\partial x_1}(x_1,x_2,x_3) &= 2x_1 + \ x_2 + 3x_3 \\ &= a_{1\,1}\,x_1 + a_{1\,2}\,x_2 + a_{1\,3}\,x_3\ . \end{aligned}$$

De même $\dfrac{1}{2}\dfrac{\partial Q}{\partial x_2}(x_1,x_2,x_3) = x_1 = a_{2\,1}\,x_1 + a_{2\,2}\,x_2 + a_{2\,3}\,x_3$ et
$\dfrac{1}{2}\dfrac{\partial Q}{\partial x_3}(x_1,x_2,x_3) = 3x_1 - x_3 = a_{3\,1}\,x_1 + a_{3\,2}\,x_2 + a_{3\,3}\,x_3$.

b) Réciproquement, en dimension 2 par exemple, donnons nous un polynôme Q homogène de degré 2 : $Q(x_1, x_2) = x_1^2 + 3x_2^2 + 10x_1x_2$.

$$\begin{aligned} \frac{1}{2}\frac{\partial Q}{\partial x_1}(x_1, x_2) &= x_1 + 5x_2 \\ \frac{1}{2}\frac{\partial Q}{\partial x_2}(x_1, x_2) &= 5x_1 + 3x_2 \end{aligned}$$

Considérons la matrice $A = \begin{pmatrix} 1 & 5 \\ 5 & 3 \end{pmatrix}$. Soit f la forme bilinéaire symétrique admettant A pour matrice relativement à la base $\mathcal{B} = (e_1, e_2)$. On a alors :
$f(x_1e_1 + x_2e_2, y_1e_1 + y_2e_2) = x_1y_1 + 3x_2y_2 + 5x_1y_2 + 5x_2y_1$ et
$q_f(x_1e_1 + x_2e_2) = x_1^2 + 3x_2^2 + 10x_1x_2 = Q((x_1, x_2))$.

2.2.5. Proposition . *Soit E un espace vectoriel de dimension finie n.*

(*i*) *$\mathcal{Q}(E)$ est un sous-espace vectoriel de l'espace vectoriel $\mathcal{F}(E, K)$.*

(*ii*) *L'application $f \longmapsto q_f$ est un isomorphisme de $\mathcal{S}(E, K)$ sur $\mathcal{Q}(E)$.*

(*iii*) $\dim(\mathcal{Q}(E)) = \dfrac{n(n+1)}{2}$.

Preuve. D'après le lemme précédent, l'application $f \longmapsto q_f$ est une bijection entre l'ensemble des formes bilinéaires symétriques sur E et l'ensemble des formes quadratiques sur E. On vérifie aisément que $q_{\lambda f + \mu g} = \lambda q_f + \mu q_g$. Il en résulte les assertions (*i*) et (*ii*). L'assertion (*iii*) découle alors du corollaire 2.1.7. ∎

2.2.6. Corollaire . *Soient $\varphi_1, \ldots, \varphi_r$ des formes linéaires sur E et $a_1, \ldots, a_r$ des scalaires. Alors $\sum_{i=1}^{r} a_i \varphi_i^2$ est une forme quadratique sur E dont la forme polaire est f définie par*

$f(x, y) = \sum_{i=1}^{r} a_i \varphi_i(x)\varphi_i(y)$, pour tout $(x, y) \in E^2$.

Preuve. Pour tout $i \in [1, r]_{\mathbb{N}}$, l'application $x \longmapsto \varphi_i^2(x)$ est une forme quadratique admettant pour forme polaire la forme bilinéaire symétrique $(x, y) \longmapsto \varphi_i(x)\varphi_i(y)$ (2.1.2). La linéarité de l'application $f \longmapsto q_f$ permet de conclure. ∎

2.3. Orthogonalité

2.3.1. Définitions. Soient E un K-espace vectoriel, q une forme quadratique sur E et f sa forme polaire.
On dit que deux vecteurs x et y de E sont *orthogonaux pour* q (ou pour f) si $f(x,y)=0$.
On dit qu'une famille $(x_1, \dots, x_k)$ de E est *orthogonale pour* q si pour tout couple (i,j) avec $i \neq j$, les vecteurs x_i et x_j sont orthogonaux pour q.
On dit que deux parties A et B de E sont *orthogonales* pour q si on a $f(a,b)=0$, pour tout $a \in A$ et tout $b \in B$.
Pour toute partie A de E on appelle *orthogonal de* A pour q, l'ensemble, noté $A^\perp$ des vecteurs de E orthogonaux à A. On a donc : $A^\perp = \{\, x \in E \,;\ \forall a \in A \quad f(a,x)=0 \,\}$.
On utilisera aussi la locution *q-orthogonal* (ou même simplement *orthogonal*) à la place de orthogonal pour q.

2.3.2. Proposition. *Soient E un K-espace vectoriel, q une forme quadratique sur E.*

(*i*) *Pour toute partie A de E, $A^\perp$ est un sous-espace vectoriel de E.*

(*ii*) $\forall (A,B) \in (\mathcal{P}(E))^2 \quad A \subset B \Rightarrow B^\perp \subset A^\perp$.

(*iii*) $\forall A \in \mathcal{P}(E) \quad A^\perp = (\mathrm{Vect}(A))^\perp$.

(*iv*) $\{0\}^\perp = E$.

(*v*) *Pour tout sous-espace vectoriel F de E, on a $F \subset F^{\perp\perp}$.*

Preuve. Soit f la forme polaire de q.

(*i*) Clairement 0 appartient à $A^\perp$. Soient x et y dans $A^\perp$, λ et μ dans K.
Pour tout $a \in A$, on a $f(\lambda x + \mu y, a) = \lambda f(x,a) + \mu f(y,a) = 0$.
Donc $\lambda x + \mu y$ appartient à $A^\perp$.

(*ii*) Soit $y \in B^\perp$. Pour tout $a \in A$, on a $f(y,a)=0$ car a appartient aussi à B. Donc $y \in A^\perp$.

(*iii*) D'après l'assertion précédente, l'inclusion $A \subset \mathrm{Vect}(A)$, implique $(\mathrm{Vect}(A))^\perp \subset A^\perp$.
Montrons l'inclusion inverse. Soit $x \in A^\perp$. Tout vecteur $y \in \mathrm{Vect}(A)$ s'écrit comme combinaison linéaire de vecteurs de A, c'est à dire qu'il existe $a_1, \dots, a_n$ dans A et $\lambda_1, \dots, \lambda_n$ dans K tels que $y = \sum_{i=1}^{n} \lambda_i\, a_i$. D'où
$f(x,y) = f(x, \sum_{i=1}^{n} \lambda_i\, a_i) = \sum_{i=1}^{n} \lambda_i\, f(x,a_i) = 0$. Donc $x \in (\mathrm{Vect}(A))^\perp$.

(*iv*) Pour tout $x \in E$, on a $f(x,0)=0$ par linéarité de f par rapport à la deuxième variable.

(*v*) Pour tout $x \in F$ et pour tout $y \in F^\perp$, on a $f(x,y)=0$. Donc $x \in F^{\perp\perp}$. ∎

2.3.3. Définitions. Soient E un K-espace vectoriel, q une forme quadratique sur E et f sa forme polaire.
On appelle *noyau de* q (ou de f), noté $N(q)$, l'orthogonal de E pour q. On a donc :

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

On dit que q (ou f) est *non dégénérée* si son noyau est réduit à $\{0\}$; dans le cas contraire on dit que q (ou f) est *dégénérée*.
On dit qu'un vecteur $x \in E$ est *isotrope* si $q(x) = 0$.
On appelle *cône isotrope* de q, l'ensemble, noté $C(q)$, des vecteurs isotropes pour q.
On dit que q (ou f) est *définie* si $C(q) = \{0\}$.

2.3.4. Exemples.

a) Soient $E = \mathbb{R}^3$ et q telle que $q((x_1, x_2, x_3)) = x_1^2 + x_2^2 + x_3^2$.
Alors $f((x_1, x_2, x_3), (y_1, y_2, y_3)) = x_1y_1 + x_2y_2 + x_3y_3$ est le produit scalaire usuel. On a $N(q) = \{0\}$. Le cône isotrope $C(q)$ est réduit à $\{0\}$. La forme quadratique q est non dégénérée et définie.

b) Soient $E = \mathbb{R}^2$ et q telle que $q((x_1, x_2)) = x_1^2 - x_2^2$.
Alors $f((x_1, x_2), (y_1, y_2)) = x_1y_1 - x_2y_2$, d'où $N(q) = \{0\}$. La forme quadratique q est non dégénérée. Le cône isotrope $C(q)$ est la réunion de la droite Δ d'équation $x_1 - x_2 = 0$ et de la droite Δ' d'équation $x_1 + x_2 = 0$. La forme quadratique q est non définie.

c) Soient $E = \mathbb{R}^3$ et q telle que $q((x_1, x_2, x_3)) = x_1^2 - x_3^2$.
Alors $f((x_1, x_2, x_3), (y_1, y_2, y_3)) = x_1y_1 - x_3y_3$; d'où $N(q) = \mathbb{R}(0, 1, 0)$. La forme quadratique q est dégénérée. On a
$C(q) = \{(\lambda, \mu, \lambda) \in \mathbb{R}^3 \, ; \; (\lambda, \mu) \in \mathbb{R}^2 \, \} \cup \{(\lambda, \mu, \, -\lambda) \in \mathbb{R}^3 \, ; \; (\lambda, \mu) \in \mathbb{R}^2 \, \}$, réunion de deux plans. La forme quadratique q est non définie.

d) Soient $E = \mathbb{R}^3$ et q telle que $q((x_1, x_2, x_3)) = 2x_1^2 + x_2^2 + x_1x_2 - x_3^2$.
Alors $f((x_1, x_2, x_3), (y_1, y_2, y_3)) = 2x_1y_1 + x_2y_2 + x_1y_2 + x_2y_1 - x_3y_3$. En remarquant que $N(q) = \{\varepsilon_1, \varepsilon_2, \varepsilon_3\}^{\perp}$ on prouve que la forme quadratique q est non dégénérée. Le vecteur $(0, 1, 1)$ étant isotrope, la forme quadratique q est non définie.

2.3.5. Remarque. On a évidemment $N(q) \subset C(q)$, mais l'inclusion peut-être stricte (Exemple (b)). De plus on remarque que $C(q)$ n'est pas en général un sous-espace vectoriel.

2.3.6. Définition. Soient E un K-espace vectoriel de dimension finie et q une forme quadratique sur E. On appelle *rang* de q (ou de sa forme polaire f), l'entier noté $\mathrm{rang}(q)$ égal à $\dim(E) - \dim(N(q))$.

2.3.7. Théorème . *Soient E un K-espace vectoriel de dimension n et q une forme quadratique sur E. Soient $\mathcal{B}$ une base de E et $A = \mathrm{Mat}_{\mathcal{B}}(q)$.*

(*i*) *Soient $x \in E$ et $X = \mathrm{Mat}_{\mathcal{B}}(x)$. Alors $x \in N(q) \Leftrightarrow AX = 0$.*

(*ii*) $\mathrm{rang}(q) = \mathrm{rang}(A)$.

(*iii*) *q est non dégénérée si et seulement si* $\det A \neq 0$.

Preuve. Soient $\mathcal{B} = (e_1, \ldots, e_n)$ et $x = \sum_{j=1}^{n} x_j e_j$.

(*i*) Par bilinéarité de f forme polaire de q, on a les équivalences :

$$\begin{aligned} x \in N(q) \quad &\Leftrightarrow \quad \forall i \in [1,n]_{\mathbb{N}} \quad f(e_i, x) = 0 \\ &\Leftrightarrow \quad \forall i \in [1,n]_{\mathbb{N}} \quad \sum_{j=1}^{n} x_j\, f(e_i, e_j) = 0 \\ &\Leftrightarrow \quad \forall i \in [1,n]_{\mathbb{N}} \quad \sum_{j=1}^{n} a_{i\,j} x_j \ = 0 \\ &\Leftrightarrow \quad AX = 0 \end{aligned}$$

(*ii*) D'après l'assertion précédente et le théorème du rang, on a :
$\dim(N(q)) = \dim(\mathrm{Ker}\,(A)) = \dim(E) - \mathrm{rang}(A)$.
D'où $\mathrm{rang}(q) = \dim(E) - \dim(N(q)) = \mathrm{rang}(A)$.

(*iii*) q est non dégénérée si et seulement si $\dim(N(q)) = 0$, autrement dit, si et seulement si $rg(A) = n$, ce qui est équivalent à $\det A \neq 0$. ∎

2.4. Formes non dégénérées

2.4.1. Théorème . *Soit f une forme bilinéaire symétrique non dégénérée sur un K-espace vectoriel E de dimension finie.*

(*i*) *Pour tout $x \in E$, l'application φ_x de E dans K, définie par $\varphi_x(y) = f(x, y)$, pour tout $y \in E$, est une forme linéaire sur E.*

(*ii*) *L'application $\Phi : x \longmapsto \varphi_x$ de E dans E^* est un isomorphisme.*

Preuve. L'assertion (i) résulte de la linéarité de f par rapport à la deuxième variable.
(ii) Pour tous x_1 et x_2 dans E et tous λ et μ dans K on a, pour tout $y \in E$:
$\Phi(\lambda x_1 + \mu x_2)(y) = f(\lambda x_1 + \mu x_2, y) = \lambda f(x_1, y) + \mu f(x_2, y) = \lambda\Phi(x_1)(y) + \mu\Phi(x_2)(y) = (\lambda\Phi(x_1) + \mu\Phi(x_2))(y)$.

D'où $\Phi(\lambda x_1 + \mu x_2) = \lambda\Phi(x_1) + \mu\Phi(x_2)$. Donc Φ est linéaire.
Montrons que Φ est injective. Soit $x \in E$ tel que $\Phi(x) = 0$. Alors, pour tout $y \in E$, on a $0 = \Phi(x)(y) = f(x, y)$. Par conséquent x appartient au noyau de f qui est réduit à $\{0\}$, car f est non dégénérée. Donc $x = 0$.
Comme $\dim(E^*) = \dim(E)$, l'application linéaire Φ injective de E dans E^* est aussi surjective, c'est donc un isomorphisme. ∎

2.4.2. Corollaire . *Soient E un espace vectoriel de dimension finie, q une forme quadratique non dégénérée sur E et F un sous-espace vectoriel de E. Alors*

(*i*) $\dim(F) + \dim(F^\perp) = \dim(E)$.

(*ii*) $(F^\perp)^\perp = F$.

Preuve. Reprenons les notations du théorème précédent.

(*i*) Soit $(e_1, \ldots, e_p)$ une base de F, que nous complétons en une base $\mathcal{B} = (e_1, \ldots, e_n)$ de E. Désignons par $(e_1^*, \ldots, e_n^*)$ la base duale de $\mathcal{B}$. Soit $x \in E$, la forme linéaire $\Phi(x) = \varphi_x$ se décompose en :

$$\Phi(x) = \sum_{i=1}^{n} \varphi_x(e_i)\, e_i^* = \sum_{i=1}^{n} f(x, e_i)\, e_i^* \ .$$

Comme $F = \mathrm{Vect}(e_1, \ldots, e_p)$, le vecteur x est q-orthogonal à F si et seulement si pour tout $i \in [1, p]_{\mathbb{N}}$, $f(x, e_i) = 0$. C'est à dire : $x \in F^\perp \Leftrightarrow \Phi(x) \in \mathrm{Vect}(e_{p+1}^*, \ldots, e_n^*)$.
D'où $\Phi(F^\perp) = \mathrm{Vect}(e_{p+1}^*, \ldots, e_n^*)$. Comme Φ est un isomorphisme, on déduit

$$\dim(F^\perp) = \dim(\Phi(F^\perp)) = \dim(\mathrm{Vect}(e_{p+1}^*, \ldots, e_n^*)) = n - p = \dim(E) - \dim(F) \ .$$

(*ii*) On a déjà prouvé que $F \subset (F^\perp)^\perp$. Comme nous sommes en dimension finie, il suffit de montrer que ces deux sous-espaces vectoriels ont même dimension pour obtenir l'égalité. Or

$$\dim((F^\perp)^\perp) = \dim(E) - \dim(F^\perp) = \dim(E) - (\dim(E) - \dim(F)) = \dim(F) \ .$$ ∎

2.4.3. Remarque . Attention, en général F et $F^\perp$ ne sont pas supplémentaires. Reprenons l'exemple b) de 2.3.4 : $E = \mathbb{R}^2$ et q telle que $q((x_1, x_2)) = x_1^2 - x_2^2$, et soit $F = \mathbb{R}(1, 1)$. Alors (x_1, x_2) appartient à $F^\perp$ si et seulement s'il est q-orthogonal à $(1, 1)$ c'est à dire si $x_1 - x_2 = 0$. Donc $F^\perp = F$.

2.5. Bases orthogonales

2.5.1. Définitions. Soient E un espace vectoriel de dimension finie et q une forme quadratique sur E. Rappelons qu'une base $\mathcal{B} = (e_1, \ldots, e_n)$ de E est *orthogonale pour* q si pour tout couple (i, j) avec $i \neq j$, les vecteurs e_i et e_j sont orthogonaux pour q.
Si en outre, on a $q(e_i) = 1$, pour tout $i \in [1, n]_{\mathbb{N}}$, on dira que la base $\mathcal{B}$ est *orthonormale pour* q.

2.5.2. Proposition. *Soient E un espace vectoriel de dimension finie muni d'une forme quadratique q, $\mathcal{B} = (e_1, \ldots, e_n)$ une base de E et $\mathcal{B}^* = (e_1^*, \ldots, e_n^*)$ la base duale. Les conditions suivantes sont équivalentes :*

(*i*) *$\mathcal{B}$ est une base q-orthogonale ;*

(*ii*) $\mathrm{Mat}_{\mathcal{B}}(q)$ *est diagonale ;*

(*iii*) *q est combinaison linéaire des carrés des formes linéaires e_i^*.*

Preuve. Soit f la forme polaire de q.
$(i) \Rightarrow (ii)$ Pour tout $i \in [1, n]_{\mathbb{N}}$, posons $a_i = q(e_i) = f(e_i, e_i)$. Par hypothèse, pour $i \neq j$, on a $f(e_i, e_j) = 0$. Il en résulte que $\mathrm{Mat}_{\mathcal{B}}(q) = \mathrm{Mat}_{\mathcal{B}}(f) = \begin{pmatrix} a_1 & & & \\ & \ddots & \mathbf{0} & \\ & \mathbf{0} & \ddots & \\ & & & a_n \end{pmatrix}$.
$(ii) \Rightarrow (iii)$ Par hypothèse $\mathrm{Mat}_{\mathcal{B}}(q) = \mathbf{Diag}(a_1, \ldots, a_n)$. Pour tout $x = \sum_{i=1}^{n} x_i e_i$ de E, on a $q(x) = \sum_{i=1}^{n} a_i\, x_i^2 = \sum_{i=1}^{n} a_i\, (e_i^*)^2(x)$, car $x_i = e_i^*(x)$; d'où $q = \sum_{i=1}^{n} a_i (e_i^*)^2$.
$(iii) \Rightarrow (i)$ Si $q = \sum_{k=1}^{n} a_k (e_k^*)^2$, alors sa forme polaire f est définie, pour tout $(x, y) \in E^2$, par $f(x, y) = \sum_{k=1}^{n} a_k\, e_k^*(x) e_k^*(y)$. Par conséquent, si $i \neq j$ on a :
$f(e_i, e_j) = \sum_{k=1}^{n} a_k\, e_k^*(e_i) e_k^*(e_j) = \sum_{k=1}^{n} a_k\, \delta_{k\,i}\, \delta_{k\,j} = 0$. La base $\mathcal{B}$ est donc q-orthogonale. ∎

2.5.3. Remarque. Si $\mathcal{B}$ est une base q-orthogonale, l'expression de q dans cette base est simple : $q\,(\sum_{i=1}^{n} x_i e_i) = \sum_{i=1}^{n} a_i x_i^2$.
Le noyau de q est engendré par les vecteurs e_i tels que $a_i = 0$ et le rang de q est le nombre de a_i non nuls.

2.5.4. Corollaire. *Soient E un espace vectoriel de dimension finie muni d'une forme quadratique q, $\mathcal{B} = (e_1, \ldots, e_n)$ une base de E et $\mathcal{B}^* = (e_1^*, \ldots, e_n^*)$ la base duale. Les conditions suivantes sont équivalentes :*

(*i*) *$\mathcal{B}$ est une base q-orthonormale ;*

(*ii*) $\mathrm{Mat}_{\mathcal{B}}(q) = I_n$;

(*iii*) $q = \sum_{i=1}^{n} (e_i^*)^2$.

Preuve. En reprenant la démonstration précédente, on remarque que l'on a en plus $1 = q(e_i) = a_i$, pour tout $i \in [1, n]_{\mathbb{N}}$. ■

2.5.5. Corollaire. *Soient E un espace vectoriel de dimension finie, $\mathcal{F} = (\varphi_1, \ldots, \varphi_r)$ une famille libre de formes linéaires de E^* et $a_1, \ldots, a_r$ des scalaires non nuls.*
Posons $q = \sum_{i=1}^{r} a_i \varphi_i^2$. Alors :

(*i*) *q est de rang r.*

(*ii*) $N(q) = \bigcap_{i=1}^{r} \mathrm{Ker}(\varphi_i)$.

(*iii*) *Si on complète $\mathcal{F}$ pour obtenir une base $\mathcal{B}'$ de E^* et si on considère la base $\mathcal{B}$ de E dont $\mathcal{B}'$ est la base duale, alors $\mathcal{B}$ est orthogonale pour q.*

Preuve. Soit $\mathcal{B} = (e_1, \ldots, e_n)$ la base dont $\mathcal{B}' = (\varphi_1, \ldots, \varphi_n)$ est la base duale. Pour tout $i \in [r+1, n]_{\mathbb{N}}$, posons $a_1 = 0$. On peut alors écrire $q = \sum_{i=1}^{n} a_i \varphi_i^2 = \sum_{i=1}^{n} a_i (e_i^*)^2$.
D'après la proposition 2.5.2, la base $\mathcal{B}$ est q -orthogonale et

$$\mathrm{Mat}_{\mathcal{B}}(q) = \begin{pmatrix} a_1 & & & & & \\ & \ddots & & & \mathbf{0} & \\ & & a_r & & & \\ & & & 0 & & \\ & \mathbf{0} & & & \ddots & \\ & & & & & 0 \end{pmatrix}.$$

Comme $a_i \neq 0$, pour $1 \leq i \leq r$, le rang de q est r.
Soit f la forme polaire de q. On a $f(x, y) = \sum_{i=1}^{r} a_i \varphi_i(x) \varphi_i(y)$. D'où $N(q) \subset \bigcap_{i=1}^{r} \mathrm{Ker}(\varphi_i)$.
De plus $\dim(\bigcap_{i=1}^{r} \mathrm{Ker}(\varphi_i)) = n - r = \dim(N(q))$. D'où l'assertion (ii). ■

2.5.6. Méthode de Gauss. *Soit E un espace vectoriel de dimension finie. Il existe un algorithme pour décomposer toute forme quadratique sur E en combinaison linéaire de carrés de formes linéaires indépendantes.*

Etudions d'abord deux exemples dans $\mathbb{R}^3$:
Désignons par $\mathcal{B}_0$ la base canonique de $\mathbb{R}^3$ et par $\mathcal{B}_0^* = (\varepsilon_1^*, \varepsilon_2^*, \varepsilon_3^*)$ sa base duale.

- Soit q_1 définie par $q_1(x_1, x_2, x_3) = x_1^2 + 2x_2^2 + 8x_3^2 - 2x_1x_2 + 4x_1x_3$.
Il y a un terme en x_1^2, considérons cette expresion comme un trinôme du second degré en x_1 :

$$\begin{aligned} q_1(x_1, x_2, x_3) &= x_1^2 + 2(-x_2 + 2x_3)x_1 + (2x_2^2 + 8x_3^2) \\ &= (x_1 - x_2 + 2x_3)^2 - (-x_2 + 2x_3)^2 + (2x_2^2 + 8x_3^2) \\ &= (x_1 - x_2 + 2x_3)^2 + x_2^2 + 4x_3^2 + 4x_2x_3 \\ &= (x_1 - x_2 + 2x_3)^2 + (x_2 + 2x_3)^2 \end{aligned}$$

Posons $\varphi_1 = \varepsilon_1^* - \varepsilon_2^* + 2\varepsilon_3^*$ et $\varphi_2 = \varepsilon_2^* + 2\varepsilon_3^*$. Ces deux formes linéaires sont indépendantes et le calcul précédent prouve que $q_1 = \varphi_1^2 + \varphi_2^2$.
On a donc $\mathrm{rang}(q_1) = 2$. Complétons (φ_1, φ_2) pour obtenir une base $\mathcal{B}'$ de $(\mathbb{R}^3)^*$, par exemple par $\varphi_3 = \varepsilon_3^*$. La matrice de passage de $\mathcal{B}_0^*$ à $\mathcal{B}'$ est $Q = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 2 & 2 & 1 \end{pmatrix}$.
Soit $\mathcal{B}$ la base de E dont $\mathcal{B}'$ est la base duale. D'après la proposition 1.4.4, la matrice de passage de $\mathcal{B}_0$ à $\mathcal{B}$ est ${}^tQ^{-1}$ que l'on calcule. On obtient $\begin{pmatrix} 1 & 1 & -4 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix}$.
Donc $e_1 = \varepsilon_1 = (1, 0, 0)$, $e_2 = \varepsilon_1 + \varepsilon_2 = (1, 1, 0)$ et $e_3 = -4\varepsilon_1 - 2\varepsilon_2 + \varepsilon_3 = (-4, -2, 1)$ constituent une base q_1-orthogonale.

- Soit q_2 définie par $q_2(x_1, x_2, x_3) = x_1x_2 + x_1x_3 + 2x_2x_3$.
Il n'y a pas de carrés dans cette expression, mais un terme en x_1x_2 ; en considérant cette expresion comme un polynôme du premier degré en x_1, puis en x_2 on peut écrire : $q_2(x_1, x_2, x_3) = (x_2 + x_3)x_1 + 2x_2x_3 = (x_1 + 2x_3)x_2 + x_1x_3$. D'où

$$\begin{aligned} q_2(x_1, x_2, x_3) &= (x_2 + x_3)(x_1 + 2x_3) - 2x_3^2 \\ &= \frac{1}{4}\Big((x_1 + x_2 + 3x_3)^2 - (x_1 - x_2 + x_3)^2\Big) - 2x_3^2 \\ &= \frac{1}{4}(x_1 + x_2 + 3x_3)^2 - \frac{1}{4}(x_1 - x_2 + x_3)^2 - 2x_3^2 \end{aligned}$$

Posons $\psi_1 = \varepsilon_1^* + \varepsilon_2^* + 3\varepsilon_3^*$, $\psi_2 = \varepsilon_1^* - \varepsilon_2^* + \varepsilon_3^*$ et $\psi_3 = \varepsilon_3^*$; ces trois formes linéaires sont indépendantes et le calcul précédent prouve que $q_2 = \frac{1}{4}\psi_1^2 - \dfrac{1}{4}\psi_2^2 - 2\psi_3^2$.
On a donc $\mathrm{rang}(q_2) = 3$. La matrice de passage de $\mathcal{B}_0^*$ à $\mathcal{C}' = (\psi_1, \psi_2, \psi_3)$ est $R = \begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 3 & 1 & 1 \end{pmatrix}$. Soit $\mathcal{C}$ la base de E dont $\mathcal{C}'$ est la base duale. D'après la

proposition 1.4.4, la matrice de passage de $\mathcal{B}_0$ à C est ${}^tR^{-1}$ que l'on calcule. On obtient $\begin{pmatrix} \frac{1}{2} & \frac{1}{2} & -2 \\ \frac{1}{2} & -\frac{1}{2} & -1 \\ 0 & 0 & 1 \end{pmatrix}$.
Donc $e_1 = \frac{1}{2}\varepsilon_1 + \frac{1}{2}\varepsilon_2$, $e_2 = \frac{1}{2}\varepsilon_1 - \frac{1}{2}\varepsilon_2$ et $e_3 = -2\varepsilon_1 - \varepsilon_2 + \varepsilon_3$ constituent une base q_2-orthogonale.

• **Démonstration dans le cas général :** La démonstration se fait par récurrence sur la dimension de E. Si $\dim(E) = 1$ il n'y a rien à faire. Supposons que, pour tout $p < n$, on ait un algorithme pour décomposer toute forme quadratique sur un espace vectoriel de dimension p en combinaison linéaire de carrés de formes linéaires indépendantes, et considérons q forme quadratique sur E de dimension n rapporté à une base $\mathcal{B} = (e_1, \ldots, e_n)$. L'expresssion de $q(x)$ est un polynôme Q homogène de degré 2 en fonction des coordonnées $x_1, \ldots, x_n$ de x.

- Premier cas : Q contient un terme en x_i^2, par exemple x_1^2 (en permutant éventuellement deux vecteurs de la base). On peut alors écrire :

$$Q(x_1, \ldots, x_n) = a_1 x_1^2 + 2L(x_2, \ldots, x_n)\, x_1 + Q_1(x_2, \ldots, x_n) \ ,$$

avec $a_1 \in \mathbb{R}^*$, L polynôme homogène de degré 1 en fonction de $x_2, \ldots, x_n$, c'est à dire de la forme $L(x_2, \ldots, x_n) = \alpha_2\, x_2 + \cdots + \alpha_n\, x_n$ et Q_1 polynôme homogène de degré 2 en fonction de $x_2, \ldots, x_n$. D'où

$$Q(x_1, \ldots, x_n) = a_1 \left(x_1 + \frac{L(x_2, \ldots, x_n)}{a_1} \right)^2 - \frac{L(x_2, \ldots, x_n)^2}{a_1} + Q_1(x_2, \ldots, x_n) \ .$$

Par hypothèse de récurrence, la forme quadratique q' définie sur $\mathbb{R}^{n-1}$ par

$$q'(x_2, \ldots, x_n) = -\frac{L(x_2, \ldots, x_n)^2}{a_1} + Q_1(x_2, \ldots, x_n)$$

se décompose en $q' = \sum_{j=2}^{r} a_j\, \ell_j^2$ où $\ell_2, \ldots, \ell_r$ sont des formes linéaires indépendantes dans $(\mathbb{R}^{n-1})^*$.

Pour tout $j \in [2, r]_{\mathbb{N}}$, notons φ_j, la forme linéaire sur E qui à $x = \sum_{i=1}^{n} x_i\, e_i$, associe $\ell_j(x_2, \ldots, x_n)$ et φ_1 la forme linéaire définie par $\varphi_1(x) = x_1 + \dfrac{L(x_2, \ldots, x_n)}{a_1}$.

On a alors $q = \sum_{j=1}^{r} a_j\, \varphi_j^2$. Il reste à montrer que la famille $(\varphi_j)_{j=1\cdots r}$ est libre dans E^*. Soit $\sum_{j=1}^{r} \lambda_j\, \varphi_j = 0$ une combinaison linéaire nulle des (φ_j). On a d'abord $0 = \sum_{j=1}^{r} \lambda_j\, \varphi_j(e_1) = \lambda_1$. Puis, pour tout $(x_2, \ldots, x_n) \in \mathbb{R}^{n-1}$, en posant $z = \sum_{i=2}^{n} x_i\, e_i$, il vient $0 = \sum_{j=2}^{r} \lambda_j\, \varphi_j(z) = \sum_{j=2}^{r} \lambda_j\, \ell_j(x_2, \ldots, x_n)$, c'est à dire $\sum_{j=2}^{r} \lambda_j\, \ell_j = 0$. Comme les formes linéaires $\ell_2, \ldots, \ell_r$ sont indépendantes, on en déduit que $\lambda_j = 0$, pour tout $j \in [2, r]_{\mathbb{N}}$.

- Deuxième cas : Q ne contient aucun terme en x_i^2. Si $q = 0$ il n'y a rien à faire, sinon on peut supposer, par exemple, que Q contient $x_1 x_2$. On peut alors écrire :

$$Q(x_1, \ldots, x_n) = a\, x_1\, x_2 + L_1(x_3, \ldots, x_n)\, x_1 + L_2(x_3, \ldots, x_n)\, x_2 + Q_1(x_3, \ldots, x_n) \;,$$

avec $a \in \mathbb{R}^*$, L_1 et L_2 polynômes homogènes de degré 1 en fonction de $x_3, \ldots, x_n$ et Q_1 polynôme homogène de degré 2 en fonction de $x_3, \ldots, x_n$. D'où en posant $y = (x_3, \ldots, x_n)$,

$$Q(x_1, \ldots, x_n) = a\left(\left(x_1 + \frac{L_2(y)}{a}\right)\left(x_2 + \frac{L_1(y)}{a}\right) - \frac{L_1(y)L_2(y)}{a^2}\right) + Q_1(y) \;.$$

Or

$$4\left(x_1+\frac{L_2(y)}{a}\right)\left(x_2+\frac{L_1(y)}{a}\right) = \left(x_1+x_2+\frac{(L_1+L_2)(y)}{a}\right)^2 - \left(x_1-x_2+\frac{(L_2-L_1)(y)}{a}\right)^2 .$$

Notons φ_1 et φ_2 les formes linéaires sur E définies par $\varphi_1(x) = x_1+x_2+\dfrac{(L_1+L_2)(y)}{a}$ et $\varphi_2(x) = x_1 - x_2 + \dfrac{(L_2-L_1)(y)}{a}$.

Par hypothèse de récurrence, la forme quadratique q' définie sur $\mathbb{R}^{n-2}$ par :
$q'(x_3, \ldots, x_n) = Q_1(x_3, \ldots, x_n) - \dfrac{L_1(x_3, \ldots, x_n)L_2(x_3, \ldots, x_n)}{a}$
se décompose en $q' = \sum_{j=3}^{r} a_j\, \ell_j^2$ où $\ell_3, \ldots, \ell_r$ sont des formes linéaires indépendantes dans $(\mathbb{R}^{n-2})^*$.

Pour tout $j \in [3, r]_{\mathbb{N}}$, notons φ_j, la forme linéaire sur E qui à $x = \sum_{i=1}^{n} x_i\, e_i$, associe $\ell_j(x_3, \ldots, x_n)$. On obtient alors $q = \dfrac{a}{4}\varphi_1^2 - \dfrac{a}{4}\varphi_2^2 + \sum_{j=3}^{r} a_j\, \varphi_j^2$. Il reste à montrer que la famille $(\varphi_j)_{j=1\cdots r}$ est libre dans E^*. Soit $\sum_{j=1}^{r} \lambda_j\, \varphi_j = 0$ une combinaison linéaire nulle des (φ_j). On a d'abord $0 = \sum_{j=1}^{r} \lambda_j\, \varphi_j(e_1) = \lambda_1 + \lambda_2$ et $0 = \sum_{j=1}^{r} \lambda_j\, \varphi_j(e_2) = \lambda_1 - \lambda_2$. D'où $\lambda_1 = \lambda_2 = 0$. Puis, pour tout $(x_3, \ldots, x_n) \in \mathbb{R}^{n-1}$, en posant $z = \sum_{i=3}^{n} x_i\, e_i$, il vient $0 = \sum_{j=3}^{r} \lambda_j\, \varphi_j(z) = \sum_{j=3}^{r} \lambda_j\, \ell_j(x_3, \ldots, x_n)$, c'est à dire $\sum_{j=3}^{r} \lambda_j\, \ell_j = 0$. Comme les formes linéaires $\ell_3, \ldots, \ell_r$ sont indépendantes, on en déduit que $\lambda_j = 0$, pour tout $j \in [3, r]_{\mathbb{N}}$. ∎

2.5.7. Théorème. *Soient E un espace vectoriel de dimension finie et q une forme quadratique sur E. Alors il existe au moins une base de E orthogonale pour q.*

Preuve. Cela résulte immédiatement de la décomposition obtenue par la méthode de Gauss et du corollaire 2.5.5. ∎

2.5.8. Corollaire. *Soit $n \in \mathbb{N}^*$. Pour toute matrice symétrique $S \in \boldsymbol{S}_n(K)$, il existe une matrice diagonale $D \in \boldsymbol{D}_n(k)$ et une matrice inversible $P \in \boldsymbol{GL}_n(K)$ telles que $S = {}^tPDP$.*

Preuve. Soit f la forme bilinéaire symétrique sur K^n admettant S pour matrice relativement à la base canonique $\mathcal{B}_0$ et q la forme quadratique associée. D'après le théorème précédent, il existe $\mathcal{B}$ base q-orthognale de K^n. Soient $D = \mathrm{Mat}_{\mathcal{B}}(f)$ et Q la matrice de passage de $\mathcal{B}_0$ à $\mathcal{B}$. Le théorème 2.1.8 de changement de bases pour une forme bilinéaire, nous donne :

$$D = \mathrm{Mat}_{\mathcal{B}}(f) = {}^tQ\,\mathrm{Mat}_{\mathcal{B}_0}(f)\;Q = {}^tQ\,S\,Q\ .$$

D'où le résultat en posant $P = Q^{-1}$. ∎

2.6. Réduction des formes quadratiques

Réduire une forme quadratique q, c'est déterminer une base q-orthogonale relativement à laquelle la matrice de q est la « plus simple possible ». La réduction dépend du corps de base. Etudions tout d'abord les formes quadratiques dans le cas complexe.

2.6.1. Proposition. *Soient E un espace vectoriel de dimension finie n sur $\mathbb{C}$ et q une forme quadratique sur E. Il existe un entier $r \in [0, n]_{\mathbb{N}}$ et une base $\mathcal{B} = (e_1, \ldots, e_n)$ orthogonale pour q tels que $q(e_i) = 1$ pour $i \leq r$ et $q(e_i) = 0$ pour $i > r$.*
L'entier r est le rang de q et ne dépend donc pas de la base $\mathcal{B}$.

Preuve. Soit $(e'_1, \ldots, e'_n)$ une base de E orthogonale pour q. Posons $a_i = q(e'_i)$ pour tout $i \in [1, n]_{\mathbb{N}}$. Nous avons déjà remarqué que le rang r de q est le nombre de a_i non nuls. Quitte à réordonner la base nous pouvons supposer que $a_i \neq 0$ pour $i \leq r$ et $a_i = 0$ pour $i > r$. Tout nombre complexe admettant une racine carrée, pour $i \leq r$, soit $b_i \in \mathbb{C}$ tel que $b_i^2 = a_i$. Posons enfin, $e_i = \dfrac{1}{b_i}e'_i$ pour $i \leq r$ et $e_i = e'_i$ pour $i > r$. Alors $(e_1, \ldots, e_n)$ est une base orthogonale pour q et de plus $q(e_i) = q\left(\dfrac{1}{b_i}e'_i\right) = \left(\dfrac{1}{b_i}\right)^2 q(e'_i) = \dfrac{a_i}{a_i} = 1$, pour $i \leq r$ et $q(e_i) = q(e'_i) = 0$, pour $i > r$. ∎

2.6.2. Remarque. L'expression de q dans une telle base est très simple puisque :

$$\boxed{q\left(\sum_{i=1}^{n} x_i e_i\right) = \sum_{i=1}^{r} x_i^2 \quad \text{et} \quad \mathrm{Mat}_{\mathcal{B}}(q) = \begin{pmatrix} I_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}}$$

2.6.3. Corollaire. *Soient E un espace vectoriel de dimension finie n sur $\mathbb{C}$ et q une forme quadratique sur E. Alors il existe une base orthonormale pour q si et seulement si q est non dégénérée.*

Preuve. Cela résulte immédiatement de la proposition précédente car q est non dégénérée si et seulement si $r = n$. ∎

Etudions maintenant le cas des formes quadratiques dans le cas réel.

2.6.4. Théorème de Sylvester. *Soient E un espace vectoriel de dimension finie n sur $\mathbb{R}$ et q une forme quadratique sur E. Il existe des entiers s et t une base $\mathcal{B} = (e_1, \ldots, e_n)$ orthogonale pour q tels que $q(e_i) = 1$ pour $1 \leq i \leq s$, $q(e_i) = -1$ pour $s < i \leq s+t$ et $q(e_i) = 0$ pour $s + t < i \leq n$.*
De plus le couple (s,t) ne dépend pas de la base choisie et $s + t = \mathrm{rang}(q)$.

Preuve. Soit $(e'_1, \ldots, e'_n)$ une base de E orthogonale pour q. Posons $a_i = q(e'_i)$, pour tout $i \in [1,n]_{\mathbb{N}}$. Nous avons déjà remarqué que le rang r de q est le nombre de a_i non nuls. Quitte à réordonner la base nous pouvons supposer que $a_i > 0$ pour $i \leq s$, $a_i < 0$ pour $s < i \leq r$ et $a_i = 0$ pour $i > r$. Posons $t = r - s$ et

$$e_i = \begin{cases} \dfrac{1}{\sqrt{a_i}}\, e'_i & \text{si } 1 \leq i \leq s\,, \\ \dfrac{1}{\sqrt{-a_i}}\, e'_i & \text{si } s < i \leq r\,, \\ e'_i & \text{si } r > i\,. \end{cases}$$

Il est clair que $(e_1, \ldots, e_n)$ est une base possédant les propriétés voulues. Reste à prouver l'unicité du couple (s,t).
Soit une autre base $(f_1, \ldots, f_n)$ orthogonale pour q telle que $q(f_i) = 1$ pour $1 \leq i \leq s'$, $q(f_i) = -1$ pour $s' < i \leq s' + t'$ et $q(f_i) = 0$ pour $i > s' + t'$. Notons F le sous-espace vectoriel engendré par les vecteurs $e_1, \ldots, e_s$ et G le sous-espace vectoriel engendré par les vecteurs $f_{s'+1}, \ldots, f_n$. Montrons que $F \cap G = \{0\}$. Soit $x \in F \cap G$. Ce vecteur se décompose en : $x = \sum_{i=1}^{s} x_i\, e_i = \sum_{j=s'+1}^{n} y_j\, f_j$. D'où, d'une part $q(x) = q\left(\sum_{i=1}^{s} x_i\, e_i\right) = \sum_{i=1}^{s} x_i^2 \geq 0$ et d'autre part $q(x) = q\left(\sum_{j=s'+1}^{n} y_j\, f_j\right) = \sum_{j=s'+1}^{s'+t'} -(y_j)^2 \leq 0$. Donc $q(x) = 0$ et par suite $x_i = 0$, pour tout $i \in [1,s]_{\mathbb{N}}$. Il en résulte que $x = 0$. Les sous-espaces vectoriels F et G étant en somme directe, on a $\dim(F) + \dim(G) \leq \dim(E)$, c'est à dire $s + n - s' \leq n$. D'où $s \leq s'$. En échangeant le rôle des deux bases, on obtient $s' \leq s$. Ainsi $s = s'$. Comme par ailleurs $s + t = \mathrm{rang}(q) = s' + t'$, il vient $t = t'$. ∎

2.6.5. Définition . Soient E un espace vectoriel de dimension finie n sur $\mathbb{R}$ et q une forme quadratique sur E. On appelle *signature* de q, le couple (s,t) défini dans le théorème précédent. On note $\operatorname{sgn}(q) = (s,t)$.

2.6.6. Remarques .

a) Soit q une forme quadratique sur un espace vectoriel réel de dimension finie n, de signature (s,t) . Il existe une base $\mathcal{B} = (e_1, \ldots, e_n)$ de E telle que

$$\boxed{q\left(\sum_{i=1}^{n} x_i\, e_i\right) = \sum_{i=1}^{s} x_i^2 - \sum_{i=s+1}^{s+t} x_i^2 \quad \text{et} \quad \operatorname{Mat}_{\mathcal{B}}(q) = \begin{pmatrix} I_s & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -I_t & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix}}$$

q est non dégénérée si et seulement si $s + t = n$.

b) *Si E est un $\mathbb{R}$-espace vectoriel de dimension finie et q une forme quadratique sur E qui se décompose en $q = \sum_{i=1}^{r} a_i\, \varphi_i^2$, avec $a_1, \ldots, a_r$ non nuls dans $\mathbb{R}$ et $\varphi_1, \ldots, \varphi_r$ formes linéaires <u>indépendantes</u> dans E^*, alors la signature de q est (s,t) où s est le nombre de a_i tels que $a_i > 0$ et t le nombre de a_i tels que $a_i < 0$.*
Pour les formes quadratiques décomposées en 2.5.6 on a donc $\operatorname{sgn}(q_1) = (2,0)$ et $\operatorname{sgn}(q_2) = (1,2)$.

2.6.7. Corollaire . *Soient E un espace vectoriel de dimension finie n sur $\mathbb{R}$ et q une forme quadratique sur E. Il existe une base orthonormale pour q si et seulement si la signature de q est $(n,0)$.*

Preuve. Immédiat d'après la remarque. ∎

2.6.8. Remarque . Pour une forme quadratique q de signature $(n,0)$, (c'est le cas de la forme quadratique associée au produit scalaire usuel de $\mathbb{R}^n$ par exemple) il existe une base $(e_1, \ldots, e_n)$ de E telle que $q\left(\sum_{i=1}^{n} x_i\, e_i\right) = \sum_{i=1}^{n} x_i^2$. Donc q est définie et pour tout x, on a $q(x) \geq 0$; cequi nous conduit à poser la définition suivante.

2.6.9. Définition . Soient E un espace vectoriel réel de dimension finie, f une forme bilinéaire symétrique sur E et q la forme quadratique associée.
On dit que f (ou q) est *positive* (respectivement *négative*), si pour tout $x \in E$ on a $q(x) \geq 0$ (respectivement $q(x) \leq 0$).

2.6.10. Remarque . Une forme quadratique q sur un $\mathbb{R}$-espace vectoriel de dimension finie est positive (respectivement négative) si et seulement si sa signature est de la forme $(r,0)$ (respectivement $(0,r)$).

3 Espace Euclidien

Dans toute cette partie le corps de base sera $\mathbb{R}$.

3.1. Produit scalaire

3.1.1. Définitions. Soit E un espace vectoriel réel. On appelle *produit scalaire* sur E, noté généralement $\langle\ ,\ \rangle$ toute forme bilinéaire symétrique définie positive sur E. On dit que E est un *espace préhilbertien réel* s'il est muni d'un produit scalaire. On appelle *espace euclidien* tout espace préhilbertien réel de dimension finie.

3.1.2. Remarque. Rappelons les propriétés de $\langle\ ,\ \rangle$ forme bilinéaire symétrique définie positive :

- $\forall (x,y) \in E^2 \quad \langle\, y\,,\,x\,\rangle = \langle\, x\,,\,y\,\rangle$.
- $\forall (x,y,z) \in E^3 \quad \forall (\lambda,\mu) \in \mathbb{R}^2 \quad \langle\, \lambda x + \mu z\,,\,y\,\rangle = \lambda\langle\, x\,,\,y\,\rangle + \mu\langle\, z\,,\,y\,\rangle$.
- $\forall x \in E \quad \langle\, x\,,\,x\,\rangle \geq 0$.
- $\forall x \in E \quad (\,\langle\, x\,,\,x\,\rangle = 0 \Rightarrow x = 0\,)$.

3.1.3. Exemples.

a) $E = \mathcal{C}([a,b],\mathbb{R})$, avec $a < b$, est un espace préhilbertien réel pour le produit scalaire défini par : $\langle\, f\,,\,g\,\rangle = \displaystyle\int_a^b f(t)g(t)dt$.
La symétrie de l'application est immédiate et le caractère bilinéaire est conséquence de la linéarité de l'intégrale. Pour tout $f \in E$ on a $\langle\, f\,,\,f\,\rangle = \displaystyle\int_a^b f^2(t)dt \geq 0$ par positivité de l'intégrale. De plus, comme f^2 est une fonction continue et positive, la nullité de l'intégrale ne peut se produire que si $f^2 = 0$, c'est à dire $f = 0$.

b) $\mathbb{R}^n$ muni du produit scalaire usuel $\langle\, x\,,\,y\,\rangle = \displaystyle\sum_{i=1}^n x_i y_i$ est un espace euclidien.
Nous avons déjà remarqué que le produit scalaire usuel est une forme bilinéaire symétrique définie ; elle est clairement positive.

c) $\boldsymbol{M}_n(\mathbb{R})$ muni de $\langle\, M\,,\,N\,\rangle = \mathrm{Tr}({}^tM\,N)$ est un espace euclidien. Pour tous M et N dans $\boldsymbol{M}_n(\mathbb{R})$, on a $\langle\, N\,,\,M\,\rangle = \mathrm{Tr}({}^tN\,M) = \mathrm{Tr}({}^t({}^tN\,M)) = \mathrm{Tr}({}^tM\,N) = \langle\, M\,,\,N\,\rangle$; d'où la symétrie. Le caractère bilinéaire est alors conséquence de la linéarité de la trace. De plus, $\mathrm{Tr}({}^tM\,M) = \displaystyle\sum_{i=1}^n ({}^tM\,M)_{i\,i} = \sum_{i=1}^n \sum_{k=1}^n {}^tM_{i\,k}\,M_{k\,i} = \sum_{i=1}^n \sum_{k=1}^n M_{k\,i}^2 \geq 0$.
La nullité de cette somme équivaut alors à $M_{k\,i} = 0$ pour tout $(i,k) \in [1,n]_{\mathbb{N}}^2$, c'est à dire $M = 0$.

d) $\mathbb{R}_n[X]$ muni de $\langle P , Q \rangle = \int_a^b P(t)Q(t)dt$, avec $a < b$, est un espace euclidien.
Les justifications sont les mêmes que pour l'exemple (a) sauf pour le caractère défini qui nécessite un argument supplémentaire. En effet, si $\langle P , P \rangle = 0$, on déduit comme précédemment que pour tout $t \in [a, b]$ on a $P(t) = 0$; il faut alors remarquer que le polynôme P admet un nombre infini de racines et est donc le polynôme nul. (Un polynôme de degré d admet au plus d racines).

3.1.4. Théorème. **(Inégalité de Cauchy-Schwarz)** *Soient x et y des vecteurs d'un espace préhilbertien réel $(E, \langle\ ,\ \rangle)$. Alors*

(*i*) $|\langle x , y \rangle| \leq \langle x , x \rangle^{\frac{1}{2}} \langle y , y \rangle^{\frac{1}{2}}$;

(*ii*) *l'inégalité précédente est une égalité si et seulement si x et y sont liés.*

Preuve. (i) Puisque le produit scalaire est une forme bilinéaire symétrique positive, pour tout $\lambda \in \mathbb{R}$ on a :

$$\begin{array}{rcl} 0 & \leq & \langle x + \lambda y , x + \lambda y \rangle \\ & \leq & \langle x , x \rangle + \langle \lambda y , \lambda y \rangle + 2\langle x , \lambda y \rangle \\ & \leq & \langle y , y \rangle \lambda^2 + 2\langle x , y \rangle \lambda + \langle x , x \rangle . \end{array}$$

Ce trinôme à coefficients réels, étant positif ou nul pour toute valeur de λ, son discriminant doit être négatif ou nul i.e. $4(\langle x , y \rangle)^2 - 4\langle y , y \rangle \langle x , x \rangle \leq 0$. D'où l'inégalité.
(ii) Si x est nul alors x et y sont liés et il est clair que l'on a égalité. Supposons maintenant $x \neq 0$; on a donc $\langle x , x \rangle \neq 0$.
Si x et y sont liés, il existe $\lambda \in \mathbb{R}$ tel que $y = \lambda x$ et on a d'une part,
$|\langle x , y \rangle| = |\lambda \langle x , x \rangle| = |\lambda| \langle x , x \rangle$,
d'autre part $\langle x , x \rangle^{\frac{1}{2}} \langle y , y \rangle^{\frac{1}{2}} = \langle x , x \rangle^{\frac{1}{2}} (\lambda^2 \langle x , x \rangle)^{\frac{1}{2}} = |\lambda| \langle x , x \rangle$. D'où une égalité.
Réciproquement, supposons que l'on ait une égalité. Posons $\alpha = \dfrac{\langle x , y \rangle}{\langle x , x \rangle}$, on a alors :

$$\begin{array}{rcl} \langle y - \alpha x , y - \alpha x \rangle & = & \langle y , y \rangle + \alpha^2 \langle x , x \rangle - 2\,\alpha \langle x , y \rangle \\ & = & \langle y , y \rangle + \dfrac{\langle x , y \rangle^2}{\langle x , x \rangle} - 2\,\dfrac{\langle x , y \rangle^2}{\langle x , x \rangle} \\ & = & \langle y , y \rangle - \dfrac{\langle x , y \rangle^2}{\langle x , x \rangle} \\ & = & 0 . \end{array}$$

Le produit scalaire étant une forme bilinéaire symétrique définie, il en résulte que $y = \alpha\, x$, c'est à dire que x et y sont liés. ∎

3.1.5. Corollaire. **(Inégalité de Minkowski)** *Soient x et y des vecteurs d'un espace préhilbertien réel $(E, \langle\ ,\ \rangle)$. Alors*

(i) $\langle x+y\,,\,x+y\rangle^{\frac{1}{2}} \leq \langle x\,,\,x\rangle^{\frac{1}{2}} + \langle y\,,\,y\rangle^{\frac{1}{2}}$;

(ii) *l'inégalité précédente est une égalité si et seulement si x et y sont positivement liés (i.e. $x = 0$ ou il existe $\lambda \in \mathbb{R}^+$ tel que $y = \lambda x$).*

Preuve. (i) On a

$$\begin{aligned} \langle x+y\,,\,x+y\rangle &= \langle x\,,\,x\rangle + \langle y\,,\,y\rangle + 2\langle x\,,\,y\rangle \\ &\leq \langle x\,,\,x\rangle + \langle y\,,\,y\rangle + 2\,|\langle x\,,\,y\rangle| \\ &\leq \langle x\,,\,x\rangle + \langle y\,,\,y\rangle + 2\langle x\,,\,x\rangle^{\frac{1}{2}}\langle y\,,\,y\rangle^{\frac{1}{2}} \qquad (\text{ Cauchy-Schwarz }) \\ &\leq \left(\langle x\,,\,x\rangle^{\frac{1}{2}} + \langle y\,,\,y\rangle^{\frac{1}{2}}\right)^2 . \end{aligned}$$

(ii) Si $x = 0$ alors x et y sont positivement liés et il est clair que l'on a égalité. Supposons donc $x \neq 0$.
Si $y = \lambda\, x$ avec $\lambda \in \mathbb{R}^+$ alors :
$\langle x\,,\,x\rangle^{\frac{1}{2}} + \langle y\,,\,y\rangle^{\frac{1}{2}} = \langle x\,,\,x\rangle^{\frac{1}{2}} + |\lambda|\langle x\,,\,x\rangle^{\frac{1}{2}} = (1+\lambda)\langle x\,,\,x\rangle^{\frac{1}{2}}$ et
$\langle x+y\,,\,x+y\rangle^{\frac{1}{2}} = \langle x+\lambda x\,,\,x+\lambda x\rangle^{\frac{1}{2}} = \left((1+\lambda)^2\langle x\,,\,x\rangle\right)^{\frac{1}{2}} = (1+\lambda)\langle x\,,\,x\rangle^{\frac{1}{2}}$.
Réciproquement, supposons que l'on ait une égalité dans l'inégalité de Minkowski. Dans la démonstration de (i), toutes les inégalités doivent être des égalités. On doit donc avoir d'une part, égalité dans l'inégalité de Cauchy-Schwarz, d'où l'existence d'un $\lambda \in \mathbb{R}$ tel que $y = \lambda x$, et d'autre part $\langle x\,,\,y\rangle = |\langle x\,,\,y\rangle|$, ce qui impose $\lambda \in \mathbb{R}^+$. ∎

3.1.6. Corollaire. *Soit $(E, \langle\ ,\ \rangle)$ un espace préhilbertien réel. L'application $\|\ \|$ de E dans $\mathbb{R}^+$ définie par $\|x\| = \langle x\,,\,x\rangle^{\frac{1}{2}}$ est une norme sur E.*

Preuve. Le produit scalaire étant une forme bilinéaire symétrique définie positive, pour tout $x \in E$, on a l'équivalence : $\|x\| = 0 \Leftrightarrow x = 0$. De plus pour tout $\lambda \in \mathbb{R}$, on a $\|\lambda x\| = \langle \lambda x\,,\,\lambda x\rangle^{\frac{1}{2}} = (\lambda^2\langle x\,,\,x\rangle)^{\frac{1}{2}} = |\lambda|\,\|x\|$. L'inégalité de Minkowski est l'inégalité triangulaire. ∎

3.1.7. Définition. Soit $(E, \langle\ ,\ \rangle)$ un espace préhilbertien. On appelle *norme euclidienne* la norme définie dans le corollaire précédent.

3.1.8. Proposition. **(Identité du parallélogramme)** *Soient x et y des vecteurs d'un espace préhilbertien réel $(E, \langle\ ,\ \rangle)$. Alors*

$$\|x+y\|^2 + \|x-y\|^2 = 2\left(\|x\|^2 + \|y\|^2\right) .$$

Preuve. On a

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

■

3.1.9. Proposition. *Soit* $(E,\langle\ \ ,\ \ \rangle)$ *un espace préhilbertien réel. Pour tous* x *et* y *dans* E *on a les formules de polarisation :*

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

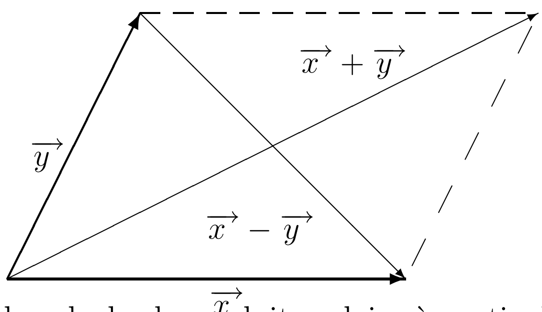
Preuve. Ce n'est que la réécriture des formules établies dans le lemme 2.2.1. ■

3.1.10. Remarques.

a) La forme quadratique q associée à un produit scalaire est le carré de la norme euclidienne : $\forall x\in E\quad q(x)=\|x\|^2$.
Dans le cas d'un espace euclidien de dimension n sa signature est $(n,0)$.

b) L'inégalité de Cauchy-Schwarz se traduit par $|\langle\, x\,,\, y\,\rangle|\le\|x\|\,\|y\|$.

c) L'identité du parallélogramme signifie que dans un parallélogramme, la somme des carrés des diagonales est égale à la somme des carrés des côtés.



d) Les formules de polarisation permettent de calculer le produit scalaire à partir de la norme ; on a par exemple : $\langle\, x\,,\, y\,\rangle=\frac{1}{4}\left(\ \|x+y\|^2-\|x-y\|^2\ \right)$.

e) L'identité du parallélogramme caractérise les normes issues d'un produit scalaire. En effet, si $(E,\|\ \|)$ est un espace vectoriel normé réel tel que l'identité du parallélogramme soit vérifiée, on démontre qu'on peut définir sur E un produit scalaire par la formule ci-dessus.

f) On appelle *espace de Hilbert réel* tout espace préhilbertien réel, complet pour la norme issue du produit scalaire. Un espace euclidien est un cas particulier d'espace de Hilbert, puisque tout espace vectoriel normé de dimension finie est complet.

Nous allons maintenant, dans le cas particulier des espaces préhilbertiens réels, étudier les propriétés spécifiques de la notion d'orthogonalité définie à la section 2.3..

3.2. Orthogonalité dans les espaces préhilbertiens réels

3.2.1. Rappels. Soit $(E, \langle\ ,\ \rangle)$ un espace préhilbertien réel.
On dit que deux vecteurs x et y de E sont *orthogonaux* si leur produit scalaire est nul.
Soit A une partie de E. On appelle *orthogonal* de A, noté $A^{\perp}$, l'ensemble des vecteurs de E orthogonaux à tout vecteur de A. $A^{\perp} = \{ x \in E \ ; \ \forall a \in A \quad \langle x , a \rangle = 0 \}$.
Soit $(e_i)_{i\in I}$ une famille de E. On dit que cette famille est *orthogonale* si pour tous i et j distincts dans I, les vecteurs e_i et e_j sont orthogonaux. Si de plus $\|e_i\| = 1$ pour tout $i \in I$, on dit que la famille est *orthonormale*.

3.2.2. Exemples.

a) Dans le cas de l'espace $E = \mathcal{C}([0, 2\pi], \mathbb{R})$ muni du produit scalaire défini par : $\langle f , g \rangle = \dfrac{1}{\pi}\displaystyle\int_0^{2\pi} f(t)g(t)dt$, considérons les fonctions $f_k : x \longmapsto \sin(kx)$, pour tout $k \in \mathbb{N}^*$. Alors la famille $(f_k)_{k\in\mathbb{N}^*}$ est orthonormale : En effet

$$\begin{aligned} \langle f_k , f_\ell \rangle &= \frac{1}{\pi}\int_0^{2\pi} \sin(kt)\sin(\ell t)dt = \frac{1}{\pi}\int_0^{2\pi} \frac{\cos(kt-\ell t) - \cos(kt+\ell t)}{2}\,dt \\ &= \frac{1}{2\pi}\left[\frac{\sin((k-\ell)t)}{k-\ell} - \frac{\sin((k+\ell)t)}{k+\ell}\right]_0^{2\pi} = 0 \quad \text{si} \quad k \neq \ell \\ &= \frac{1}{2\pi}\left[t - \frac{\sin(2kt)}{2k}\right]_0^{2\pi} = 1 \quad \text{si} \quad k = \ell\ . \end{aligned}$$

b) Dans l'espace euclidien $\mathbb{R}^n$ grâce au produit scalaire usuel, la base canonique est orthonormale.

c) Dans l'espace euclidien $\boldsymbol{M}_n(\mathbb{R})$ grâce au produit scalaire $\langle M , N \rangle = \mathrm{Tr}({}^tM\,N)$, la base canonique $(E_{i\,j})_{(i,j)\in[1,n]_{\mathbb{N}}^2}$ est orthonormale.

En effet : $\langle E_{i\,j} , E_{k\,\ell} \rangle = \mathrm{Tr}(E_{j\,i}\,E_{k\,\ell}) = \mathrm{Tr}(\delta_{i\,k}\,E_{j\,\ell}) = \delta_{i\,k}\,\delta_{j\,\ell} = \delta_{(i,j)\,(k,\ell)}$.

3.2.3. Théorème de Pythagore. *Soient x et y des vecteurs d'un espace préhilbertien réel E. Alors x et y sont orthogonaux si et seulement si $\|x+y\|^2 = \|x\|^2 + \|y\|^2$.*

Preuve. On a $\|x+y\|^2 = \|x\|^2 + \|y\|^2 + 2\langle x , y \rangle$. D'où l'équivalence :
$\|x+y\|^2 = \|x\|^2 + \|y\|^2 \Leftrightarrow \langle x , y \rangle = 0$. ∎

3.2.4. Proposition. *Soit $(e_j)_{j=1\cdots d}$ une famille orthogonale de vecteurs non nuls d'un espace préhilbertien réel. Alors cette famille est libre.*

Preuve. Si $\displaystyle\sum_{j=1}^{d} \lambda_j e_j$ est une combinaison linéaire nulle, alors pour tout $i \in [1, d]_{\mathbb{N}}$ on a

$0 = \langle\, e_i\,,\, \displaystyle\sum_{j=1}^{d} \lambda_j e_j \,\rangle = \displaystyle\sum_{j=1}^{d} \lambda_j \langle\, e_i\,,\, e_j\,\rangle = \lambda_i \langle\, e_i\,,\, e_i\,\rangle = \lambda_i \|e_i\|^2$. D'où $\lambda_i = 0$ car $\|e_i\| \neq 0$. ■

La réciproque de la proposition précédente est fausse. Cependant à partir d'une famille libre, nous pouvons construire une famille orthogonale ; c'est l'objet du théorème suivant.

3.2.5. Théorème. (Procédé d'orthogonalisation de Schmidt)
Soit $(f_i)_{i=1\cdots d}$ *une famille libre d'un espace préhilbertien réel* E.
Posons $e_1 = f_1$ *et* $e_{k+1} = f_{k+1} - \displaystyle\sum_{i=1}^{k} \frac{\langle\, e_i\,,\, f_{k+1}\,\rangle}{\langle\, e_i\,,\, e_i\,\rangle} e_i$, *pour tout* $k \in [1, d-1]_{\mathbb{N}}$.
Alors la famille $(e_i)_{i=1\cdots d}$ *est orthogonale et de plus pour tout entier* $k \in [1, d]_{\mathbb{N}}$, *on a* $\mathrm{Vect}\{e_1, \ldots, e_k\} = \mathrm{Vect}\{f_1, \ldots, f_k\}$.

Preuve. Montrons que nous pouvons construire e_k possédant les propriétés voulues par récurrence. Pour $k = 1$ c'est évident. Supposons la construction faite jusqu'au rang k. Puisque $\dim(\mathrm{Vect}\{e_1, \ldots, e_k\}) = \dim(\mathrm{Vect}\{f_1, \ldots, f_k\}) = k$, la famille $(e_i)_{i=1\cdots k}$ est libre et donc $\langle\, e_i\,,\, e_i\,\rangle \neq 0$. Le vecteur e_{k+1} est bien défini par la formule de l'énoncé. Pour $j \in [1, k]_{\mathbb{N}}$ on a :

$$\begin{aligned}
\langle\, e_j\,,\, e_{k+1}\,\rangle &= \langle\, e_j\,,\, f_{k+1} - \sum_{i=1}^{k} \frac{\langle\, e_i\,,\, f_{k+1}\,\rangle}{\langle\, e_i\,,\, e_i\,\rangle} e_i \,\rangle \\
&= \langle\, e_j\,,\, f_{k+1}\,\rangle - \sum_{i=1}^{k} \frac{\langle\, e_i\,,\, f_{k+1}\,\rangle}{\langle\, e_i\,,\, e_i\,\rangle} \langle\, e_j\,,\, e_i\,\rangle \\
&= \langle\, e_j\,,\, f_{k+1}\,\rangle - \langle\, e_j\,,\, f_{k+1}\,\rangle \\
&= 0\,.
\end{aligned}$$

La famille $(e_i)_{i=1\cdots k+1}$ est orthogonale.
Par définition e_{k+1} appartient à $\mathrm{Vect}\{f_{k+1}, e_1, \ldots, e_k\}$ qui, par hypothèse de récurrence, est égal à $\mathrm{Vect}\{f_{k+1}, f_1, \ldots, f_k\}$. Donc $\mathrm{Vect}\{e_1, \ldots, e_{k+1}\} \subset \mathrm{Vect}\{f_1, \ldots, f_{k+1}\}$. Comme on a également $f_{k+1} = \displaystyle\sum_{i=1}^{k} \frac{\langle\, e_i\,,\, f_{k+1}\,\rangle}{\langle\, e_i\,,\, e_i\,\rangle} e_i + e_{k+1}$, on déduit de même l'inclusion inverse. ■

3.2.6. Corollaire. *Tout sous-espace vectoriel de dimension finie d'un espace préhilbertien réel admet une base orthonormale. En particulier tout espace euclidien possède (au moins) une base orthonormale.*

Preuve. Soit $(f_i)_{i=1\cdots d}$ une base du sous-espace vectoriel F de dimension finie. Par le procédé d'orthogonalisation de Schmidt, nous obtenons $(e_i)_{i=1\cdots d}$ qui est une base orthogonale. Il suffit de poser $e'_i = \dfrac{1}{\|e_i\|} e_i$, pour obtenir une base orthonormale de F.
Dans le cas d'un espace euclidien E, il suffit d'appliquer le résultat à E lui-même. ■

3.2.7. Remarques .

a) La deuxième assertion de ce corollaire peut aussi se déduire du corolllaire 2.6.7. Mais le procédé d'orthogonalisation de Schmidt nous donne un algorithme, facilement programmable, pour construire une base orthonormale.

b) Dans le corollaire précédent la matrice de passage de la base (f_i) à la base orthonormale (e'_i) est triangulaire supérieure, car $\mathrm{Vect}\{e'_1, \ldots, e'_k\} = \mathrm{Vect}\{f_1, \ldots, f_k\}$.

c) Dans une base orthonormale $(e_i)_{i=1\cdots n}$, le produit scalaire des vecteurs $x = \sum_{i=1}^{n} x_i\, e_i$ et $y = \sum_{i=1}^{n} y_i\, e_i$ a pour expression $\boxed{\langle\, x \,,\, y \,\rangle = \sum_{i=1}^{n} x_i\, y_i \quad \text{et} \quad \|x\|^2 = \sum_{i=1}^{n} x_i^2 \,.}$

3.2.8. Théorème . *Soit F un sous-espace vectoriel de* <u>*dimension finie*</u> *d'un espace préhilbertien réel E. Alors*

(*i*) $F^\perp$ *est un supplémentaire de* F : $E = F \oplus F^\perp$.

(*ii*) $F^{\perp\perp} = F$.

Preuve. (i) D'après l'assertion(i) de la proposition 2.3.2, $F^\perp$ est un sous-espace vectoriel de E. Si $x \in F \cap F^\perp$ on a $\langle\, x \,,\, x \,\rangle = 0$, soit $x = 0$. Les sous-espaces vectoriels F et $F^\perp$ sont en somme directe.
Soit $(f_1, \ldots, f_p)$ une base de F. Grâce au procédé d'orthonormalisation de Schmidt, nous obtenons une base orthonormale $(e_1, \ldots, e_p)$ de F.
Pour tout $x \in E$, posons $y = \sum_{i=1}^{p} \langle\, e_i \,,\, x \,\rangle\, e_i \qquad (*)$ et $z = x - y$. On a évidemment $x = y + z$ avec $y \in F$. Vérifions que z appartient à $F^\perp$. D'après l'assertion (iii) de la proposition 2.3.2, il suffit de vérifier qu'il est orthogonal à tous les e_k. Pour $k \in [1, p]_{\mathbb{N}}$,

$$\begin{aligned} \langle\, e_k \,,\, z \,\rangle &= \langle\, e_k \,,\, x \,\rangle - \langle\, e_k \,,\, \sum_{i=1}^{p} \langle\, e_i \,,\, x \,\rangle\, e_i \,\rangle \\ &= \langle\, e_k \,,\, x \,\rangle - \sum_{i=1}^{p} \langle\, e_i \,,\, x \,\rangle \langle\, e_k \,,\, e_i \,\rangle \\ &= \langle\, e_k \,,\, x \,\rangle - \langle\, e_k \,,\, x \,\rangle \\ &= 0 \,. \end{aligned}$$

Remarquons que la formule $(*)$ nous donne à partir d'une base orthonormale de F, une expression de la projection y de x sur F et parallèlement à $F^\perp$.

(ii) L'inclusion $F \subset F^{\perp\perp}$ résulte de l'assertion (v) de la proposition 2.3.2. Montrons l'inclusion inverse. Soit $x \in F^{\perp\perp}$. Ce vecteur se décompose en $x = y + z$ avec $y \in F$ et $z \in F^\perp$. D'où $0 = \langle\, x \,,\, z \,\rangle = \langle\, y \,,\, z \,\rangle + \langle\, z \,,\, z \,\rangle = \|z\|^2$. Il en résulte que x est égal à y et appartient à F. ∎

3.2.9. Définitions. Soit F un sous-espace vectoriel de dimension finie d'un espace préhilbertien réel E. On dit que $F^\perp$ est *le supplémentaire orthogonal* de F dans E. On appelle *projection (vectorielle) orthogonale* sur F, la projection sur F parallèlement à $F^\perp$.
On dit que des sous-espaces vectoriels de E sont en *somme directe orthogonale* s'ils sont en somme directe et deux à deux orthogonaux.

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

(i) *Si $(e_1, \ldots, e_p)$ est une base orthonormale de F alors, pour tout vecteur x de E, on a* $P_F(x) = \sum_{i=1}^{p} \langle e_i , x \rangle e_i$.

(ii) *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* $\|x - P_F(x)\| = \min_{y \in F}\{\|x - y\|\}$.

Preuve. L'assertion (i) résulte de la remarque faite dans la démonstration du théorème précédent.
(ii) Soit $x \in E$. Alors $x - P_F(x)$ appartient à $F^\perp$ et pour tout $y \in F$ on a, d'après le théorème de Pythagore :

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

Donc $P_F(x)$ 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 $\|P_F(x)-y\|^2 = 0$, c'est à dire $y = P_F(x)$. ∎

Nous allons maintenant et dans toute la suite, nous placer en dimension finie, c'est à dire dans le cadre des espaces euclidiens. Reformulons d'abord les résultats obtenus.

3.2.11. Corollaire. *Soit E un espace euclidien.*

(i) *Si A est une partie de E alors $A^\perp$ est un sous-espace vectoriel de E ; de plus on a* $A^{\perp\perp} = \mathrm{Vect}(A)$.

(ii) *Si F est un sous-espace vectoriel de E alors $F^\perp$ est un supplémentaire de F. On a :* $E = F \oplus F^\perp$ *et* $\dim(E) = \dim(F) + \dim(F^\perp)$.

Preuve. La première partie de l'assertion (i) et l'assertion (ii) découlent immédiatement du théorème 3.2.8. Pour obtenir l'égalité de (i), posons $F = \mathrm{Vect}(A)$. On a alors $F^\perp = A^\perp$ et $F = F^{\perp\perp} = A^{\perp\perp}$. ∎

3.3. Adjoint d'un endomorphisme

3.3.1. Théorème de représentation de Riesz . *Soit E un espace euclidien.*

(*i*) *Pour tout $x \in E$, l'application $\varphi_x : y \longmapsto \langle\, x\,,\, y\,\rangle$ est une forme linéaire sur E.*

(*ii*) *L'application $\Phi : x \longmapsto \varphi_x$ est un isomorphisme de E dans E^*.*

Preuve. Le produit scalaire étant une forme bilinéaire symétrique non dégénérée, on peut appliquer le théorème 2.4.1. ∎

3.3.2. Corollaire . *Soit E un espace euclidien.*
Si $\mathcal{B} = (e_i)_{i=1\cdots n}$ est une base orthonormale de E, alors sa base duale est $(\langle\, e_i\,,\, \cdot\,\rangle)_{i=1\cdots n}$. D'où

$$\forall x \in E \quad x = \sum_{i=1}^{n} \langle\, e_i\,,\, x\,\rangle\, e_i \quad \text{et} \quad \forall \psi \in E^* \quad \psi = \sum_{i=1}^{n} \psi(e_i)\, \langle\, e_i\,,\, \cdot\,\rangle\ .$$

Preuve. D'après le théorème de Riesz, $\langle\, e_i\,,\, \cdot\,\rangle = \varphi_{e_i}$ est une forme linéaire sur E qui coïncide avec e_i^* sur la base $\mathcal{B}$ car $\langle\, e_i\,,\, e_j\,\rangle = \delta_{i\,j}$. ∎

3.3.3. Corollaire . *Soient E et F des espaces euclidiens et $u \in \mathcal{L}(E, F)$. Il existe une unique application linéaire u^* de F dans E telle que :*

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

Preuve. Soit y appartenant à F. L'application $x \longmapsto \langle\, y\,,\, u(x)\,\rangle$ appartient à E^*. D'après le théorème de Riesz, il existe un unique vecteur, que nous noterons $u^*(y)$ tel que pour tout $x \in E$ on ait $\langle\, u^*(y)\,,\, x\,\rangle = \langle\, y\,,\, u(x)\,\rangle$, soit encore en utilisant la symétrie $\langle\, u(x)\,,\, y\,\rangle = \langle\, x\,,\, u^*(y)\,\rangle$. Ceci prouve l'existence et l'unicité de u^*. Il reste à montrer que u^* est linéaire. Soient y et z dans F, λ et μ dans $\mathbb{R}$. Pour tout $x \in E$ on a

$$\begin{aligned}
\langle\, x\,,\, u^*(\lambda y + \mu z)\,\rangle &= \langle\, u(x)\,,\, \lambda y + \mu z\,\rangle \\
&= \lambda\langle\, u(x)\,,\, y\,\rangle + \mu\langle\, u(x)\,,\, z\,\rangle \\
&= \lambda\langle\, x\,,\, u^*(y)\,\rangle + \mu\langle\, x\,,\, u^*(z)\,\rangle \\
&= \langle\, x\,,\, \lambda u^*(y) + \mu u^*(z)\,\rangle
\end{aligned}$$

D'où $u^*(\lambda y + \mu z) = \lambda u^*(y) + \mu u^*(z)$. ∎

3.3.4. Définition . L'application linéaire u^* définie dans le corollaire précédent est appelée *adjoint* de u.

3.3.5. Proposition. *Soient E, F et G des espaces euclidiens.*

(*i*) *Pour tout $u \in \mathcal{L}(E,F)$ on a $u^{**} = u$.*

(*ii*) *L'application $u \longmapsto u^*$ est linéaire bijective de $\mathcal{L}(E,F)$ dans $\mathcal{L}(F,E)$.*

(*iii*) *Pour tout $u \in \mathcal{L}(E,F)$ et tout $v \in \mathcal{L}(F,G)$ on a $(v \circ u)^* = u^* \circ v^*$.*

(*iv*) *Pour tout $u \in \mathcal{L}(E,F)$ on a : $\mathrm{Ker}\,(u^*) = (\mathrm{Im}\,(u))^\perp$ et $\mathrm{Im}\,(u^*) = (\mathrm{Ker}\,(u))^\perp$.*

(*v*) *Soient $u \in \mathcal{L}(E)$ et $\mathcal{B}$ une base orthonormale de E. Relativement à la base $\mathcal{B}$, la matrice de u^* est la transposée de la matrice de u : $\mathrm{Mat}_{\mathcal{B}}\,(u^*) = {}^t(\mathrm{Mat}_{\mathcal{B}}\,(u))$.*

Preuve. (i) Pour tous x dans E et y dans F on a :

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

D'où $u^{**} = u$.
(ii) Soient u et v dans $\mathcal{L}(E,F)$, λ et μ dans $\mathbb{R}$. On a pour tous x et y dans F :

$$\begin{aligned}
\langle\, x\,,\, (\lambda u + \mu v)^*(y)\,\rangle &= \langle\, (\lambda u + \mu v)(x)\,,\, y\,\rangle \\
&= \lambda\,\langle\, (u(x)\,,\, y\,\rangle + \mu\,\langle\, v(x)\,,\, y\,\rangle \\
&= \lambda\,\langle\, x\,,\, u^*(y)\,\rangle + \mu\,\langle\, x\,,\, v^*(y)\,\rangle \\
&= \langle\, x\,,\, \lambda\, u^*(y) + \mu\, v^*(y)\,\rangle\,.
\end{aligned}$$

D'où : $(\lambda u + \mu v)^* = \lambda\, u^* + \mu\, v^*$.
Le caractère surjectif résulte de l'assertion (i), de plus $\mathcal{L}(E,F)$ et $\mathcal{L}(F,E)$ ont même dimension et par suite l'application $u \longmapsto u^*$ est bijective.
(iii) Pour tous $x \in E$ et $z \in G$ on a :

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

(iv) Pour tout $y \in F$ on a les équivalences :

$$\begin{aligned}
y \in \mathrm{Ker}\,(u^*) &\Leftrightarrow \forall x \in E \quad \langle\, u^*(y)\,,\, x\,\rangle = 0 \\
&\Leftrightarrow \forall x \in E \quad \langle\, y\,,\, u(x)\,\rangle = 0 \\
&\Leftrightarrow y \in (\mathrm{Im}\,(u))^\perp\,.
\end{aligned}$$

D'où $\mathrm{Ker}\,(u^*) = (\mathrm{Im}\,(u))^\perp$.
En l'appliquant cette égalité à u^*, on obtient $\mathrm{Ker}\,(u) = (\mathrm{Im}\,(u^*))^\perp$.
D'où $(\mathrm{Ker}\,(u))^\perp = (\mathrm{Im}\,(u^*))^{\perp\perp} = \mathrm{Im}\,(u^*)$.
(v) Soient $A = \mathrm{Mat}_{\mathcal{B}}\,(u)$ et $A' = \mathrm{Mat}_{\mathcal{B}}\,(u^*)$. Pour tous i, $j \in [1,n]_{\mathbb{N}}$, $a'_{i\,j}$ est la $\mathrm{i}^{\text{ème}}$-composante de $u^*(e_j)$ dans la base $(e_i)_{i=1\cdots n}$, c'est à dire :

$$a'_{i\,j} = e_i^*(u^*(e_j)) = \langle\, e_i\,,\, u^*(e_j)\,\rangle = \langle\, u(e_i)\,,\, e_j\,\rangle = \langle\, e_j\,,\, u(e_i)\,\rangle = e_j^*(u(e_i)) = a_{j\,i}\,. \qquad \blacksquare$$

3.4. Endomorphisme symétrique

3.4.1. Définitions . Soit u un endomorphisme d'un espace euclidien E. On dit que u est *autoadjoint* ou *symétrique* si $u^* = u$.
On note $\boldsymbol{S}(E)$ l'espace vectoriel des endomorphismes symétriques sur E.

3.4.2. Remarques .

a) *Soit u un endomorphisme d'un espace euclidien. Alors u est symétrique si et seulement si dans une (ou toute) base orthonormale, sa matrice est symétrique.*

b) *Un projecteur orthogonal est autoadjoint.*
En effet, soient F un sous-espace vectoriel de E espace euclidien et p le projecteur orthogonal sur F. Pour tout $x \in E$, on a $p(x) \in F$ et $(x - p(x)) \in F^\perp$. Donc pour x et y dans E, on a :

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

∎

3.4.3. Proposition . *Soit u un endomorphisme symétrique d'un espace euclidien E.*

(*i*) *Les sous-espaces propres de u sont deux à deux orthogonaux.*

(*ii*) *Si F est un sous-espace vectoriel stable par u, alors $F^\perp$ est stable par u.*

Preuve. (i) Soient λ et μ deux valeurs propres distinctes de u. Pour tous $x \in E_\lambda(u)$ et $y \in E_\mu(u)$, on a

$$\mu\,\langle\, x\,,\, y\,\rangle = \langle\, x\,,\, u(y)\,\rangle = \langle\, u^*(x)\,,\, y\,\rangle = \langle\, u(x)\,,\, y\,\rangle = \lambda\langle\, x\,,\, y\,\rangle\ .$$

Comme $\lambda \neq \mu$ on a donc $\langle\, x\,,\, y\,\rangle = 0$.
(ii) Supposons que $u(F) \subset F$. Soit $y \in F^\perp$. Pour tout $x \in F$ on a :
$\langle\, u(y)\,,\, x\,\rangle = \langle\, y\,,\, u(x)\,\rangle = 0$ car $u(x) \in F$ et $y \in F^\perp$.
Donc $u(F^\perp) \subset F^\perp$. ∎

3.4.4. Lemme . *Soient E un espace vectoriel réel de dimension finie, $n \geq 1$ et $u \in \mathcal{L}(E)$. Alors il existe dans E une droite ou un plan stable par u.*

Preuve. Si u possède un vecteur propre x alors la droite engendrée par x est stable par u. Sinon u ne possède pas de valeur propre et son polynôme minimal π_u n'a pas de racines réelles. Par conséquent π_u se factorise dans $\mathbb{R}$ en produit de polynômes irréductibles de

degré 2. On peut donc écrire $\pi_u = (X^2 + aX + b)Q$. Par minimalité du degré de π_u, le polynôme Q n'est pas annulateur pour u et il existe un vecteur $y \in E$ tel que $Q(u)(y) \neq 0$. Posons $z = Q(u)(y)$ et $P = \mathrm{Vect}\{\, z, u(z)\,\}$. Comme u ne possède pas de vecteur propre, $u(z)$ n'est pas colinéaire à z et P est un plan. On a :
$0 = \pi_u(u)(y) = (u^2 + au + b\mathrm{Id}_E)Q(u)(y) = (u^2 + au + b\mathrm{Id}_E)(z)$.
D'où $u^2(z) = -au(z) - bz$ appartient à P et par suite P est stable par u. ∎

3.4.5. Théorème. *Soient E un espace euclidien et $u \in \mathcal{L}(E)$. Les conditions suivantes sont équivalentes :*

(*i*) *L'endomorphisme u est symétrique.*

(*ii*) *Il existe une base orthonormale qui diagonalise u.*

Preuve. $(i) \Rightarrow (ii)$ La démonstration se fait par récurrence sur la dimension n de E.
Si $n = 1$ c'est évident.
Si $n = 2$. Soit $\mathcal{B}$ une base orthonormale de E. Alors $\mathrm{Mat}_{\mathcal{B}}(u)$ est de la forme $\begin{pmatrix} a & b \\ b & d \end{pmatrix}$ et le polynôme caractéristique de u est $X^2 - (a+d)X + ad - b^2$. Son discrimant Δ vaut $(a+d)^2 - 4(ad - b^2) = (a-d)^2 + 4b^2$. Si $b = 0$ la propriété est clairement établie, sinon $\Delta > 0$ et u admet deux valeurs propres distinctes. Par conséquent u est diagonalisable ; d'après la proposition 3.4.3 les sous-espaces propres de u sont orthogonaux et il existe donc une base orthonormale qui diagonalise u.
Soit $n > 2$ et supposons la propriété vraie jusqu'au rang $n-1$. Soient E un espace euclidien de dimension n et $u \in \mathcal{L}(E)$ un endomorphisme symétrique. D'après le lemme précédent u admet au moins un sous-espace stable F de dimension $d = 1$ ou 2. Désignons par v l'endomorphisme de F défini par restriction de u. Cet endomorphisme est symétrique car :
$\forall (x,y) \in F^2 \quad \langle\, v(x)\,,\, y\,\rangle = \langle\, u(x)\,,\, y\,\rangle = \langle\, x\,,\, u^*(y)\,\rangle = \langle\, x\,,\, u(y)\,\rangle = \langle\, x\,,\, v(y)\,\rangle$.
Il existe donc une base orthonormale $\mathcal{B}' = (ei)_{1 \leq i \leq d}$ de F qui diagonalise v. Autrement dit, il existe $(\lambda_i)_{1 \leq i \leq d}$ tel que $u(e_i) = v(e_i) = \lambda_i e_i$, pour $1 \leq i \leq d$.
Posons $G = F^\perp$. D'après la Proposition 3.4.3, G est stable par u. Désignons par w l'endomorphisme de G défini par restriction de u. Comme pour v, cet endomorphisme est symétrique. Puisque $\dim(G) = n - d \leq n - 1$, l'hypothèse de récurrence s'applique à w et il existe une base orthonormale $\mathcal{B}'' = (e_j)_{d+1 \leq j \leq n}$ de G qui diagonalise w. Autrement dit, il existe $\lambda_{d+1}, \ldots, \lambda_n$ réels tels que $u(e_j) = w(e_j) = \lambda_j e_j$, pour $d+1 \leq j \leq n$.
Il est clair que $\mathcal{B} = (e_1, e_2, \ldots, e_n)$ est une base orthonormale de E qui diagonalise u.
$(ii) \Rightarrow (i)$ Soit $\mathcal{B}$ une base orthonormale qui diagonalise u. La matrice de u relativement à la base $\mathcal{B}$ étant diagonale, elle est symétrique et par conséquent u est symétrique. ∎

3.4.6. Corollaire. *Toute matrice réelle symétrique est diagonalisable.*

Preuve. Soit $M \in \boldsymbol{S}_n(\mathbb{R})$. Considérons $E = \mathbb{R}^n$ muni du produit scalaire usuel et $\mathcal{B}$ la base canonique. Désignons par u l'endomorphisme de E tel que $\mathrm{Mat}_{\mathcal{B}}(u) = M$. Cet endomorphisme est symétrique ; d'après le théorème précédent, il existe une base orthonormale $\mathcal{B}'$ qui diagonalise u. Il existe donc une matrice P inversible telle que $P^{-1}MP$ soit diagonale. ∎

3.4.7. Remarque. Le résultat précédent est faux si le corps de base est $\mathbb{C}$: La matrice $A = \begin{pmatrix} 2 & i \\ i & 0 \end{pmatrix} \in \boldsymbol{M}_2(\mathbb{C})$ est symétrique mais A n'est pas diagonalisable ; en effet son polynôme caractéristique est $X^2 - 2X + 1 = (X-1)^2$ et le sous-espace propre relatif à la valeur propre 1 est de dimension 1, car $A \neq I_2$.

3.5. Endomorphisme orthogonal

3.5.1. Proposition. *Soient E un espace euclidien et $u \in \mathcal{L}(E)$. Les conditions suivantes sont équivalentes :*

(*i*) *u est isométrique : $\forall x \in E \quad \|u(x)\| = \|x\|$.*

(*ii*) *u préserve le produit scalaire : $\forall (x,y) \in E^2 \quad \langle u(x) , u(y) \rangle = \langle x , y \rangle$.*

(*iii*) $u^* \circ u = \mathrm{Id}_E$.

(*iv*) *u est bijective et $u^* = u^{-1}$.*

(*v*) *u transforme toute base orthonormale de E en une base orthonormale.*

(*vi*) *Il existe une base orthonormale $(e_i)_{i\in I}$ telle que $(u(e_i))_{i\in I}$ soit une base orthonormale.*

Preuve. $(i) \Rightarrow (ii)$ Pour tous x et y dans E, on a d'après la formule de polarisation :

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

$(ii) \Rightarrow (iii)$ Pour tous x et y dans E on a $\langle u^*u(x) , y \rangle = \langle u(x) , u(y) \rangle = \langle x , y \rangle$.
Le vecteur $u^*u(x) - x$ appartient donc à $E^\perp$ qui est réduit à $\{0\}$; d'où le résultat.

$(iii) \Rightarrow (iv)$ D'après (iii) l'endomorphisme u est injectif. Comme E est de dimension finie, u est bijectif. Donc u^{-1} existe et $u^{-1} = (u^* \circ u) \circ u^{-1} = u^*$.

$(iv) \Rightarrow (v)$ Soit $(e_i)_{i=1\cdots n}$ une base orthonormale de E. Alors pour tous i et j dans $[1,n]_{\mathbb{N}}$, on a :

$$\delta_{ij} = \langle e_i , e_j \rangle = \langle u^* \circ u(e_i) , e_j \rangle = \langle u(e_i) , u(e_j) \rangle .$$

La famille $(u(e_i))_{i=1\cdots n}$ est orthonormale, donc libre ; elle est constituée de n vecteurs, c'est donc une base.

$(v) \Rightarrow (vi)$ Immédiat car il existe des bases orthonormales.

$(vi) \Rightarrow (i)$ Soit $x \in E$. Il se décompose en $x = \sum_{i=1}^{n} \lambda_i \, e_i$. On a alors $u(x) = \sum_{i=1}^{n} \lambda_i \, u(e_i)$ et comme $(u(e_i))$ et (e_i) sont des bases orthonormales, $\|u(x)\|^2 = \sum_{i=1}^{n} |\lambda_i|^2 = \|x\|^2$. ∎

3.5.2. Définitions . Un endomorphisme $u \in \mathcal{L}(E)$ pour lequel les conditions de la proposition précédente sont satisfaites est dit *orthogonal*. L'ensemble des endomorphismes orthogonaux de E est noté $\boldsymbol{O}(E)$.
Une matrice $A \in \boldsymbol{M}_n(\mathbb{R})$ est dite *orthogonale* si l'endomorphisme de $\mathbb{R}^n$ représenté par A dans la base canonique, est un endomorphisme orthogonal de $\mathbb{R}^n$ muni du produit scalaire usuel. L'ensemble des matrices orthogonales de $\boldsymbol{M}_n(\mathbb{R})$ est noté $\boldsymbol{O}_n(\mathbb{R})$.

3.5.3. Corollaire . *Soit $A \in \boldsymbol{M}_n(\mathbb{R})$. Les conditions suivantes sont équivalentes :*

(*i*) $A \in \boldsymbol{O}_n(\mathbb{R})$.

(*ii*) ${}^tAA = I_n$.

(*iii*) *A est inversible et $A^{-1} = {}^tA$.*

(*iv*) *Les vecteurs colonnes de A constituent une base orthonormale de $\mathbb{R}^n$.*

Preuve. Le corollaire se déduit immédiatement de la proposition précédente. ∎

3.5.4. Remarques .

a) Une projection orthogonale distincte de Id_E n'est pas un endomorphisme orthogonal.

b) Le corollaire 3.4.6 de diagonalisation des matrices réelles symétriques, peut être précisé comme suit :
Pour toute matrice $A \in \boldsymbol{S}_n(\mathbb{R})$, il existe $P \in \boldsymbol{O}_n(\mathbb{R})$ telle que ${}^tPAP$ soit diagonale.

3.5.5. Proposition . *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.*

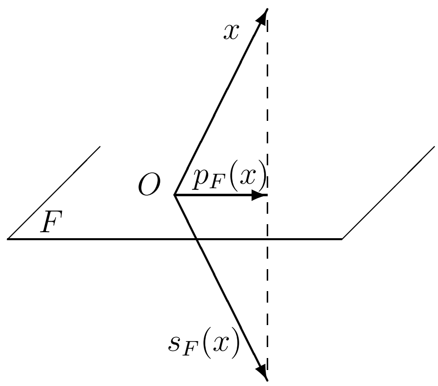
Preuve. On a $s_F^* = 2p_F^* - \mathrm{Id}_E = 2p_F - \mathrm{Id}_E = s_F$. Donc s_F est un endomorphisme symétrique.
Il est orthogonal car $s_F^* \circ s_F = (2p_F - \mathrm{Id}_E)^2 = 4p_F - 2p_F - 2p_F + \mathrm{Id}_E = \mathrm{Id}_E$. ∎

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

3.5.6. Définitions. Soit F un sous-espace vectoriel d'un espace euclidien E. On appelle *symétrie orthogonale* ou *symétrie vectorielle orthogonale* par rapport à F l'application s_F définie dans la proposition précédente.
Si F est un hyperplan on dit que s_F est une *réflexion* d'hyperplan F.

$p_F(x)$ projection orthogonale de x sur F

$s_F(x)$ symétrique orthogonal de x par rapport à F.



3.5.7. Proposition. *Soient E un espace euclidien et $u \in \boldsymbol{O}(E)$. Alors*

(*i*) $\mathrm{Sp}\,(u) \subset \{-1, 1\}$ *et les sous-espaces vectoriels (propres)* $E_1(u) = \mathrm{Ker}\,(u - \mathrm{Id}_E)$ *et* $E_{-1}(u) = \mathrm{Ker}\,(u + \mathrm{Id}_E)$ *sont orthogonaux.*

(*ii*) $\det(u) \in \{-1, 1\}$.

(*iii*) *Si F est un sous-espace vectoriel stable par u, alors $F^{\perp}$ est stable par u.*

Preuve. (i) Soit x un vecteur propre de u relatif à la valeur propre λ. On a donc $u(x) = \lambda x$. Comme u est isométrique, il vient $|\lambda| \|x\| = \|\lambda x\| = \|u(x)\| = \|x\|$. D'où $|\lambda| = 1$.
Soient maintenant $x \in E_1(u)$ et $y \in E_{-1}(u)$.
On a : $\langle\, x\,,\, y\,\rangle = \langle\, u(x)\,,\, u(y)\,\rangle = \langle\, x\,,\, -y\,\rangle = -\langle\, x\,,\, y\,\rangle$. D'où $\langle\, x\,,\, y\,\rangle = 0$.
(ii) On a $1 = \det(\mathrm{Id}_E) = \det(u^* u) = \det(u^*) \times \det(u) = (\det u)^2$.
(iii) Supposons que $u(F) \subset F$. Comme u est bijective, on a $\dim(u(F)) = \dim(F)$ et par suite $u(F) = F$. Soit $y \in F^{\perp}$, pour tout $x \in F$, il existe $z \in F$ tel que $x = u(z)$ et on a : $\langle\, u(y)\,,\, x\,\rangle = \langle\, u(y)\,,\, u(z)\,\rangle = \langle\, y\,,\, z\,\rangle = 0$. Il en résulte que $u(y)$ appartient à $F^{\perp}$. ∎

3.5.8. Définitions. Soient E un espace euclidien et $u \in \boldsymbol{O}(E)$. On dit que u est un endomorphisme orthogonal *direct* (resp. *indirect*) si $\det u = 1$ (resp. $\det u = -1$). On note $\boldsymbol{SO}(E)$ l'ensemble des endomorphismes orthogonaux directs de E.
Soit $n \in \mathbb{N}^*$. On pose $\boldsymbol{SO}_n(\mathbb{R}) = \boldsymbol{O}_n(\mathbb{R}) \cap \boldsymbol{SL}_n(\mathbb{R})$.
(On rappelle que $\boldsymbol{SL}_n(\mathbb{R}) = \{\, A \in \boldsymbol{M}_n(\mathbb{R})\ ;\ \det A = 1\,\}$).

3.5.9. Proposition.

(*i*) *Soit E un espace euclidien. Alors $\boldsymbol{O}(E)$ est un sous-groupe de $\boldsymbol{GL}(E)$ et $\boldsymbol{SO}(E)$ est un sous-groupe de $\boldsymbol{O}(E)$.*

(*ii*) *Soit* $n \in \mathbb{N}^*$. *Alors* $\boldsymbol{O}_n(\mathbb{R})$ *est un sous-groupe de* $\boldsymbol{GL}_n(\mathbb{R})$ *et* $\boldsymbol{SO}_n(\mathbb{R})$ *est un sous-groupe de* $\boldsymbol{O}_n(\mathbb{R})$.

Preuve. (i) Utilisons les résultats de la proposition 3.5.1. Tout élément de $\boldsymbol{O}(E)$ est inversible donc appartient à $\boldsymbol{GL}(E)$; de manière évidente Id_E appartient à $\boldsymbol{O}(E)$; si u et v sont dans $\boldsymbol{O}(E)$ alors pour tout $x \in E$ on a $\|v \circ u(x)\| = \|u(x)\| = \|x\|$, d'où $v \circ u \in \boldsymbol{O}(E)$, et $\|u^{-1}(x)\| = \|u(u^{-1}(x))\| = \|x\|$, d'où $u^{-1} \in \boldsymbol{O}(E)$. Donc $\boldsymbol{O}(E)$ est un sous-groupe de $\boldsymbol{GL}(E)$.
Il est clair que $\boldsymbol{SO}(E)$ est inclus dans $\boldsymbol{O}(E)$ et que Id_E appartient à $\boldsymbol{SO}(E)$; si u et v sont dans $\boldsymbol{SO}(E)$ alors $\det(v \circ u) = \det v \times \det u = 1$ et $\det(u^{-1}) = \dfrac{1}{\det u} = 1$, d'où $v \circ u$ et u^{-1} appartiennent à $\boldsymbol{SO}(E)$. Donc $\boldsymbol{SO}(E)$ est un sous-groupe de $\boldsymbol{O}(E)$.
L'assertion (ii) n'est que la reformulation en terme de matrices de l'assertion (i). ■

3.5.10. Définitions. Soit E un espace euclidien. Le groupe $\boldsymbol{O}(E)$ est appelé *groupe orthogonal* de E et le groupe $\boldsymbol{SO}(E)$ est appelé *groupe spécial orthogonal* de E.

3.6. Forme réduite d'un endomorphisme orthogonal

Avant d'établir le résultat en dimension n, nous allons faire quelques remarques et rappels généraux et nous intéresser tout particulièrement à la dimension 2.

3.6.1. Remarque. Si E est un espace euclidien de dimension 1 alors $\boldsymbol{SO}(E) = \{\mathrm{Id}_E\}$ et $\boldsymbol{O}(E) = \{\mathrm{Id}_E, -\mathrm{Id}_E\}$.

3.6.2. Rappel. Soit E un espace vectoriel réel de dimension finie. Des bases $\mathcal{B}$ et $\mathcal{B}'$ de E sont dites de même sens si $\det_{\mathcal{B}}(\mathcal{B}') > 0$. La relation ainsi définie est une relation d'équivalence sur l'ensemble des bases de E et elle définit 2 classes d'équivalence. L'espace E sera dit *orienté* si l'on choisit l'une des classes d'équivalence dont les éléments sont appelés *bases directes* ; les bases de l'autre classe sont appelées *bases indirectes*.

3.6.3. Lemme. *Soit* P *un plan vectoriel euclidien. Si* e_1 *et* f_1 *sont des vecteurs de* P *de norme 1, alors il existe un unique* $u \in \boldsymbol{SO}(P)$ *tel que* $u(e_1) = f_1$.

Preuve. Orientons P. Il existe un unique vecteur e_2 (resp. f_2) tel que $\mathcal{B} = (e_1, e_2)$ (resp. $\mathcal{B}' = (f_1, f_2)$) soit une base orthonormale directe de P.
Il existe $u \in \boldsymbol{SO}(P)$ transformant $\mathcal{B}$ en $\mathcal{B}'$; d'où l'existence.
Soit $v \in \boldsymbol{SO}(P)$ tel que $v(e_1) = f_1$. Alors v doit transformer toute base orthonormale directe de premier vecteur e_1 en une base orthonormale directe de premier vecteur f_1. Donc $v(\mathcal{B}) = \mathcal{B}'$ et $v = u$. D'où l'unicité. ■

3.6.4. Théorème. *Soit P un plan vectoriel euclidien orienté.*

(*i*) *Soient $u \in \boldsymbol{SO}(P)$ et $\mathcal{B}$ une base orthonormale directe. Il existe $\theta \in \mathbb{R}$ tel que* $\mathrm{Mat}_{\mathcal{B}}(u) = R_\theta$ *avec* $R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$.

(*ii*) *La matrice de u et par suite, la classe de θ modulo $2\pi\mathbb{Z}$, sont indépendantes de la base orthonormale directe.*

(*iii*) *Le groupe $\boldsymbol{SO}(P)$ est abélien.*

Preuve. (i) Soit $\begin{pmatrix} a & c \\ b & d \end{pmatrix} = \mathrm{Mat}_{\mathcal{B}}(u)$. C'est une matrice orthogonale de déterminant 1. D'après le corollaire 3.5.3 on a

$$\begin{aligned} a^2+b^2 &= 1 \quad (1)\\ c^2+d^2 &= 1 \quad (2)\\ ac+bd &= 0 \quad (3)\\ ad-bc &= 1 \quad (4) \end{aligned}$$

D'après (1), il existe $\theta \in \mathbb{R}$ tel que $a = \cos\theta$ et $b = \sin\theta$.
D'après (2), il existe $\theta' \in \mathbb{R}$ tel que $c = \cos\theta'$ et $d = \sin\theta'$.
(3) et (4) s'écrivent alors :
$0 = \cos\theta\cos\theta' + \sin\theta\sin\theta' = \cos(\theta'-\theta)$ et $1 = \cos\theta\sin\theta' - \sin\theta\cos\theta' = \sin(\theta'-\theta)$.
D'où $\theta' - \theta = \dfrac{\pi}{2} + 2k\pi$ et $c = \cos(\theta + \frac{\pi}{2} + 2k\pi) = -\sin\theta$ et $d = \sin(\theta + \frac{\pi}{2} + 2k\pi) = \cos\theta$.
(iii) Le caractère abélien du groupe se vérifie aisément à partir de l'expression matricielle établie en (i). En effet

$$R_\theta\, R_{\theta'} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \cos\theta' & -\sin\theta' \\ \sin\theta' & \cos\theta' \end{pmatrix} = \begin{pmatrix} \cos(\theta+\theta') & -\sin(\theta+\theta') \\ \sin(\theta+\theta') & \cos(\theta+\theta') \end{pmatrix}$$

(ii) Soient $\mathcal{B} = (e_1, e_2)$ et $\mathcal{B}' = (e'_1, e'_2)$ des bases orthonormales directes. Il existe v dans $\boldsymbol{SO}(P)$ transformant $\mathcal{B}$ en $\mathcal{B}'$. On a alors pour tous i et j dans $\{1,2\}$:

$$\langle\, u(e'_j)\,,\, e'_i \,\rangle = \langle\, u(v(e_j))\,,\, v(e_i) \,\rangle = \langle\, v(u(e_j))\,,\, v(e_i) \,\rangle = \langle\, u(e_j)\,,\, e_i \,\rangle \;.$$

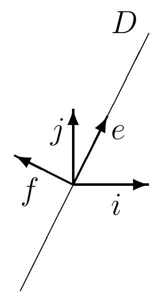
D'où l'invariance de la matrice de u dans toute base orthonormale directe. ∎

3.6.5. Définitions. Soit P un plan vectoriel euclidien orienté. Pour tout $\theta \in \mathbb{R}$, on appelle *rotation (vectorielle) d'angle* θ, notée Rot_θ, l'endomorphisme orthogonal dont la matrice dans toute base orthonormale directe est R_θ.
Si x et y sont des vecteurs de P de norme 1, d'après la proposition 3.6.3, il existe une unique rotation transformant x en y ; on appelle *angle* de x et y, noté $\widehat{(x,y)}$, tout réel θ (unique modulo $2\pi\mathbb{Z}$) tel que $\mathrm{Rot}_\theta(x) = y$.

3.6.6. Corollaire. *Soient P un plan vectoriel euclidien orienté et $u \in \boldsymbol{O}(P) \setminus \boldsymbol{SO}(P)$. Alors u est une symétrie orthogonale par rapport à une droite vectorielle D (réflexion par rapport à D). Si e est un vecteur directeur de norme 1 de D et $\mathcal{B} = (i, j)$ une base orthonormale directe de P, alors*

$$\mathrm{Mat}_{\mathcal{B}}(u) = \begin{pmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{pmatrix} \quad \textit{où} \quad \varphi = 2(\widehat{i, e}) .$$

Preuve. Comme $\det u = -1$, son polynôme caractéristique, $\chi_u(X) = X^2 - \mathrm{Tr}(u)X - 1$, admet 2 racines réelles de signes opposés. D'après la proposition 3.5.7, ces racines ne peuvent être que -1 et 1. Par conséquent u est diagonalisable et c'est la symétrie par rapport à $E_1(u)$ et parallèlement à $E_{-1}(u)$. Comme les sous-espaces propres de u sont orthogonaux (Proposition 3.5.7), il en résulte que u est la symétrie orthogonale par rapport à la droite vectorielle $D = E_1(u)$. En dimension 2, les hyperplans sont les droites et u est donc la réflexion par rapport à D.



Soit f le vecteur de P tel que $\mathcal{B}' = (e, f)$ soit une base orthonormale directe de P. Posons $\theta = (\widehat{i, e})$. La matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$ est R_θ. D'où

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \mathrm{Mat}_{\mathcal{B}'}(u) = R_\theta^{-1} \mathrm{Mat}_{\mathcal{B}}(u) \, R_\theta .$$

Il en résulte :

$$\begin{aligned} \mathrm{Mat}_{\mathcal{B}}(u) &= \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \\ &= \begin{pmatrix} \cos^2\theta - \sin^2\theta & 2\sin\theta\cos\theta \\ 2\sin\theta\cos\theta & \sin^2\theta - \cos^2\theta \end{pmatrix} \\ &= \begin{pmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{pmatrix} . \end{aligned}$$

■

3.6.7. Théorème. *Soient E un espace euclidien et $u \in \boldsymbol{O}(E)$. Alors E est somme directe orthogonale de $E_1(u) = \mathrm{Ker}\,(u - \mathrm{Id}_E)$, de $E_{-1}(u) = \mathrm{Ker}\,(u + \mathrm{Id}_E)$ et de plans stables sur lesquels u agit par rotation d'angle non multiple de π. Autrement dit, il existe une base orthonormale de E relativement à laquelle la matrice de u soit constituée de blocs sur la diagonale : $I_k, -I_\ell, R_{\theta_1}, \cdots, R_{\theta_m}$ avec $\theta_i \notin \pi\mathbb{Z}$ pour $i \in [1, m]_{\mathbb{N}}$, $(k, \ell, m) \in \mathbb{N}^3$ et $k + \ell + 2m = \dim(E)$.*

Preuve. On a déjà vu (Proposition 3.5.7) que $E_1(u)$ et $E_{-1}(u)$ sont en somme directe orthogonale. Soit $F_0 = (E_1(u) \oplus E_{-1}(u))^\perp$. Si $F_0 = \{0\}$, le théorème est démontré.
Sinon, comme $E_1(u)$ et $E_{-1}(u)$ sont stables par u, leur somme directe l'est également

et par suite F_0 est stable par u (Proposition 3.5.7) ; de plus la restriction u_0 de u à F_0 est orthogonale et ne possède pas de vecteurs propres. Posons $v_0 = u_0 + u_0^*$. C'est un opérateur symétrique de F_0. Il admet donc au moins une valeur propre λ et un vecteur propre associé w. Posons $P_1 = \mathrm{Vect}\{w, u_0(w)\}$. Comme u_0 est sans vecteur propre $u_0(w)$ et w ne sont pas liés et P_1 est un plan. Appliquons u_0 aux deux membres de l'égalité $(u_0 + u_0^*)(w) = \lambda w$, il vient $u_0^2(w) = -w + \lambda u_0(w)$. Par conséquent P_1 est stable par u_0. La restriction de u_0, donc de u, à P_1 appartient à $\boldsymbol{O}(P_1)$ et est sans vecteur propre ; d'après l'étude faite en 3.6.4 et 3.6.6, c'est donc une rotation d'angle θ_1 avec $\theta_1 \notin \pi\mathbb{Z}$. Soit F_1 l'orthogonal de P_1 dans F_0. Si $F_1 = \{0\}$, le théorème est démontré, sinon on itère le procédé qui s'arrête, car la dimension des espaces F_j construits décroît strictement. ■

3.6.8. Application : Classification en dimension 3.

	$\boldsymbol{SO}(\mathrm{E})$			$\boldsymbol{O}(\mathrm{E})\setminus\boldsymbol{SO}(\mathrm{E})$		
$k = \dim(E_1)$	3	1	1	2	0	0
$\ell = \dim(E_{-1})$	0	0	2	1	1	3
Nature	Id_E	rotation d'angle $\theta \notin \pi\mathbb{Z}$	symétrie orthogonale par rapport à une droite	symétrie orthogonale par rapport à un plan	rotation-symétrie d'angle $\theta \notin \pi\mathbb{Z}$	$-\mathrm{Id}_E$
Trace	3	$2\cos\theta+1$	-1	1	$2\cos\theta-1$	-3
Symétrique	oui	non	oui	oui	non	oui

Méthode pratique : Soit E un espace euclidien orienté de dimension 3 rapporté à une base orthonormale directe $\mathcal{B} = (e_1, e_2, e_3)$. Déterminons la nature géométrique des endomorphismes f et g admettant respectivement pour matrice dans $\mathcal{B}$:

$$A = \frac{1}{7}\begin{pmatrix} -2 & 6 & -3 \\ 6 & 3 & 2 \\ -3 & 2 & 6 \end{pmatrix} \quad \text{et} \quad B = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

- Les vecteurs colonnes de A forment une base orthonormale :

$$\frac{1}{49}((-2)^2 + 6^2 + (-3)^2) = \frac{1}{49}(6^2 + 3^2 + 2^2) = \frac{1}{49}((-3)^2 + 2^2 + 6^2) = 1$$

$$(-2)\times 6 + 6\times 3 + (-3)\times 2 = (-2)\times(-3) + 6\times 2 + (-3)\times 6 = 6\times(-3) + 3\times 2 + 2\times 6 = 0 .$$

Donc A appartient à $\boldsymbol{O}_3(\mathbb{R})$ et $\underline{f \text{ appartient à } \boldsymbol{O}(E)}$.

- A est une matrice symétrique. Donc $\underline{f \text{ est une symétrie orthogonale.}}$

- $\mathrm{Tr}(A) = 1$. Donc f est une symétrie orthogonale par rapport à un plan.

- Déterminons le sous-espace vectoriel propre de f associé à la valeur propre 1 :
$\frac{1}{7}\begin{pmatrix} -2 & 6 & -3 \\ 6 & 3 & 2 \\ -3 & 2 & 6 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$, soit encore $\begin{cases} -9x + 6y - 3z & = & 0 \\ 6x - 4y + 2z & = & 0 \\ -3x + 2y - z & = & 0 \end{cases}$.
C'est donc le plan d'équation : $3x - 2y + z = 0$.
f est la symétrie orthogonale par rapport au plan d'équation : $3x - 2y + z = 0$.

- Les vecteurs colonnes de B forment une base orthonormale :
Donc B appartient à $\boldsymbol{O}_3(\mathbb{R})$ et g appartient à $\boldsymbol{O}(E)$.

- B n'est pas symétrique, donc g n'est pas une symétrie orthogonale. L'endomorphisme g est une rotation ou une rotation-symétrie.

- Déterminons s'il existe des vecteurs invariants par g :
$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$, d'où le système d'équations : $\begin{cases} -x + y & = & 0 \\ -y + z & = & 0 \\ x - z & = & 0 \end{cases}$.
Le sous-espace vectoriel des vecteurs invariants est la droite engendrée par le vecteur u de coordonnées $(1, 1, 1)$; comme ce sous-espace vectoriel n'est pas réduit à $\{0\}$, g est une rotation d'axe D dirigé par u.

- $\mathrm{Tr}(B) = 0$; d'où $\cos\theta = -\dfrac{1}{2}$.
Donc g est une rotation d'axe D dirigé par u et d'angle $\theta = \pm\frac{2\pi}{3}$.

- Choisissons un vecteur v, non colinéaire à u et calculons $\Delta = \det_{\mathcal{B}}(v, g(v), u)$.
Prouvons que Δ et $\sin\theta$ sont de même signe. Posons $e_1' = \dfrac{1}{\|u\|}\, u$ et désignons par p la projection orthogonale sur le plan $D^\perp$. Comme v n'est pas colinéaire à u, le vecteur $p(v)$ est non nul et nous pouvons poser $e_2' = \dfrac{1}{\|p(v)\|}\, p(v)$. On complète le système orthonormal (e_1', e_2') pour obtenir une base orthonormale directe $\mathcal{B}' = (e_1', e_2', e_3')$.
$$\mathrm{Mat}_{\mathcal{B}'}(v) = \begin{pmatrix} \lambda \\ \mu \\ 0 \end{pmatrix},\ \mathrm{Mat}_{\mathcal{B}'}(g) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix},\ \mathrm{Mat}_{\mathcal{B}'}(g(v)) = \begin{pmatrix} \lambda \\ \mu\cos\theta \\ \mu\sin\theta \end{pmatrix}$$
et $\det_{\mathcal{B}'}(v, g(v), u) = \|u\|\,\mu^2\sin\theta$ avec $\mu \neq 0$. Comme $\mathcal{B}$ et $\mathcal{B}'$ sont des bases orthonormales directes la matrice passage de $\mathcal{B}$ à $\mathcal{B}'$ appartient à $\boldsymbol{SO}(3)$ et est donc de déterminant 1. Il en résulte que $\det_{\mathcal{B}}(v, g(v), u) = \det_{\mathcal{B}'}(v, g(v), u)$ est de même signe que $\sin\theta$.
Prenons ici $v = e_1$, on a alors $\Delta = \begin{vmatrix} 1 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \end{vmatrix} = -1$.
Donc g est la rotation d'axe D dirigé par u de coordonnées $(1, 1, 1)$ et d'angle $\theta = -\frac{2\pi}{3}$.

4 Applications

4.1. Endomorphismes symétriques et formes quadratiques

Dans toute cette section, E désignera un espace euclidien de dimension $n \geq 1$, $\boldsymbol{S}(E)$ l'espace vectoriel des endomorphismes symétriques sur E et $\mathcal{Q}(E)$ l'espace vectoriel des formes quadratiques sur E.

4.1.1. Proposition. *Soit E un espace euclidien. L'application qui à tout $S \in \boldsymbol{S}(E)$, associe q_S définie par : $\forall x \in E \quad q_S(x) = \langle\, x\,,\, S(x)\,\rangle$ est un isomorphisme entre $\boldsymbol{S}(E)$ et $\mathcal{Q}(E)$.*
De plus, si $\mathcal{B}$ est une base orthonormale de E alors $\mathrm{Mat}_{\mathcal{B}}(S) = \mathrm{Mat}_{\mathcal{B}}(q_S)$.

Preuve. Soit $S \in \boldsymbol{S}(E)$. Puisque S est symétrique, l'application $(x,y) \longmapsto \langle\, x\,,\, S(y)\,\rangle$ de $E \times E$ dans $\mathbb{R}$ est une forme bilinéaire symétrique dont la forme quadratique associée est q_S.
Pour tous S et T dans $\boldsymbol{S}(E)$ et tous λ et μ réels on a, pour tout $x \in E$:

$$\langle\, x\,,\, (\lambda S + \mu T)(x)\,\rangle = \lambda\langle\, x\,,\, S(x)\,\rangle + \mu\langle\, x\,,\, T(x)\,\rangle = (\lambda q_S + \mu q_T)(x).$$

D'où $q_{\lambda S + \mu T} = \lambda q_S + \mu q_T$.
Soient $\mathcal{B} = (e_1, \ldots, e_n)$ une base orthonormale de E et $A = \mathrm{Mat}_{\mathcal{B}}(S)$. On a alors $a_{i\,j} = \langle\, e_i\,,\, S(e_j)\,\rangle = f_{q_S}(e_i, e_j)$, pour tous i et j dans $[1,n]_{\mathbb{N}}$. D'où
$\mathrm{Mat}_{\mathcal{B}}(S) = \mathrm{Mat}_{\mathcal{B}}\left(f_{q_S}\right) = \mathrm{Mat}_{\mathcal{B}}(q_S)$.
Le caractère bijectif de l'application en résulte immédiatement. ∎

4.1.2. Corollaire. *Soit q une forme quadratique sur un espace euclidien E. Alors il existe $\mathcal{B}$ base orthonormale de E (pour le produit scalaire) qui est également q-orthogonale.*

Preuve. Soit $S \in \boldsymbol{S}(E)$ tel que $q_S = q$. D'après le théorème 3.4.5, il existe $\mathcal{B}$ base orthonormale qui diagonalise S. D'après la proposition précédente $\mathrm{Mat}_{\mathcal{B}}(q) = \mathrm{Mat}_{\mathcal{B}}(S)$ est diagonale. ∎

4.1.3. Corollaire. *Soient E un espace euclidien et $S \in \boldsymbol{S}(E)$. Alors*

(*i*) $\mathrm{rang}(S) = \mathrm{rang}(q_S)$. *En particulier : S bijectif $\quad\Leftrightarrow\quad$ q_S non dégénérée.*

(*ii*) *Les conditions suivantes sont équivalentes :*

(a) $\forall x \in E \quad \langle\, x\,,\, S(x)\,\rangle \geq 0$.

(b) Toutes les valeurs propres de S sont positives ou nulles.

(c) q_S est une forme quadratique positive.

Preuve. L'assertion (i) résulte immédiatement de la proposition précédente.
(ii) L'équivalence (a) $\Leftrightarrow$ (c) résulte également immédiatement de la proposition précédente.
(a) $\Rightarrow$ (b) Soit λ une valeur propre de S. Il existe $y \neq 0$ tel que $S(y) = \lambda y$. Alors $0 \leq \langle y\,,\,S(y)\,\rangle = \langle\, y\,,\,\lambda y\,\rangle = \lambda\,\|y\|^2$. D'où $\lambda \geq 0$.
(b) $\Rightarrow$ (a) Soit $\mathcal{B} = (e_1, \ldots, e_n)$ une base orthonormale qui diagonalise S. D'après l'hypothèse $\mathrm{Mat}_{\mathcal{B}}(S) = \mathrm{Diag}(\lambda_1, \ldots, \lambda_n)$ avec $\lambda_i \geq 0$, pour tout $i \in [1, n]_{\mathbb{N}}$. Tout $x \in E$ se décompose en $x = \sum_{i=1}^{n} x_i\, e_i$, d'où $S(x) = \sum_{i=1}^{n} x_i\, \lambda_i\, e_i$ et $\langle\, x\,,\,S(x)\,\rangle = \sum_{i=1}^{n} \lambda_i\, x_i^2 \geq 0$. ∎

4.1.4. Corollaire. *Soit q une forme quadratique sur $\mathbb{R}^n$ dont la matrice dans la base canonique est A. Alors :*

q positive	$\Leftrightarrow$	$\mathrm{Sp}\,(A) \subset \mathbb{R}_+$.
q définie positive	$\Leftrightarrow$	$\mathrm{Sp}\,(A) \subset \mathbb{R}_+^*$.
q définie négative	$\Leftrightarrow$	$\mathrm{Sp}\,(A) \subset \mathbb{R}_-^*$.
q ni positive ni négative	$\Leftrightarrow$	*A a deux valeurs propres non nulles de signes opposés.*

Preuve. Immédiat. ∎

4.1.5. Applications. *Soit q une forme quadratique sur $\mathbb{R}^2$ dont la matrice dans la base canonique est $A = \begin{pmatrix} a & b \\ b & d \end{pmatrix}$. Alors q est définie positive (resp. négative) si et seulement si* $\det(A) = ad - b^2 > 0$ *et* $\mathrm{Tr}(A) = a + d > 0$ (*resp* $\mathrm{Tr}(A) < 0$).

Preuve. La matrice A possède deux valeurs propres λ et μ, distinctes ou non, racines du polynôme caractéristique $X^2 - \mathrm{Tr}(A)X + \det(A)$. Donc λ et μ sont les réels dont la somme est $\mathrm{Tr}(A)$ et le produit $\det(A)$. ∎

4.1.6. Remarque. En analyse on démontre le théorème suivant :
Soient U un ouvert de $\mathbb{R}^n$, f une application de classe C^2 de U dans $\mathbb{R}$, $a \in U$ un point critique pour f (i.e. $\forall i \in [1, n]_{\mathbb{N}}$ $\dfrac{\partial f}{\partial x_i}(a) = 0$) et q la forme quadratique sur $\mathbb{R}^n$ définie par :

$$q(h_1, \ldots, h_n) = \sum_{1 \leq i, j \leq n} \frac{\partial^2 f}{\partial x_i\, \partial x_j}(a)\, h_i\, h_j \ .$$

(*i*) *Si q est définie positive, alors f admet un minimum local strict en a.*

(*ii*) *Si q est définie négative, alors f admet un maximum local strict en a.*

(*iii*) *Si q n'est ni positive ni négative, alors f n'admet pas d'extremum local en a.*

(q est la forme quadratique dont la matrice dans la base canonique est la *hessienne* de f en a, c'est à dire la matrice symétrique H telle que $h_{i\,j} = \dfrac{\partial^2 f}{\partial x_i\, \partial x_j}(a)$).
Il est donc important de pouvoir déterminer la nature d'une forme quadratique de plusieurs manières : méthode de Gauss ou recherhe des valeurs propres.

4.2. Polynômes orthogonaux

4.2.1. Proposition. *Soient $I =]a, b[$ un intervalle non vide de $\mathbb{R}$, avec éventuellement $a = -\infty$ ou $b = +\infty$, et ω une application continue strictement positive de I dans $\mathbb{R}$, telle que l'intégrale $\int_a^b t^n\, \omega(t)\, dt$ soit absolument convergente, pour tout $n \in \mathbb{N}$.*
Pour tout $(P, Q) \in \mathbb{R}[X]$, posons $\langle\, P\, ,\, Q\,\rangle = \int_a^b P(t)Q(t)\omega(t)\, dt$.
Alors $E(I, \omega) = (\mathbb{R}[X], \langle\ ,\ \rangle)$ est un espace préhilbertien réel.

Preuve. Pour tout $n \in \mathbb{N}$, par hypothèse l'intégrale $\int_a^b t^n\, \omega(t)\, dt$ est absolument convergente, par linéarité il en résulte que l'intégrale $\int_a^b P(t)Q(t)\omega(t)\, dt$ est absolument convergente donc convergente.
Les caractères symétrique et positif ainsi que la bilinéarité de $\langle\ ,\ \rangle$ sont immédiats.
Soit $P \in \mathbb{R}[X]$ tel que $\langle\, P\, ,\, P\,\rangle = 0$. Choisissons c et d tels que $[c, d] \subset I$ avec $c < d$. On a alors $0 \leq \int_c^d P^2(t)\, \omega(t)\, dt \leq \int_a^b P^2(t)\, \omega(t)\, dt = \langle\, P\, ,\, P\,\rangle = 0$. La fonction $t \longmapsto P^2(t)\, \omega(t)$ étant continue, positive et d'intégrale nulle sur $[c, d]$, il en résulte que $P^2(t)\, \omega(t) = 0$, pour tout $t \in [c, d]$. D'où $P(t) = 0$, pour tout $t \in [c, d]$, car ω est à valeurs strictement positives. Le polynôme P ayant une infinité de racines est donc nul.
Par conséquent $(E, \langle\ ,\ \rangle)$ est un espace préhilbertien réel. ∎

4.2.2. Définition. Reprenons les données de la proposition précédente, on appelle *suite de polynômes orthogonaux associés au poids ω sur I*, la suite $(P_k)_{k\in\mathbb{N}}$ de polynômes obtenue par le procédé d'orthogonalisation de Schmidt à partir de la base canonique $(X^k : t \longmapsto t^k)_{k\in\mathbb{N}}$.

4.2.3. Remarque. Rappelons que : $P_0 = 1$ et $P_k = X^k - \sum_{i=0}^{k-1} \frac{\langle\, P_i\, ,\, X^k\,\rangle}{\langle\, P_i\, ,\, P_i\,\rangle} P_i$, pour tout $k \in \mathbb{N}^*$, d'où $P_k = X^k - \sum_{i=0}^{k-1} \frac{\int_a^b P_i(t)\, t^k\, \omega(t)\, dt}{\int_a^b P_i^2(t)\, \omega(t)\, dt} P_i$.
Par construction $(P_0, \ldots, P_k)$ *est une base de $\mathbb{R}_k[X]$ et (P_k) est un polynôme de degré k dont le terme de plus haut degré est 1 et tel que pour tout $Q \in \mathbb{R}_{k-1}[X]$ on ait* $\int_a^b P_k(t)\, Q(t)\, \omega(t)\, dt = 0$.

4.2.4. Théorème. *Soit $(P_k)_{k\in\mathbb{N}}$ une suite de polynômes orthogonaux associés au poids ω sur I. Pour tout $k \in \mathbb{N}$, le polynôme P_k est simplement scindé sur $\mathbb{R}$ et toutes ses racines appartiennent à I.*

Preuve. Soient $x_1, \ldots, x_p$ les racines de P_k, de multiplicité impaire et appartenant à $]a, b[$, et $x_{p+1}, \ldots, x_q$ les racines de P_k, de multiplicité paire et appartenant à $]a, b[$. Posons $Q(X) = \prod_{i=1}^{p} (X - x_i)$. (Si $p = 0$ alors $Q = 1$). La fonction $f : t \longmapsto P_k(t)Q(t)\omega(t)$ est continue sur I et d'intégrale absolument convergente. De plus cette fonction ne s'annule sur I qu'en $x_1, \ldots, x_q$ et sans changer de signe car $(X - x_i)$ est en facteur avec une puissance paire. La fonction f reste de signe constant sur I et $\int_a^b P_k(t)Q(t)\omega(t)\, dt \neq 0$. D'après la remarque précédente le degré du polynôme Q doit être au moins k. Or $d°(Q) = p \leq d°(P_k) = k$. Donc le degré de Q est k. Par conséquent le polynôme P_k admet dans I, k racines distinctes de multiplicité impaire, comme il est de degré k nécessairement ces racines sont simples et il n'en a pas d'autre. ∎

4.2.5. Exemples .

a) Pour $I =]-1, 1[$ et w constante égale à 1 on obtient les polynômes de *Legendre*.

b) Pour $I =]-1, 1[$ et w définie par $w(t) = \dfrac{1}{\sqrt{1-t^2}}$ on obtient les polynômes de *Tchebychev*.

c) Pour $I = \mathbb{R}$ et w définie par $w(t) = e^{-t^2}$ on obtient les polynômes de *Hermite*.

d) Pour $I = \mathbb{R}_+$ et w définie par $w(t) = e^{-t}$ on obtient les polynômes de *Laguerre*.

4.2.6. Remarque . Ces exemples n'ont pas été choisis au hasard. Les polynômes orthogonaux apparaissent comme des vecteurs propres d'opérateurs différentiels issus de la physique : équation de la chaleur, électromagnétisme, équation de Schrodinger, ...

4.3. Coniques et quadriques

4.3.1. Définitions . Soient n un entier appartenant à $\{2, 3\}$ et P un polynôme à n variables $x_1, \ldots, x_n$ de degré 2 par rapport à l'ensemble de ces variables. On se place dans $\mathbb{R}^n$ muni du produit scalaire usuel et du repère orthonormal $\mathcal{R} = (0, \mathcal{B}_0)$ où $\mathcal{B}_0$ est la base canonique de $\mathbb{R}^n$. L'ensemble $S = \{\, (x_1, \ldots, x_n) \in \mathbb{R}^n \,;\ P(x_1, \ldots, x_n) = 0 \,\}$ est appelé *conique* (respectivement *quadrique*) si $n = 2$ (respectivement si $n = 3$).

4.3.2. Remarque .
On peut écrire P sous la forme : $P(x_1, \ldots, x_n) = P(x) = Q(x) + L(x) + k$ où Q est un polynôme homogène de degré 2 qui définit une forme quadratique q sur $\mathbb{R}^n$, L un

polynôme homogène de degré 1 qui définit une forme linéaire ℓ sur $\mathbb{R}^n$ et k une constante. Soient $A = (a_{i\,j}) = \mathrm{Mat}_{\mathcal{B}_0}(q) \in \boldsymbol{S}_n(\mathbb{R})$ et $B = (b_i) = \mathrm{Mat}_{\mathcal{B}_0}(\ell) \in \boldsymbol{M}_{1\,n}(\mathbb{R})$. Alors

$$P(x) = P(x_1, \ldots, x_n) = \sum_{i=1}^{n} a_{i\,i}\, x_i^2 \; + 2 \sum_{1 \leq i < j \leq n} a_{i\,j}\, x_i x_j \; + \sum_{i=1}^{n} b_i\, x_i + k \; .$$

Remarquons que $\dfrac{\partial P}{\partial x_k}(x) = 2 \displaystyle\sum_{i=1}^{n} a_{i,\,k}\, x_i \; + \; b_k$.

4.3.3. Proposition. *Reprenons les données précédentes. Si q est non dégénérée alors le polynôme P admet un unique point critique ω qui est centre de symétrie pour S.*
Dans ces conditions, en prenant comme nouvelle origine le point ω, l'équation de S dans $\mathcal{R}' = (\omega, \mathcal{B}_0)$, devient $Q(x') = c^{\text{te}}$.

Preuve. x est un point critique pour P si et seulement si, pour tout $k \in [1, n]_{\mathbb{N}}$, on a $2 \displaystyle\sum_{i=1}^{n} a_{i,\,k}\, x_i \; + \; b_k = 0$, ce qui se traduit matriciellement par $A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = -\dfrac{1}{2} \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$.

Comme q est non dégénérée, la matrice A est inversible et le système admet une unique solution ω. On a donc

$$A \begin{pmatrix} \omega_1 \\ \vdots \\ \omega_n \end{pmatrix} + \frac{1}{2} \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} .$$

En prenant les transposées on obtient $(\omega_1, \ldots, \omega_n)A + \dfrac{1}{2}B = (0, \ldots, 0)$, car A est symétrique. D'où, pour tout $x \in \mathbb{R}^n$,

$$(\omega_1, \ldots, \omega_n)A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \frac{1}{2}B \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = 0 \quad \text{c'est à dire} \quad f_q(\omega, x) + \frac{1}{2}L(x) = 0 \quad (*) \; .$$

Or ω est centre de symétrie pour S si et seulement si pour tout $x \in \mathbb{R}^n$, $(w + x) \in S$ implique $(\omega - x) \in S$, soit encore : $\forall x \in \mathbb{R}^n \quad P(\omega + x) = 0 \Rightarrow P(\omega - x) = 0$.
Remarquons que :

$$\begin{aligned} P(\omega + x) - P(\omega - x) &= Q(\omega + x) - Q(\omega - x) + L(\omega + x) - L(\omega - x) \\ &= 4 f_q(\omega, x) + 2L(x) \\ &= 0 \; . \end{aligned}$$

Donc ω est centre de symétrie pour S.
Prenons ω comme nouvelle origine et désignons par $x' = (x'_1, \ldots, x'_n)$ les coordonnées de x dans le repère $\mathcal{R}' = (\omega, \mathcal{B}_0)$. On a $x' = x - \omega$, d'où $P(x) = 0 \iff P(x' + \omega) = 0$. Or

$$P(x'+\omega) = Q(x'+\omega)+L(x'+\omega)+k = Q(x')+Q(\omega)+2f_q(x', \omega)+L(x')+L(\omega)+k = Q(x')+k' ,$$

en utilisant $(*)$ et en posant $k' = Q(\omega) + L(\omega) + k$. ∎

4.3.4. Classification des coniques.

Il existe une base orthonormale $\mathcal{B}$ qui diagonalise q : $\mathrm{Mat}_{\mathcal{B}}(q) = \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}$.

• Coniques à centre

Dans le cas où q est non dégénérée, on a $\lambda\mu \neq 0$ et dans le repère $(\omega, \mathcal{B})$ l'équation de S devient : $\lambda X^2 + \mu Y^2 = c$.

- Si $c = 0$, S a pour équation : $Y^2 = -\dfrac{\lambda}{\mu} X^2$

  Si λ et μ de même signe alors $\underline{S = \{\omega\}}$,

  sinon $\underline{S \text{ est la réunion de 2 droites concourantes en } \omega}$.

- Si $c \neq 0$, S a pour équation $\dfrac{\lambda}{c} X^2 + \dfrac{\mu}{c} Y^2 = 1$.

  Si $\dfrac{\lambda}{c} < 0$ et $\dfrac{\mu}{c} < 0$ alors $\underline{S = \emptyset}$

  Si $\dfrac{\lambda}{c} > 0$ et $\dfrac{\mu}{c} > 0$ alors $\underline{S \text{ est une } \textit{ellipse}}$ d'équation réduite $\dfrac{X^2}{a^2} + \dfrac{Y^2}{b^2} = 1$

  Si $\lambda\mu < 0$ alors $\underline{S \text{ est une } \textit{hyperbole}}$ d'équation réduite $\dfrac{X^2}{a^2} - \dfrac{Y^2}{b^2} = 1$ si $\dfrac{\lambda}{c} > 0$, ou $\dfrac{X^2}{a^2} - \dfrac{Y^2}{b^2} = -1$ sinon.

• Autres coniques

Dans le cas où q est dégénérée, une des valeurs propres, par exemple μ, est nulle et λ n'est pas nulle car P est de degré 2. Donc dans le repère $(0, \mathcal{B})$, S admet une équation de la forme $X^2 + 2dX + 2eY + f = 0$

- Si $e = 0$, S a pour équation : $X^2 + 2dX + f = 0$

  Si $d^2 - f < 0$ alors $\underline{S = \emptyset}$

  Si $d^2 - f > 0$ alors $\underline{S \text{ est la réunion de 2 droites parallèles}}$

  Si $d^2 - f = 0$ alors $\underline{S \text{ est une droite (double)}}$

- Si $e \neq 0$, S a pour équation : $(X + d)^2 = -2e\left(Y - \dfrac{d^2 - f}{2e}\right)$.

  En prenant pour origine le point $\omega' = (-d, \dfrac{d^2 - f}{2e})$ l'équation de S est de la forme $X'^2 = 2pY'$ avec $p = -e \neq 0$.
  $\underline{S \text{ est une } \textit{parabole}}$ de sommet ω'.

4.3.5. Classification des quadriques à centre.
Nous supposons q non dégénérée, donc S admet un centre de symétrie ω et il existe une base orthonormale $\mathcal{B}$ qui diagonalise q : $\mathrm{Mat}_{\mathcal{B}}(q) = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \mu & 0 \\ 0 & 0 & \nu \end{pmatrix}$ avec $\lambda\mu\nu \neq 0$. Quitte à multiplier les deux membres de l'équation par -1 nous pouvons supposer $\lambda > 0$ et $\mu > 0$. Dans le repère $(\omega, \mathcal{B})$ une équation de S est de la forme $\dfrac{X^2}{a^2} + \dfrac{Y^2}{b^2} + \varepsilon\,\dfrac{Z^2}{c^2} = \delta$ avec $\varepsilon \in \{1, -1\}$ et $\delta \in \{0, 1, -1\}$. D'où la classification

- $S = \{\omega\}$ équation réduite : $\dfrac{X^2}{a^2} + \dfrac{Y^2}{b^2} + \dfrac{Z^2}{c^2} = 0$.
- S est un *cône* de sommet ω d'équation réduite : $\dfrac{X^2}{a^2} + \dfrac{Y^2}{b^2} - \dfrac{Z^2}{c^2} = 0$.
- S est un *ellipsoïde* d'équation réduite : $\dfrac{X^2}{a^2} + \dfrac{Y^2}{b^2} + \dfrac{Z^2}{c^2} = 1$.
- $S = \emptyset$ équation réduite : $\dfrac{X^2}{a^2} + \dfrac{Y^2}{b^2} + \dfrac{Z^2}{c^2} = -1$.
- S est un *hyperboloïde à une nappe* d'équation réduite : $\dfrac{X^2}{a^2} + \dfrac{Y^2}{b^2} - \dfrac{Z^2}{c^2} = 1$.
- S est un *hyperboloïde à deux nappes* d'équation réduite : $\dfrac{X^2}{a^2} + \dfrac{Y^2}{b^2} - \dfrac{Z^2}{c^2} = -1$.

4.3.6. Remarque. Dans les quadriques dégénérées, citons l'exemple du *cylindre elliptique* d'équation réduite : $\dfrac{X^2}{a^2} + \dfrac{Y^2}{b^2} = 1$. Si $a = b$ on a un *cylindre de révolution*.

Index