UNIVERSITÉ D'ANGERS — Année 2024-2025

L3 MATH - MA - DL ME - DL MI

Algèbre linéaire et bilinéaire

Table des matières

Dans tout ce qui suit $\mathbb{K}$ désigne $\mathbb{R}$ ou $\mathbb{C}$, et E ,F, G seront des $\mathbb{K}$-espaces vectoriels.

1 Quelques rappels d'algèbre linéaire

Cette partie rappelle quelques notions déjà vues en L2 qui sont indispensables pour la suite.

1.1 Sous-espaces vectoriels

Un sous-espace vectoriel de E est un sous-ensemble non-vide X de E stable par combinaisons linéaires :

$$\forall (x,y) \in X^2,\ \forall (\lambda,\mu) \in \mathbb{K}^2,\ \lambda x + \mu y \in X.$$

Un sous-espace vectoriel de E, muni de l'addition et de la multiplication par un scalaire est un espace vectoriel.

Si X et Y sont deux sous-espaces vectoriels de E, $X \cap Y$ l'est aussi (mais en général $X \cup Y$ ne l'est pas !).

Si X et Y sont deux sous-espaces vectoriels de E, la somme $X+Y$ est définie par

$$X + Y = \{x + y \mid x \in X, y \in Y\}.$$

C'est un sous-espace vectoriel de E.

Si $x_1,\dots,x_r$ sont des éléments de E, l'ensemble de toutes les combinaisons lináires $\alpha_1 x_1 + \dots + \alpha_r x_r$ de $x_1,\dots,x_r$ est un sous-espace vectoriel de E. On l'appelle le sous-espace vectoriel engendré par $x_1,\dots,x_r$, et on le note $\mathrm{Vect}(x_1,\dots,x_r)$. Si $r=1$ on note $\mathrm{Vect}(x_1) = \mathbb{K}x_1$.

Deux sous-espaces vectoriels E_1, E_2 sont dits en somme directe si $E_1 \cap E_2 = \{0\}$. Dans ce cas, un élément $x \in E_1 + E_2$ s'écrit de manière unique comme $u+v$ avec $u \in E_1$ et $v \in E_2$. La somme $E_1 + E_2$ est alors notée $E_1 \oplus E_2$.

De façon plus générale, si $(E_k)_{1\le k\le r}$ est une famille finie de sous-espaces vectoriels de E, on dit qu'ils sont en somme directe si

$$\forall (x_k)_{1\le k\le r} \in \prod_{1\le k\le r} E_k,\quad \sum_{1\le k\le r} x_k = 0 \Longrightarrow \forall k \in [[1,r]], x_k = 0.$$

Dans ce cas, on note leur somme $\bigoplus_{1\le k\le r} E_k$. Il ne suffit pas d'imposer que les E_k sont en somme directe deux à deux.

1.2 Familles libres, familles génératrices, bases

On dit qu'une famille $(x_1,\dots,x_r)$ d'éléments de E est une famille libre, ou encore que les éléments $x_1,\dots,x_r$ sont linéairement indépendants, si les droites $(\mathbb{K}x_k)_{1\le k\le r}$ sont en somme directe.

Une famille (peut être infinie) $(x_\alpha)_{\alpha\in I}$ d'éléments de E est dite libre si toute sous-famille finie de $(x_\alpha)_{\alpha\in I}$ est libre.

On dit qu'une famille $(x_1,\dots,x_r)$ d'éléments de E est une famille génératrice, ou encore que les éléments $x_1,\dots,x_r$ engendrent linéairement E, si $\mathrm{Vect}(x_1,\dots,x_r) = E$.

Une famille d'éléments de E est une base de E si elle est libre et génératrice.

Théorème 1.1 (Base incomplète)**.** *Si $(x_1,\dots,x_r)$ est une famille libre d'un espace vectoriel de dimension finie E, alors il existe des éléments $x_{r+1},\dots,x_n$ tels que $(x_1,\dots,x_n)$ soit une base de E.*

Théorème 1.2. *Si $(x_1,\dots,x_r)$ et $(y_1,\dots,y_s)$ sont deux bases d'un espace vectoriel E, alors elles ont le même nombre d'éléments $r=s$. On appelle ce nombre la* dimension *de E, que l'on note* $\dim E$.

Ainsi, chaque élément de E s'écrit de manière unique comme combinaison linéaire des vecteurs d'une base donnée.

Proposition 1.3. *Soit E un espace vectoriel de dimension n, et soit $\mathscr{E} = (e_1,\dots,e_n)$ une famille de n vecteurs de E. Alors les assertions suivantes sont équivalentes :*

(i) $\mathscr{E}$ est une base de E,

(ii) $\mathscr{E}$ *est une famille libre de* E*,*

(iii) $\mathscr{E}$ *est une famille génératrice de* E*.*

Le théorème de la base incomplète entraîne que si $X \subset Y$ sont deux sous-espaces vectoriels de dimension finie, alors $\dim X \leq \dim Y$ et qu'on a $X = Y$ si, et seulement si, $\dim X = \dim Y$.

Théorème 1.4. *Soient* X *et* Y *deux sous-espaces vectoriels de* E*. Alors*

$$\dim X + \dim Y = \dim(X+Y) + \dim(X \cap Y).$$

Soit $\mathscr{E} = (e_1, \ldots, e_n)$ une base de E. Pour tout vecteur $x \in E$, on note $[x]_{\mathscr{E}}$ le vecteur colonne dont les coefficients sont les coordonnées de x dans la base $\mathscr{E}$.

Soit $\mathscr{F} = (f_1, \ldots, f_n)$ une autre base de E. On note $\mathscr{P}_{\mathscr{E}}^{\mathscr{F}} = \left[[f_1]_{\mathscr{E}}, \ldots, [f_n]_{\mathscr{E}}\right]$ la matrice de passage de la base $\mathscr{E}$ à la base $\mathscr{F}$.

Alors les coordonnées d'un vecteur x dans les bases $\mathscr{E}$ et $\mathscr{F}$ sont liées par

$$[x]_{\mathscr{E}} = \mathscr{P}_{\mathscr{E}}^{\mathscr{F}} [x]_{\mathscr{F}}.$$

Par ailleurs, cette égalité entraîne que

$$\mathscr{P}_{\mathscr{E}}^{\mathscr{F}} = \left(\mathscr{P}_{\mathscr{F}}^{\mathscr{E}}\right)^{-1}.$$

1.3 Applications linéaires

Une application $u : E \to F$ est dite linéaire si elle vérifie

$$\forall (\alpha, \beta) \in \mathbb{K}^2,\ \forall (x, y) \in E^2,\ u(\alpha x + \beta y) = \alpha u(x) + \beta u(y).$$

Dans le cas où $E = F$, une application linéaire $u : E \to E$ est appelée un endomorphisme de E. L'ensemble des applications linéaires $u : E \to F$ est un $\mathbb{K}$-espace vectoriel noté par $\mathscr{L}(E, F)$; si $E = F$, on notera simplement $\mathscr{L}(E)$ au lieu de $\mathscr{L}(E, E)$.

On définit alors le noyau d'une application linéaire $u : E \to F$ par

$$\ker(u) = \{x \in E \mid u(x) = 0\},$$

et l'image de u

$$\mathrm{Im}\,(u) = \{y \in F \mid \exists x \in E,\ u(x) = y\}.$$

On déduit immédiatement de la linéarité de u que $\ker u$ est un sous-espace vectoriel de E et $\mathrm{Im}\, u$ est un sous-espace vectoriel de F. La dimension de $\mathrm{Im}\, u$ est appelée le rang de u et est notée $\mathrm{rg}(u)$.

Théorème 1.5. *Soit* $u : E \to F$ *une application linéaire. Alors*

$$\dim \ker(u) + \mathrm{rg}(u) = \dim E.$$

On se rappellera bien ici que c'est la dimension de l'espace de départ de u qui compte.

Corollaire 1.6. *Soit* $u : E \to F$ *une application linéaire avec* $\dim E = \dim F$*. Alors les trois assertions suivantes sont équivalentes :*

— u *est injective*

— u *est surjective*

— u *est bijective.*

1.4 Matrice d'une application linéaire

Soit $u : E \to F$ une application linéaire, et soient $\mathscr{E}$ une base de E et $\mathscr{F}$ une base de F. La matrice de u relativement aux bases $\mathscr{E}$ et $\mathscr{F}$ est la matrice dont la colonne d'indice j est le vecteur colonne des coordonnées dans la base $\mathscr{F}$ de l'image par u du j-ième vecteur de $\mathscr{E}$. On note :

$$[u]_{\mathscr{E}}^{\mathscr{F}} = [[u(e_1)]_{\mathscr{F}}, \dots, [u(e_n)]_{\mathscr{F}}].$$

Dans le cas où $E = F$, on remarque que

$$\mathscr{P}_{\mathscr{E}}^{\mathscr{F}} = [\mathrm{id}_E]_{\mathscr{F}}^{\mathscr{E}}.$$

Dans le cas où $u \in \mathscr{L}(E)$ et $\mathscr{E}$ est une base de E, on note simplement

$$[u]_{\mathscr{E}} = [u]_{\mathscr{E}}^{\mathscr{F}}.$$

Les coordonnées de l'image par u d'un vecteur x sont exprimées par

$$[u(x)]_{\mathscr{F}} = [u]_{\mathscr{E}}^{\mathscr{F}} [x]_{\mathscr{E}}.$$

Proposition 1.7. *Soient E, F, G des espaces vectoriels munis de bases $\mathscr{E}, \mathscr{F}, \mathscr{G}$, et soient des applications linéaires $f : E \to F$ et $g : F \to G$. Alors*

$$[g \circ f]_{\mathscr{E}}^{\mathscr{G}} = [g]_{\mathscr{F}}^{\mathscr{G}} [f]_{\mathscr{E}}^{\mathscr{F}}.$$

Démonstration. En effet, pour tout $x \in E$, on a

$$[g \circ f]_{\mathscr{E}}^{\mathscr{G}} [x]_{\mathscr{E}} = [g \circ f(x)]_{\mathscr{G}} = [g]_{\mathscr{F}}^{\mathscr{G}} [f(x)]_{\mathscr{F}} = [g]_{\mathscr{F}}^{\mathscr{G}} [f]_{\mathscr{E}}^{\mathscr{F}} [x]_{\mathscr{E}}. \qquad \square$$

Corollaire 1.8. *Soit $u \in \mathscr{L}(E)$, et $P = \mathscr{P}_{\mathscr{E}}^{\mathscr{E}'}$ la matrice de passage d'une base $\mathscr{E}$ à une autre base $\mathscr{E}'$. On a*

$$[u]_{\mathscr{E}'} = P^{-1} [u]_{\mathscr{E}} P.$$

1.5 Déterminant d'une matrice, déterminant d'un endomorphisme

Le déterminant d'une matrice carrée $A = (a_{ij})_{1 \le i,j \le n}$ est défini par

$$\det(A) = \sum_{\sigma \in \mathscr{S}_n} \varepsilon(\sigma) a_{1\sigma(1)} \dots a_{n\sigma(n)},$$

où $\mathscr{S}_n$ est le groupe des permutations de l'ensemble $\{1, \dots, n\}$ et $\varepsilon(\sigma)$ est la signature de σ.

Proposition 1.9 (Déterminant d'un produit)**.** *Si A et B sont deux matrices carrées de même taille taille, alors*

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

Il résulte de cette propriété que, si P est une matrice carrée inversible, on a

$$\det(P^{-1}) = (\det P)^{-1},$$

et que, si A et P sont des matrices carrées de même taille avec P inversible, on a

$$\det(P^{-1}AP) = \det A.$$

Corollaire 1.10. *Soit $u : E \to E$ un endomorphisme et soient $\mathscr{E}, \mathscr{F}$ deux bases de E. Alors*

$$\det[u]_{\mathscr{E}} = \det[u]_{\mathscr{F}}.$$

En d'autres termes, cette valeur est indépendante de la base qu'on utilise pour la calculer et ne dépend que de u. On l'appelle donc le déterminant de u, et on la note $\det(u)$. L'expression du déterminant d'un produit entraîne que si u et v sont des endomorphismes de E alors

$$\det(u \circ v) = \det(u)\det(v).$$

Proposition 1.11. *Un endomorphisme est inversible si et seulement si son déterminant est non nul.*

Démonstration. Si u est inversible alors il existe v tel que $u \circ v = \mathrm{Id}_E$. Il suit que $1 = \det(u \circ v) = \det(u)\det(v)$, donc $\det(u) \neq 0$. La réciproque est vraie grâce à la formule de la comatrice. □

Proposition 1.12 (Déterminant d'une matrice par blocs)**.** *Soit A une matrice qui s'écrit*

$$A = \left(\begin{array}{c|c} B & D \\ \hline 0 & C \end{array}\right),$$

où B et C sont des blocs carrés. Alors

$$\det(A) = \det(B)\det(C).$$

2 Sous-espaces stables par un endomorphisme

Définition 2.1. Soit $u \in \mathscr{L}(E)$. Un sous-espace vectoriel $F \subset E$ est dit stable par u si $u(F) \subset F$. On note alors $u_F = u|_F^F \in \mathscr{L}(F)$ l'endomorphisme induit défini par $u_F(x) = u(x)$ pour tout $x \in F$.

Si $F \subseteq E$ est stable par u, alors en complétant une base $\mathscr{F}$ de F en une base $\mathscr{E}$ de E, on obtient une matrice triangulaire supérieure par blocs de la forme

$$[u]_{\mathscr{E}} = \left(\begin{array}{c|c} B & D \\ \hline 0 & C \end{array}\right).$$

où $B = [u_F]_{\mathscr{F}}$ et C sont carrées de taille $\dim F$ et $\operatorname{codim} F := \dim E - \dim F$.

Définition 2.2. Soit $u \in \mathscr{L}(E)$ et $\lambda \in \mathbb{K}$. On note $E_\lambda(u) = \ker(u - \lambda \mathrm{id}_E)$ l'espace propre associé à u et λ. On dit que λ est une valeur propre pour u si $E_\lambda(u) \neq \{0\}$, et on appelle vecteurs propres les vecteurs non nuls de $E_\lambda(u)$.

Théorème 2.3. *Soit $u \in \mathscr{L}(E)$. Les espaces propres de u sont en somme directe.*

Démonstration. On démontre par récurrence sur r que toute relation de dépendance linéaire $\sum_{1\leq k\leq r} x_k = 0$, où $x_k \in E_{\lambda_k}(u)$ pour une famille de scalaires $(\lambda_k)_{1\leq k\leq r}$ deux à deux distincts, est triviale.

Si $r = 1$, il n'y a rien à démontrer. On suppose que le résultat est vrai pour un $r \geq 1$ et on montre qu'il reste vrai pour $r+1$. Soit donc une relation

$$\sum_{1\leq k\leq r+1} x_k = 0, \tag{1}$$

où $x_k \in E_{\lambda_k}(u)$ pour des λ_k deux à deux distincts. On applique u pour obtenir

$$\sum_{1\leq k\leq r+1} \lambda_k x_k = 0. \tag{2}$$

On retranche λ_{r+1} fois (1) dans (2) :

$$\sum_{1\leq k\leq r+1} (\lambda_k - \lambda_{r+1}) x_k = 0.$$

Comme le dernier terme est nul, on peut utiliser l'hypothèse de récurrence pour obtenir

$$\forall 1 \leq k \leq r, \quad (\lambda_k - \lambda_{r+1}) x_k = 0.$$

Comme $\lambda_k \neq \lambda_{r+1}$ on déduit que $x_k = 0$ pour tout $1 \leq k \leq r$. Grâce à (1), il suit également que $x_{r+1} = 0$. □

Comme conséquence immédiate on a :

Théorème 2.4. *Un endomorphisme d'un espace vectoriel de dimension n a au plus n valeurs propres distinctes.*

En dimension finie on notera par fois $\mathrm{Sp}(u)$ le spectre de u, soit l'ensemble de ses valeurs propres.

Définition 2.5. Un endomorphisme u de E est dit diagonalisable si E est la somme (directe) des espaces propres de u.

Supposons notre endomorphisme u défini par une matrice M dans une base $\mathscr{E}$. Si u est diagonalisable, il existe une base $\mathscr{F}$ de vecteurs propres pour u. La matrice D de u dans la base $\mathscr{F}$ est donc diagonale, d'où la terminologie. On obtient alors, en utilisant la formule de changement de bases :

$$D = P^{-1}MP$$

où $P = \mathscr{P}_{\mathscr{E}}^{\mathscr{F}}$.

Théorème 2.6. *Soit E un espace vectoriel de dimension n. Un endomorphisme de E qui possède n valeurs propres distinctes est diagonalisable et ses espaces propres sont des droites.*

Démonstration. On a

$$\dim \bigoplus_{\lambda\in\mathrm{Sp}(u)} E_\lambda(u) = \sum_{\lambda\in\mathrm{Sp}(u)} \dim E_\lambda(u) \geq \sum_{\lambda\in\mathrm{Sp}(u)} 1 = n.$$

Ainsi $\oplus_{\lambda\in\mathrm{Sp}(u)} E_\lambda(u) = E$ et l'inégalité est une égalité : les espaces propres sont tous de dimension 1. □

Définition 2.7. On dit qu'une matrice carrée M est diagonalisable si l'endomorphisme de $\mathscr{M}_{n,1}(\mathbb{K}) \simeq \mathbb{K}^n$ associé $u_M : x \mapsto Mx$ est diagonalisable. Une matrice A est semblable à une matrice B s'il existe une matrice inversible P telle que $B = P^{-1}AP$. Un scalaire λ est une valeur propre de M s'il est une valeur propre de l'endomorphisme associé u_M.

Remarque 2.8. Une matrice carrée est diagonalisable si et seulement si elle est semblable à une matrice diagonale.

Remarque 2.9. La condition dans le théorème précédent n'est que suffisante. Ainsi, la matrice

$$M = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

est diagonalisable mais elle n'a que deux valeurs propres distinctes.

Remarque 2.10. Il existe des matrices carrées non diagonalisables. Ainsi la matrice

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

n'est pas diagonalisable. En effet, on voit facilement que la seule valeur propre de M est 1. Si M était diagonalisable, M serait semblable à I_2, ce qui force $M = I_2$, ce qui est manifestement faux.

3 Polynôme caractéristique

3.1 Quelques rappels sur les polynômes

Un polynôme à coefficients dans $\mathbb{K}$ est une expression $P = a_0 + a_1X + \cdots + a_nX^n$, où les a_j sont des éléments de $\mathbb{K}$. Si P est non nul, on appelle degré de P le plus grand entier r tel que $a_r \neq 0$. On le note $\deg P$.

L'ensemble des polynômes à coefficients dans $\mathbb{K}$ est noté par $\mathbb{K}[X]$. C'est un $\mathbb{K}$-espace vectoriel de dimension infinie.

On note $\mathbb{K}[X]_n$ l'ensemble des polynômes à coefficients dans k et de degré inférieur ou égale à n. C'est un sous-espace vectoriel de $\mathbb{K}[X]$, de dimension $n+1$. On peut prendre pour base de $\mathbb{K}[X]_n$ la famille $X^0 = 1, X, X^2, \ldots, X^n$.

On dit qu'un élément $a \in \mathbb{K}$ est une racine de P si $P(a) = 0$.

On dit que A divise B ($A, B \in \mathbb{K}[X]$) s'il existe $C \in \mathbb{K}[X]$ tel que $B = AC$ (dans ce cas on note $A|B$).

Le polynôme dérivé de P est

$$P'(X) = a_1 + 2a_2X + \cdots + na_nX^{n-1}.$$

On note $P^{(k)}$ le polynôme obtenu en appliquant k fois l'application de dérivation à P. On vérifie que

$$(X^n)^{(k)} = \frac{n!}{(n-k)!}X^{n-k}$$

si $k \leq n$, et que $P^{(k)} = 0$ dès que $k > \deg P$.

Proposition 3.1 (Division euclidienne)**.** *Soient $A, B \in \mathbb{K}[X]$, avec $B \neq 0$. Alors il existe des polynômes $Q, R \in \mathbb{K}[X]$ uniques tels que*

$$A = BQ + R \quad \text{et} \quad \deg R < \deg B.$$

Définition 3.2 (Racine multiple)**.** Un élément $a \in \mathbb{K}$ est dit racine de P de multiplicité r si $(X-a)^r$ divise P et $(X-a)^{r+1}$ ne divise pas P.

On a la caractérisation suivante :

Proposition 3.3. *Le scalaire $a \in \mathbb{K}$ est une racine de multiplicité r de P si et seulement si $P(a) = P'(a) = \cdots = P^{(r-1)}(a) = 0$ et $P^{(r)}(a) \neq 0$.*

Un polynôme $P = a_0 + a_1X + \cdots + a_nX^n$ est unitaire si son coefficient dominant $a_n = 1$.

Théorème 3.4. *Soient $A, B \in \mathbb{K}[X]$ non tous les deux nuls. Alors il existe un unique polynôme D unitaire tel que*

$$\{UA + VB \mid U, V \in \mathbb{K}[X]\} = \mathbb{K}[X]D.$$

Ce polynôme est le plus grand diviseur commun à A et B : il divise A et B et si E divise A et B alors E divise D.

Pour calculer explicitement le PGCD de deux polynômes, on peut utiliser l'algorithme d'Euclide, basé sur une suite de division euclidiennes :

$$\begin{aligned} A &= P_1B + R_1 \\ B &= P_2R_1 + R_2 \\ R_1 &= P_3R_2 + R_2 \\ &\vdots \\ R_{n-2} &= P_nR_{n-1} + R_n. \end{aligned}$$

Le PGCD est alors le dernier reste non nul. Par ailleurs, cette méthode permet de déterminer les polynômes U, V tels que $UA + VB = D$.

Définition 3.5. Deux polynômes A et B sont dits premiers entre eux si et seulement si les seuls diviseurs communs à A et B sont les constantes non nulles. De manière équivalente, A et B sont premiers entre eux si le PGCD de A et B est 1.

Théorème 3.6. *Deux polynômes A et B sont premiers entre eux si et seulement si il existe deux polynômes U et V tels que $UA + VB = 1$.*

3.2 Le polynôme caractéristique

Dire que λ est une valeur propre d'un endomorphisme $u : E \to E$ d'un $\mathbb{K}$-espace vectoriel de dimension finie revient à dire que $\det(u - \lambda id_E) = 0$. C'est pour cette raison qu'on introduit, pour toute matrice M carrée de taille n, l'expression

$$\chi_M(\lambda) = \det(M - \lambda I_n), \quad \forall \lambda \in \mathbb{K}.$$

Remarque 3.7. La définition du déterminant implique que $\chi_M(\lambda)$ est une expression polynômiale en λ, et on dispose donc bien d'un polynôme $\chi_M \in \mathbb{K}[X]$. Il est de degré n, de coefficient dominant $(-1)^n$, et de coefficient constant $\det(M)$. Le coefficient de degré 1 de χ_M est $(-1)^{n-1}\mathrm{tr}(M)$, où $\mathrm{tr}(M)$ est la trace de M, i.e. la somme des termes sur la diagonale de M.

Remarque 3.8. Si M et N sont les matrices d'un endomorphisme $u : E \to E$ d'un $\mathbb{K}$-espace vectoriel de dimension n dans deux bases $\mathscr{E}$ et $\mathscr{F}$ alors

$$M = P^{-1}NP,$$

où $P = \mathscr{P}_{\mathscr{F}}^{\mathscr{E}}$. Il suit que

$$\chi_M(X) = \det(M - XI_n) = \det(P^{-1}(N - XI_n)P) = \det(N - XI_n) = \chi_N(X).$$

On résume ceci en un lemme :

Lemme 3.9. *Deux matrices semblables ont même polynôme caractéristique.*

On dispose donc pour tout $u \in \mathscr{L}(E)$ d'un polynôme caractéristique $\chi_u \in \mathbb{K}[X]$ qui satisfait $\chi_u(\lambda) = \det(u - \lambda\mathrm{id}_E)$ pour tout $\lambda \in \mathbb{K}$. En particulier :

Proposition 3.10. *Un scalaire λ est une valeur propre de $u \in \mathscr{L}(E)$ si et seulement si $\chi_u(\lambda) = 0$.*

Remarque 3.11. Si $k = \mathbb{C}$, tout $f \in \mathscr{L}(E)$ a au moins une valeur propre $\lambda \in \mathbb{K}$.

Pour montrer que λ est une valeur propre d'une matrice A, on peut montrer que $\mathrm{rg}(A - \lambda I_n) < n$, ce qui est parfois plus simple que de montrer $\chi_A(\lambda) = 0$. Par exemple si $n \geq 2$ et

$$A = \begin{pmatrix} a & 1 & \cdots & 1 \\ 1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 1 & \cdots & 1 & a \end{pmatrix} \in \mathscr{M}_n(\mathbb{R}),$$

on voit que $\mathrm{rg}(A - (a-1)I_n) = 1 < n$, donc $a - 1$ est une valeur propre de A et $\dim E_{a-1} = n - 1$.

Dans la pratique, pour déterminer $E_\lambda(A)$ lorsque λ est une valeur propre d'une matrice $A = (a_{i,j})_{1\leq i,j\leq n} \in M_n(k)$, on résout le système $AX = \lambda X$.

Proposition 3.12. *Soit $u \in \mathscr{L}(E)$, et F un sous-espace vectoriel de E stable par u. Alors $\chi_{u_F} | \chi_u$.*

Démonstration. Soit $\mathscr{F}$ une base de F complétée en une base $\mathscr{E}$ de E. La matrice M de f dans la base $\mathscr{E}$ a la forme

$$M = \left(\begin{array}{c|c} A & C \\ \hline 0 & B \end{array}\right),$$

où $A = [u_F]_{\mathscr{F}}$. Mais alors, si $r = \dim F$,

$$\begin{aligned} \chi_u(X) = \det(M - XI_n) &= \det\left(\begin{array}{c|c} A - XI_r & C \\ \hline 0 & B - XI_{n-r} \end{array}\right) \\ &= \det(A - XI_r)\det(B - XI_{n-r}) \\ &= \chi_{u_F}(X)\det(B - XI_{n-r}). \end{aligned}$$

□

Définition 3.13. Un endomorphisme $f \in \mathscr{L}(E)$ est dit nilpotent si il existe un entier positif r tel que $f^r := f^{\circ r} = f \circ \cdots \circ f = 0$.

Proposition 3.14. *Soit $f \in \mathscr{L}(E)$ un endomorphisme nilpotent. Alors $\chi_f(X) = (-1)^n X^n$, où $n = \dim E$.*

Démonstration. On va montrer ce résultat par récurrence sur $n \in \mathbb{N}^*$. Pour $n = 1$, $f^r = 0$ implique $f = 0$, d'où le résultat. On suppose que le résultat est vrai au rang $r - 1$, et on le montre au rang r. On a $(\det f)^r = \det(f^r) = 0$, donc $\det(f) = 0$ et donc $\exists e_1 \in \ker f \setminus \{0\}$. On complète e_1 en une base $\mathscr{E}$ de E si bien que par blocs

$$[f]_{\mathscr{E}} = \left(\begin{array}{c|c} 0 & L \\ \hline 0 & B \end{array}\right)$$

où $L \in \mathscr{M}_{1,n}(\mathbb{K})$ et $B \in \mathscr{M}_{n-1}(\mathbb{K})$ vérifient

$$0 = [f^r]_{\mathscr{E}} = \left(\begin{array}{c|c} 0 & LB^{r-1} \\ \hline 0 & B^r \end{array}\right),$$

donc $B^r = 0$. D'après l'hypothèse de récurrence $\chi_B(X) = (-1)^{n-1}X^{n-1}$. Il suit que

$$\chi_f(X) = \det\left(\begin{array}{c|c} -X & L \\ \hline 0 & B - XI_{n-1} \end{array}\right) = (-X)\chi_B(X) = (-1)^n X^n. \qquad \square$$

Proposition 3.15 (Matrice compagnon)**.** *Soit $P = \sum_{0\le k\le n} a_k X^k \in \mathbb{K}[X]_n$ un polynôme unitaire de degré n ($a_n = 1$). Soit*

$$A = \begin{pmatrix} 0 & \cdots & \cdots & 0 & -a_0 \\ 1 & \ddots & & \vdots & -a_1 \\ 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & -a_{n-2} \\ 0 & \cdots & 0 & 1 & -a_{n-1} \end{pmatrix} \in \mathscr{M}_n(\mathbb{K}).$$

Alors $\chi_A = (-1)^n P$.

Démonstration. On va montrer le résultat par récurrence sur n. Si $n = 1$, $A = (-a_0)$ est un scalaire, Donc $\chi_A = -X - a_0$. Si $n = 2$, on a

$$A = \begin{pmatrix} 0 & -a_0 \\ 1 & -a_1 \end{pmatrix},$$

donc $\chi_A = (0 - X)(-a_1 - X) + a_0 = X^2 + a_1X + a_0$.

On suppose que le résultat est vrai au rang n, et on le montre au rang $n+1$. En développant par rapport à la première ligne, on obtient

$$\chi_A = -X \det\begin{pmatrix} -X & \cdots & \cdots & 0 & -a_1 \\ 1 & \ddots & & \vdots & -a_2 \\ 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & -X & -a_{n-1} \\ 0 & \cdots & 0 & 1 & -a_n - X \end{pmatrix} + (-1)^{n+1} a_0 \det\begin{pmatrix} 1 & -X & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & -X \\ 0 & \cdots & \cdots & 0 & 1 \end{pmatrix}.$$

Donc par l'hypothèse de récurrence, on a

$$\chi_A(X) = -X(-1)^n(X^n + \sum_{1\le k\le n} a_k X^{k-1}) + (-1)^{n+1}a_0 = (-1)^{n+1}(X^{n+1} + \sum_{0\le k\le n} a_k X^k). \qquad \square$$

Théorème 3.16. *Un endomorphisme $u \in \mathscr{L}(E)$ d'un espace vectoriel de dimension finie est diagonalisable si et seulement si son polynôme caractéristique est scindé et la multiplicité de chaque valeur propre est égale à la dimension de l'espace propre correspondant, c'est à dire si et seulement si*

$$\chi_u(X) = (-1)^n \prod_{\lambda \in \mathrm{Sp}(u)} (X - \lambda)^{\dim E_\lambda(u)}.$$

Démonstration. Si u est diagonalisable on peut trouver une base constituée de vecteurs propres de u. Dans cette base la matrice de u est diagonale et sur sa diagonale chaque valeur propre λ apparaît $\dim E_\lambda(u)$ fois donc χ_u est de la forme désirée. Réciproquement, si χ_u est de cette forme, alors $n = \deg \chi_u = \sum_{\lambda\in\mathrm{Sp}(u)} \dim E_\lambda(u)$, donc E est la somme directe des espaces propres de u, et u est diagonalisable. $\square$

4 Trigonalisation

Définition 4.1. Un endomorphisme $u \in \mathscr{L}(E)$ est dit trigonalisable s'il existe une base $\mathscr{E}$ de E dans laquelle la matrice de u est triangulaire supérieure : $([u]_{\mathscr{E}})_{i,j} = 0$ si $i > j$. On dit alors que la base $\mathscr{E}$ trigonalise u. Une matrice $A \in \mathscr{M}_n(\mathbb{K})$ est dite trigonalisable si elle est semblable à une matrice triangulaire supérieure.

Remarque 4.2. Un endomorphisme $u \in \mathscr{L}(E)$ est trigonalisable si et seulement si sa matrice dans une base quelconque de E est trigonalisable.

Théorème 4.3. *Un endomorphisme $u \in \mathscr{L}(E)$ est trigonalisable si et seulement si son polynôme caractéristique est scindé sur $\mathbb{K}$.*

Démonstration. On suppose que u est trigonalisable. Soit $\mathscr{E}$ une base de E telle que $M_{\mathscr{E}}(u)$ est triangulaire supérieure :

$$[u]_{\mathscr{E}} = \begin{pmatrix} \lambda_1 & \times & \cdots & \times \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \times \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}.$$

Alors $\chi_u(X) = (\lambda_1 - X)\dots(\lambda_n - X)$ est scindé sur $\mathbb{K}$.

Supposons maintenant que le polynôme caractéristique χ_u de u est scindé sur $\mathbb{K}$. Montrons par récurrence sur $n = \dim E$ que u est trigonalisable. Pour $n = 1$ c'est évident. Supposons le résultat au rang $n-1$ et montrons le au rang n. Comme le polynôme caractéristique est scindé, Il existe un vecteur propre e_1 de u, que l'on complète en une base $\mathscr{E} = (e_1, e_2, \dots, e_n)$ de E. On a

$$M = [u]_{\mathscr{E}} = \left(\begin{array}{c|ccc} \lambda_1 & \times & \cdots & \times \\ \hline 0 & & & \\ \vdots & & N & \\ 0 & & & \end{array}\right).$$

Alors

$$\chi_M = (\lambda_1 - X)\chi_N,$$

donc χ_N est scindé sur $\mathbb{K}$. Par l'hypothèse de récurrence, il existe P inversible telle que $P^{-1}NP$ est triangulaire supérieure. On peut alors vérifier que $Q^{-1}MQ$ est triangulaire supérieure avec Q définie par blocs par

$$Q = \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & P \end{array}\right).$$

□

5 Polynômes d'endomorphismes

Définition 5.1. Si u est un endomorphisme de E, on note $u^k = u^{\circ k} = u \circ \cdots \circ u$ (k fois) avec la convention $u^0 = \mathrm{id}_E$. Pour tout polynôme $P = \sum_{0 \le k \le d} a_k X^k$, on note

$$P(u) = \sum_{0 \le k \le d} a_k u^k \in \mathscr{L}(E).$$

Un calcul direct donne :

Proposition 5.2. *Soient $(P, Q) \in \mathbb{K}[X]^2$ et $u \in \mathscr{L}(E)$. Alors $(P+Q)(u) = P(u) + Q(u)$ et $P(u) \circ Q(u) = (PQ)(u) = (QP)(u)$.*

Si $A = Q^{-1}BQ$ avec Q inversible, alors $A^k = Q^{-1}B^kQ$ pour tout $k \ge 0$, donc pour tout $P \in \mathbb{K}[X]$, on a $P(A) = Q^{-1}P(B)Q$.

Un polynôme d'une matrice triangulaire supérieure est une matrice triangulaire supérieure. Plus précisément, si $P \in \mathbb{K}[X]$ alors

$$M = \begin{pmatrix} \lambda_1 & \times & \cdots & \times \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \times \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix} \Longrightarrow P(M) = \begin{pmatrix} P(\lambda_1) & \times & \cdots & \times \\ 0 & P(\lambda_2) & \ddots & \vdots \\ \vdots & \ddots & \ddots & \times \\ 0 & \cdots & 0 & P(\lambda_n) \end{pmatrix}.$$

Proposition 5.3. *Soit $P \in \mathbb{K}[X]$ et $u \in \mathscr{L}(E)$ tels que $P(u) = 0$. Si λ est une valeur propre de u alors $P(\lambda) = 0$.*

Démonstration. Soit $x \neq 0$ un vecteur propre associé à la valeur propre λ. Alors $u^k(x) = \lambda^k x$, donc $0 = P(u)(x) = P(\lambda)x$. Il suit que $P(\lambda) = 0$. □

La réciproque est fausse. Par exemple $P = X(X-1)$ et $u = id_E$. Alors $P(0) = 0$ mais 0 n'est pas une valeur propre de u.

5.1 Décomposition des noyaux

Théorème 5.4 (Lemme des noyaux)**.** *Soient $u \in \mathscr{L}(E)$ et $(P_k)_{1\le k\le r}$ une famille de polynôme deux à deux premiers entre eux. Soit $P = \prod_{1\le k\le r} P_k$. Alors*

$$\ker P(u) = \bigoplus_{1\le k\le r} \ker P_k(u).$$

Démonstration. On procède par récurrence sur $k \ge 2$. On commence par $k = 2$. Les polynômes P_1, P_2 sont premiers entre eux donc il existe deux polynômes Q_1, Q_2 tels que $Q_1P_1 + Q_2P_2 = 1$. Si $x \in \ker P_1(u) \cap \ker P_2(u)$, alors

$$x = Q_1(u) \circ P_1(u)(x) + Q_2(u) \circ P_2(u)(x) = 0.$$

Il suit que $\ker P_1(u) \cap \ker P_2(u) = \{0\}$. Si $x \in \ker P(u)$, on écrit $x = P_1(u)(Q_1(u)(x)) + P_2(u)(Q_2(u)(x)) =: y + z$. On a

$$P_2(u)(y) = (P_2(u) \circ P_1(u) \circ Q_1(u))(x) = Q_1(u) \circ P(u)(x) = 0,$$

et de même $P_1(u)(z) = 0$. Ainsi le résultat est vrai pour $r = 2$.

On suppose que le résultat est vrai au rang r, et on le montre au rang $r+1$. On note $Q = P_1 \cdots P_r$. Les polynômes Q et P_{r+1} sont premiers entre eux donc d'après le cas $r = 2$, on a

$$\ker P(u) = \ker Q(u) \oplus \ker P_{r+1}(u).$$

On conclut en utilisant l'hypothèse de récurrence. □

Théorème 5.5. *Un endomorphisme $u \in \mathscr{L}(E)$ est diagonalisable si et seulement s'il existe un polynôme $P \in \mathbb{K}[X]$ scindé sur $\mathbb{K}$ ayant toutes ses racines simples tel que $P(u) = 0$.*

Démonstration. On suppose que u est diagonalisable, alors

$$E = \bigoplus_{\lambda\in \mathrm{Sp}(u)} E_\lambda(u) = \bigoplus_{\lambda\in \mathrm{Sp}(u)} \ker((X-\lambda)(u)) = \ker P(u)$$

où $P = \prod_{\lambda\in\mathrm{Sp}(u)}(X-\lambda)$ grâce au lemme des noyaux. Ainsi $P(u) = 0$ avec P simplement scindé.

Réciproquement, on suppose qu'il existe un polynôme $P = \prod_{1\le k\le r}(X-\lambda_k)$ simplement scindé tel que $P(u) = 0$. Alors par le lemme des noyaux on a

$$E = \ker P(u) = \bigoplus_{1\le k\le r} E_{\lambda_k}(u)$$

ce qui implique que u est diagonalisable (quitte à enlever les indices pour lesquels $E_{\lambda_k}(u) = \{0\}$). □

Corollaire 5.6. *Soit $u \in \mathscr{L}(E)$ un endomorphisme diagonalisable et F un sous-espace stable par u. Alors u_F est aussi diagonalisable.*

Démonstration. D'après le théorème précédent, il existe un polynôme $P \in \mathbb{K}[X]$ simplement scindé tel que $P(u) = 0$. Il suit que $P(u_F) = P(u)_F = 0$. On applique encore une fois le théorème 5.5 pour conclure que u_F est diagonalisable. □

5.2 Théorème de Cayley–Hamilton

Si $u \in \mathscr{L}(E)$, on note I_u l'ensemble des polynômes $P \in \mathbb{K}[X]$ tels que $P(u) = 0$.

Proposition 5.7. *Si $u \in \mathscr{L}(E)$ alors l'ensemble I_u vérifie les conditions suivantes :*

— $0 \in I_u$,

— *pour tout $(P, Q) \in I_u^2$, $P + Q \in I_u$,*

— *pour tous $P \in I_u$ et $Q \in \mathbb{K}[X]$, $PQ \in I_u$.*

Les propriétés indiquées ci-dessus se traduisent en disant que I_u est un idéal de $\mathbb{K}[X]$. On l'appelle l'idéal annulateur de u.

L'idéal annulateur d'un endomorphisme u d'un espace vectoriel E de dimension n contient nécessairement un polynôme non nul de degré au plus n^2. En effet, la famille $(1, u, \ldots, u^{n^2})$ est liée vu que $\dim \mathscr{L}(E) = n^2$.

En réalité on a un résultat bien meilleur :

Théorème 5.8 (Cayley–Hamilton)**.** *Soit $u \in \mathscr{L}(E)$ un endomorphisme d'un espace vectoriel de dimension finie. Alors $\chi_u(u) = 0$.*

Démonstration. On va montrer que $\chi_u(u)(x) = 0$ pour tout $x \in E$. On fixe un vecteur $x \in E$ non nul. Soit r le plus grand nombre naturel tel que la famille $\mathscr{F} = (x, u(x), \ldots, u^{r-1}(x))$ soit libre. Alors $u^r(x) \in F = \mathrm{Vect}(x, u(x), \ldots, u^{r-1}(x))$ et il existe P unitaire de degré r tel que $P(u)(x) = 0$. Aussi, F est stable par u. Complétons $\mathscr{F}$ en une base de $\mathscr{E}$. Dans cette base la matrice de u s'écrit par blocs

$$[u]_{\mathscr{E}} = \left(\begin{array}{c|c} A & \star \\ \hline 0 & N \end{array}\right),$$

avec N carrée et $A = [u_F]_{\mathscr{F}}$ matrice compagnon du polynôme P. D'après la proposition 3.15, on a $\chi_A(X) = (-1)^r P$ et donc $\chi_u = \chi_N P$. Il suit que $\chi_u(u)(x) = \chi_N(u)(P(u)(x)) = \chi_N(u)(0) = 0$. $\square$

5.3 Polynôme minimal

Proposition 5.9. *Soit $u \in \mathscr{L}(E)$ un endomorphisme d'un espace vectoriel de dimension finie. Soit $P \in I_u$ non nul de degré minimal. Alors P divise tout élément de I_u.*

Démonstration. Si $Q \in I_u$, effectuons la division euclidienne de Q par P :

$$Q = PG + R,$$

avec $\deg R < \deg P$. Comme $R(u) = Q(u) - P(u) \circ G(u) = 0$, $R \in I_u$, et donc $R = 0$. $\square$

Définition 5.10. Le polynôme unitaire non nul de degré minimal dans I_u est appélé le polynôme minimal de u et est noté μ_u.

Théorème 5.11. *Un endomorphisme $u \in \mathscr{L}(E)$ d'un espace vectoriel de dimension finie est diagonalisable si et seulement si son polynôme minimal est simplement scindé.*

Démonstration. Si le polynôme minimal de u est simplement scindé, alors le théorème 5.5 assure que u est diagonalisable. Réciproquement, si u est diagonalisable, le théorème 5.5 implique qu'il existe un polynôme P annulateur de u simplement scindé. Le polynôme minimal de u divise P, donc il est aussi simplement scindé. $\square$

Proposition 5.12. *Soit $u \in \mathscr{L}(E)$ un endomorphisme d'un espace vectoriel de dimension finie. Alors, un scalaire λ est une valeur propre de u si et seulement s'il est une racine du polynôme minimal de u.*

Démonstration. D'après la proposition 5.3 toute valeur propre de u est racine de μ_u. Réciproquement, si $\mu_u(\lambda) = 0$, alors $\chi_u(\lambda) = 0$ car $\mu_u | \chi_u$. Donc λ est une valeur propre de u. $\square$

Remarque 5.13. Le polynôme minimal de u divise le polynôme caractéristique de u, et ils ont les mêmes racines. Ainsi

$$\chi_u = (-1)^n \prod_{\lambda \in \mathrm{Sp}(u)} (X-\lambda)^{n_\lambda} \Longrightarrow \mu_u = \prod_{\lambda \in \mathrm{Sp}(u)} (X-\lambda)^{m_\lambda}$$

avec $1 \leq m_\lambda \leq n_\lambda$. On a u diagonalisable si et seulement si tous les m_λ sont égaux à 1.

Exemple 5.14. Soit la matrice

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

Le polynôme caractéristique de M est $\chi_M = (X-1)^2(X-2)^2$. On vérifie que $(M-I)(M-2I) \neq 0$ et $(M-I)^2(M-2I) = 0$, donc le polynôme minimal de M est $(X-1)^2(X-2)$.

6 Réduction d'endomorphisme

On suppose que E est un $\mathbb{K}$-espace vectoriel de dimension finie n.

6.1 Décomposition de Dunford

Lemme 6.1. *Soient $u, v \in \mathscr{L}(E)$ diagonalisables tels que $u \circ v = v \circ u$. Alors il existe une base $\mathscr{E}$ de E telle que $[u]_{\mathscr{E}}$ et $[v]_{\mathscr{E}}$ sont toutes les deux diagonales. On dit que u et v sont codiagonalisables.*

Démonstration. Soit $\lambda \in \mathrm{Sp}(u)$. Alors on vérifie que $E_\lambda(u)$ est stable par v. D'après le corollaire 5.6 il existe une base $\mathscr{E}_\lambda$ de $E_\lambda(u)$ telle que la matrice de $v_{E_\lambda(u)}$ dans cette base est diagonale. La matrice de $u_{E_\lambda(u)}$ dans cette base est λI_{n_λ} où $n_\lambda = \dim E_\lambda(u)$. Une concaténation des $\mathscr{E}_\lambda$ convient puisque E est la somme directe des $E_\lambda(u)$. □

Définition 6.2. Soit $u \in \mathscr{L}(E)$ un endomorphisme, λ une valeur propre pour u, et n_λ sa multiplicité dans χ_u. On appelle

$$N_\lambda(u) = \ker(u - \lambda id_E)^{n_\lambda}$$

le sous-espace caractéristique de u associé à la valeur propre λ.

Proposition 6.3. *Soit $u \in \mathscr{L}(E)$ de polynôme caractéristique scindé. Alors*

$$E = \bigoplus_{\lambda \in \mathrm{Sp}(u)} N_\lambda(u)$$

et $\dim N_\lambda(u)$ est la multiplicité n_λ de λ dans χ_u.

Démonstration. La première décomposition découle du lemme des noyaux. On en déduit déjà que $\sum_{\lambda \in \mathrm{Sp}(u)} n_\lambda = \dim E = \sum_{\lambda \in \mathrm{Sp}(u)} \dim N_\lambda(u)$, et il suffit pour conclure de montrer que pour toute valeur propre λ on a $n_\lambda \geq \dim N_\lambda(u)$.

Soit λ une valeur propre. On remarque déjà que $N_\lambda(u)$ est stable par u (c'est le noyau d'un polynôme en u). Si on note u_λ l'endomorphisme induit, il vérifie

$$(u_\lambda - \lambda \mathrm{id}_{N_\lambda(u)})^{n_\lambda} = 0$$

par définition de $N_\lambda(u)$. D'après la proposition 3.14, le polynôme caractéristique de $u_\lambda - \lambda \mathrm{id}_{N_\lambda(u)}$ est $(-X)^{\dim N_\lambda(u)}$ donc celui de u_λ est $(\lambda - X)^{\dim N_\lambda(u)}$. D'après la proposition 3.12, χ_{u_λ} divise χ_u, donc $n_\lambda \geq \dim N_\lambda(u)$ comme souhaité. □

Théorème 6.4 (Décomposition de Dunford)**.** *Soit $u \in \mathscr{L}(E)$ de polynôme caractéristique scindé. Il existe un unique couple $(d, v) \in (\mathscr{L}(E))^2$ avec d diagonalisable et v nilpotente tel que $u = d + v$ et $d \circ v = v \circ d$.*

Démonstration. Pour l'existence il suffit de définir d et v sur chaque $N_\lambda(u)$ par $d_{N_\lambda(u)} = \lambda \mathrm{id}_{N_\lambda(u)}$ et $v_{N_\lambda(u)} = u_{N_\lambda(u)} - d_{N_\lambda(u)}$. Alors d commute avec u, puis v aussi, et v est nilpotent d'après la preuve précédente.

On montre l'unicité de cette construction. On suppose que $(d', v') \in \mathscr{L}(E)^2$ tels que d' est diagonalisable, v' est nilpotent, $u = d' + v'$, et $d' \circ v' = v' \circ d'$. On veut montrer que $d' = d$ et $v' = v$.

Déjà d commute avec d'. En effet d' stabilise les sous-espaces caractéristiques car d' commute avec u, et d est scalaire sur ces sous-espaces. Il suit que $v = u - d$ commute avec $v' = u - d'$. Si on choisit p, q tels que $v^p = 0$ et $v'^q = 0$, on a

$$(v - v')^{p+q} = \sum_{i+j=p+q} \binom{p+q}{i} v^i (v')^j = 0$$

car dans tous les membres de la somme on a $i \geq p$ ou $j \geq q$. Mais alors $d' - d = v - v'$ est nilpotent et diagonalisable d'après le lemme 6.1, donc $d' - d = 0$, puis $v' = v$. □

6.2 Réduction de Jordan

On utilise la notation $A \subsetneq B$ lorsque $A \subset B$ et $A \neq B$.

Proposition 6.5. *Pour tout $u \in \mathscr{L}(E)$, il existe un unique naturel r tel que*

$$\{0\} = \ker(u^0) \subsetneq \ker(u) \subsetneq \cdots \subsetneq \ker(u^r) = \ker(u^{r+1}) = \ker(u^{r+2}) = \cdots$$

L'entier r s'appelle l'indice de u. C'est aussi le plus petit entier naturel tel que $\ker(u^r) = \ker(u^{r+1})$.

Démonstration. On pose $n_p := \dim \ker(u^p)$. Comme $\ker(u^p) \subset \ker(u^{p+1})$, la suite $(n_p)_{p \in \mathbb{N}}$ est croissante, à valeur dans $\{0, 1, \ldots, n\}$. On peut donc définir

$$r = \min\{p \in \mathbb{N} \mid n_p = n_{p+1}\}.$$

Pour tout $p \geq r$, et tout $x \in \ker(u^{p+1})$ on a $u^{r+1}(u^{p-r}(x)) = 0$, donc $u^{p-r}(x) \in \ker(u^{r+1}) = \ker(u^r)$, donc $u^p(x) = u^r(u^{p-r}(x) = 0$, donc $x \in \ker(u^p)$. Il suit que $\ker(u^p) = \ker(u^{p+1})$. □

Remarque 6.6. — L'indice r de u est aussi l'unique entier vérifiant

$$\forall q < r,\ \ker(u^q) \subsetneq \ker(u^r),\ \text{et}\ \forall q \geq r,\ \ker(u^q) = \ker(u^r).$$

— Si u est nilpotent, l'indice r de u n'est autre que $\min\{p \in \mathbb{N} \mid u^p = 0\}$.

— Si r est l'indice de u, alors

$$\ker(u^r) \oplus \mathrm{Im}(u^r) = E.$$

— L'indice r de u vérifie aussi

$$E = \mathrm{Im}(u^0) \supsetneq \mathrm{Im}(u) \supsetneq \ldots () \supsetneq \mathrm{Im}(u^r) = \mathrm{Im}(u^{r+1}) = \mathrm{Im}(u^{r+2}) = \cdots$$

La démonstration de ces propriétés est laissée en exercice.

Théorème 6.7. *Soit $u \in \mathscr{L}(E)$ un endomorphisme de polynôme caractéristique scindé. La multiplicité d'une valeur propre λ dans le polynôme minimal est donnée par l'indice de l'endomorphisme $u - \lambda id_E$.*

Démonstration. On écrit $\mu_u = (X - \lambda)^m Q$ où m est la multiplicité cherchée, de sorte que $(X - \lambda)^m$ et Q sont premiers entre eux. D'après le lemme des noyaux, on a

$$E = \ker(u - \lambda \mathrm{id}_E)^m \oplus \ker(Q(u)).$$

Pour chaque naturel k, on pose $Q_k = (X - \lambda)^k Q$. On a

$$\ker(Q_k(u)) = \ker(u - \lambda \mathrm{id}_E)^k \oplus \ker(Q(u)).$$

Par minimalité de μ_u, si $k < m$ on a $\ker(Q_k(u)) \subsetneq E$ donc $\ker(u - \lambda \mathrm{id}_E)^k \subsetneq \ker(u - \lambda \mathrm{id}_E)^m$, tandis que si $k \geq m$, $\ker(Q_k(u)) = E$ et $\ker(u - v)^k = \ker(u - \lambda \mathrm{id}_E)^m$. □

Définition 6.8. Un bloc de Jordan est une matrice $J_k(\lambda) \in \mathscr{M}_k(\mathbb{R})$ de la forme $J_1(\lambda) = \lambda$ et

$$J_k(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \lambda & 1 \\ 0 & \cdots & \cdots & 0 & \lambda \end{pmatrix}.$$

Une matrice de Jordan est une matrice de la forme

$$J = \begin{pmatrix} J_{n_1}(\mu_1) & 0 & \cdots & 0 \\ 0 & J_{n_2}(\mu_2) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & J_{n_r}(\mu_r) \end{pmatrix},$$

où chaque $J_{n_i}(\mu_i)$ est un bloc de Jordan.

Les redondances sont autorisées parmi les μ_i.

Théorème 6.9 (Réduction de Jordan d'un endomorphisme nilpotent)**.** *Soit $u \in \mathscr{L}(E)$ un endomorphisme nilpotent. Alors, il existe une base $\mathscr{E}$ de E telle que*

$$[u]_{\mathscr{E}} = \begin{pmatrix} J_{n_1}(0) & 0 & \cdots & 0 \\ 0 & J_{n_2}(0) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & J_{n_r}(0) \end{pmatrix},$$

est une matrice nilpotente de Jordan.

Démonstration. On va démontrer le résultat par récurrence sur $n = \dim E$. Si $n = 1$, il y a rien à faire. On suppose le résultat vrai au rang n, et on le montre au rang $n+1$. Soit $F = \mathrm{Im}(u)$ l'image de E par u. Comme u n'est pas surjective, la dimension de F est strictement inférieure à $n+1$. De plus, F est stable par u, donc on peut appliquer l'hypothèse de récurrence à u_F. Il existe donc une base de F composée de vecteurs $u^{n_1-1}(f_1), \ldots, f_1; \ldots; u^{n_r-1}(f_r), \ldots, f_r$, avec $u^{n_i}(f_i) = 0$, pour tout $1 \le i \le r$.

Pour chaque $1 \le i \le r$, il existe $e_i \in E$ tel que $u(e_i) = f_i$.

On montre que $\mathrm{Im}(u) \cap \ker(u) = \mathrm{Vect}(u^{n_1}(e_1), \ldots, u^{n_r}(e_r))$. En effet, soit $x \in \mathrm{Im}(u) \cap \ker(u)$. Alors, on peut écrire

$$x = \sum_{j=1}^{r} \sum_{k=0}^{n_j-1} \mu_{j,k} u^k(f_j),$$

avec $\mu_{j,k} \in \mathbb{K}$. En appliquant u, on obtient

$$0 = \sum_{j=1}^{r} \sum_{k=1}^{n_j-2} \mu_{j,k} u^{k+1}(f_j).$$

La famille $(u^k(f_j))_{1 \le j \le r, 1 \le k \le n_r-1}$ étant libre, on déduit que $\mu_{j,k} = 0$ dès que $k \le n_j - 2$, donc $x \in \mathrm{Vect}(u^{n_1}(e_1), \ldots, u^{n_r}(e_r))$.

On ajoute $e_{r+1}, \ldots, e_s$ à $u^{n_1}(e_1), \ldots, u^{n_r}(e_r)$ pour avoir une base de $\ker(u)$. Soit $\mathscr{E}$ la concaténation de $(u^k(e_j))_{1 \le j \le r, 0 \le k \le n_j}$ et $(e_{r+1}, \ldots, e_s)$.

Montrons $\mathscr{E}$ est libre. Soit une relation de dépendance linéaire

$$0 = \sum_{i=r+1}^{s} \lambda_i e_i + \sum_{j=1}^{r} \sum_{k=0}^{n_j} \mu_{j,k} u^k(e_j).$$

En appliquant u, on obtient

$$0=\sum_{j=1}^{r}\sum_{k=0}^{n_j-1}\mu_{j,k}u^{k+1}(e_j)$$

donc $\mu_{j,k}=0$ si $k\leq n_j-1$, et

$$0=\sum_{j=1}^{r}\mu_{j,n_j}u^{n_j}(e_j)+\sum_{i=r+1}^{s}\lambda_i e_i.$$

La famille $(e_{r+1},\dots,e_s,u^{n_1}(e_1),\dots,u^{n_r}(e_r))$ étant libre, on obtient ce qu'on veut.

On a $\dim\operatorname{Im}(u)=n_1+\cdots+n_r$, $\dim\ker(u)=s$, donc $n_1+\cdots+n_r+s=\dim E$. Le cardinal de $\mathscr{E}$ est $(1+n_1)+\cdots+(1+n_r)+s-r=n_1+\cdots+n_r+s=\dim E$, donc $\mathscr{E}$ est une base de E. □

Théorème 6.10 (Réduction de Jordan d'un endomorphisme)**.** *Soit $u\in\mathscr{L}(E)$ un endomorphisme tel que son polynôme caractéristique χ_u soit scindé sur $\mathbb{K}$. Alors, il existe une base $\mathscr{E}$ de E telle que*

$$[u]_{\mathscr{E}}=J=\begin{pmatrix} J_{n_1}(\mu_1) & 0 & \cdots & 0\\ 0 & J_{n_2}(\mu_2) & \ddots & \vdots\\ \vdots & \ddots & \ddots & 0\\ 0 & \cdots & 0 & J_{n_r}(\mu_r)\end{pmatrix},$$

est de Jordan. La matrice J est unique modulo permutation de blocs.

Ici les l'ensemble des μ_i est le spectre de u, avec éventuelles redondances.

Démonstration. D'après le théorème précédent, pour toute valeur propre λ, il existe une base $\mathscr{E}_\lambda$ de $N_\lambda(u)$ dans laquelle la matrice de $u_{N_\lambda(u)}-\lambda id_{N_\lambda(u)}$ est de Jordan nilpotente. On en déduit l'existence par concaténation des $\mathscr{E}_\lambda$.

L'unicité de la matrice J modulo permutation de blocs se traduit en disant que pour chaque paire (μ_i,n_i), le nombre de blocs de Jordan $J_{n_i}(\mu_i)$ est uniquement déterminé par u.

On remarque que pour tout $m\geq 0$, la matrice $J_k(0)^m$ est donnée par $(J_k(0)^m)_{i,j}=\delta_{j,i+m}$. Il suit que

$$\operatorname{rg}(J_k(0)^m)=\max(k-m,0).$$

Quitte à permuter les blocs de Jordan, on peut supposer que $\mu_1=\cdots=\mu_k$ et $\mu_1\neq\mu_j$ si $j>k$. La matrice $J-\mu_1 I_n$ se décompose en blocs dont seuls les k premiers sont non inversibles. Ainsi, pour tout $m\geq 0$,

$$\operatorname{rg}(J-\mu_1 I_n)^m=\max(n_1-m,0)+\cdots+\max(n_k-m,0)+n_{k+1}+\cdots+n_r.$$

Il suit que

$$\operatorname{rg}(u-\mu_1\mathrm{id}_E)^{m-1}-\operatorname{rg}(u-\mu_1\mathrm{id}_E)^m=d_{\geq m},$$

où $d_{\geq m}=\#\{1\leq j\leq k\mid n_j\geq m\}$ est le nombre de blocs de Jordan de taille $\geq m$ et d'éléments diagonaux μ_1. Par conséquent, $d_{\geq m}$ est uniquement déterminé par u (le rang ne dépend pas de la base). Comme le nombre de blocs de Jordan de taille m et d'élément diagonaux λ_1 est $d=d_{\geq m}-d_{\geq m+1}$, il suit que d est uniquement déterminé par u. □

7 Formes bilinéaires

Définition 7.1. Soit E un $\mathbb{K}$-espace vectoriel. Une application $\Phi:E\times E\to\mathbb{K}$ est une forme bilinéaire si pour tout $x\in E$ les applications $y\mapsto\Phi(x,y)$ et $y\mapsto\Phi(y,x)$ de E dans $\mathbb{K}$ sont des formes linéaires.

7.1 Écriture dans une base

- Soit $\Phi : E \times E \to \mathbb{K}$ une forme bilinéaire et $\mathscr{E} = (e_1, \dots, e_n)$ une base de E. Par bilinéarité, on a
$$\Phi(x_1 e_1 + \cdots + x_n e_n, y_1 e_1 + \cdots + y_n e_n) = \sum_{1 \le i,j \le n} x_i y_j \Phi(e_i, e_j).$$
Considérons la matrice $M = (\Phi(e_i, e_j))_{i,j}$. Soient $x, y \in E$, $X = [x]_{\mathscr{E}}$, $Y = [y]_{\mathscr{E}}$. La formule précédente s'écrit
$$\Phi(x, y) = {}^tXMY.$$
La matrice M est appelée la matrice de Φ dans la base $\mathscr{E}$ et sera notée $[\Phi]_{\mathscr{E}}$.
- Soit $\mathscr{E}'$ une autre base de E et soit $P = \mathscr{P}_{\mathscr{E}}^{\mathscr{E}'}$ la matrice de passage de la base $\mathscr{E}$ à la base $\mathscr{E}'$. Alors $X = PX'$ et $Y = PY'$, où $X' = [x]_{\mathscr{E}'}$ et $Y' = [y]_{\mathscr{E}'}$. Soit M' la matrice de Φ dans la base $\mathscr{E}'$. On obtient ainsi
$${}^tX'M'Y' = \Phi(x, y) = {}^tXMY = {}^tX'{}^tPMPY',$$
donc
$$[\Phi]_{\mathscr{E}'} = {}^tP[\Phi]_{\mathscr{E}}P.$$

7.2 Dualité

Rappelons que si E est un $\mathbb{K}$-espace vectoriel de dimension n, l'ensemble E^* des formes linéaires sur E (i.e. des applications linéaires $E \to \mathbb{K}$) est un $\mathbb{K}$-espace vectoriel de dimension n. Si $\mathscr{E} = (e_1, \dots, e_n)$ est une base de E, on définit une base $\mathscr{E}^* = (e_1^*, \dots, e_n^*)$ de E^*, en posant
$$e_i^*(x_1 e_1 + \cdots + x_n e_n) = x_i.$$

- Toute forme linéaire u sur E s'écrit
$$u = \sum_{i=1}^n u(e_i) e_i^*$$
et tout vecteur f de E s'écrit
$$f = \sum_{i=1}^n e_i^*(f) e_i.$$
Si $\mathscr{F} = (f_1, \dots, f_n)$ est une autre base de E on en déduit que
$$(\mathscr{P}_{\mathscr{E}}^{\mathscr{F}})_{i,j} = e_i^*(f_j).$$
- Il existe une application $\mathrm{ev} : E \to E^{**}$ définie par $\mathrm{ev}(f) = [\varphi \mapsto \varphi(f)]$ qui est clairement injective. Comme $\dim E^{**} = \dim E^* = \dim E$, θ est un isomorphisme. Si $\Xi = (\xi_1, \dots, \xi_n)$ une base de E^*, on peut donc définir sa base *antéduale* $\mathscr{F} = (f_1, \dots, f_n)$ par $f_j = \mathrm{ev}^{-1}(\xi_j^*)$, si bien que $\xi_i(f_j) = \mathrm{ev}(f_j)(\xi_i) = \xi_j^*(\xi_i) = \delta_{i,j}$ et donc $\mathscr{F}^* = \Xi$. Comme
$$e_j^* = \sum_{i=1}^n e_j^*(f_i) f_i^*$$
on en déduit que
$$\mathscr{P}_{\mathscr{F}^*}^{\mathscr{E}^*} = {}^t\mathscr{P}_{\mathscr{E}}^{\mathscr{F}}.$$
- Soit Φ une forme bilinéaire sur E. On lui associe une application linéaire $\theta_\Phi : E \to E^*$ en posant
$$\theta_\Phi(x)(y) = \Phi(x, y).$$
On note $[\Phi]_{\mathscr{E}} = (a_{i,j})$. On a
$$\Phi(x_1 e_1 + \cdots + x_n e_n, y_1 e_1 + \cdots + y_n e_n) = \sum_{1 \le i,j \le n} a_{i,j} x_i y_j = \sum_{1 \le i,j \le n} a_{i,j} x_i e_j^*(y),$$
donc
$$\theta_\Phi(x) = \sum_{1 \le i,j \le n} a_{i,j} x_i e_j^*.$$
En particulier $\theta_\Phi(e_i) = \sum_{j=1}^n a_{i,j} e_j^*$ et donc
$${}^t[\Phi]_{\mathscr{E}} = [\theta_\Phi]_{\mathscr{E}}^{\mathscr{E}^*}. \tag{3}$$

7.3 Formes bilinéaires symétriques et formes quadratiques

Définition 7.2. On dit qu'une forme bilinéaire Φ sur E est symétrique si

$$\forall (x,y) \in E^2,\ \Phi(x,y) = \Phi(y,x).$$

Définition 7.3. Une application $Q : E \to \mathbb{K}$ est une forme quadratique si il existe une forme bilinéaire symétrique Φ sur E telle que

$$\forall x \in E,\ Q(x) = \Phi(x,x).$$

Q est appelée la forme quadratique associée à la forme bilinéaire Φ.

Proposition 7.4 (Formule de polarisation)**.** *Soit Φ est une forme bilinéaire symétrique, et Q la forme quadratique associée à Φ. Alors*

$$\forall (x,y) \in E^2,\ \Phi(x,y) = \frac{1}{2}(Q(x+y) - Q(x) - Q(y)) = \frac{1}{4}(Q(x+y) - Q(x-y)).$$

Démonstration. Développer les termes de droite, en utilisant la bilinéarité et la symétrie de Φ. □

Cette proposition veut donc dire que la forme quadratique Q détermine la forme bilinéaire symétrique Φ, que l'on appelle *forme polaire* de Q.

Remarque 7.5. Si Φ est une forme bilinéaire symétrique sur $\mathbb{K}^n$, alors sa matrice $A = (a_{i,j})$ dans la base canonique est symétrique. Pour tout $x \in \mathbb{K}^n$ on a

$$Q(x) = \Phi(x,x) = \sum_{i=1}^{n} a_{i,i} x_i^2 + \sum_{i \neq j} a_{i,j} x_i x_j = \sum_{i=1}^{n} a_{i,i} x_i^2 + 2 \sum_{i<j} a_{i,j} x_i x_j.$$

par symétrie de A. Cette formule permet réciproquement d'écrire la matrice A, à partir de la forme quadratique. Par exemple, considérons la forme quadratique sur $\mathbb{K}^3$ définie par

$$Q(x) = x_1^2 + 2x_2^2 + 3x_3^2 + 4x_1x_2 + 5x_2x_3 + 6x_3x_1.$$

La matrice de la forme bilinéaire associée est alors

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 2 & 5/2 \\ 3 & 5/2 & 3 \end{pmatrix}.$$

7.4 Formes quadratiques définies

Définition 7.6. Une forme bilinéaire symétrique Φ sur E est dite non dégénérée si, pour tout $x \in E$,

$$(\forall y \in E, \Phi(x,y) = 0) \Longrightarrow x = 0.$$

Remarque 7.7. Dire que Φ est non dégénérée se traduit en disant que l'application linéaire induite $\theta_\Phi : E \to E^*$ est injective (et donc bijective). D'après (3), cela équivaut à dire que la matrice de Φ dans une base quelconque est inversible, ou encore que son déterminant est non nul.

Définition 7.8. Une forme bilinéaire symétrique Φ est dite définie si

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

Proposition 7.9. *Si une forme bilinéaire symétrique Φ est définie alors elle est non-dégénérée.*

Démonstration. Soit $x \in E$ tel que $\theta_\Phi(x) = 0$. Alors $\Phi(x,x) = \theta_\Phi(x)(x) = 0$, donc $x = 0$. □

Remarque 7.10. La réciproque est fausse. Ainsi, la forme bilinéaire symétrique Φ sur $\mathbb{R}^2$ définie par

$$\Phi((x_1,x_2),(y_1,y_2)) = x_1 y_2 + x_2 y_1$$

est non dégénérée, mais elle n'est pas définie, puisque $\Phi((1,0),(1,0)) = 0$.

Proposition 7.11. *Soit Q une forme quadratique définie sur un $\mathbb{R}$-espace vectoriel. Alors*

— *soit $\forall x \in E$, $Q(x) > 0$ (on dit alors que Q est définie positive),*

— *soit $\forall x \in E$, $Q(x) < 0$ (on dit alors que Q est définie négative).*

Démonstration. On suppose par l'absurde qu'il existe $x, y \in E$ tel que $Q(x)Q(y) < 0$. On remarque d'abord que x et y ne sont pas colinéaires. En effet si $x = \lambda y$ alors $Q(x) = \lambda^2 Q(y)$, donc ils sont de même signe.

On considère la fonction $\gamma(t) = Q(tx + (1-t)y)$, $t \in \mathbb{R}$. C'est un polynôme en t et on a $\gamma(0)\gamma(1) < 0$. D'après le théorème des valeurs intermédiaires, il existe $t \in \mathbb{R}$ tel que $\gamma(t) = 0$. Comme Q est non dégénérée, il suit que $tx + (1-t)y = 0$, donc x et y sont colinéaire, une contradiction. □

7.5 Réduction d'une forme quadratique en somme des carrés

Traitons d'abord des exemples.

Exemple 7.12. Considérons la forme quadratique sur $\mathbb{R}^3$ définie par

$$Q(x_1, x_2, x_3) = x_1^2 + x_1 x_2 + x_2 x_3 + x_3 x_1.$$

On va essayer de l'écrire sous une forme plus maniable. Pour cela, commence à voir l'expression $x_1^2 + x_1(x_2 + x_3)$ comme le début d'un carré.

$$x_1^2 + x_1(x_2 + x_3) = \left(x_1 + \frac{x_2 + x_3}{2}\right)^2 - \frac{1}{4}(x_2 + x_3)^2.$$

Il suit donc que

$$\begin{aligned} Q(x_1, x_2, x_3) &= \left(x_1 + \frac{x_2 + x_3}{2}\right)^2 - \frac{1}{4}(x_2 + x_3)^2 + x_2 x_3 \\ &= \left(x_1 + \frac{x_2 + x_3}{2}\right)^2 - \frac{1}{4}(x_2 - x_3)^2. \end{aligned}$$

Exemple 7.13. On considère la forme quadratique

$$L(x_1, x_2, x_3) = x_1 x_2 + x_2 x_3 + x_3 x_1.$$

Ici, on ne peut pas faire apparaître le début d'un carré ; on va procéder autrement en faisant apparaître un produit

$$L(x) = (x_1 + x_3)(x_2 + x_3) - x_3^2 = \frac{1}{4}(x_1 + x_2 + 2x_3)^2 - \frac{1}{4}(x_1 - x_2)^2 - x_3^2.$$

Dans chacun des deux exemples, on a transformé la forme quadratique en combinaisons linéaires de carrés de formes linéaires. Le théorème suivant formalise ceci.

Théorème 7.14. *Soit Q une forme quadratique sur un $\mathbb{K}$-espace vectoriel de dimension n. Alors il existe des formes linéaires $f_1, \dots, f_r$ sur E, qui sont linéairement indépendantes, et des éléments non nuls $a_1, \dots, a_r \in \mathbb{K}$ tels que*

$$\forall x \in E,\ Q(x) = a_1 f_1(x)^2 + \cdots + a_r f_r(x)^2.$$

Remarque 7.15. Remarquons tout de suite que, puisque les formes linéaires f_i sont linéairement indépendantes, on a $r \leq n$.

Remarque 7.16. Il faut aussi noter que la famille de formes linéaires qui interviennent n'est pas unique, comme le montrent les deux écritures suivantes d'une forme quadratique sur $\mathbb{R}^2$:

$$2x_1^2 + 2x_2^2 = (x_1 + x_2)^2 + (x_1 - x_2)^2.$$

Démonstration. La preuve se fait par récurrence forte sur la dimension n.

Pour $n = 1$, c'est évident puisque une forme quadratique sur $\mathbb{K}$ s'écrit $Q(x) = ax^2$.

Supposons le résultat vrai pour toute dimension $k < n$. Considérons une forme quadratique non nulle $Q(x) = \sum_{i=1}^n a_{i,i}x_i^2 + 2\sum_{i<j} a_{i,j}x_ix_j$ décomposée dans une base quelconque. On distingue deux cas.

Cas 1 : au moins l'un des coefficients $a_{i,i}$ est non nul. On choisit le premier, que l'on suppose être $a_{1,1} \neq 0$ sans perte de généralité, et écrivons

$$Q(x) = a_{1,1}x_1^2 + 2x_1 \sum_{j=2}^n a_{1,j}x_j + R(x_2, \dots, x_n).$$

On voit dans les deux premiers termes le début d'un carré, que l'on fait donc apparaître, en regroupant tous les termes en trop dans un terme $S(x_2, \dots, x_n)$:

$$Q(x) = a_{1,1}\left(x_1 + \sum_{j=2}^n \frac{a_{1,j}}{a_{1,1}}x_j\right)^2 + S(x_2, \dots, x_n).$$

$S(x_2, \dots, x_n)$ est une forme quadratique en $(n-1)$ variables, qu'on peut donc décomposer en

$$S(x_2, \dots, x_n) = a_2 f_2(x_2, \dots, x_n)^2 + \cdots + a_r f_r(x_2, \dots, x_n)^2,$$

où $f_2, \dots, f_r$ sont linéairement indépendantes. En posant

$$f_1(x) = x_1 + \sum_{j=2}^n \frac{a_{1,j}}{a_{1,1}}x_j,$$

les formes $f_1, \dots, f_r$ sont indépendantes et on a

$$Q(x) = a_{1,1}f_1(x)^2 + \cdots + a_r f_r(x)^2.$$

Cas 2 : Tous les coefficients $a_{i,i}$ sont nuls. Alors au moins un coefficient $a_{i,j}$ est non nul. Supposons donc sans perdre de généralité $a_{1,2} \neq 0$, et écrivons

$$Q(x) = a_{1,2}\left(x_1 + \sum_{j=3}^n \frac{a_{2,j}}{a_{1,2}}x_j\right)\left(x_2 + \sum_{j=3}^n \frac{a_{1,j}}{a_{1,2}}x_j\right) + R(x_3, \dots, x_n).$$

Le produit qui intervient dans cette expression peut se transformer, par la formule

$$ab = \frac{1}{4}(a+b)^2 - \frac{1}{4}(a-b)^2,$$

en combinaison linéaire des carrés de deux formes linéaires indépendantes. En appliquant l'hypothèse de récurrence à $R(x_3, \dots, x_n)$, on obtient le résultat recherché. □

Remarque 7.17. Cette preuve est constructive : l'algorithme qui en découle s'appelle la *réduction de Gauss*. Nous verrons plus tard dans le cas réel une autre manière d'obtenir une décomposition en combinaison linéaire de carrés de formes linéaires deux à deux *orthogonales*. Ce sera une conséquence de l'orthodiagonalisation des matrices symétriques réelles.

7.6 Invariants d'une forme quadratique

On a vu que l'écriture

$$Q(x) = a_1 f_1(x)^2 + \cdots + a_r f_r(x)^2$$

où les a_i sont non nuls, et les f_i sont des formes linéaires linéairement indépendantes, n'est pas unique. Cependant certaines données numériques de cette décomposition ne dépendent que de Q, et pas de la décomposition choisie. Grâce à 7.1, on peut parler du *rang* d'une forme bilinéaire.

Théorème 7.18 (Théorème d'inertie de Sylvester)**.** *Soit Q une forme quadratique sur un $\mathbb{K}$-espace vectoriel de dimension n. Le nombre r de formes linéaires qui interviennent dans une décomposition de Q est égale au rang de la forme polaire de Q.*

Démonstration. Soit Q une forme quadratique sur un $\mathbb{K}$-espace vectoriel de dimension n, et considérons des formes linéaires, linéairement indépendantes, $f_1,\dots,f_r$ telles que $Q = a_1 f_1^2 + \cdots + a_r f_r^2$, avec les a_i non nuls. Complétons $f_1,\dots,f_r$ en une base $f_1,\dots,f_n$ de E^*. On peut trouver alors une base $\mathscr{E} = (e_1,\dots,e_n)$ de E tel que $f_i(e_j) = \delta_{i,j}$ (base antéduale, cf 7.2). Soit Φ la forme polaire de Q. Alors on peut vérifier que

$$\Phi(e_i, e_j) = \frac{1}{2}(Q(e_i + e_j) - Q(e_i) - Q(e_j)) = \delta_{i,j} a_i.$$

La matrice de Φ dans la base $\mathscr{E}$ est alors

$$M = \left(\begin{array}{c|c} A & 0 \\ \hline 0 & 0 \end{array}\right) \quad \text{où} \quad A = \begin{pmatrix} a_1 & & 0 \\ & \ddots & \\ 0 & & a_r \end{pmatrix}.$$

L'entier r est donc le rang de Φ. □

Théorème 7.19 (Théorème d'inertie de Sylvester sur $\mathbb{R}$)**.** *Soit Q une forme quadratique sur un $\mathbb{R}$-espace vectoriel E de dimension n. Soit*

$$Q(x) = a_1 f_1(x)^2 + \cdots + a_s f_s(x)^2 - a_{s+1} f_{s+1}(x)^2 - \cdots - f_{s+t}(x)^2$$

une décomposition de Q en combinaison linéaire de carrés de formes linéaires linéairement indépendantes, avec $\forall i,\ a_i > 0$. Alors les nombres s et t ne dépendent que de Q et pas de la décomposition choisie.

Définition 7.20. Le couple (s, t) est appelé la signature de Q.

Démonstration. On cherche à exprimer s et t comme des invariants dépendant uniquement de Q. Complétons la famille libre $f_1,\dots,f_{s+t}$ en une base $f_1,\dots,f_n$ de E^*. Soit $\mathscr{E} = (e_1,\dots,e_n)$ la base de E antéduale. Introduisons les sous-espaces vectoriels $F = \mathrm{Vect}(e_1,\dots,e_s)$ et $G = \mathrm{Vect}(e_{s+1},\dots,e_n)$.

D'abord, la restriction $Q|_F$ est définie positive, et on a $\dim F = s$. D'autre part, si F' est un sous-espace vectoriel de E tel que $Q|_{F'}$ est définie positive, alors $\dim F' \le s$. En effet, dans le cas contraire, on aurait $\dim F' + \dim G > n$, donc $F' \cap G \neq \{0\}$. Soit x un élément non nul de $F' \cap G$. On a $Q(x) > 0$ puisque $x \in F'$, mais $Q(x) \le 0$ puisque $x \in G$, ce qui conduit à une contradiction.

Il résulte du raisonnement précédent que

$$s = \max\{\dim V \mid V \text{ sev de } E \text{ tel que } Q|_V \text{ définie positive}\}$$

Un argument analogue montre que

$$t = \max\{\dim W \mid W \text{ sev de } E \text{ tel que } Q|_W \text{ définie négative}\}.$$ □

Remarque 7.21. Nous verrons plus tard que s (resp. t) est le nombre de valeurs propres strictement positives (resp. négatives) d'une matrice représentant la forme polaire de Q.aa

8 Espaces euclidiens

Définition 8.1. Soit E un $\mathbb{R}$-espace vectoriel. Un produit scalaire sur E est une forme bilinéaire symétrique définie positive. On note $\langle x, y\rangle$ le produit scalaire de x et y. Un espace vectoriel de dimension finie muni d'un produit scalaire est appelé un espace euclidien.

Définition 8.2. Si E est muni d'un produit scalaire, pour tout $x \in E$ on note $\|x\| = \sqrt{\langle x, x\rangle}$.

Exemple 8.3. Dans $\mathbb{R}^n$, le produit scalaire usuel (ou canonique) est défini par

$$\langle x, y\rangle = \sum_{i=1}^{n} x_i y_i \quad \text{si} \quad x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix},\ y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

Exemple 8.4. Sur $\mathbb{R}^3$ considérons l'application $\langle -, -\rangle$ définie par

$$\langle x, y\rangle = x_1 y_1 + 2x_2 y_2 + 3x_3 y_3 + x_1 y_2 + x_2 y_1.$$

Vérifier que cette application définit un produit scalaire sur $\mathbb{R}^3$.

8.1 Norme d'un vecteur, angle non orienté

Soit $(E, \langle -, -\rangle)$ un espace euclidien. La norme d'un vecteur $x \in E$ est définie par

$$\|x\| = \sqrt{\langle x, x\rangle}.$$

Un calcul direct donne

$$2\langle x, y\rangle = \|x+y\|^2 - \|x\|^2 - \|y\|^2.$$

Lemme 8.5 (Inégalité de Cauchy-Schwarz)**.** *Pour tout $x, y \in E$ on a*

$$\left|\langle x, y\rangle\right| \leq \|x\| \|y\|,$$

l'égalité ayant lieu si et seulement si x et y sont liés.

Démonstration. Si $y = 0$ alors c'est évident. On suppose maintenant que $y \neq 0$. Pour tout $\lambda \in \mathbb{R}$, on a

$$\|x + \lambda y\|^2 \geq 0,$$

et donc

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

En considérant le trinôme en λ : $f(\lambda) = \|y\|^2\lambda^2 + 2\langle x, y\rangle\lambda + \|x\|^2 \geq 0$, on en déduit que le discriminant simplifié est ≤ 0, donc

$$\left|\langle x, y\rangle\right|^2 \leq \|x\|^2 \|y\|^2,$$

d'où l'inégalité de Cauchy-Schwarz. Si $\left|\langle x, y\rangle\right| = \|x\| \|y\|$, alors il existe un $\lambda \in \mathbb{R}$ tel que $\|x + \lambda y\| = 0$, donc $x + \lambda y = 0$, donc x et y sont liés. □

Proposition 8.6. *L'application $\|.\|$ définie par $\|x\| = \sqrt{\langle x, x\rangle}$ est une norme, appelée norme associée au produit scalaire $\langle, -, -\rangle$, i.e. elle vérifie les propriétés suivantes :*

1. *$\|\lambda x\| = |\lambda| \|x\|$, pour tous $x \in E$, $\lambda \in \mathbb{R}$;*
2. *$\|x\| = 0 \Leftrightarrow x = 0$;*
3. *inégalité triangulaire : pour tout $x, y \in E$,*

$$\|x + y\| \leq \|x\| + \|y\|,$$

de plus, l'égalité a lieu si et seulement si il existe $\lambda \geq 0$ tel que $x = \lambda y$ ou $y = \lambda x$.

Démonstration. Les propriétés 1) et 2) sont évidentes. Montrons l'inégalité triangulaire. Par l'inégalité de Cauchy-Schwarz, on a

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

donc $\|x + y\| \leq \|x\| + \|y\|$. Supposons que $\|x + y\| = \|x\| + \|y\|$. Si $y = 0$ alors $y = \lambda x$ avec $\lambda = 0$. On suppose que $y \neq 0$. Alors on a égalité dans l'inégalité de Cauchy Schwarz, donc $x = \lambda y$ avec $\lambda \in \mathbb{R}$. Or $\langle x, y\rangle = \|x\| \|y\|$, donc $\lambda \|y\|^2 = |\lambda| \|y\|^2$, et comme $\|y\| \neq 0$, on déduit que $\lambda = |\lambda| \geq 0$. □

8.2 Orthogonalité, bases orthogonales, bases orthonormées

Soit $(E, \langle -, - \rangle)$ un espace vectoriel euclidien.

Définition 8.7. Deux vecteurs x et y sont dits orthogonaux si $\langle x, y \rangle = 0$. On note alors $x \perp y$. Si A est un sous-ensemble de E, on appelle orthogonal de A, noté $A^\perp$, l'ensemble

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

Si $u \in E$, on note $u^\perp = \{u\}^\perp$. Un argument direct donne $0^\perp = E$ et $E^\perp = \{0\}$.

Proposition 8.8. *Soit A un sous-ensemble de E. Alors $A^\perp$ est un sous-espace vectoriel de E.*

La proposition suivante donne une caractérisation de l'orthogonalité des vecteurs en fonction de leurs normes. C'est une version vectorielle du théorème de Pythagore.

Proposition 8.9. *Pour $x, y \in E$, on a*

$$x \perp y \Longleftrightarrow \|x + y\|^2 = \|x\|^2 + \|y\|^2 \Longleftrightarrow \|x - y\|^2 = \|x\|^2 + \|y\|^2.$$

Proposition 8.10. *Soit $u \in E$, $u \neq 0$. On a* $\mathrm{Vect}(u)^\perp = u^\perp$ *et*

$$E = u^\perp \oplus \mathrm{Vect}(u).$$

En particulier $\dim u^\perp = \dim E - 1$.

Démonstration. Un vecteur x est orthogonal à u si et seulement si il est orthogonal à λu pour tout $\lambda \in \mathbb{R}$, donc $u^\perp = \mathrm{Vect}(u)^\perp$.

Si $x \in u^\perp \cap \mathrm{Vect}(u)$, alors $x = \lambda u$ avec $\lambda \in \mathbb{R}$ et $\langle x, u \rangle = 0$, donc $\lambda = 0$, puis $x = 0$. Pour tout $x \in E$, on écrit $x = (x - \lambda u) + \lambda u$ avec $\lambda = \|u\|^{-2} \langle x, u \rangle$. On a $\langle u, x - \lambda u \rangle = 0$, donc $x - \lambda u \in u^\perp$, donc $E = u^\perp \oplus \mathrm{Vect}(u)$. □

Définition 8.11. Soit $\mathscr{E} = (e_1, \dots, e_n)$ une base de E.

1. La base $\mathscr{E}$ est dite orthogonale si $\langle e_i, e_j \rangle = 0$ pour tous $i \neq j$.
2. La base $\mathscr{E}$ est dite orthonormée si $\langle e_i, e_j \rangle = \delta_{i,j}$ pour tous i, j.

Remarque 8.12. Si $(e_1, \dots, e_n)$ est une base orthogonale de E alors $\left(\frac{e_1}{\|e_1\|}, \cdots, \frac{e_n}{\|e_n\|}\right)$ est une base orthonormée de E.

Exemple 8.13. La base canonique de $\mathbb{R}^n$ est orthonormée pour le produit scalaire usuel.

Théorème 8.14. *Dans un espace euclidien E il existe toujours une base orthonormée.*

Démonstration. Il est suffisant de montrer qu'il existe une base orthogonale. On le montre par récurrence sur la dimension n de E. Si $n = 1$ toute base est orthogonale. On suppose $n \geq 2$ et que tout espace vectoriel euclidien de dimension $n-1$ admet une base orthogonale. Soit E un espace euclidien de dimension n, et soit $v \in E$, $v \neq 0$. Le sous-espace $v^\perp$ admet une base orthogonale $(e_1, \dots, e_{n-1})$. La base $(e_1, \dots, e_{n-1}, v)$ est bien orthogonale. □

Remarque 8.15. Soit $\mathscr{E}$ orthonormée, $(x, y) \in E^2$, et $X = [x]_{\mathscr{E}}$, $Y = [y]_{\mathscr{E}}$. Alors

$$\langle x, y \rangle = {}^t X Y.$$

Proposition 8.16. *Pour tout sous-espace vectoriel F d'un espace euclidien E, on a*

1. $E = F \oplus F^\perp$, $\dim F^\perp = \dim E - \dim F$;
2. $(F^\perp)^\perp = F$.

Démonstration. Si $x \in F \cap F^{\perp}$ alors $x \perp y$ pour tout $y \in F$, en particulier, $x \perp x$, donc $x = 0$. Cela implique que $F \cap F^{\perp} = \{0\}$. Soit $(e_1, \dots, e_k)$ une base orthonormée de F, et soit $x \in E$. Soit $y = \sum_{j=1}^{k} \langle x, e_j \rangle e_j$, on calcule $\langle x - y, e_j \rangle = 0$, pour tout j, donc $x - y \in F^{\perp}$, donc $x = y + (x - y) \in F + F^{\perp}$. Il suit que $E = F \oplus F^{\perp}$.

Pour le deuxième point, on remarque d'abord que $F \subset (F^{\perp})^{\perp}$. De plus, d'après le premier point on a

$$\dim (F^{\perp})^{\perp} = \dim E - \dim F^{\perp} = \dim E - (\dim E - \dim F) = \dim F,$$

donc $F = (F^{\perp})^{\perp}$. □

Théorème 8.17 (Procédé d'orthogonalisation de Gram–Schmidt)**.** *Soit* $(x_1, \dots, x_n)$ *une base d'un espace euclidien* E*. Alors il existe une base orthonormée* $(y_1, \dots, y_n)$ *de* E *telle que*

$$\forall k = 1, \dots, n, \ \mathrm{Vect}(x_1, \dots, x_k) = \mathrm{Vect}(y_1, \dots, y_k).$$

Démonstration. Par récurrence sur k. On pose

$$y_1 = \frac{x_1}{\|x_1\|}.$$

Supposons qu'on ait déjà trouvé une famille orthonormée $(y_1, \dots, y_k)$ telle que

$$\mathrm{Vect}(x_1, \dots, x_k) = \mathrm{Vect}(y_1, \dots, y_k).$$

On pose

$$y'_{k+1} = x_{k+1} - \sum_{1 \le i \le k} \langle x_{k+1}, y_i \rangle y_i$$

et on vérifie $\langle y'_{k+1}, y_j \rangle = 0$ pour tout $1 \le j \le k$. On définit ensuite

$$y_{k+1} = \frac{y'_{k+1}}{\|y'_{k+1}\|}$$

si bien que la famille $(y_1, \dots, y_{k+1})$ ainsi obtenue est orthonormée et

$$\mathrm{Vect}(x_1, \dots, x_{k+1}) = \mathrm{Vect}(y_1, \dots, y_{k+1}).$$ □

8.3 Le groupe orthogonal

Définition 8.18. On dit qu'une matrice $A \in \mathrm{GL}_n(\mathbb{R})$ est orthogonale si $A^{-1} = {}^tA$. On note

$$O_n(\mathbb{R}) = \{A \in \mathrm{GL}_n(\mathbb{R}) \mid {}^tAA = A\,{}^tA = I_n\}.$$

Exemple 8.19. Soient $\mathscr{E}$ et $\mathscr{F} = (f_1, \dots, f_n)$ deux bases orthonormées de $\mathbb{R}^n$. Alors $P = \mathscr{P}_{\mathscr{E}}^{\mathscr{F}} \in O_n(\mathbb{R})$. En effet

$$({}^tPP)_{i,j} = {}^t[f_i]_{\mathscr{E}} [f_j]_{\mathscr{E}} = \langle f_i, f_j \rangle = \delta_{i,j}.$$

Exemple 8.20. La matrice

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

est orthogonale car ${}^tAA = I_3$ ou simplement car les vecteurs colonnes forment une base orthonormée de $\mathbb{R}^3$.

Proposition 8.21. *L'ensemble* $O_n(\mathbb{R})$ *est un sous-groupe de* $\mathrm{GL}_n(\mathbb{R})$.

Démonstration. Soient $A, B \in O_n(\mathbb{R})$. On a $(AB)^{-1} = B^{-1}A^{-1} = {}^tB\,{}^tA = {}^t(AB)$ donc $AB \in O_n(\mathbb{R})$. Aussi $(A^{-1})^{-1} = A = {}^t(A^{-1})$ donc $A^{-1} \in O_n(\mathbb{R})$. □

Remarque 8.22. Si $A \in O_n(\mathbb{R})$, on a $\det(A)^2 = 1$ donc $\det(A) \in \{\pm 1\}$. Le sous-groupe

$$SO(n, \mathbb{R}) = O_n(\mathbb{R}) \cap \det^{-1}(1)$$

est appelé groupe spécial orthogonal.

Définition 8.23. Soit $f \in \mathscr{L}(E)$ un endomorphisme. On dit que f est un endomorphisme orthogonal (ou une transformation orthogonale ou une isométrie vectorielle) si

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

On note $O(E)$ l'ensemble des endomorphisme orthogonaux de E.

Proposition 8.24. *Soit $f \in \mathscr{L}(E)$ un endomorphisme. Alors les propriétés suivantes sont équivalentes :*

(i) *f est orthogonal;*

(ii) *$\|f(x)\| = \|x\|$, pour tout $x \in E$;*

(iii) *$\forall \mathscr{E}$ base orthonormée de E, $[f]_{\mathscr{E}} \in O_n(\mathbb{R})$;*

(iv) *$\exists \mathscr{E}$ base orthonormée de E telle que $[f]_{\mathscr{E}} \in O_n(\mathbb{R})$.*

Démonstration. $(i) \Rightarrow (ii)$ est triviale. Montrons $(ii) \Rightarrow (i)$. Soient $x, y \in E$. On a par polarisation puis linéarité de f,

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

comme souhaité.

Montrons $(i) \Rightarrow (iii)$. Soit $\mathscr{E} = (e_1, \ldots, e_n)$ une base orthonormée de E, et $(x, y) \in E^2$. On note X, Y, A les matrices respectives de x, y, f dans $\mathscr{E}$. Alors

$$^tXY = \langle x, y\rangle = \langle f(x), f(y)\rangle = {}^t(AX)AY = {}^tX({}^tAA)Y. \tag{4}$$

C'est vrai pour tous x, y donc par non dégénérescence du produit scalaire canonique sur $\mathbb{R}^n$ on obtient $I_n = {}^tAA$ comme attendu. Il est clair que $(iii) \Rightarrow (iv)$, et $(iv) \Rightarrow (i)$ grâce à (4). □

Proposition 8.25. *L'ensemble $O(E)$ est un sous-groupe de* $\mathrm{GL}(E)$.

Démonstration. Pour toute base $\mathscr{E}$ on a un isomorphisme de groupes $\varphi_{\mathscr{E}} : \mathrm{GL}(E) \to \mathrm{GL}_n(\mathbb{R})$ donné par $f \mapsto [f]_{\mathscr{E}}$. D'après la proposition précédente on a $O(E) = \varphi_{\mathscr{E}}^{-1}(O_n(\mathbb{R}))$, d'où le résultat. □

Remarque 8.26. On note $SO(E) = \{f \in O(E) \mid \det f = 1\}$ le groupe spécial orthogonal de E.

Exercice 8.27 (Décomposition QR)**.** Montrer grâce à l'algorithme de Gram-Schmidt que toute matrice inversible A peut s'écrire comme le produit QR d'une matrice $Q \in O_n(\mathbb{R})$ et d'une matrice triangulaire supérieure réelle R à coefficients diagonaux strictement positifs.

8.4 Endomorphismes symétriques d'un espace euclidien

Soit $(E, \langle -, -\rangle)$ un espace vectoriel euclidien de dimension n.

Notation 8.28. Pour tout $u \in \mathscr{L}(E)$ on définit $\theta_u : E \to E^*$ par $\theta_u(y) = \langle u(-), y\rangle$, c'est-à-dire

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

Remarque 8.29. Par non dégénérescence du produit scalaire on a θ_{id} inversible.

Définition 8.30. Pour tout $u \in \mathscr{L}(E)$ on définit $u^* = \theta_{\mathrm{id}}^{-1} \circ \theta_u \in \mathscr{L}(E)$.

Proposition 8.31. *Pour tous $u \in \mathscr{L}(E)$ et $x, y \in E$ on a*

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

Pour toute base orthonormée $\mathscr{E}$, on a $[u^*]_{\mathscr{E}} = {}^t[u]_{\mathscr{E}}$.

Démonstration. Pour tous $x, y \in E$ on a

$$\langle x, u^*(y)\rangle = \theta_{\mathrm{id}}(u^*(y))(x) = (\theta_{\mathrm{id}} \circ u^*)(y)(x) = \theta_u(y)(x) = \langle u(x), y\rangle.$$

Soit ensuite une base orthonormée $\mathscr{E}$, on note A, A^*, X, Y les matrices respectives dans $\mathscr{E}$ de u, u^*, x, y. L'égalité précédente s'écrit

$$^tXA^*Y = {}^t(AX)Y = {}^tX^tAY,$$

vrai pour tous X, Y donc on conclut à nouveau par non dégénérescence du produit scalaire canonique sur $\mathbb{R}^n$. □

Définition 8.32. Un endomorphisme $u : E \to E$ est dit symétrique (ou auto-adjoint) si $u^* = u$.

Exemple 8.33. Toute projection orthogonale est symétrique. En effet soit p la projection sur un sous-espace F parallèlement à $F^\perp$. Pour tous $x, y \in E$ on a $x - p(x), y - p(y) \in F^\perp$ donc

$$\langle p(x), y\rangle = \langle p(x), p(y)\rangle = \langle x, p(y)\rangle.$$

Remarque 8.34. Un endomorphisme de E est symétrique si, et seulement si, sa matrice dans une base orthonormée est symétrique.

Proposition 8.35. *Toutes les valeurs propres complexes d'une matrice symétrique réelle sont en fait réelles.*

Démonstration. Soit λ une valeur propre (à priori complexe) d'une matrice symétrique réelle M. Il existe un vecteur $X \in \mathbb{C}^n$, $X \neq 0$ tel que $MX = \lambda X$. En prenant le conjugué, on obtient $M\bar{X} = \bar{\lambda}\bar{X}$, puisque M est réelle. En prenant ensuite la transposée, on obtient $^t\bar{X}M = \bar{\lambda}{}^t\bar{X}$, puisque M est symétrique. En multipliant ceci avec X, on obtient

$$\lambda({}^t\bar{X}X) = {}^t\bar{X}(\lambda X) = {}^t\bar{X}(MX) = ({}^t\bar{X}M)X = \bar{\lambda}({}^t\bar{X}X).$$

Comme $^t(\bar{X}X) = |x_1|^2 + \cdots + |x_n|^2 > 0$, il suit que $\lambda = \bar{\lambda}$ est un nombre réel. □

Théorème 8.36. *Tout endomorphisme symétrique est diagonalisable dans une base orthonormée.*

Démonstration. On raisonne par récurrence sur la dimension n de E. Si $n = 1$, il y a rien à démontrer.

On suppose que le résultat est vrai pour tout espace euclidien de dimension $\leq n - 1$.

Soit $(E, \langle -, -\rangle)$ un espace euclidien de dimension n, et soit $u \in \mathscr{L}(E)$ un endomorphisme symétrique. D'après la proposition 8.35, il existe une valeur propre réelle λ_1 et un vecteur propre non nul e_1, choisi de norme 1. L'espace $F = e_1^\perp$ est stable par u. En effet, si $x \in F$ alors

$$\langle u(x), e_1\rangle = \langle x, u(e_1)\rangle = \langle x, \lambda e_1\rangle = 0,$$

donc $u(x) \in F$. L'espace F étant de dimension $n-1$, l'hypothèse de récurrence s'applique à F : il existe une base orthonormée $(e_2, \ldots, e_n)$ de F dans laquelle la matrice de u_F est diagonale. Par conséquent, u est diagonalisable dans la base orthonormée $(e_1, \ldots, e_n)$. □

Ce théorème se reformule en termes matriciels si on considère $\mathbb{R}^n$ muni de sa structure euclidienne canonique.

Théorème 8.37. *Pour toute matrice symétrique réelle M, il existe une matrice orthogonale P telle que ^tPMP soit diagonale.*

Remarque 8.38. Soit Q une forme quadratique sur E euclidien, et Φ sa forme polaire. Soit $\mathscr{E}$ une base orthonormée. Alors le spectre de $[\Phi]_\mathscr{E}$ ne dépend pas du choix de cette base orthonormée. En effet, si $\mathscr{F}$ en est une autre et P désigne la matrice de passage de $\mathscr{E}$ à $\mathscr{F}$, on sait que $P \in O(E)$ donc

$$[\Phi]_\mathscr{F} = {}^tP[\Phi]_\mathscr{E}P = P^{-1}[\Phi]_\mathscr{E}P.$$

En fait il existe une manière intrinsèque d'associer un endomorphisme symétrique u à Q. Soit Φ sa forme polaire et $\theta_\Phi : E \to E^*$, $x \mapsto \Phi(x, -)$ comme précédemment. On pose alors $u = \theta^{-1} \circ \theta_\Phi$ où $\theta : E \to E^*$, $x \mapsto \langle x, -\rangle$. Alors pour tous $x, y \in E$ on a

$$\langle u(x), y\rangle = \theta(u(x))(y) = \theta_\Phi(x)(y) = \Phi(x, y)$$

donc u est symétrique, $Q(x) = \langle x, u(x)\rangle$, et $[u]_\mathscr{E} = [\Phi]_\mathscr{E}$ pour toute base orthonormée.

Remarque 8.39. Soit Q une forme quadratique sur $E = \mathbb{R}^n$. Soit M la matrice symétrique de sa forme polaire dans la base canonique, puis P orthogonale telle que ${}^tPMP = D = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$ avec $\lambda, \dots, \lambda_s > 0$ et $\lambda_{s+1}, \dots, \lambda_{s+t} < 0$ (éventuellement s ou t peut être nul). Soit $X = (x_1, \dots, x_n) \in E$, on a $Q(X) = {}^tXMX = {}^tXPD{}^tPX$. On définit des formes linéaires φ_i indépendantes par ${}^tPX = (\varphi_1(X), \dots, \varphi_n(X))$ si bien que

$$Q = \sum_{1 \le i \le n} \lambda_i \varphi_i^2$$

et la signature de Q est donnée par (s, t). Par ailleurs les matrices des $f\varphi_i$ dans la base canonique sont les transposées des colonnes de P, qui forment une famille orthonormée.

Pour Q quadratique sur E euclidien quelconque, soit u l'endomorphisme symétrique de E défini dans 8.38. Soit $\mathscr{F} = (f_1, \dots, f_n)$ une base orthonormée de diagonalisation, où $u(f_i) = \lambda_i f_i$. Alors pour tout $x = \sum_{1 \le i \le n} x_i f_i$ on a $Q(x) = \langle x, u(x) \rangle = \sum_{1 \le i \le n} \lambda_i x_i^2$ car $\mathscr{F}$ orthonormée, et donc

$$Q = \sum_{1 \le i \le n} \lambda_i \langle f_i, - \rangle^2 = \sum_{1 \le i \le n} \lambda_i (f_i^*)^2.$$

En effet être orthonormée signifie précisément que $\langle f_i, - \rangle = f_i^*$.

Quand Q est définie positive dans un espace euclidien on dispose de deux normes que l'on sait équivalentes. On peut expliciter les constantes de comparaison :

Corollaire 8.40. *Soit Q une forme quadratique définie positive sur un espace euclidien E. Soit $\lambda_1 \le \cdots \le \lambda_n$ son spectre ordonné. Alors pour tout $x \in E$ on a*

$$\lambda_1 \|x\|^2 \le Q(x) \le \lambda_n \|x\|^2$$

et ces inégalités sont optimales.

Démonstration. Méthode 1. Soit $\mathscr{E}$ une base orthonormée de E. et M la matrice de la forme polaire de Q dans cette base. Elle est symétrique donc il existe $P \in O_n(\mathbb{R})$, où $n = \dim E$, telle que ${}^tPMP = D = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$. Soit alors $x \in E$ de matrice colonne X dans $\mathscr{E}$ et $Y = {}^tPX = (y_i)_{1 \le i \le n}$. On a

$$Q(x) = {}^tXMX = {}^tXPD{}^tPX = {}^tYDY = \sum_{1 \le i \le n} \lambda_i y_i^2.$$

On a donc

$$\lambda_1 \sum_{1 \le i \le n} y_i^2 \le Q(x) \le \lambda_n \sum_{1 \le i \le n} y_i^2$$

mais comme P est orthogonale, $\sum_{1 \le i \le n} y_i^2 = \sum_{1 \le i \le n} x_i^2$, et comme $\mathscr{E}$ est orthonormée, cette quantité est bien $\|x\|^2$. On a égalité à gauche (resp. droite) si $y_i = \delta_{i,1}$ (resp. $y_i = \delta_{i,n}$).
Méthode 2. Soit u l'endomorphisme symétrique de E défini dans 8.38. Soit $\mathscr{F} = (f_1, \dots, f_n)$ une base orthonormée de diagonalisation, où $u(f_i) = \lambda_i f_i$. Alors pour tout $x = \sum_{1 \le i \le n} x_i f_i$ on a $Q(x) = \langle x, u(x) \rangle = \sum_{1 \le i \le n} \lambda_i x_i^2$ car $\mathscr{F}$ orthonormée, et on conclut car pour la même raison $\|x\|^2 = \sum_{1 \le i \le n} x_i^2$. □

8.5 Polynômes orthogonaux

8.5.1 Généralités

Soit I un intervalle de $\mathbb{R}$ et $w : I \to \mathbb{R}_+^*$ une fonction continue telle que pour tout $P \in \mathbb{R}[X]$ l'intégrale $\int_I Pw$ converge. Dans la suite on note $E = \mathbb{R}[X]$ et $E_n = \mathbb{R}[X]_n$ pour tout entier $n \ge 0$.

Proposition 8.41. *La forme bilinéaire symétrique $\langle -, - \rangle : E \times E \to \mathbb{R}$ définie par*

$$(P, Q) \mapsto \int_I PQw$$

est un produit scalaire.

Démonstration. La fonction $P^2 w$ est continue et positive sur I donc si $\langle P, P \rangle = 0$, on a $P(t)^2 w(t) = 0$ pour tous $t \in I$. Mais $w(t) > 0$ donc P a une infinité de racines, et donc P est le polynôme nul. □

Définition 8.42. On parle d'une famille $(P_n)_{n\in\mathbb{N}}$ de polynômes orthogonaux si par ailleurs elle est échelonnée : $\deg(P_n) = n$.

Remarque 8.43. Toute famille de polynôme orthogonale est en particulier une base.

Proposition 8.44. *Il existe une famille de polynômes orthogonaux. Si $(P_n)_{n\in\mathbb{N}}$ et $(Q_n)_{n\in\mathbb{N}}$ en sont deux, alors P_n et Q_n sont colinéaires pour tout $n \in \mathbb{N}$.*

Démonstration. On obtient l'existence en appliquant le procédé de Gram-Schmidt à la base canonique. Ensuite, on a P_n et Q_n qui sont dans la droite $E_n \cap E_{n-1}^{\perp}$. □

Corollaire 8.45. *Soit u un endomorphisme symétrique de E qui stabilise tous les E_n, avec un spectre simple sur chaque E_n. Alors une base est propre pour u si et seulement si c'est une famille de polynômes orthogonaux.*

Démonstration. Soit n donné. La matrice de u dans la base canonique de E_n est triangulaire supérieure à diagonale $(\lambda_1, \dots, \lambda_n)$ où les λ_k sont tous distincts. Ceci implique que tout vecteur propre associé à λ_k est dans $\mathrm{Vect}(1, X, \dots, X^k) = E_k$, avec un coefficient non nul selon X^k. Ainsi la base propre pour u est non seulement orthogonale, mais aussi échelonnée, et on conclut grâce à la proposition précédente. □

La proposition suivante est claire (et vraie en remplaçant X par n'importe quel polynôme donné).

Proposition 8.46. *L'endomorphisme de E donné par la multiplication par X est symétrique.*

8.5.2 Racines

Définition 8.47. Pour tout $n \in \mathbb{N}$ on note $p_n \in \mathscr{L}(E_n)$ la projection orthogonale de E_n sur E_{n-1} et $T_n \in \mathscr{L}(E_{n-1})$ défini par $T_n(P) = p_n(XP)$.

Remarque 8.48. (i) En particulier pour toute famille $(P_n)_{n\in\mathbb{N}}$ de polynômes orthogonaux on a $p_n(P_n) = 0$ pour tout $n \geq 0$.

(ii) D'après 8.33, p_n est symétrique pour le produit scalaire induit sur E_n.

Proposition 8.49. *Soit $(P_n)_{n\in\mathbb{N}}$ une famille de polynômes orthogonaux. L'endomorphisme T_n est symétrique, son spectre est simple et ses valeurs propres sont les racines de P_n. L'espace propre associé à λ est engendré par le quotient de P_n par $X - \lambda$.*

Remarque 8.50. En particulier P_n est simplement scindé et on retrouve la "pseudo-unicité" de 8.44.

Démonstration. On sait que p_n est symétrique, donc pour tous $P, Q \in E_{n-1}$ on a

$$\langle T_n(P), Q\rangle = \langle p_n(XP), Q\rangle = \langle XP, \underbrace{p_n(Q)}_{=Q}\rangle \underset{8.46}{=} \langle P, XQ\rangle = \langle p_n(P), XQ\rangle = \langle P, T_n(Q)\rangle.$$

Soit λ une valeur propre de T_n et Q propre. Déjà $\deg(Q) = n-1$, sinon $\deg(Q) < n-1$ et $\lambda Q = p_n(XQ) = XQ$ ce qui n'est possible que si $Q = 0$. Mais alors il existe $S \in E_{n-1}$ et $\alpha \in \mathbb{R}^*$ tels que $XQ = \alpha P_n + S$, et vu que $p_n(P_n) = 0$, on obtient $\lambda Q = T_n(Q) = p_n(S) = S$, donc $(X - \lambda)Q = \alpha P_n$. Les espaces propres sont en particulier des droites et on obtient le résultat. □

Remarque 8.51. Attention dans la preuve ci-dessus la multiplication par X n'est *pas* un endomorphisme de E_n. Par ailleurs la composée de deux endomorphismes symétriques n'est en général pas symétrique.

8.5.3 L'exemple des polynômes d'Hermite

On considère $I=\mathbb{R}$ et $w(t)=e^{-t^2/2}$ qui vérifie bien que pour tout $P\in E$, $\int_I Pw$ converge, donc on dispose d'un produit scalaire associé sur E. Soit $u\in\mathscr{L}(E)$ défini par $u(P)=P''-XP'$. Alors pour tout $t\in\mathbb{R}$ on a

$$u(P)(t)=e^{t^2/2}\frac{d}{dt}\left(e^{-t^2/2}P'(t)\right).$$

Proposition 8.52. *L'endomorphisme u est symétrique, et stabilise tous les E_n où son spectre est simple, donné par $\{-k\mid k\in[\![0,n]\!]\}$.*

Démonstration. Pour tous $P,Q\in E$ on a

$$\langle u(P),Q\rangle=\int_{\mathbb{R}} w^{-1}\left(wP'\right)'Qw=\int_{\mathbb{R}}\left(wP'\right)'Q=\underbrace{\left[wP'Q\right]_{-\infty}^{+\infty}}_{=0}-\int_{\mathbb{R}} wP'Q'$$

et comme l'expression de droite est symétrique en P et Q on a bien u symétrique. La stabilité de E_n est claire et implique que dans sa base canonique la matrice de u est triangulaire supérieure (vu que u stabilies les E_k pour $k<n$). On lit donc son spectre sur la diagonale, et on obtient le résultat souhaité puisque le coefficient de X^k dans $u(X^k)$ est $-k$. □

Proposition 8.53. *Soit $h_n=w^{(n)}$. Alors $H_n=w^{-1}h_n$ est une fonction polynômiale de degré n, de coefficient dominant $(-1)^n$, et on a les relations suivantes pour tout $n\in\mathbb{N}$:*

(i) $H_n'=XH_n+H_{n+1}$;

(ii) $H_{n+1}=-XH_n-nH_{n-1}$;

(iii) $u(H_n)=-nH_n$.

Démonstration. Pour tout $t\in\mathbb{R}$ on a $(w^{-1})'(t)=tw^{-1}(t)$ donc $H_n'(t)=tH_n(t)+H_{n+1}(t)$, ce qui prouve par récurrence que H_n est polynômiale de degré n, de coefficient dominant $(-1)^n$, et (i). Par Leibniz on a

$$h_{n+1}(t)=\frac{d^n}{dt^n}\left(-te^{-t^2/2}\right)=-th_n(t)-nh_{n-1}$$

et donc (ii). Mais alors

$$\begin{aligned}u(H_n)(t)&=e^{t^2/2}\frac{d}{dt}\left(H_n'(t)e^{-t^2/2}\right)\\&\underset{\text{(i)}}{=}e^{t^2/2}\frac{d}{dt}\left(\left(tH_n(t)+H_{n+1}(t)\right)e^{-t^2/2}\right)\\&\underset{\text{(ii)}}{=}-ne^{t^2/2}\frac{d}{dt}\left(H_{n-1}(t)e^{-t^2/2}\right)\\&=-ne^{t^2/2}\frac{d}{dt}h_{n-1}(t)=-ne^{t^2/2}h_n(t)=-nH_n(t).\end{aligned}$$

□

Corollaire 8.54. *Les polynômes dits d'Hermite $(H_n)_{n\in\mathbb{N}}$ forment une famille de polynômes orthogonaux. Chaque H_n est solution de l'équation différentielle $y''-xy'+ny=0$.*

Remarque 8.55. L'orthogonalité peut se montrer par le calcul. Par intégration par parties on obtient pour tout $n\geq 1$ et tout $P\in E$

$$\langle H_n,P\rangle=-\langle H_{n-1},P'\rangle.$$

Ainsi par récurrence, si $k<n$,

$$\langle H_n,H_k\rangle=(-1)^{k+1}\langle H_{n-k-1},\underbrace{H_k^{(k+1)}}_{=0}\rangle=0.$$