CHAPITRE

1

Espaces vectoriels, compléments

Sommaire

$\mathbb{K}$ désigne l'un des corps $\mathbb{R}$ ou $\mathbb{C}$; E et F sont des $\mathbb{K}$-espaces vectoriels ; on note I_E (resp. I_F) l'application identique de E (resp. de F).

1 Somme directe

Dans ce paragraphe, q est un entier au moins égal à 2, et $E_1, E_2, \dots, E_q$ désignent des $\mathbb{K}$-sous-espaces vectoriels de E.

1.1 Somme

Définition 1.1 (Somme de sous-espaces vectoriels).
On note $\displaystyle\sum_{i=1}^{q} E_i$ l'ensemble $\displaystyle\left\{ \sum_{i=1}^{q} \mathbf{x}_i \ / \ (\mathbf{x}_1, \dots, \mathbf{x}_q) \in E_1 \times \cdots \times E_q \right\}$;
$\sum_{i=1}^q E_i$ est appelé *somme* des sous-espaces vectoriels E_i.

Remarques.
$\sum_{i=1}^q E_i$ est l'image de l'application linéaire Ψ

$$\Psi : (\mathbf{x}_1, \dots, \mathbf{x}_q) \in E_1 \times \cdots \times E_q \mapsto \mathbf{x}_1 + \cdots + \mathbf{x}_q$$

Ainsi, $\sum_{i=1}^q E_i$ est un $\mathbb{K}$-sous-espace vectoriel de E et $\Psi : E_1 \times \cdots \times E_q \to \sum_{i=1}^q E_i = \operatorname{Im}\Psi$ est une application linéaire *surjective*.
Si $\mathcal{G}_i$ est une famille génératrice de E_i, $\bigcup_{i=1}^q \mathcal{G}_i$ est une famille génératrice de $\sum_{i=1}^q E_i$.

1.2 Somme directe

Définition 1.2 (Somme directe de sous-espaces vectoriels).
La somme $\sum_{i=1}^q E_i$ est une *somme directe* si, et seulement si, l'application linéaire Ψ

$$\Psi : (\mathbf{x}_1, \dots, \mathbf{x}_q) \mapsto \mathbf{x}_1 + \cdots + \mathbf{x}_q$$

est un *isomorphisme* d'espaces vectoriels entre $E_1 \times \cdots \times E_q$ et $\sum_{i=1}^q E_i$; dans ce cas, la somme est notée $\displaystyle\bigoplus_{i=1}^{q} E_i$.

Théorème 1.1 (Caractérisation d'une somme directe).
Les propriétés suivantes sont équivalentes :

(i) *la somme $\sum_{i=1}^q E_i$ est directe ;*
(ii) *la seule décomposition de $\mathbf{0}_E$ dans $\sum_{i=1}^q E_i$ est $\mathbf{0}_E = \sum_{i=1}^q \mathbf{0}_{E_i}$;*
(iii) $\forall \mathbf{x} \in \sum_{i=1}^q E_i,\ \exists!(\mathbf{x}_1, \dots, \mathbf{x}_q) \in E_1 \times \cdots \times E_q,\ \mathbf{x} = \sum_{i=1}^q \mathbf{x}_i$
(iv) $\forall k \in [\![1, q-1]\!],\ \left(\sum_{i=1}^k E_i\right) \cap E_{k+1} = \{\mathbf{0}_E\}$

Démonstration. Puisque Ψ est une application linéaire surjective, Ψ est un isomorphisme si, et seulement si, Ψ est injective.
$(ii) \iff \ker\Psi = \{\mathbf{0}_E\} \iff \Psi$ est injective $\iff (i)$.

$(iii) \iff \Psi$ est bijective $\iff (i)$.

$(i) \implies (iv)$ Soient $k \in [\![1, q-1]\!]$ et $\mathbf{x} \in \left(\sum_{i=1}^k E_i\right) \cap E_{k+1}$; il existe donc $(\mathbf{x}_1, \dots, \mathbf{x}_k) \in E_1 \times \cdots \times E_k$ tel que $\mathbf{x} = \mathbf{x}_1 + \cdots + \mathbf{x}_k$, *i.e.*

$$\Psi(\mathbf{x}_1, \dots, \mathbf{x}_k, \mathbf{0}_{E_{k+1}}, \dots, \mathbf{0}_{E_q}) = \Psi(\mathbf{0}_{E_1}, \dots, \mathbf{0}_{E_k}, \mathbf{x}, \mathbf{0}_{E_{k+2}}, \dots, \mathbf{0}_{E_q})$$

et, puisque Ψ est une application injective

$$(\mathbf{x}_1, \dots, \mathbf{x}_k, \mathbf{0}_{E_{k+1}}, \dots, \mathbf{0}_{E_q}) = (\mathbf{0}_{E_1}, \dots, \mathbf{0}_{E_k}, \mathbf{x}, \mathbf{0}_{E_{k+2}}, \dots, \mathbf{0}_{E_q})$$

ce qui montre que $\mathbf{x} = \mathbf{0}_E$.

$(iv) \Longrightarrow (i)$ Supposons l'application Ψ non injective ; il existe alors $(\mathbf{x}_1, \dots, \mathbf{x}_q) \in \ker \Psi \setminus \{\mathbf{0}_E\}$, *i.e.* $\mathbf{x}_1 + \dots + \mathbf{x}_q = \mathbf{0}_E$ avec $(\mathbf{x}_i)_{1 \leqslant i \leqslant q}$ non tous nuls. On pose $i_0 = \max\{i \ / \ \mathbf{x}_i \neq \mathbf{0}\}$; i_0 est un entier au moins égal à 2 et

$$\mathbf{x}_{i_0} = -(\mathbf{x}_1 + \dots + \mathbf{x}_{i_0-1}) \in E_{i_0} \cap \Big(\sum_{i=1}^{i_0-1} E_i\Big)$$

ce qui est en contradiction avec l'hypothèse ; ainsi $\ker \Psi$ est réduit à $\{\mathbf{0}_E\}$ et Ψ est une application injective. cqfd

1.3 Supplémentaire

Définition 1.3 (Sous-espace supplémentaire).
Deux sous-espaces vectoriels V et W de E sont dits *supplémentaires* si, et seulement si, E est la somme directe de V et W, soit

$$E = V \oplus W$$

Théorème 1.2 (Caractérisation des supplémentaires).
Les propriétés suivantes sont équivalentes

(i) *V et W sont supplémentaires (dans E) ;*
(ii) *tout $\mathbf{x}$ de E s'écrit de manière unique $\mathbf{x} = \mathbf{v} + \mathbf{w}$ avec $(\mathbf{v}, \mathbf{w}) \in V \times W$;*
(iii) *$V + W = E$ et $V \cap W = \{\mathbf{0}_E\}$.*

Démonstration. C'est la démonstration précédente pour $q = 2$ et $V + W = E$. cqfd

Exemple 1.1. Si P est un polynôme de $\mathsf{K}[X]$ de degré $n+1$, l'ensemble $(P) = \{AP \ / \ A \in \mathsf{K}[X]\}$ des multiples de P et l'ensemble $\mathsf{K}_n[X]$ des polynômes de degré au plus n, sont supplémentaires dans $\mathsf{K}[X]$.

$$\boxed{\mathsf{K}[X] = (P) \oplus \mathsf{K}_n[X] \qquad \text{avec } \deg P = n+1}$$

Démonstration. (P) et $\mathsf{K}_n[X]$ sont des sous-espaces vectoriels de $\mathsf{K}[X]$ et la division euclidienne des polynômes donne, pour tout $B \in \mathsf{K}[X]$, l'existence d'un couple unique $(A, R) \in \mathsf{K}[X] \times \mathsf{K}_n[X]$ tel que $B = AQ + R$. cqfd

1.4 Cas de la dimension finie

Théorème 1.3 (Somme directe et dimension).
Si E est un espace vectoriel de dimension finie, la somme $\sum_{i=1}^q E_i$ est directe si, et seulement si,

$$\dim\Big(\sum_{i=1}^q E_i\Big) = \sum_{i=1}^q \dim E_i$$

Démonstration. $\Psi : (\mathbf{x}_1, \dots, \mathbf{x}_q) \in E_1 \times \dots \times E_q \mapsto \mathbf{x}_1 + \dots + \mathbf{x}_q \in \sum_{i=1}^q E_i$ est une application linéaire et surjective ; Ψ est bijective si, et seulement si les espaces vectoriels $E_1 \times \dots \times E_q$ et $\sum_{i=1}^q E_i$ ont même dimension ; or $\dim(E_1 \times \dots \times E_q) = \sum_{i=1}^q \dim E_i$, ce qui démontre l'équivalence annoncée. cqfd

Corollaire (Cas des supplémentaires).
Soient E un espace vectoriel de dimension finie n et V un sous-espace vectoriel de dimension p $(p \leqslant n)$;

(i) *tout supplémentaire de V est de dimension $n-p$;*
(ii) *soit W un sous-espace vectoriel de dimension $n-p$; W est un supplémentaire de V si, et seulement si, $V \cap W = \{\mathbf{0}_E\}$, ou bien si, et seulement si, $V + W = E$.*

Démonstration.

(i) $E = V \oplus W \implies \dim E = \dim(V \oplus W) = \dim V + \dim W$;

(ii) seule les réciproques méritent une démonstration. Si $V \cap W = \{\mathbf{0}_E\}$, $\dim(V+W) = \dim V + \dim W = n = \dim E$, et donc $V + W = E$ et $V \oplus W = E$. Si $V + W = E$, $\dim(V+W) = \dim E = n = \dim V + \dim W$, ce qui montre que $V \oplus W = E$.

cqfd

Remarque. La connaissance d'un renseignement sur la dimension permet de diviser le travail par deux !

2 Décomposition de E en somme directe

2.1 Applications linéaires

Théorème 2.1 (Contruction d'applications linéaires).
Si $E = \bigoplus_{i=1}^{q} E_i$ et si, pour tout $i \in [\![1,q]\!]$, $u_i \in \mathcal{L}(E_i, F)$, il existe une unique application linéaire $u \in \mathcal{L}(E,F)$ telle que

$$\forall i \in [\![1,q]\!], \quad u|_{E_i} = u_i$$

Démonstration.

Analyse. Si $\mathbf{x} = \mathbf{x}_1 + \cdots + \mathbf{x}_q$ est la décomposition d'un vecteur $\mathbf{x} \in E$ suivant les facteurs E_i,

$$\begin{aligned} u(\mathbf{x}) &= u(\mathbf{x}_1) + \cdots + u(\mathbf{x}_q) && \text{par linéarité} && (2.1)\\ &= u_1(\mathbf{x}_1) + \cdots + u_q(\mathbf{x}_q) && \text{car } u|_{E_i} = u_i && (2.2)\end{aligned}$$

ce qui montre l'unicité de u.

Synthèse. Pour $\mathbf{x} \in E$, on pose $u(\mathbf{x}) = u_1(\mathbf{x}_1) + \cdots + u_q(\mathbf{x}_q)$, ce qui définit u puisque la décomposition $\mathbf{x} = \mathbf{x}_1 + \cdots + \mathbf{x}_q$ est unique. Par construction $u|_{E_i} = u_i$ et u est une application linéaire car pour tout $(\lambda, \mu) \in \mathsf{K}^2$

$$\mathbf{x} = \mathbf{x}_1 + \cdots + \mathbf{x}_q,\ \mathbf{y} = \mathbf{y}_1 + \cdots + \mathbf{y}_q \implies \lambda\mathbf{x} + \mu\mathbf{y} = (\lambda\mathbf{x}_1 + \mu\mathbf{y}_1) + \cdots + (\lambda\mathbf{x}_q + \mu\mathbf{y}_q)$$

et

$$\begin{aligned} u(\lambda\mathbf{x} + \mu\mathbf{y}) &= u_1(\lambda\mathbf{x}_1 + \mu\mathbf{y}_1) + \cdots + u_q(\lambda\mathbf{x}_q + \mu\mathbf{y}_q)\\ &= \big(\lambda u_1(\mathbf{x}_1) + \mu u_1(\mathbf{y}_1)\big) + \cdots + \big(\lambda u_q(\mathbf{x}_q) + \mu u_q(\mathbf{y}_q)\big)\\ &= \lambda\big(u_1(\mathbf{x}_1) + \cdots + u_q(\mathbf{x}_q)\big) + \mu\big(u_1(\mathbf{y}_1) + \cdots + u_q(\mathbf{y}_q)\big)\\ &= \lambda u(\mathbf{x}) + \mu u(\mathbf{y})\end{aligned}$$

cqfd

A) Projecteur

Soient $E = V \oplus W$ et $\mathbf{x} = \mathbf{v} + \mathbf{w}$ la décomposition de $\mathbf{x}$ suivant V et W ; on pose

$$p_V : \mathbf{x} \mapsto \mathbf{v} \quad \text{et} \quad p_W : \mathbf{x} \mapsto \mathbf{w}$$

On a alors les propriétés suivantes :

$$p_V \circ p_V = p_V, \quad p_V + p_W = I_E, \quad p_V \circ p_W = p_W \circ p_V = 0_{\mathcal{L}(E)}$$
$$\ker p_V = W, \quad \operatorname{Im} p_V = \ker(I_E - p_V) = V, \quad p_V|_V = I_V, \quad p_V|_W = 0_{\mathcal{L}(V)}$$

Réciproquement, si p est un endomorphisme de E qui vérifie $p \circ p = p$, on a

(i) $\operatorname{Im} p = \ker(I_E - p)$;
(ii) $\operatorname{Im} p$ et $\ker p$ sont supplémentaires (dans E) ;
(iii) p est le projecteur d'image $\operatorname{Im} p = \ker(I_E - p)$ parallèlement à $\ker p$, *i.e.*

$$p|_{\operatorname{Im} p} = I_{\operatorname{Im} p} \quad \text{et} \quad p|_{\ker p} = 0$$

B) Symétrie vectorielle

Si $E = V \oplus W$, la *symétrie vectorielle par rapport à V et parallèlement à W* est l'automorphisme s de E tel que

$$s(\mathbf{x}) = \mathbf{v} - \mathbf{w} \quad \textit{i.e.} \quad s = p_V - p_W = 2p_V - I_E$$

où $\mathbf{x} = \mathbf{v} + \mathbf{w}$ est la décomposition de $\mathbf{x}$ suivant $V \oplus W$.
On a les propriétés suivantes :

$$s|_V = I_V, \quad s|_W = -I_W, \quad s = p_V - p_W, \quad I_E + s = 2p_V$$
$$s \circ s = I_E, \quad s^{-1} = s$$

Réciproquement, tout endomorphisme s de E qui vérifie $s \circ s = I_E$ est la symétrie vectorielle par rapport à $\ker(s - I_E)$ et parallèlement à $\ker(I_E + s)$.

C) Affinité vectorielle

Si $E = V \oplus W$, l'*affinité vectorielle de direction V, de rapport $\lambda \in \mathsf{K}^*$, parallèlement à W* est l'automorphisme a de E tel que

$$a(\mathbf{x}) = \lambda\mathbf{v} + \mathbf{w}$$

où $\mathbf{x} = \mathbf{v} + \mathbf{w}$ est la décomposition de $\mathbf{x}$ suivant $V \oplus W$; a est caractérisé par $a|_V = \lambda I_V$ et $a|_W = I_W$ et on peut encore écrire $a = \lambda p_V + p_W$.

D) Projecteurs associés à une décomposition en somme directe

Si $E = \bigoplus_{i=1}^q E_i$ et si $\mathbf{x} = \mathbf{x}_1 + \cdots + \mathbf{x}_q$ est la décomposition de $\mathbf{x}$ suivant les facteurs E_i, on pose :

$$p_i : \mathbf{x} \mapsto \mathbf{x}_i$$

p_i est le projecteur sur E_i parallèlement à $\bigoplus_{j \neq i} E_j$ et on a les propriétés suivantes

$$p_i \circ p_i = p_i, \quad i \neq j \implies p_i \circ p_j = p_j \circ p_i = 0, \quad \sum_{i=1}^q p_i = I_E,$$
$$\operatorname{Im} p_i = E_i, \quad \ker p_i = \bigoplus_{\substack{j=1 \\ j \neq i}}^q E_j$$

L'application linéaire de (2.2) s'écrit $u = u_1 \circ p_1 + \cdots + u_q \circ p_q$.

2.2 Cas de la dimension finie

Définition 2.1 (Base adaptée à une somme directe).
Si $E = \bigoplus_{i=1}^q E_i$ et si $\mathcal{B}_i$ est une base de E_i, la famille $\mathcal{B} = \bigcup_{i=1}^q \mathcal{B}_i$ est une base de E ; elle est dite *adaptée à la décomposition en somme directe* de E.

Démonstration. $\#\mathcal{B} = \sum_i \dim E_i = \dim E$; il suffit de montrer que $\mathcal{B}$ est une famille génératrice, ce qui est le cas puisque $\bigcup_i \mathcal{B}_i$ engendre $\sum_i E_i$. cqfd

Définition 2.2 (Base adaptée à un sous-espace vectoriel).
Si F est un sous-espace vectoriel de E, toute base de E dont les premiers éléments constituent une base de F, est dite *adaptée à F*.

Remarques.
Les bases adaptées permettent de donner des matrices « simples » d'endomrphismes. Quelles sont les matrices des projecteurs, des symétries et des affinités dans des bases adaptées ?
Si $\mathcal{B}$ est une base de E, $\mathcal{B}_1, \dots, \mathcal{B}_q$ une partition de $\mathcal{B}$, E_i le sous-espace vectoriel engendré par $\mathcal{B}_i$, E est la somme directe des E_i et $\mathcal{B}$ est une base adaptée à cette décomposition.

3 Applications linéaires

3.1 Isomorphisme de tout supplémentaire du noyau avec l'image

Théorème 3.1 (Théorème du noyau et de l'image).
Si $u \in \mathcal{L}(E, F)$ et V est un supplémentaire de $\ker u$, *l'application*

$$\bar{u} : \mathbf{x} \in V \mapsto u(\mathbf{x}) \in \operatorname{Im} u$$

définit un isomorphisme de V sur $\operatorname{Im} u$.

Démonstration. $\bar{u}$ est une application linéaire et

$$\ker \bar{u} = \{\mathbf{x} \in V \ / \ u(\mathbf{x}) = \mathbf{0}_F\} = \ker u \cap V = \mathbf{0}_E$$

ce qui montre l'injectivité de $\bar{u}$.
Pour tout $\mathbf{y} \in \operatorname{Im} u$, il existe $\mathbf{x} \in E$ tel que $u(\mathbf{x}) = \mathbf{y}$; puisque $E = V \oplus \ker u$, $\mathbf{x}$ se décompose en $\mathbf{v} + \mathbf{w}$ avec $(\mathbf{v}, \mathbf{w}) \in V \times \ker u$, et

$$\mathbf{y} = u(\mathbf{x}) = u(\mathbf{v}) + u(\mathbf{w}) = u(\mathbf{v}) = \bar{u}(\mathbf{v})$$

et montre la surjectivité de $\bar{u}$. cqfd

Corollaire (Théorème du rang). *Si E est de* dimension finie

$$\boxed{\dim E = \operatorname{rg} u + \dim(\ker u)}$$

Démonstration. $\bar{u}$ est un isomorphisme donc conserve la dimension et $\dim V = \dim(\operatorname{Im} u) = \operatorname{rg} u$; d'autre part, V et $\ker u$ sont supplémentaires dans E, donc $\dim E = \dim V + \dim(\ker u)$. cqfd

Corollaire (Caractérisation des isomorphismes en dimension finie).
Si E et F sont deux espaces vectoriels de même dimension finie *n et u une application linéaire de E vers F, alors*

$$\boxed{\begin{array}{c} u \textit{ est un isomorphisme} \iff u \textit{ est injective} \iff \dim(\ker u) = 0 \\ \iff u \textit{ est surjective} \iff \operatorname{rg} u = n \end{array}}$$

Démonstration. C'est une conséquence immédiate de $n = \dim E = \dim F = \operatorname{rg} u + \dim(\ker u)$ et des équivalences

$$u \text{ est injective} \iff \ker u = \{\mathbf{0}_E\} \iff \dim(\ker u) = 0$$

ainsi que des équivalences

$$u \text{ est surjective} \iff \operatorname{Im} u = F \iff \dim(\operatorname{Im} u) = \dim F$$

cqfd

Corollaire (Caractérisation des automorphismes en dimension finie).
Si E est un espace vectoriel de dimension finie *n et u un endomorphisme de E*

$$\boxed{u \in \mathcal{GL}(E) \iff u \text{ est injective} \iff u \text{ est surjective} \iff \dim(\ker u) = 0 \iff \operatorname{rg} u = n}$$

3.2 Applications

A) Projection

Si V_1 et V_2 sont deux supplémentaires dans E du même sous-espace vectoriel W, la projection p_{V_1} de E sur V_1 parallèlement à W induit un isomorphisme de V_2 sur V_1.
C'est une application du théorème du noyau et de l'image : p_{V_1} est une application linéaire d'image V_1 et de noyau W ; elle induit un isomorphisme de V_2, supplémentaire de W sur son image. En classe de troisième, ce théorème porte le nom de « Théorème de Thalès ».

B) Formule de Grassmann

Si V et W sont deux sous-espaces vectoriels de dimension finie, on a

$$\boxed{\dim(V \cap W) + \dim(V + W) = \dim V + \dim W}$$

L'application $u : (\mathbf{v}, \mathbf{w}) \in V \times W \mapsto \mathbf{v} + \mathbf{w} \in V + W$ est linéaire et surjective. Son noyau vérifie

$$\ker u = \{(\mathbf{v}, \mathbf{w}) \in V \times W \;/\; \mathbf{v} + \mathbf{w} = \mathbf{0}\} = \{(\mathbf{v}, -\mathbf{v}) \;/\; \mathbf{v} \in V \cap W\}$$

$\ker u$ est donc isomorphe à $V \cap W$, ce qui donne

$$\dim(\ker u) = \dim(V \times W) - \operatorname{rg} u = \dim V + \dim W - \dim(V + W)$$

C) Polynômes d'interpolation de Lagrange

Comment déterminer les polynômes P qui prennent des valeurs données sur une famille $(a_i)_{i=0}^n$ d'éléments de K distincts deux à deux ? En utilisant l'application linéaire

$$u : P \in \mathsf{K}[X] \mapsto \big(P(a_0), \dots, P(a_n)\big) \in \mathsf{K}^{n+1}$$

Le noyau de u est constitué des polynômes qui admettent pour racines les scalaires a_i, $i \in [\![0, n]\!]$; $\ker u$ est donc l'ensemble des multiples du polynôme $N = \prod_{i=0}^n (X - a_i)$. Puisque N est un polynôme de degré $n + 1$, $\mathsf{K}_n[X]$ est un supplémentaire de $(N) = \ker u$, donc est isomorphe à $\operatorname{Im} u$. Ainsi $\operatorname{Im} u$ est un sous-espace vectoriel de K^{n+1} de dimension $\dim \mathsf{K}_n[X] = n + 1$, donc $\operatorname{Im} u = \mathsf{K}^{n+1}$ et $P \mapsto \big(P(a_0), \dots, P(a_n)\big)$ réalise un isomorphisme de $\mathsf{K}_n[X]$ sur K^{n+1}.
Pour $i \in [\![0, n]\!]$, on pose $L_i = \prod_{j \in [\![0,n]\!] \setminus \{j\}} \frac{X - a_j}{a_i - a_j}$. Les L_i sont sont des polynômes de degré n qui vérifient

$$\forall (i, j) \in [\![0, n]\!]^2,\; L_i(a_j) = \delta_{ij}$$

Puisque $\sum_{i=0}^{n} \lambda_i L_i(a_k) = \sum_{i=0}^{n} \lambda_i \delta_{ik} = \lambda_k$, la famille $(L_0, \dots, L_n)$ est une famille libre et maximale, donc une base de $\mathsf{K}_n[X]$ et

$$\forall P \in \mathsf{K}_n[X],\ P = \sum_{i=0}^{n} P(a_i) L_i$$

$$\left(u|_{\mathsf{K}_n[X]}\right)^{-1}(\lambda_0, \dots, \lambda_n) = \sum_{i=0}^{n} \lambda_i L_i$$

$$u(P) = (\lambda_0, \dots, \lambda_n) \iff \exists A \in \mathsf{K}_n[X],\ P = \sum_{i=0}^{n} \lambda_i L_i + AN$$

Si $\mathsf{K}_n[X]$ et K^{n+1} sont munies de leurs bases canoniques, déterminer la matrice de $u|_{\mathsf{K}_n[X]}$ et son inverse.

4 Dualité

E est un K-espace vectoriel, que l'on supposera non réduit à $\{\mathbf{0}_E\}$.

4.1 Espace dual

Définition 4.1 (Forme linéaire).
Une *forme linéaire sur* E est une application linéaire de E vers le corps des scalaires K.

Définition 4.2 (Espace dual ou dual).
Le *dual de* E est l'ensemble des formes linéaires sur E ; il est noté E^*. Rappelons que $E^* = \mathcal{L}(E, \mathsf{K})$ est un K-espace vectoriel.

Remarques.
Toute forme linéaire non nulle sur E est de rang 1, donc surjective.
Si E est de dimension finie, E^* l'est aussi et

$$\dim E^* = \dim \mathcal{L}(E, \mathsf{K}) = \dim E$$

Exemples 4.1.
Les formes linéaires sur K^n sont du type

$$\varphi : (x_1, \dots, x_n) \in \mathsf{K}^n \mapsto a_1 x_1 + \cdots + a_n x_n = (a_1, \dots, a_n) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

$\delta_a : f \mapsto f(a)$, $f \mapsto \int_a^b f$ et $\delta'_a : f \mapsto f'(a)$ sont des formes linéaires sur, respectivement, $\mathcal{C}(\mathsf{R})$, $\mathcal{C}([a, b])$ et $\mathcal{C}^1(\mathsf{R})$.
Si E est de dimension finie n, si $\mathcal{B} = (\mathbf{e}_1, \dots, \mathbf{e}_n)$ est une base de E et $\mathbf{x} = \sum_{j=1}^{n} x_j \mathbf{e}_j$ la décomposition de $\mathbf{x}$ suivant $\mathcal{B}$, les applications

$$\varphi_j : \mathbf{x} \in E \mapsto x_j$$

qui au vecteur $\mathbf{x}$ associent sa composante suivant $\mathbf{e}_j$ relativement à $\mathcal{B}$, sont des formes linéaires sur E ; on les appelle les *formes linéaires coordonnées*.

4.2 Hyperplan

Définition 4.3 (Hyperplan).
Un *hyperplan* de E est un sous-espace vectoriel qui admet une droite (vectorielle) pour supplémentaire.

$$\boxed{\mathcal{H} \text{ est un hyperplan} \iff \exists\, \mathbf{e} \in E,\ E = \mathcal{H} \oplus \mathcal{D} = \mathcal{H} \oplus \mathsf{K}\mathbf{e}}$$

Remarques.
Tous les supplémentaires d'un hyperplan sont des droites (vectorielles) puisqu'ils sont isomorphes.
Si E est de dimension finie n, les hyperplans de E sont les sous-espaces vectoriels de E de dimension $n-1$.

A) Caractérisation à l'aide des formes linéaires

Proposition 4.1. *$\mathcal{H}$ est un hyperplan si, et seulement si, $\mathcal{H}$ est le noyau d'une forme linéaire non nulle.*

Démonstration.
$\boxed{\Longrightarrow}$ Soit $\mathbf{e}$ tel que $\mathcal{H} \oplus \mathsf{K}\mathbf{e} = E$; ainsi $\forall \mathbf{x} \in E,\ \exists!(\mathbf{h}, \lambda) \in \mathcal{H} \times \mathsf{K},\ \mathbf{x} = \mathbf{h} + \lambda\mathbf{e}$ et l'application $\varphi : \mathbf{x} \in E \mapsto \lambda \in \mathsf{K}$ est une forme linéaire sur E dont le noyau est $\mathcal{H}$.
$\boxed{\Longleftarrow}$ Soient $\varphi \in E^* \setminus \{0_{E^*}\}$ et $\mathbf{e} \in E$ tel que $\varphi(\mathbf{e}) \neq 0$; pour tout $\mathbf{x} \in E$, on a

$$\mathbf{x} - \lambda\mathbf{e} \in \ker\varphi \iff 0 = \varphi(\mathbf{x} - \lambda\mathbf{e}) = \varphi(\mathbf{x}) - \lambda\varphi(\mathbf{e}) \iff \lambda = \frac{\varphi(\mathbf{x})}{\varphi(\mathbf{e})}$$

ce qui montre que $\exists!(\mathbf{h}, \lambda) = \left(\mathbf{x} - \lambda\mathbf{e}, \frac{\varphi(\mathbf{x})}{\varphi(\mathbf{e})}\right) \in \mathcal{H} \times \mathsf{K}$ et $E = \ker\varphi \oplus \mathsf{K}\mathbf{e}$, *i.e.* $\ker\varphi$ est un hyperplan. cqfd

Remarque. Si $\mathcal{H}$ est un hyperplan et φ une forme linéaire sur E de noyau $\mathcal{H}$, on a

$$E = \mathcal{H} \oplus \mathsf{K}\mathbf{e} \iff \varphi(\mathbf{e}) \neq 0 \iff \mathbf{e} \notin \mathcal{H}$$

Ainsi, toute droite (vectorielle) $\mathcal{D}$ non contenue dans l'hyperplan $\mathcal{H}$ est un supplémentaire de $\mathcal{H}$.

B) Équation d'un hyperplan

Proposition 4.2. *Si $\mathcal{H}$ est un hyperplan noyau de la forme linéaire φ, toute autre forme linéaire ψ s'annule sur $\mathcal{H}$ si, et seulement si, (φ, ψ) est liée.*

Démonstration.
$\boxed{\Longleftarrow}$ Puisque la famille (φ, ψ) est liée et $\varphi \neq 0$, il existe $\lambda \in \mathsf{K}$ tel que $\psi = \lambda\varphi$; ainsi $\ker\psi \supset \ker\varphi = \mathcal{H}$.
$\boxed{\Longrightarrow}$ Si ψ est une forme linéaire dont le noyau contient $\mathcal{H} = \ker\varphi$, $\varphi|_{\mathcal{H}} = 0 = \psi|_{\mathcal{H}}$. Si $\mathbf{e} \notin \mathcal{H}$; alors $\psi(\mathbf{e}) = \alpha\varphi(\mathbf{e})$ en posant $\alpha = \frac{\psi(\mathbf{e})}{\varphi(\mathbf{e})}$, ce qui montre que $\psi|_{\mathsf{K}\mathbf{e}} = \alpha\varphi|_{\mathsf{K}\mathbf{e}}$. Par conséquent, $\psi = \alpha\varphi$ puisque cette égalité a lieu sur $\mathcal{H}$ et $\mathsf{K}\mathbf{e}$ supplémentaires dans E. cqfd

Remarque. φ est *une équation* de $\mathcal{H}$; cette équation est unique à un *coefficient multiplicatif non nul près*.

C) Formes linéaires et vecteurs de E

Proposition 4.3. *Si $\mathbf{e}$ est un vecteur non nul de E, il existe une forme linéaire sur E qui vaut 1 en $\mathbf{e}$.*

Démonstration. Appelons $\mathcal{H}$ un hyperplan supplémentaire de la droite $\mathsf{K}\mathbf{e}$ et notons ψ une équation de $\mathcal{H}$; puisque $\mathbf{e} \notin \mathcal{H}$, $\psi(\mathbf{e}) \neq 0$ et $\left(\psi(\mathbf{e})\right)^{-1}\psi$ convient. cqfd

Corollaire. *Le seul élément de E qui annule toutes les formes linéaires sur E est $\mathbf{0}_E$*, i.e.

$$\bigcap_{\varphi \in E^*} \ker\varphi = \{\mathbf{0}_E\}$$

4.3 Base duale

Dans ce paragraphe, E désigne un K-espace vectoriel de dimension finie n muni d'une base $\mathcal{B} = (\mathbf{e}_1, \dots, \mathbf{e}_n)$; les composantes relatives à $\mathcal{B}$ d'un vecteur $\mathbf{x}$ de E sont notées $(x_1, \dots, x_n)$. Rappelons que E^* est un espace vectoriel de même dimension que E.
Les formes linéaire coordonnées $\varphi_i : \mathbf{x} \mapsto x_i$ vérifient les relations d'orthogonalité de Kronecker

$$\forall (i, j) \in [\![1, n]\!]^2,\ \varphi_i(\mathbf{e}_j) = \delta_{ij} \tag{4.1}$$

Elles constituent une base de E^* car elles forment une famille maximale et libre :

$$\sum_{i=1}^{n} \lambda_i \varphi_i = 0_{E^*} \implies \forall j \in [\![1, n]\!],\ 0 = \Big(\sum_{i=1}^{n} \lambda_i \varphi_i\Big)(\mathbf{e}_j) = \sum_{i=1}^{n} \lambda_i \varphi_i(\mathbf{e}_j) = \sum_{i=1}^{n} \lambda_i \delta_{ij} = \lambda_j$$

Définition 4.4 (Base duale).
La famille des formes linéaires coordonnées $(\varphi_i)_{1 \leqslant i \leqslant n}$ relatives à une base $\mathcal{B}$ constitue une base de E^* ; on la note $\mathcal{B}^*$ et on l'appelle *base duale de $\mathcal{B}$*.

Remarques. Si φ est une forme linéaire sur E, la coordonnée de φ suivant φ_j est $\varphi(\mathbf{e}_j)$ car

$$\varphi = \sum_{i=1}^{n} \lambda_i \varphi_i \implies \varphi(\mathbf{e}_j) = \sum_{i=1}^{n} \lambda_i \varphi_i(\mathbf{e}_j) = \sum_{i=1}^{n} \lambda_i \delta_{ij} = \lambda_j$$

Si $\mathcal{B}$ est une base de E, $\mathcal{B}^*$ est l'unique base de E^* qui vérifie les relations 4.1.

4.4 Équation d'un hyperplan en dimension finie

On note $(x_1, \dots, x_n)$ les coordonnées d'un vecteur $\mathbf{x}$ de E relatives à la base $\mathcal{B} = (\mathbf{e}_1, \dots, \mathbf{e}_n)$.

Théorème 4.4. *$\mathcal{H}$ est un hyperplan de E si, et seulement si, il existe $(a_1, \dots, a_n) \in \mathsf{K}^n \setminus \{\mathbf{0}\}$ tel que*

$$\mathbf{x} \in \mathcal{H} \iff a_1 x_1 + \cdots + a_n x_n = 0$$

Démonstration.
$\boxed{\Longrightarrow}$ Soit φ une forme linéaire non nulle sur E de noyau $\mathcal{H}$; posons $a_j = \varphi(\mathbf{e}_j)$ ce qui donne $\varphi = \sum_{i=1}^{n} a_i \varphi_i$ et : $\mathbf{x} \in \mathcal{H} = \ker \varphi \iff 0 = \varphi(\mathbf{x}) = \sum_{i=1}^{n} a_i \varphi_i(\mathbf{x}) = \sum_{i=1}^{n} a_i x_i$
$\boxed{\Longleftarrow}$ Si $a_1 x_1 + \cdots + a_n x_n = 0$ désigne une équation de $\mathcal{H}$, posons $\varphi = \sum_{i=1}^{n} a_i \varphi_i$; φ est une forme linéaire non nulle puisque les coefficients a_i ne sont pas tous nuls et $\mathcal{H} = \ker \varphi$. cqfd

Remarque. L'équation $a_1 x_1 + \cdots + a_n x_n = 0$ est une équation de $\mathcal{H}$, et toute autre équation $b_1 x_1 + \cdots + b_n x_n = 0$ vérifie

$$\exists \lambda \neq 0,\ (b_1, \dots, b_n) = \lambda (a_1, \dots, a_n)$$

5 Trace

5.1 Trace d'une matrice

Définition 5.1 (Trace d'une matrice).
À toute matrice carrée $A = [a_{ij}]$ d'ordre n, on associe le nombre $\sum_{i=1}^{n} a_{ii}$ que l'on nomme *trace* de la matrice A.

$$\boxed{\operatorname{tr} A = \sum_{i=1}^{n} a_{ii}}$$

Proposition 5.1 (Trace d'une combinaison linéaire de matrices).
tr *est une forme linéaire non nulle sur* $\mathcal{M}_n(\mathsf{K})$.

Démonstration. Elle est laissée aux soins du lecteur. cqfd

Remarque. On note $E^{i,j}$ la matrice élémentaire de $\mathcal{M}_n(\mathsf{K})$ dont tous les éléments sont nuls excepté celui à l'intersection de la $i^{\text{ème}}$ ligne et de la $j^{\text{ème}}$ colonne qui vaut 1 ; la famille $(E^{i,j})_{1\leqslant i,j\leqslant n}$ est une base de $\mathcal{M}_n(\mathsf{K})$ (appelée base canonique de $\mathcal{M}_n(\mathsf{K})$). La base duale $\mathcal{B}^* = (\varphi^{i,j})_{1\leqslant i,j\leqslant n}$ vérifie

$$\mathrm{tr} = \sum_{i=1}^{n} \varphi^{i,i}$$

ce qui montre que tr est une forme linéaire non nulle.

Corollaire. ker(tr) *est un hyperplan de* $\mathcal{M}_n(\mathsf{K})$ *dont un supplémentaire est la droite (vectorielle) des matrices scalaires,* i.e. *la droite dirigée par* I_n.

$$\boxed{\mathcal{M}_n(\mathsf{K}) = \ker(\mathrm{tr}) \oplus \mathsf{K}I_n}$$

Démonstration. Puisque tr est une forme linéaire non nulle, son noyau est un hyperplan ; puisque $\mathrm{tr}\, I_n = n \neq 0$, la droite dirigée par I_n est un supplémentaire de ker(tr). cqfd

Théorème 5.2 (Trace d'un produit de matrices).
Si A et B sont deux matrices, A de taille $n \times p$ et B de taille $p \times n$, les traces de AB et BA sont égales.

$$\boxed{\forall (A, B) \in \mathcal{M}_{n,p}(\mathsf{K}) \times \mathcal{M}_{p,n}(\mathsf{K}),\ \mathrm{tr}\, AB = \mathrm{tr}\, BA}$$

Démonstration. AB (resp. BA) est une matrice carrée d'ordre n (resp. p) et

$$\mathrm{tr}\, AB = \sum_{i=1}^{n} (AB)_{ii} = \sum_{i=1}^{n} \Big(\sum_{s=1}^{p} a_{is} b_{si} \Big) = \sum_{s=1}^{p} \sum_{i=1}^{n} a_{is} b_{si}$$
$$\mathrm{tr}\, BA = \sum_{j=1}^{p} (BA)_{jj} = \sum_{j=1}^{p} \Big(\sum_{r=1}^{n} b_{jr} a_{rj} \Big) = \sum_{j=1}^{p} \sum_{r=1}^{n} a_{jr} b_{rj}$$

cqfd

5.2 Matrices semblables

Définition 5.2 (Matrices semblables).
Deux matrices carrées d'ordre n A et B sont *semblables* s'il existe une matrice carrée d'ordre n et inversible P telle que $B = P^{-1}AP$.

$$\boxed{A \in \mathcal{M}_n(\mathsf{K}) \text{ et } B \in \mathcal{M}_n(\mathsf{K}) \text{ sont semblables} \iff \exists P \in \mathcal{GL}_n(\mathsf{K}),\ B = P^{-1}AP}$$

Remarques.
La matrice unité d'ordre n I_n n'est semblable qu'à elle même.
La relation « être semblable à » est une relation d'équivalence.

Théorème 5.3 (Endomorphisme et matrices semblables).
Si u est un endomorphisme d'un espace vectoriel E de dimension finie et si $\mathcal{B}$ et $\mathcal{B}'$ sont deux bases de E, les matrices $\mathcal{M}\mathrm{at}_{\mathcal{B}}(u)$ et $\mathcal{M}\mathrm{at}_{\mathcal{B}'}(u)$ sont semblables.

Démonstration. Notons P la matrice de changement de bases de la base $\mathcal{B}$ à la base $\mathcal{B}'$, matrice dont les colonnes sont les composantes des vecteurs de $\mathcal{B}'$ relativement à $\mathcal{B}$, soit $P = \mathcal{M}\mathrm{at}_{\mathcal{B}}(\mathcal{B}')$. Dans ce cas, on a :

$$\mathcal{M}\mathrm{at}_{\mathcal{B}'}(u) = \mathcal{M}\mathrm{at}_{\mathcal{B}'}(\mathcal{B})\mathcal{M}\mathrm{at}_{\mathcal{B}}(u)\mathcal{M}\mathrm{at}_{\mathcal{B}}(\mathcal{B}') = P^{-1}\mathcal{M}\mathrm{at}_{\mathcal{B}}(u)P \tag{5.1}$$

cqfd

Corollaire (Trace de deux matrices semblables).
Deux matrices semblables ont même trace.

$$\boxed{\forall (A, P) \in \mathcal{M}_n(\mathbb{K}) \times \mathcal{GL}_n(\mathbb{K}),\ \mathrm{tr}(P^{-1}AP) = \mathrm{tr}\, A}$$

Démonstration.
$\mathrm{tr}(P^{-1}AP) = \mathrm{tr}\big(P^{-1}(AP)\big) = \mathrm{tr}\big((AP)P^{-1}\big) = \mathrm{tr}\big(A(PP^{-1})\big) = \mathrm{tr}\, A$ cqfd

Remarque. $\mathrm{tr}(ABC) = \mathrm{tr}(BCA)$ mais $\mathrm{tr}(ABC) \neq \mathrm{tr}(BAC)$ en général. Voici un contre-exemple :

$$\begin{aligned} E^{1,2}E^{2,1}E^{1,1} = E^{1,1}E^{1,1} = E^{1,1} &\implies \mathrm{tr}(E^{1,2}E^{2,1}E^{1,1}) = \mathrm{tr}\, E^{1,1} = 1 \\ E^{2,1}E^{1,2}E^{1,1} = E^{2,2}E^{1,1} = 0 &\implies \mathrm{tr}(E^{1,2}E^{2,1}E^{1,1}) = 0 \end{aligned}$$

5.3 Trace d'un endomorphisme

Si u est un endomorphisme d'un espace vectoriel E de dimension finie et $\mathcal{B}$ et $\mathcal{B}'$ deux bases de E, les matrices $\mathcal{M}\mathrm{at}_{\mathcal{B}}(u)$ et $\mathcal{M}\mathrm{at}_{\mathcal{B}'}(u)$ sont semblables et donc leurs traces sont égales, ce qui permet la

Définition 5.3 (Trace d'un endomorphisme).
On appelle *trace d'un endomorphisme* la trace de l'une quelconque de ses matrices.

$$\boxed{\mathrm{tr}\, u = \mathrm{tr}\big(\mathcal{M}\mathrm{at}_{\mathcal{B}}(u)\big)}$$

Remarque. tr est une forme linéaire non nulle sur $\mathcal{L}(E)$ et $\mathcal{L}(E) = \ker(\mathrm{tr}) \oplus \mathbb{K}I_E$

Proposition 5.4 (Trace d'un projecteur).
La trace d'un projecteur est égale à son rang.

Démonstration. Si p est un projecteur et $\mathcal{B}$ une base adaptée à la décomposition $E = \ker(I_E - p) \oplus \ker p$, la matrice de p relativement à cette base est

$$\mathcal{M}\mathrm{at}_{\mathcal{B}}(p) = \left(\begin{array}{c|c} I_r & 0_{r,n-r} \\ \hline 0_{n-r,r} & 0_{n-r,n-r} \end{array}\right)$$

ce qui montre que $\mathrm{tr}\, p = r = \mathrm{rg}\, p$. cqfd

Cette proposition ne caractérise pas les projecteurs : $A = \begin{pmatrix} 0 & 1 \\ 1 & 2 \end{pmatrix}$ a une trace et un rang égaux à 2, mais ne peut être un projecteur car tout projecteur de rang maximum est l'identité.

CHAPITRE

2

Déterminant

Sommaire

1 Groupe symétrique

n désigne un entier au moins égal à 2.

1.1 Généralités

Définition 1.1 (Permutation).
Une *permutation* d'un ensemble X est une bijection de X sur lui-même.

Définition 1.2 (Groupe symétrique).
$\mathfrak{S}_n$ désigne l'ensemble des permutations de $[\![1, n]\!]$; on l'appelle le *groupe symétrique* d'ordre n. Une permutation $s \in \mathfrak{S}_n$ est notée

$$s = \begin{pmatrix} 1 & 2 & \dots & n-1 & n \\ s(1) & s(2) & \dots & s(n-1) & s(n) \end{pmatrix}$$

Définition 1.3 (Transposition).
La permutation qui échange i et j et laisse les autres éléments invariants est appelée *transposition* et est notée $\tau_{i,j}$.

Définition 1.4 (Cycle).
Un *cycle de longueur* q ($q \leqslant n$) est une permutation c telle qu'il existe un sous-ensemble à q éléments de $\{a_1, \cdots, a_q\}$ vérifiant

$$c(a_1) = a_2,\ c(a_2) = a_3,\ \dots, c(a_{q-1}) = a_q,\ c(a_q) = a_1$$

les autres éléments restant invariants.

Exemples 1.1.
$\mathfrak{S}_2$ contient deux éléments : la permutation identique et la transposition qui échange 1 et 2.
Les six éléments de $\mathfrak{S}_3$ sont

- la permutation identique : $e = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}$;
- trois transpositions : $\tau_1 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}$, $\tau_2 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$ et $\tau_3 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$;
- deux cycles de longueur 3 : $c_1 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$ et $c_2 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$.

Théorème 1.1 (Structure de groupe pour $\mathfrak{S}_n$).
Muni de la composition des applications, $\mathfrak{S}_n$ est un groupe de cardinal $n!$ et non commutatif pour $n \geqslant 3$.

Démonstration. Rappelons qu'une structure de groupe nécessite une loi de composition interne, associative, possédant un élément neutre et un symétrique pour tout élément.

- $(s_1, s_2) \mapsto s_1 \circ s_2$ est une loi de composition interne : la composée de deux bijections est une bijection ;
- la composition des applications est associative ;
- l'identité de $[\![1, n]\!]$, que l'on note e, est une permutation ; c'est l'élément neutre pour la composition ;
- si s est une permutation, s^{-1} en est une.

$(\mathfrak{S}_n, \circ)$ est donc un groupe. cqfd

Remarques. $\mathfrak{S}_n$ contient exactement $\binom{n}{2}$ transpositions. Toute transposition est une involution : $\tau \circ \tau = e$; une transposition est un cycle de longueur 2.

1.2 Structure d'une permutation

Théorème 1.2 (Décomposition d'une permutation).
Toute permutation de $\mathfrak{S}_n$ est un produit d'au plus n transpositions.

Démonstration. Effectuons une récurrence sur n.
$\mathfrak{S}_2$ est un groupe à 2 éléments τ et $e = \tau \circ \tau$; la propriété est donc vraie pour $n = 2$.
On suppose la propriété vraie au rang n. Soit $s \in \mathfrak{S}_{n+1}$;

- si $s(n+1) = n+1$, on note s' la permutation induite par la restriction de s à $[\![1, n]\!]$; s' est une permutation de $[\![1, n]\!]$ et grâce à l'hypothèse de récurrence, $s' = \tau'_1 \circ \cdots \circ \tau'_k$ avec $k \leqslant n$, où les τ'_j sont des transpositions de $\mathfrak{S}_n$. On pose $\tau_j(i) = \tau'_j(i)$ si $i \in [\![1, n]\!]$ et $\tau_j(n+1) = n+1$; les τ_j sont des transpositions de $\mathfrak{S}_{n+1}$ et $s = \tau_1 \circ \cdots \circ \tau_k$;
- si $s(n+1) = p \in [\![1, n]\!]$, $\tau_{p,n+1} \circ s$ est une permutation qui laisse $n+1$ invariant et on retrouve le cas précédent ; on peut écrire que $\tau_{p,n+1} \circ s = \tau_1 \circ \cdots \circ \tau_k$ avec $k \leqslant n$, et $s = \tau_{p,n+1} \circ \tau_1 \circ \cdots \circ \tau_k$

Ainsi toute permutation $s \in \mathfrak{S}_{n+1}$ est un produit d'au plus $n+1$ transpositions.
Le théorème de récurrence montre que la propriété est vraie pour tout n. cqfd

Remarque.
La décomposition en produit de transpositions n'est pas unique ; par exemple

$$s = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 4 & 2 \end{pmatrix} = \tau_{2,4} \circ \tau_{1,2} \circ \tau_{1,3} = \tau_{1,3} \circ \tau_{3,4} \circ \tau_{2,4}$$

1.3 Signature d'une permutation

Définition 1.5 (Signature).
Soit $s \in \mathfrak{S}_n$; on dit que le couple (i, j) avec $i < j$, présente une *inversion* pour s si $s(i) > s(j)$. On note $I(s)$ le nombre d'inversions de s et on appelle *signature de la permutation* s le nombre $\varepsilon(s) \in \{-1, 1\}$ défini par

$$\varepsilon(s) = (-1)^{I(s)}$$

La signature de l'identité est 1 : $\varepsilon(e) = 1$.

Théorème 1.3 (Signature d'une transposition).
La signature d'une transposition est -1.

Démonstration. Soit τ la transposition qui échange k et l, avec $k < l$:

$$\tau = \begin{pmatrix} 1 & 2 & \ldots & k-1 & k & k+1 & \ldots & l-1 & l & l+1 & \ldots & n \\ 1 & 2 & \ldots & k-1 & l & k+1 & \ldots & l-1 & k & l+1 & \ldots & n \end{pmatrix}$$

Comptons le nombre d'inversions de τ :

- les couples (i, j) avec $i \in [\![1, k-1]\!] \cup [\![l, n]\!]$ et $i < j$ ne présentent pas d'inversion ;
- le couple (k, j) avec $k < j$ présente une inversion si, et seulement si, j appartient à $[\![k+1, l]\!]$, ce qui fait $l - k$ inversion(s) ;
- si $i \in [\![k+1, l-1]\!]$ et $i < j$, (i, j) présente une inversion si, et seulement si, $j = l$, ce qui fait $l - 1 - k$ inversion(s).

ce qui donne $I(\tau) = (l-k) + (l-1-k) = 2(l-k-1) + 1$; ce nombre est impair et $\varepsilon(\tau) = (-1)^{I(\tau)} = -1$.
cqfd

Considérons maintenant le produit

$$V_n = \prod_{1\leqslant i<j\leqslant n} (j-i) = \big[(2-1)\big] \times \big[(3-1)(3-2)\big] \times \cdots \times \big[(n-1)(n-2)\dots\big(n-(n-1)\big)\big]$$

Pour $s \in \mathfrak{S}_n$, posons $s \cdot V_n = \prod_{1\leqslant i<j\leqslant n}\big(s(j)-s(i)\big)$. Puisque s est une bijection, les $\binom{n}{2}$ facteurs de V_n se retrouvent, au signe près, une et une seule fois dans $s \cdot V_n$ et $s \cdot V_n = (-1)^{I(s)}V_n = \varepsilon(s)V_n$. Ainsi :

Proposition 1.4 (Expression de la signature).
La signature d'une permutation $s \in \mathfrak{S}_n$ est donnée par : $\varepsilon(s) = \displaystyle\prod_{1\leqslant i<j\leqslant n} \frac{s(j)-s(i)}{j-i}$

Théorème 1.5 (Signature d'un produit de permutations).
ε est un morphisme du groupe $(\mathfrak{S}_n, \circ)$ sur le groupe $\{1,-1\}$ muni de la multiplication, i.e.

$$\forall (s_1, s_2) \in \mathfrak{S}_n^2,\ \varepsilon(s_1 \circ s_2) = \varepsilon(s_1) \times \varepsilon(s_2)$$

Démonstration. On a : $(s_1 \circ s_2) \cdot V_n = s_1 \cdot (s_2 \cdot V_n) = \varepsilon(s_1)(s_2 \cdot V_n) = \varepsilon(s_1)\varepsilon(s_2)V_n$. Puisque $(s_1 \circ s_2) \cdot V_n = \varepsilon(s_1 \circ s_2)V_n$, la formule annoncée est démontrée. cqfd

Corollaire. *La signature d'une permutation produit de p transpositions est $(-1)^p$.*

2 Applications multilinéaires

2.1 Applications bilinéaires

Définition 2.1 (Bilinéarité). Soient E_1, E_2 et F trois K-espaces vectoriels ; une application f de $E_1 \times E_2$ à valeurs dans F est *bilinéaire* si

(i) pour tout $\mathbf{x} \in E_1$, l'application $\mathbf{y} \mapsto f(\mathbf{x}, \mathbf{y})$ est linéaire ;
(ii) pour tout $\mathbf{y} \in E_2$, l'application $\mathbf{x} \mapsto f(\mathbf{x}, \mathbf{y})$ est linéaire.

Si $F = \mathsf{K}$, on parle de *forme bilinéaire*.

Exemples 2.1.
Le produit scalaire est une forme bilinéaire sur $\mathsf{R}^3 \times \mathsf{R}^3$; le produit vectoriel est une application bilinéaire de $\mathsf{R}^3 \times \mathsf{R}^3$ sur R^3.
Soit $\mathcal{A}$ une K-algèbre (par exemple K, $\mathsf{K}[X]$, $\mathcal{M}_n(\mathsf{K})$, $\mathcal{L}(E)$) ; les applications

$$f_1 : (x, y) \mapsto x.y, \qquad f_2 : (x, y) \mapsto x.y + y.x, \qquad f_1 : (x, y) \mapsto x.y - y.x$$

sont des applications bilinéaires de $\mathcal{A} \times \mathcal{A}$ vers $\mathcal{A}$ (le démontrer).
Si $\mathcal{A} = \mathsf{K}$ ou $\mathsf{K}[X]$, $f_2 = 2f_1$ et $f_3 = 0$.
Si $\mathcal{A} = \mathcal{M}_n(\mathsf{K})$ ou $\mathcal{L}(E)$, f_3 n'est pas l'application nulle.
$\varphi : (f, g) \mapsto \int_a^b f(t)g(t)\,dt$ est une forme bilinéaire sur $\mathcal{C}([a, b], \mathsf{K}) \times \mathcal{C}([a, b], \mathsf{K})$.
$\psi : (A, B) \mapsto \operatorname{tr}({}^tAB)$ est une forme bilinéaire sur $\mathcal{M}_n(\mathsf{K}) \times \mathcal{M}_n(\mathsf{K})$.

Définition 2.2 (Symétrie, antisymétrie, alternance).
Soient E et F deux K-espaces vectoriels et f une application bilinéaire de $E \times E$ à valeurs dans F ; on dit que

(i) f est *symétrique* si $\forall (\mathbf{x}, \mathbf{y}) \in E^2,\ f(\mathbf{x}, \mathbf{y}) = f(\mathbf{y}, \mathbf{x})$;
(ii) f est *antisymétrique* si $\forall (\mathbf{x}, \mathbf{y}) \in E^2,\ f(\mathbf{x}, \mathbf{y}) = -f(\mathbf{y}, \mathbf{x})$;
(iii) f est *alternée* si $\forall \mathbf{x} \in E,\ f(\mathbf{x}, \mathbf{x}) = \mathbf{0}_F$;

Exemples 2.2.
Le produit scalaire est symétrique, le produit vectoriel est antisymétrique.
f_2 est symétrique, f_3 est antisymétrique et f_1 est symétrique si, et seulement si, $\mathcal{A}$ est une algèbre *commutative*.
φ est symétrique.
ψ est symétrique.

Théorème 2.1 (Antisymétrie et alternance).
Si f est une application bilinéaire de $E \times E$ dans F, f est antisymétrique si, et seulement si, f est alternée.

Démonstration.
$\boxed{\Longrightarrow}$ Pour tout $\mathbf{x} \in E$, $f(\mathbf{x}, \mathbf{x}) = -f(\mathbf{x}, \mathbf{x})$ puisque f est antisymétrique, donc $f(\mathbf{x}, \mathbf{x}) = \mathbf{0}_F$.
$\boxed{\Longleftarrow}$ Pour tout $(\mathbf{x}, \mathbf{y}) \in E^2$, on peut écrire

$$\begin{aligned} f(\mathbf{x}+\mathbf{y}, \mathbf{x}+\mathbf{y}) = \mathbf{0}_F & & \text{puisque } f \text{ est alternée} \\ = f(\mathbf{x}, \mathbf{x}+\mathbf{y}) + f(\mathbf{y}, \mathbf{x}+\mathbf{y}) & & \\ = f(\mathbf{x}, \mathbf{x}) + f(\mathbf{x}, \mathbf{y}) + f(\mathbf{y}, \mathbf{x}) + f(\mathbf{y}, \mathbf{y}) & & \text{puisque } f \text{ est bilinéaire} \\ = f(\mathbf{x}, \mathbf{y}) + f(\mathbf{y}, \mathbf{x}) & & \text{puisque } f \text{ est alternée} \end{aligned}$$

ce qui montre que $f(\mathbf{x}, \mathbf{y}) = -f(\mathbf{y}, \mathbf{x})$. cqfd

2.2 Applications p-linéaires

p désigne un entier au moins égal à 2.

Définitions 2.3 (Application et forme p-linéaires).
Soient E et F deux K-espaces vectoriels ; une application f de E^p à valeurs dans F est dite *p-linéaire* si elle est linéaire par rapport à chacune de ses variables, *i.e.* si pour tout $j \in [\![1, p]\!]$ et pour tout $\mathbf{a}_k \in E$, $k \in [\![1, p]\!] \setminus \{j\}$, les applications

$$\mathbf{x} \in E \mapsto f(\mathbf{a}_1, \ldots, \mathbf{a}_{j-1}, \mathbf{x}, \mathbf{a}_{j+1}, \ldots, \mathbf{a}_p) \in F$$

sont linéaires.
L'ensemble des applications p-linéaires de E dans F est noté $\mathcal{L}_p(E, F)$.
Les applications p-linéaires de E vers le corps des scalaires K sont appelées *formes p-linéaires* sur E.

Exemples 2.3.

(i) Si $\varphi_1, \ldots, \varphi_p$ sont des formes linéaires sur E, l'application

$$f : (\mathbf{x}_1, \ldots, \mathbf{x}_p) \in E^p \mapsto \varphi_1(\mathbf{x}_1) \times \cdots \times \varphi_p(\mathbf{x}_p)$$

est une forme p-linéaire sur E.

(ii) L'application déterminant

$$(\mathbf{x}, \mathbf{y}) \in \mathsf{K}^2 \times \mathsf{K}^2 \mapsto \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} = x_1 y_2 - x_2 y_1$$

est une forme 2-linéaire (ou bilinéaire) sur K^2.

(iii) L'application déterminant

$$(\mathbf{x}, \mathbf{y}, \mathbf{z}) \in \mathsf{K}^3 \times \mathsf{K}^3 \times \mathsf{K}^3 \mapsto \begin{vmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{vmatrix}$$
$$= (x_1 y_2 z_3 + x_2 y_3 z_1 + x_3 y_1 z_2) - (x_3 y_2 z_1 + x_1 y_3 z_2 + x_2 y_1 z_3)$$

est une forme 3-linéaire (ou trilinéaire) sur K^3.

Proposition 2.2.

(i) *L'application nulle de E^p vers F est à la fois p-linéaire et linéaire.*
(ii) *Si f est une application p-linéaire de E dans F, et si l'un des vecteurs $\mathbf{x}_i$ est nul, le vecteur $f(\mathbf{x}_1, \dots, \mathbf{x}_p)$ est nul.*
(iii) *$\mathcal{L}_p(E, F)$ est un K-espace vectoriel.*

Démonstration.

(ii) f est une application linéaire par rapport à sa i^{e} variable, et l'image de $\mathbf{0}_E$ par une application linéaire est $\mathbf{0}_F$.
(iii) $\mathcal{L}_p(E, F)$ est un sous-espace vectoriel de l'espace de toutes les applications de E^p vers F.

cqfd

Définitions 2.4 (Symétrie, antisymétrie et alternance).
Soient E et F deux K-espaces vectoriels et f une application p-linéaire sur E à valeurs dans F ; on dit que

– f est *symétrique* si $\forall (\mathbf{x}_1, \dots, \mathbf{x}_p) \in E^p,\ \forall (i, j) \in [\![1, p]\!]$,

$$i < j \implies f(\dots, \mathbf{x}_i, \dots, \mathbf{x}_j, \dots) = f(\dots, \mathbf{x}_j, \dots, \mathbf{x}_i, \dots)$$

– f est *antisymétrique* si $\forall (\mathbf{x}_1, \dots, \mathbf{x}_p) \in E^p,\ \forall (i, j) \in [\![1, p]\!]$,

$$i < j \implies f(\dots, \mathbf{x}_i, \dots, \mathbf{x}_j, \dots) = -f(\dots, \mathbf{x}_j, \dots, \mathbf{x}_i, \dots)$$

– f est *alternée* si $\forall (\mathbf{x}_1, \dots, \mathbf{x}_p) \in E^p,\ \forall (i, j) \in [\![1, p]\!]$,

$$i < j \text{ et } \mathbf{x}_i = \mathbf{x}_j \implies f(\dots, \mathbf{x}_i, \dots, \mathbf{x}_j, \dots) = \mathbf{0}_F$$

Exemples 2.4. Reprenons les exemples précédents.

(i) f est symétrique
(ii) et *iii*. Les déterminants sont antisymétriques et alternés. Pourquoi ?

Proposition 2.3. *L'ensemble des applications n-linéaires symétriques et l'ensemble des applications n-linéaires alternées de E à valeurs dans F sont des sous-espaces vectoriels de $\mathcal{L}_p(E, F)$.*

Démonstration.

Prenez vos ardoises et vos crayons.
Écrivez-moi cette démonstration.

cqfd

Théorème 2.4 (Symétrie et permutation).
Soit f une application p-linéaire de E vers F ; les propriétés suivantes sont équivalentes :

(i) *f est symétrique ;*
(ii) *pour toute permutation $s \in \mathfrak{S}_p$ et tout $(\mathbf{x}_1, \dots, \mathbf{x}_p) \in E^p$*

$$f(\mathbf{x}_{s(1)}, \dots, \mathbf{x}_{s(p)}) = f(\mathbf{x}_1, \dots, \mathbf{x}_p)$$

Démonstration.

$\boxed{\Longleftarrow}$ La transposition $\tau_{i,j} \in \mathfrak{S}_p$ qui échange i et j donne la symétrie.
$\boxed{\Longrightarrow}$ La symétrie montre que la propriété est vraie pour toute transposition, et, puisqu'une permutation est un produit de transpositions, une récurrence sur le nombre de transpositions montre le résultat. cqfd

Théorème 2.5 (Antisymétrie, alternance et permutation).
Soit f une application p-linéaire de E vers F ; les propriétés suivantes sont équivalentes :

(i) *f est alternée ;*
(ii) *f est antisymétrique ;*
(iii) *pour toute permutation $s \in \mathfrak{S}_p$ et tout $(\mathbf{x}_1, \ldots, \mathbf{x}_p) \in E^p$*

$$f(\mathbf{x}_{s(1)}, \ldots, \mathbf{x}_{s(p)}) = \varepsilon(s) f(\mathbf{x}_1, \ldots, \mathbf{x}_p)$$

Démonstration.

i. $\Longleftrightarrow$ *ii.* Soient $1 \leqslant i < j \leqslant p$ et l'application

$$g_{i,j} : (\mathbf{x}_i, \mathbf{x}_j) \in E^2 \mapsto f(\ldots, \mathbf{x}_i, \ldots, \mathbf{x}_j, \ldots)$$

$g_{i,j}$ est bilinéaire puisque f est p-linéaire ; donc $g_{i,j}$ est antisymétrique si, et seulement si, $g_{i,j}$ est alternée, ce qui montre l'équivalence.

iii. $\Longleftarrow$ *ii.* Il suffit de prendre une transposition et f est antisymétrique puisque la signature d'une transposition est -1.

ii. $\Longrightarrow$ *iii.* La formule est vraie pour les transpositions. Puisque toute permutation $s \in \mathfrak{S}_p$ se décompose en un produit de transpositions $s = \tau_1 \circ \cdots \circ \tau_k$ et puisque la signature de s vaut $(-1)^k$, une récurrence, sur le nombre k de transpositions, achève la démonstration. cqfd

Théorème 2.6 (Règle de calcul).
On ne change pas la valeur prise par une application p-linéaire alternée sur un p-uple de E^p en ajoutant à l'un des vecteurs une combinaison linéaire des autres vecteurs.
En particulier, toute application p-linéaire alternée prend la valeur $\mathbf{0}_F$ sur tout p-uple qui constitue une famille liée.

Démonstration. Soit f une application p-linéaire alternée sur E. En ajoutant la combinaison linéaire $\sum_{j \neq k} \lambda_j \mathbf{x}_j$ au vecteur $\mathbf{x}_k$, on obtient

$$\begin{aligned} f(\ldots, \mathbf{x}_k + \sum_{j \neq k} \lambda_j \mathbf{x}_j, \ldots) &= f(\ldots, \mathbf{x}_k, \ldots) + \sum_{j \neq k} \lambda_j f(\ldots, \mathbf{x}_j, \ldots, \mathbf{x}_j, \ldots) \\ &= f(\ldots, \mathbf{x}_k, \ldots) \qquad \text{car } f \text{ est alternée} \end{aligned}$$

Si la famille $(\mathbf{x}_1, \ldots, \mathbf{x}_p)$ est liée, l'un des vecteurs, par exemple $\mathbf{x}_k$, est combinaison des autres vecteurs, et

$$f(\ldots, \mathbf{x}_k, \ldots) = f(\ldots, \sum_{j \neq k} \lambda_j \mathbf{x}_j, \ldots) = \sum_{j \neq k} \lambda_j f(\ldots, \mathbf{x}_j, \ldots, \mathbf{x}_j, \ldots) = \mathbf{0}_F$$

puisque f est alternée. cqfd

Remarques.
Si E est un espace vectoriel de dimension finie, la seule application p-linéaire alternée, avec $p > \dim E$, est l'application nulle, car, dans ce cas, toute famille de p vecteurs est liée. Les seuls cas intéressants sont ceux où $p \leqslant \dim E$. Le programme nous demande d'étudier le cas où $p = \dim E = n$.
Soient E un espace vectoriel de dimension 2, $\mathcal{B} = (\mathbf{e}_1, \mathbf{e}_2)$ une base de E et f une application bilinéaire alternée sur E. Décomposons $\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2$ et $\mathbf{y} = y_1\mathbf{e}_1 + y_2\mathbf{e}_2$ sur la base $\mathcal{B}$. On obtient :

$$\begin{aligned} f(\mathbf{x}, \mathbf{y}) &= x_1y_1 f(\mathbf{e}_1, \mathbf{e}_1) + x_1y_2 f(\mathbf{e}_1, \mathbf{e}_2) + x_2y_1 f(\mathbf{e}_2, \mathbf{e}_1) + x_2y_2 f(\mathbf{e}_2, \mathbf{e}_2) \\ &= (x_1y_2 - x_2y_1) f(\mathbf{e}_1, \mathbf{e}_2) \qquad \text{car } f \text{ est alternée} \\ &= \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} f(\mathbf{e}_1, \mathbf{e}_2) \end{aligned}$$

et on retrouve le déterminant des deux vecteurs $\mathbf{x}$ et $\mathbf{y}$ dans la base $\mathcal{B}$.

Envisageons le cas d'un espace vectoriel E de dimension 3, muni d'une base $\mathcal{B} = (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$. Si f est une application trilinéaire alternée sur E, on obtient, de manière analogue en ne notant que les termes non nuls :

$$\begin{aligned} f(\mathbf{x}, \mathbf{y}, \mathbf{z}) &= x_1y_2z_3\, f(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) + x_1y_3z_2\, f(\mathbf{e}_3, \mathbf{e}_3\mathbf{e}_2) + x_2y_1z_3\, f(\mathbf{e}_2, \mathbf{e}_1, \mathbf{e}_3) \\ &\quad + x_2y_3z_1\, f(\mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_1) + x_3y_1z_2\, f(\mathbf{e}_3, \mathbf{e}_1, \mathbf{e}_2) + x_3y_2z_1\, f(\mathbf{e}_3, \mathbf{e}_2, \mathbf{e}_1) \\ &= \big((x_1y_2z_3 + x_2y_3z_1 + x_3y_1z_2) - (x_3y_2z_1 + x_1y_3z_2 + x_2y_1z_3)\big)\, f(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) \\ &= \begin{vmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{vmatrix} f(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) \end{aligned}$$

On retrouve le déterminant des trois vecteurs $\mathbf{x}$, $\mathbf{y}$ et $\mathbf{z}$ dans la base $\mathcal{B}$.

3 Déterminant de n vecteurs

Dans cette section, E désigne un $\mathbb{K}$-espace vectoriel de dimension finie n, muni d'une base $\mathcal{B} = (\mathbf{e}_1, \dots, \mathbf{e}_n)$; la base duale $\mathcal{B}^* = (\varphi_1, \dots, \varphi_n)$ est définie par $\varphi_i(\mathbf{x}) = x_i$ la coordonnée de rang i relative à $\mathcal{B}$. Nous étudions l'ensemble $\Lambda_n^*(E)$ des *formes n-linéaires alternées* définies sur E et f désigne une telle forme.

3.1 Déterminant de n vecteurs dans la base $\mathcal{B}$

Soit $(\mathbf{x}_1, \dots, \mathbf{x}_n) \in E^n$; les composantes de $\mathbf{x}_j$ dans $\mathcal{B}$ sont notées $a_{i,j}$ et pour $f \in \Lambda_n^*(E)$, on écrit

$$\begin{aligned} f(\mathbf{x}_1, \dots, \mathbf{x}_n) &= f\Big(\sum_{i_1=1}^{n} a_{i_1,1}\mathbf{e}_{i_1}, \mathbf{x}_2, \dots, \mathbf{x}_n\Big) \\ &= \sum_{i_1=1}^{n} a_{i_1,1}\, f(\mathbf{e}_{i_1}, \mathbf{x}_2, \dots, \mathbf{x}_n) \quad \text{linéarité de } f \text{ par rapport à } \mathbf{x}_1 \\ &= \sum_{1 \leqslant i_1, \dots, i_n \leqslant n} a_{i_1,1} \dots a_{i_n,n}\, f(\mathbf{e}_{i_1}, \dots, \mathbf{e}_{i_n}) \quad n\text{-linéarité de } f \end{aligned}$$

Étant donnés $(i_1, \dots, i_n) \in [\![1, n]\!]^n$, on pose $s(k) = i_k$ pour $k \in [\![1, n]\!]$. Si s n'est pas injective, deux vecteurs $\mathbf{e}_{s(k)}$ sont égaux et, puisque que f est alternée,

$$f(\mathbf{e}_{i_1}, \dots, \mathbf{e}_{i_n}) = f(\mathbf{e}_{s(1)}, \dots, \mathbf{e}_{s(n)}) = 0$$

Sinon s est bijective et appartient à $\mathfrak{S}_n$, et

$$f(\mathbf{e}_{i_1}, \dots, \mathbf{e}_{i_n}) = f(\mathbf{e}_{s(1)}, \dots, \mathbf{e}_{s(n)}) = \varepsilon(s)\, f(\mathbf{e}_1, \dots, \mathbf{e}_n)$$

Ainsi

$$\begin{aligned} f(\mathbf{x}_1, \dots, \mathbf{x}_n) &= \sum_{1 \leqslant i_1, \dots, i_n \leqslant n} a_{i_1,1} \dots a_{i_n,n}\, f(\mathbf{e}_{i_1}, \dots, \mathbf{e}_{i_n}) \\ &= \sum_{s \in \mathfrak{S}_n} a_{s(1),1} \dots a_{s(n),n}\, f(\mathbf{e}_{s(1)}, \dots, \mathbf{e}_{s(n)}) \\ &= \Big(\sum_{s \in \mathfrak{S}_n} \varepsilon(s) a_{s(1),1} \dots a_{s(n),n}\Big) f(\mathbf{e}_1, \dots, \mathbf{e}_n) \\ &= \Big(\sum_{s \in \mathfrak{S}_n} \varepsilon(s) \prod_{j=1}^{n} a_{s(j),j}\Big) f(\mathbf{e}_1, \dots, \mathbf{e}_n) \end{aligned}$$

Posons $i = s(j)$ soit $j = s^{-1}(i)$; on obtient $\prod_{j=1}^{n} a_{s(j),j} = \prod_{i=1}^{n} a_{i,s^{-1}(i)}$; comme l'application $s \to s^{-1}$ est une bijection de $\mathfrak{S}_n$ (c'est même une involution) et $\varepsilon(s^{-1}) = \varepsilon(s)^{-1} = \varepsilon(s)$, on peut écrire, en effectuant le changement d'indice de sommation $\sigma = s^{-1}$

$$\sum_{s\in\mathfrak{S}_n} \varepsilon(s) \prod_{j=1}^{n} a_{s(j),j} = \sum_{s\in\mathfrak{S}_n} \varepsilon(s^{-1}) \prod_{i=1}^{n} a_{i,s^{-1}(i)} = \sum_{\sigma\in\mathfrak{S}_n} \varepsilon(\sigma) \prod_{i=1}^{n} a_{i,\sigma(i)}$$

Définition 3.1 (Déterminant de n vecteurs dans une base).
Soient E un $\mathbb{K}$-espace vectoriel de dimension finie n et $\mathcal{B} = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ une base de E ; on appelle *déterminant dans la base $\mathcal{B}$* et l'on note $\det_{\mathcal{B}}$, l'unique forme n-linéaire alternée sur E qui prend la valeur 1 sur $\mathcal{B}$.
Si pour tout $j \in [\![1, n]\!]$, $\mathbf{x}_j = \sum_{i=1}^{n} a_{i,j}\mathbf{e}_i$, le scalaire $\det(\mathbf{x}_1, \ldots, \mathbf{x}_n)$, appelé *déterminant dans la base $\mathcal{B}$ du n-uple* $(\mathbf{x}_1, \ldots, \mathbf{x}_n) \in E^n$, admet deux expressions

$$\det_{\mathcal{B}}(\mathbf{x}_1, \ldots, \mathbf{x}_n) = \sum_{s\in\mathfrak{S}_n} \varepsilon(s) \prod_{j=1}^{n} a_{s(j),j} = \sum_{s\in\mathfrak{S}_n} \varepsilon(s) \prod_{i=1}^{n} a_{i,s(i)} \tag{3.1}$$

Démonstration. L'essentiel a été vu avant l'énoncé du théorème ; reste à prouver l'existence de ce gros machin.
En utilisant les éléments φ_i de la base duale, on obtient

$$\prod_{j=1}^{n} a_{s(j),j} = \prod_{j=1}^{n} \varphi_{s(j)}(\mathbf{x}_j) = \varphi_{s(1)}(\mathbf{x}_1) \times \cdots \times \varphi_{s(n)}(\mathbf{x}_n)$$

ce qui montre la n-linéarité de ces produits par rapport à $(\mathbf{x}_1, \ldots, \mathbf{x}_n)$ et la n-linéarité de $\det B$ par combinaison linéaire.
Soit τ une transposition de $\mathfrak{S}_n$; comme $s \to s \circ \tau$ est une bijection de $\mathfrak{S}_n$ (c'est même une involution) et puisque $\varepsilon(s \circ \tau) = \varepsilon(s)\varepsilon(\tau) = -\varepsilon(s)$, on peut écrire

$$\begin{aligned}\det_{\mathcal{B}}(\mathbf{x}_{\tau(1)}, \ldots, \mathbf{x}_{\tau(n)}) &= \sum_{s\in\mathfrak{S}_n} \varepsilon(s) \prod_{i=1}^{n} a_{i,s(\tau(i))} = \sum_{s\in\mathfrak{S}_n} -\varepsilon(s\circ\tau) \prod_{i=1}^{n} a_{i,s(\tau(i))} \\ &= -\sum_{\sigma\in\mathfrak{S}_n} \varepsilon(\sigma) \prod_{i=1}^{n} a_{i,\sigma(i)}\end{aligned}$$

en utilisant le changement d'indice $\sigma = s \circ \tau$. cqfd

Théorème 3.1 (Dimension de $\Lambda_n^*(E)$).
L'espace vectoriel $\Lambda_n^*(E)$ des formes n-linéaires alternées sur un espace vectoriel E de dimension n est de dimension 1 *; il admet pour base* $(\det_{\mathcal{B}})$ *et*

$$\forall f \in \Lambda_n^*(E), \quad f = f(\mathbf{e}_1, \ldots, \mathbf{e}_n) \det_{\mathcal{B}} = f(\mathcal{B}) \det_{\mathcal{B}} \tag{3.2}$$

Théorème 3.2 (Déterminant d'un système triangulaire de vecteurs).
Si chaque vecteur $\mathbf{x}_j$ est combinaison linéaire des vecteurs $(\mathbf{e}_1, \ldots, \mathbf{e}_j)$, autrement dit, si $a_{i,j} = 0$ pour $i > j$,

$$\det_{\mathcal{B}}(\mathbf{x}_1, \ldots, \mathbf{x}_n) = \begin{vmatrix} a_{1,1} & a_{1,2} & \ldots & a_{1,n} \\ 0 & a_{2,2} & \ldots & a_{2,n} \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \ldots & a_{n,n} \end{vmatrix} = \prod_{i=1}^{n} a_{i,i} \tag{3.3}$$

Démonstration. Soit $s \in \mathfrak{S}_n$ telle qu'il existe j_0 avec $s(j_0) > j_0$; alors $a_{s(j_0),j_0} = 0$ et $\prod_{j=1}^{n} a_{s(j),j} = 0$. Dans la somme qui définit $\det(\mathbf{x}_1, \ldots, \mathbf{x}_n)$, seuls les permutations s pour lesquelles $s(j) \leqslant j$ pour tout $j \in [\![1, n]\!]$ peuvent données un produit non nul, ce qui impose à s d'être la permutation identique et donne le résultat. cqfd

3.2 Caractérisation des bases

Théorème 3.3 (Passage d'une base à une autre).
Si $\mathcal{B}$ et $\mathcal{B}'$ sont deux bases de E et $\mathcal{V}$ un n-uple de E^n, on a

$$\det_{\mathcal{B}}(\mathcal{B}') \times \det_{\mathcal{B}'}(\mathcal{B}) = 1 \tag{3.4}$$

$$\det_{\mathcal{B}'}(\mathcal{V}) = \det_{\mathcal{B}'}(\mathcal{B}) \times \det_{\mathcal{B}}(\mathcal{V}) \tag{3.5}$$

Démonstration. $\det_{\mathcal{B}'}$ est une forme n-linéaire alternée à qui on applique la formule 3.2, soit

$$\det_{\mathcal{B}'} = \det_{\mathcal{B}'}(\mathcal{B}) \det_{\mathcal{B}}$$

En prenant la valeur de ces expressions en $\mathcal{B}'$ et en $\mathcal{V}$, on obtient les résultats. cqfd

Théorème 3.4 (Caractérisation des bases).
Si $\mathcal{V} = (\mathbf{v}_1, \dots, \mathbf{v}_n)$ est une famille de n vecteurs d'un espace vectoriel de dimension n et si $\mathcal{B}$ est une base de E,

$$\boxed{\mathcal{V} \text{ est une base de } E \iff \det_{\mathcal{B}}(\mathcal{V}) \neq 0}$$

Démonstration.
$\boxed{\Longrightarrow}$ Si $\mathcal{V}$ est une base de E, $\det_{\mathcal{B}}(\mathcal{V})$ n'est pas nul, car $\det_{\mathcal{B}}(\mathcal{V}) \times \det_{\mathcal{V}}(\mathcal{B}) = 1$.
$\boxed{\Longleftarrow}$ Par contraposée ; si $\mathcal{V}$ n'est pas une base de E, $\mathcal{V}$ est une famille liée ($\mathcal{V}$ est maximale) et donc $\det_{\mathcal{B}}(\mathcal{V}) = 0$ (image d'une famille liée par une forme n-linéaire alternée). cqfd

Exemples 3.1.

(i) $(1, j)$ est une base du $\mathbb{R}$-espace vectoriel $\mathbb{C}$.
On note $\mathcal{B} = (1, i)$ la base canonique du $\mathbb{R}$-espace vectoriel $\mathbb{C}$ et

$$\det_{\mathcal{B}}(1, j) = \begin{vmatrix} 1 & -\frac{1}{2} \\ 0 & \frac{\sqrt{3}}{2} \end{vmatrix} = \frac{\sqrt{3}}{2} \neq 0$$

(ii) Toute famille de $(n+1)$ polynômes de $\mathbb{K}_n[X]$ et échelonnés en degré est une base de $\mathbb{K}_n[X]$.
On note $\mathcal{B} = (1, X, \dots, X^n)$ la base canonique de $\mathbb{K}_n[X]$; si P_k est un polynôme de degré $k \leqslant n$, on pose $P_k = \sum_{i=0}^{k} a_{i,k} X^i$ avec $a_{k,k} \neq 0$; la famille $(P_0, P_1, \dots, P_n)$ est un système triangulaire de vecteurs et la formule (3.3) donne

$$\det_{\mathcal{B}}(P_0, P_1, \dots, P_n) = \prod_{k=0}^{n} a_{k,k} \neq 0$$

3.3 Orientation d'un $\mathbb{R}$-espace vectoriel

Dans ce paragraphe, E est un espace vectoriel réel ; on ne peut orienter que des espaces sur le corps des réels.

Définition 3.2 (Orientation de deux bases).
Deux bases ordonnées $\mathcal{B}$ et $\mathcal{B}'$ du même $\mathbb{R}$-espace vectoriel E ont *même orientation* si

$$\det_{\mathcal{B}}(\mathcal{B}') > 0$$

Cette propriété définit une relation d'équivalence sur les bases de E :

– *réflexivité* : $\det_{\mathcal{B}}(\mathcal{B}) = 1 > 0$;

– *symétrie* : $\det_{\mathcal{B}}(\mathcal{B}') > 0$ implique $\det_{\mathcal{B}'}(\mathcal{B}) = \left(\det_{\mathcal{B}}(\mathcal{B}')\right)^{-1} > 0$

– *transitivité* : $\det_{\mathcal{B}}(\mathcal{B}') > 0$ et $\det_{\mathcal{B}'}(\mathcal{B}'') > 0$ implique $\det_{\mathcal{B}}(\mathcal{B}'') = \det_{\mathcal{B}}(\mathcal{B}') \times \det_{\mathcal{B}'}(\mathcal{B}'') > 0$

Une base $\mathcal{C}$ de E étant choisie, cette relation d'équivalence définit une partition de l'ensembles des bases de E en deux classes : $\mathcal{C}^+$ et $\mathcal{C}^-$:

– $\mathcal{C}^+$ est l'ensemble des bases $\mathcal{B}$ de E de même orientation que $\mathcal{C}$: $\det_{\mathcal{C}}(\mathcal{B}) > 0$;
– $\mathcal{C}^-$ est l'ensemble des bases $\mathcal{B}$ de E ayant l'orientation opposée à celle de $\mathcal{C}$: $\det_{\mathcal{C}}(\mathcal{B}) < 0$.

Choisir une de ces classes, c'est, par définition, orienter l'espace vectoriel réel E : toutes les bases appartenant à cette classe sont appelées *directes* ou *positivement orientées*, les bases appartenant à l'autre classe sont appelées *indirectes*, *rétrogrades* ou *négativement orientées*.
Dans R^2, on décide que la base canonique $(\mathbf{e}_1, \mathbf{e}_2)$ est directe et que la base $(\mathbf{e}_2, \mathbf{e}_1)$ est rétrograde.
Dans R^3, on décide que la base canonique $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ est directe ; les bases $(\mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_1)$ et $(\mathbf{e}_3, \mathbf{e}_1, \mathbf{e}_2)$ sont directes ; les bases $(\mathbf{e}_1, \mathbf{e}_3, \mathbf{e}_2)$, $(\mathbf{e}_2, \mathbf{e}_1, \mathbf{e}_3)$ et $(-\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ sont rétrogrades (utilisez un calcul de déterminant ou un petit coup de tire-bouchon, technique bien connue de tous les amateurs de physique, ... et de bon vin).

4 Déterminant d'un endomorphisme

Soient E un K-espace vectoriel de dimension finie $n > 0$ muni d'une base $\mathcal{B} = (\mathbf{e}_1, \dots, \mathbf{e}_n)$ et u un endomorphisme de E. À toute forme n-linéaire alternée f sur E, on associe l'application $\varphi_u(f)$ définie par :

$$\varphi_u(f) : (\mathbf{x}_1, \dots, \mathbf{x}_n) \in E^n \mapsto f\big(u(\mathbf{x}_1), \dots, u(\mathbf{x}_n)\big) \in \mathsf{K} \tag{4.1}$$

L'application $\varphi_u(f)$ est une forme n-linéaire, car u est linéaire et f est n-linéaire alternée. D'autre part, l'application $\varphi_u : f \mapsto \varphi_u(f)$ est linéaire ; φ_u est donc un endomorphisme de la droite vectorielle $\Lambda_n^*(E)$, donc une homothétie, ce qui donne le

Lemme 4.1. *Si u est un endomorphisme d'un K-espace vectoriel E de dimension finie n, il existe un unique scalaire λ tel que*

$$\forall f \in \Lambda_n^*(E),\ \forall(\mathbf{x}_1, \dots, \mathbf{x}_n),\ f\big(u(\mathbf{x}_1), \dots, u(\mathbf{x}_n)\big) = \lambda f(\mathbf{x}_1, \dots, \mathbf{x}_n) \tag{4.2}$$

Définition 4.1 (Déterminant d'un endomorphisme).
Ce scalaire est appelé *déterminant* de u et noté $\det u$.

Corollaire (Expression du déterminant d'un endomorphisme).
Si $\mathcal{B}$ est une base (quelconque) de E, le déterminant de u se calcule par

$$\boxed{\det u = \det_{\mathcal{B}}\big(u(\mathbf{e}_1), \dots, u(\mathbf{e}_n)\big)}$$

Démonstration. On utilise $f = \det_{\mathcal{B}}$ et $(\mathbf{x}_1, \dots, \mathbf{x}_n) = (\mathbf{e}_1, \dots, \mathbf{e}_n) = \mathcal{B}$ dans la formule (4.2). cqfd

Remarque. L'expression du déterminant dans la formule précédente, est indépendante de la base $\mathcal{B}$ choisie.

Théorème 4.2 (Propriétés du déterminant d'un endomorphisme).
Si E est un K-espace vectoriel de dimension n, on a

(i) $\det I_E = 1$;
(ii) $\forall u \in \mathcal{L}(E),\ \forall \lambda \in \mathsf{K},\ \det(\lambda u) = \lambda^n \det(u)$;
(iii) $\forall (u, v) \in \mathcal{L}(E)^2,\ \det(u \circ v) = \det u \times \det v$.

Démonstration. Utilisons l'expression du déterminant d'un endomorphisme relativement à une base ;

(i) $\det(I_E) = \det_{\mathcal{B}}(\mathcal{B}) = 1$;
(ii) $\det(\lambda u) = \det_{\mathcal{B}}\big(\lambda u(\mathbf{e}_1), \dots, \lambda u(\mathbf{e}_n)\big) = \lambda^n \det\big(u(\mathbf{e}_1), \dots, u(\mathbf{e}_n)\big) = \lambda^n \det u$;

(iii) la formule (4.2) appliquée à $f = \det_{\mathcal{B}}$ et $(\mathbf{x}_1, \ldots, \mathbf{x}_n) = \big(v(\mathbf{e}_1), \ldots, v(\mathbf{e}_n)\big)$ donne

$$\det(u \circ v) = \det_{\mathcal{B}}\Big(u\big(v(\mathbf{e}_1)\big), \ldots, u\big(v(\mathbf{e}_n)\big)\Big) = \det u \ \det_{\mathcal{B}}\big(v(\mathbf{e}_1), \ldots, v(\mathbf{e}_n)\big) = \det u \ \det v$$

cqfd

Remarque. En général, $\det(u + v)$ est différent de $\det u + \det v$. Donnons un exemple : si $n \geqslant 2$, $\det(I_E + I_E) = \det(2I_E) = 2^n \neq \det I_E + \det I_E = 2$.

Théorème 4.3 (Caractérisation des automorphismes).
Soit u est un endomorphisme d'un K*-espace vectoriel E de dimension finie $n > 0$; u est inversible si, et seulement si, son déterminant n'est pas nul, et, dans ce cas,* $\det u^{-1} = (\det u)^{-1}$.

$$\boxed{u \in \mathcal{GL}(E) \iff \det u \neq 0 \textit{ et, dans ce cas, } \det(u^{-1}) = (\det u)^{-1}}$$

Démonstration. Soit $\mathcal{B} = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ une base de E ; u est inversible si, et seulement si, $n = \operatorname{rg} u = \operatorname{rg}\big(u(\mathbf{e}_1), \ldots, u(\mathbf{e}_n)\big)$, *i.e.* si, et seulement si, $u(\mathcal{B})$ est une base de E, soit si, et seulement si, $0 \neq \det_{\mathcal{B}} u(\mathcal{B}) = \det u$.
Si u est inversible, $u^{-1} \circ u = I_E$; on obtient $1 = \det I_E = \det(u \circ u^{-1}) = \det(u^{-1}) \det u$, ce qui montre que $\det u^{-1} = (\det u)^{-1}$. cqfd

5 Déterminant d'une matrice carrée

Considérons une matrice carrée M d'ordre $n \geqslant 1$ à coefficients dans K ; on note $a_{i,j}$ le terme général de cette matrice et $C_1, \ldots, C_n$ ses vecteurs colonnes ; on écrira indifféremment : $M = [a_{i,j}] = (C_1, \ldots, C_n)$.
Appelons $\mathcal{E} = (E_1, \ldots, E_n)$ la base canonique de $\mathcal{M}_{n,1}(\mathsf{K})$; le scalaire $a_{i,j}$ s'interprète comme la composante du vecteur colonne C_j suivant E_i relativement à la base canonique $\mathcal{C}$, ce qui donne la

Définition 5.1 (Déterminant d'une matrice carrée).
On appelle *déterminant* de la matrice carrée $M = [a_{i,j}]$, et l'on note $\det M$, le déterminant $\det_{\mathcal{E}}(C_1, \ldots, C_n)$ de la famille de ses vecteurs colonnes dans la base canonique de $\mathcal{M}_{n,1}(\mathsf{K})$.

Théorème 5.1 (Déterminant de la transposée d'une matrice).
Si $M = [a_{i,j}] \in \mathcal{M}_n(\mathsf{K})$, on a

$$\det M = \det {}^tM = \sum_{s \in \mathfrak{S}_n} \varepsilon(s) \prod_{j=1}^{n} a_{s(j),j} = \sum_{s \in \mathfrak{S}_n} \varepsilon(s) \prod_{i=1}^{n} a_{i,s(i)}$$

Démonstration. Le vecteur colonne C_j a pour coordonnées $(a_{1,j}, \ldots, a_{n,j})$ dans la base canonique de $\mathcal{M}_{n,1}(\mathsf{K})$. En changeant M en ${}^tM$, on passe de l'une des expressions du déterminant à l'autre. cqfd

5.1 Propriétés

Théorème 5.2 (Déterminant d'une matrice triangulaire).
Le déterminant d'une matrice triangulaire est le produit des éléments de sa diagonale principale.

Démonstration. Si M est une matrice triangulaire supérieure, on utilise l'expression (3.3) du déterminant d'un système triangulaire de vecteurs.
Si M est une matrice triangulaire inférieure, ${}^tM$ est une matrice triangulaire supérieure ; l'égalité $\det M = \det {}^tM$ donne le résultat. cqfd

Théorème 5.3 (Déterminant de la matrice d'un endomorphisme).
Si u est un endomorphisme d'un K*-espace vectoriel E de dimension finie $n \geqslant 1$, le déterminant de u est le déterminant de la matrice de u dans une base quelconque de E.*

$$\boxed{\textit{Pour toute base } \mathcal{B} \textit{ de } E, \quad \det u = \det\bigl(\mathcal{M}\mathrm{at}_{\mathcal{B}}(u)\bigr)}$$

Démonstration. Soit $\mathcal{B} = (\mathbf{e}_1, \dots, \mathbf{e}_n)$ une base de E ; les composantes du vecteur colonne C_j de $\mathcal{M}\mathrm{at}_{\mathcal{B}}(u)$ sont les composantes de $u(\mathbf{e}_j)$ dans $\mathcal{B}$, ce qui donne l'égalité $\det\bigl(u(\mathbf{e}_1), \dots, u(\mathbf{e}_n)\bigr) = \det_{\mathcal{E}}(C_1, \dots, C_n) = \det\bigl(\mathcal{M}\mathrm{at}_{\mathcal{B}}(u)\bigr)$. cqfd

Théorème 5.4. *Si n est un entier positif, on a*

(i) $\det I_n = 1$;
(ii) $\forall M \in \mathcal{M}_n(\mathsf{K}),\ \forall \lambda \in \mathsf{K},\ \det(\lambda M) = \lambda^n \det(M)$;
(iii) $\forall (M, N) \in \mathcal{M}_n(\mathsf{K})^2,\ \det(MN) = \det M \times \det N$;
(iv) $M \in \mathcal{GL}_n(\mathsf{K}) \iff \det M \neq 0$ *et, dans ce cas,* $\det(M^{-1}) = (\det M)^{-1}$.

Démonstration. C'est la traduction matricielles des théorèmes 4.2 et 4.3 consacrés aux déterminants d'endomorphismes. cqfd

Proposition 5.5 (Déterminant de matrices semblables).
Deux matrices semblables ont même déterminant.

Démonstration. Si A et B sont semblables, il existe une matrice inversible P telle que $B = P^{-1}AP$ et donc $\det(P^{-1}AP) = \det(P^{-1}) \det A \det P = \det A$. cqfd

Remarques.
Si M et N sont deux matrices carrées d'ordre $n \geqslant 2$, $\det(M+N)$ est (presque) toujours différent de $\det M + \det N$.
L'application det est une application continue sur l'espace vectoriel normé $\mathcal{M}_n(\mathsf{K})$, car application polynomiale en les composantes des matrices. On en déduit que $\mathcal{GL}_n(\mathsf{K})$ est une partie ouverte de $\mathcal{M}_n(\mathsf{K})$ comme image réciproque de la partie ouverte $\mathsf{K} \setminus \{0\}$ par l'application continue det.

5.2 Règles de calcul du déterminant d'une matrice

Le déterminant d'une matrice carrée est une application n-linéaire alternée des *vecteurs colonnes*, et donc

– si on échange deux colonnes d'une matrice, le déterminant se change en son opposé ;
– le déterminant d'une matrice dépend linéairement de chacun de ses vecteurs colonnes ;
– on ne change pas la valeur du déterminant d'une matrice en ajoutant à l'un de ses vecteurs colonnes, une combinaison linéaire des *autres* vecteurs colonnes ;
– le déterminant d'une matrice est nul si l'un des vecteurs colonnes est nul, ou si l'un des vecteurs colonnes est combinaison linéaire des autres vecteurs colonnes.

Puisque le déterminant d'une matrice est égal au déterminant de sa transposée, on peut remplacer « vecteur colonne » par « vecteur ligne » dans les propriétés précédentes.

5.3 Déterminant d'une matrice triangulaire par blocs

Soient $\mathcal{C} = (\varepsilon_1, \dots, \varepsilon_n)$ est la base canonique de K^n, E' le sous-espace vectoriel de K^n de base $\mathcal{C}' = (\varepsilon_2, \dots, \varepsilon_n)$ et E'' le sous-espace vectoriel de K^n de base $\mathcal{C}'' = (\varepsilon_1, \dots, \varepsilon_{n-1})$. L'application

$$f : (\mathbf{x}_1, \dots, \mathbf{x}_{n-1}) \mapsto \det_{\mathcal{C}}(\varepsilon_1, \mathbf{x}_1, \dots, \mathbf{x}_{n-1})$$

est une forme $(n-1)$-linéaire alternée sur E' ; elle est donc proportionnelle à $\det_{\mathcal{C}'}$, et comme $f(\mathcal{C}') = \det_{\mathcal{C}}(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n) = 1$, on a l'égalité

$$\forall (\mathbf{x}_1, \dots, \mathbf{x}_{n-1}) \in E',\ \det_{\mathcal{C}}(\varepsilon_1, \mathbf{x}_1, \dots, \mathbf{x}_{n-1}) = \det_{\mathcal{C}'}(\mathbf{x}_1, \dots, \mathbf{x}_{n-1})$$

De même, en utilisant la $(n-1)$-forme linéaire alternée sur E''

$$g : (\mathbf{x}_1, \dots, \mathbf{x}_{n-1}) \mapsto \det_{\mathcal{C}}(\mathbf{x}_1, \dots, \mathbf{x}_{n-1}, \varepsilon_n)$$

on obtient

$$\forall (\mathbf{x}_1, \dots, \mathbf{x}_{n-1}) \in E'',\ \det_{\mathcal{C}}(\mathbf{x}_1, \dots, \mathbf{x}_{n-1}, \varepsilon_n) = \det_{\mathcal{C}''}(\mathbf{x}_1, \dots, \mathbf{x}_{n-1})$$

Nous venons de démontrer le

Lemme 5.6. *Si* $A \in \mathcal{M}_{n-1}(\mathsf{K})$*, alors*

$$\begin{vmatrix} 1 & \vdots & 0_{1,n-1} \\ \cdots & \cdots & \cdots \\ 0_{n-1,1} & \vdots & A \end{vmatrix} = \det A = \begin{vmatrix} A & \vdots & 0_{n-1,1} \\ \cdots & \cdots & \cdots \\ 0_{1,n-1} & \vdots & 1 \end{vmatrix}$$

Lemme 5.7. *Si* $A \in \mathcal{M}_{n-1}(\mathsf{K})$ *et* $B \in \mathcal{M}_{n-1}(\mathsf{K})$*, alors*

$$\begin{vmatrix} 1 & \vdots & B \\ \cdots & \cdots & \cdots \\ 0_{n-1,1} & \vdots & A \end{vmatrix} = \det A = \begin{vmatrix} A & \vdots & 0_{1,n-1} \\ \cdots & \cdots & \cdots \\ {}^tB & \vdots & 1 \end{vmatrix}$$

Démonstration. On pose $B = (b_2, \dots, b_n)$; on effectue les transformations $C_j \leftarrow C_j - b_j C_1$ pour $j = 2..n$; ces transformations laissent le déterminant invariant et donnent :

$$\begin{vmatrix} 1 & \vdots & B \\ \cdots & \cdots & \cdots \\ 0_{n-1,1} & \vdots & A \end{vmatrix} = \begin{vmatrix} 1 & \vdots & 0_{1,n-1} \\ \cdots & \cdots & \cdots \\ 0_{n-1,1} & \vdots & A \end{vmatrix} = \det A$$

De même, en effectuant les transformations $C_j \leftarrow C_j - b_j C_n$ pour $j = 1..n-1$, on obtient :

$$\begin{vmatrix} A & \vdots & 0_{n-1,1} \\ \cdots & \cdots & \cdots \\ {}^tB & \vdots & 1 \end{vmatrix} = \begin{vmatrix} A & \vdots & 0_{n-1,1} \\ \cdots & \cdots & \cdots \\ 0_{1,n-1} & \vdots & 1 \end{vmatrix} = \det A$$

cqfd

En utilisant ce lemme et une démonstration par récurrence, on retrouve le

Théorème 5.8 (Déterminant d'une matrice triangulaire).
Le déterminant d'une matrice triangulaire est le produit des éléments de sa diagonale principale.

Lemme 5.9. *Soient* $A \in \mathcal{M}_p(\mathsf{K})$*,* $B \in \mathcal{M}_{p,q}(\mathsf{K})$ *et* $C \in \mathcal{M}_{q,p}(\mathsf{K})$ *; alors*

$$\begin{vmatrix} I_p & \vdots & B \\ \cdots & \cdots & \cdots \\ 0 & \vdots & A \end{vmatrix} = \det A = \begin{vmatrix} A & \vdots & 0 \\ \cdots & \cdots & \cdots \\ C & \vdots & I_q \end{vmatrix}$$

Démonstration. La démonstration s'effectue par récurrence sur p (ou sur q) en utilisant le lemme précédent. cqfd

Théorème 5.10 (Déterminant d'une matrice triangulaire par blocs).
Si $A \in \mathcal{M}_p(\mathsf{K})$, $B \in \mathcal{M}_q(\mathsf{K})$ et $C \in \mathcal{M}_{p,q}(\mathsf{K})$, on a

$$\begin{vmatrix} A & \vdots & C \\ \cdots & \cdots & \cdots \\ 0 & \vdots & B \end{vmatrix} = \det A \times \det B$$

Démonstration. On écrit la matrice dont on veut calculer le déterminant, comme un produit de deux matrices du type précédent, en utilisant le produit matriciel par blocs

$$M = \begin{pmatrix} A & \vdots & C \\ \cdots & \cdots & \cdots \\ 0 & \vdots & B \end{pmatrix} = \begin{pmatrix} I_p & \vdots & 0 \\ \cdots & \cdots & \cdots \\ 0 & \vdots & B \end{pmatrix} \begin{pmatrix} A & \vdots & C \\ \cdots & \cdots & \cdots \\ 0 & \vdots & I_q \end{pmatrix} = NR$$

Le lemme précédent et $\det M = \det N \det R$ donnent le résultat. cqfd

6 Développement d'un déterminant suivant une rangée

6.1 Mise en place

Soient $M = [a_{i,j}] = (C_1, \dots, C_n)$ une matrice carrée d'ordre $n \geqslant 2$, $\mathcal{E} = (E_1, \dots, E_n)$ la base canonique de $\mathcal{M}_{n,1}(\mathsf{K})$; pour $j \in [\![1, n]\!]$, $C_j = \sum_{i=1}^{n} a_{i,j} E_i$. Le déterminant de M se développe suivant son j^{e} argument en utilisant la linéarité et l'on a

$$\det M = \det_{\mathcal{E}}(C_1, \dots, \sum_{k=1}^{n} a_{k,j} E_k, \dots, C_n) = \sum_{k=1}^{n} a_{k,j} \det_{\mathcal{E}}(C_1, \dots, E_k, \dots, C_n) = \sum_{k=1}^{n} a_{k,j} A_{k,j}$$

où les déterminants $A_{k,j}$ s'exprime par

$$A_{k,j} = \begin{vmatrix} a_{1,1} & \cdots & a_{1,j-1} & 0 & a_{1,j+1} & \cdots & a_{1,n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{k-1,1} & \cdots & a_{k-1,j-1} & 0 & a_{k-1,j+1} & \cdots & a_{k-1,n} \\ a_{k,1} & \cdots & a_{k,j-1} & 1 & a_{k,j+1} & \cdots & a_{k,n} \\ a_{k+1,1} & \cdots & a_{k+1,j-1} & 0 & a_{k+1,j+1} & \cdots & a_{k+1,n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n,1} & \cdots & a_{n,j-1} & 0 & a_{n,j+1} & \cdots & a_{n,n} \end{vmatrix}$$

soit

$$A_{k,j} = (-1)^{j-1} \begin{vmatrix} 0 & a_{1,1} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1,n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & a_{k-1,1} & \cdots & a_{k-1,j-1} & a_{k-1,j+1} & \cdots & a_{k-1,n} \\ 1 & a_{k,1} & \cdots & a_{k,j-1} & a_{k,j+1} & \cdots & a_{k,n} \\ 0 & a_{k+1,1} & \cdots & a_{k+1,j-1} & a_{k+1,j+1} & \cdots & a_{k+1,n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & a_{n,1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{n,n} \end{vmatrix}$$

en effectuant $(j-1)$ transpositions de colonnes ce qui donne

$$A_{k,j} = (-1)^{(j-1)+(k-1)} \begin{vmatrix} 1 & a_{k,1} & \cdots & a_{k,j-1} & a_{k,j+1} & \cdots & a_{k,n} \\ 0 & a_{1,1} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1,n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & a_{k-1,1} & \cdots & a_{k-1,j-1} & a_{k-1,j+1} & \cdots & a_{k-1,n} \\ 0 & a_{k+1,1} & \cdots & a_{k+1,j-1} & a_{k+1,j+1} & \cdots & a_{k+1,n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & a_{n,1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{n,n} \end{vmatrix}$$

en effectuant $(k-1)$ transpositions de lignes

$$= (-1)^{j+k} \begin{vmatrix} a_{1,1} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1,n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{k-1,1} & \cdots & a_{k-1,j-1} & a_{k-1,j+1} & \cdots & a_{k-1,n} \\ a_{k+1,1} & \cdots & a_{k+1,j-1} & a_{k+1,j+1} & \cdots & a_{k+1,n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n,1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{n,n} \end{vmatrix} = (-1)^{j+k} \det M_{k,j}$$

Ainsi $A_{k,j} = (-1)^{k+j} \det M_{k,j}$ où $M_{k,j}$ est la matrice déduite de M par suppression de la k^{e} ligne et de la j^{e} colonne.

Définitions 6.1 (Mineur, cofacteur).
Si $M = [a_{i,j}]$ est une K-matrice carrée d'ordre $n \geqslant 2$, on appelle

– *mineur* relatif à l'élément $a_{i,j}$, le déterminant de la matrice carrée $M_{i,j}$ d'ordre $(n-1)$ et déduite de M par la suppression de la i^{e} ligne et de la j^{e} colonne ;

– *cofacteur* de $a_{i,j}$, le scalaire $(-1)^{i+j} \det M_{i,j}$.

Théorème 6.1 (Développement du déterminant suivant une rangée).
Si $M = [a_{i,j}]$ est une K-matrice carrée d'ordre $n \geqslant 2$ et $A_{i,j}$ le cofacteur de $a_{i,j}$, alors, pour tout i et tout j dans $[\![1, n]\!]$, on a

– $\det M = \displaystyle\sum_{k=1}^{n} a_{k,j} A_{k,j}$, *développement du déterminant suivant la j^{e} colonne ;*

– $\det M = \displaystyle\sum_{k=1}^{n} a_{i,k} A_{i,k}$, *développement du déterminant suivant la i^{e} ligne.*

Démonstration. La première formule a été démontrée. Pour la seconde, on utilise l'égalité du déterminant de M et de sa transposée, et le développement de $\det {}^tM$ par rapport à la i^{e} colonne de ${}^tM$, *i.e.* la i^{e} ligne de M. cqfd

6.2 Matrice des cofacteurs

Définition 6.2 (Matrice des cofacteurs).
Si $M = [a_{i,j}]$ est une K-matrice carrée d'ordre $n \geqslant 2$, on appelle *matrice des cofacteurs* ou *comatrice* de M, et on note $\operatorname{Com} M$, la matrice de terme général $A_{i,j}$, le cofacteur relatif à $a_{i,j}$.

$$\boxed{\operatorname{Com} M = [A_{i,j}] = \left[(-1)^{i+j} \det M_{i,j}\right]}$$

Théorème 6.2. *Pour toute K-matrice carrée M d'ordre $n \geqslant 2$, on a*

$$\boxed{M\,{}^t(\operatorname{Com} M) = {}^t(\operatorname{Com} M)\, M = (\det M) I_n}$$

Démonstration. Si, dans M, on remplace la j^e colonne $(a_{k,j})_k$ par $(b_k)_k$, le déterminant de cette nouvelle matrice s'écrit $\sum_{k=1}^n b_k A_{k,j}$: c'est le développement du déterminant par rapport à sa j^e colonne.
Si la nouvelle colonne $(b_k)_{1\leqslant k\leqslant n}$ est la i^e colonne de M, le déterminant est nul si $i \neq j$, et vaut $\det M$ si $i = j$, ce qui s'écrit

$$\forall (i,j) \in [\![1,n]\!]^2, \qquad \sum_{k=1}^n a_{k,i} A_{k,j} = \delta_{j,i} (\det M)$$

ou encore, puisque $({}^t(\operatorname{Com} M)\, M)_{j,i} = \sum_{k=1}^n A_{k,j} a_{k,i}$,

$${}^t(\operatorname{Com} M)\, M = (\det M) I_n$$

Par une méthode analogue, que je vous encourage à rédiger, on montre que

$$\forall (i,j) \in [\![1,n]\!]^2, \qquad \sum_{k=1}^n a_{i,k} A_{j,k} = \delta_{i,j} (\det M)$$

ce qui revient à écrire

$$M\, {}^t(\operatorname{Com} M) = (\det M) I_n$$

cqfd

Remarque. $M \mapsto \operatorname{Com} M$ est une application continue de $\mathcal{M}_n(\mathsf{K})$dans $\mathcal{M}_n(\mathsf{K})$, car les composantes de $\operatorname{Com} M$ sont polynomiales en les coefficients de M.

Corollaire (Inverse d'une matrice carrée).
Si M est une matrice carrée inversible d'ordre $n \geqslant 2$, on a

$$\boxed{M \in \mathcal{GL}_n(\mathsf{K}) \implies M^{-1} = \frac{1}{\det M}\, {}^t(\operatorname{Com} M)}$$

Remarques. Exceptés les cas $n = 2$ et $n = 3$, cette formule ne peut servir au calcul numérique de l'inverse car elle comporte trop d'opérations.

$$M = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \in \mathcal{GL}_2(\mathsf{K}) \implies M^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -c \\ -b & a \end{pmatrix}$$

Par contre, elle est utile dans des questions théoriques ; par exemple, $M \mapsto M^{-1}$ est une bijection continue de $\mathcal{GL}_n(\mathsf{K})$ (c'est même une involution) car produit de deux applications continues.

7 Exemple de calcul de déterminant

7.1 Déterminant de Vandermonde

Soient $n > 1$ et $(a_1, \dots, a_n) \in \mathsf{K}^n$; alors

$$\boxed{V(a_1,\dots,a_n) = \begin{vmatrix} 1 & a_1 & \cdots & a_1^{n-1} \\ 1 & a_2 & \cdots & a_2^{n-1} \\ \cdots & \cdots & \cdots & \cdots \\ 1 & a_n & \cdots & a_n^{n-1} \end{vmatrix} = \prod_{1\leqslant i<j\leqslant n} (a_j - a_i)}$$

Démonstration. S'il existe $i < j$ avec $a_i = a_j$, le déterminant possède deux lignes identiques ; il est donc nul et la formule est vérifiée. On envisage maintenant le cas où les a_i sont distincts deux à deux et on effectue une démonstration par récurrence sur n.
Pour $n = 2$, $V(a_1, a_2) = a_2 - a_1$.
En développant $P(x) = V(a_1, \dots, a_n, x)$ par rapport à la dernière ligne, on remarque que P est un polynôme de degré n et de coefficient dominant $V(a_1, \dots, a_n)$; on remarque aussi que $P(a_i) = 0$ (si $x = a_i$, le déterminant possède deux lignes identiques) et P admet n racines distinctes. Ainsi,

$$P(X) = V(a_1, \dots, a_n) \prod_{k=1}^{n} (X - a_k)$$

d'où le résultat en remplaçant X par a_{n+1}.
Le passage du rang n au rang $n+1$ se démontre aussi en manipulant les colonnes de la façon suivante :

$$C_{k+1} \leftarrow C_{k+1} - a_1 C_k \quad \text{pour } k \text{ variant de } n \text{ à } 1$$

On a donc

$$V(a_1, \dots, a_n, a_{n+1}) =$$
$$\begin{vmatrix} 1 & a_1 & \cdots & a_1^{n-1} & a_1^n \\ 1 & a_2 & \cdots & a_2^{n-1} & a_2^n \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & a_{n+1} & \cdots & a_{n+1}^{n-1} & a_{n+1}^n \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & a_2 - a_1 & \cdots & a_2^{n-1} - a_1 a_2^{n-2} & a_2^n - a_1 a_2^{n-1} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & a_{n+1} - a_1 & \cdots & a_{n+1}^{n-1} - a_1 a_{n+1}^{n-2} & a_{n+1}^n - a_1 a_{n+1}^{n-1} \end{vmatrix} =$$
$$\prod_{k=2}^{n} (a_k - a_1) \begin{vmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & a_2 & \cdots & a_2^{n-1} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & 1 & a_{n+1} & \cdots & a_{n+1}^{n-1} \end{vmatrix} = \prod_{k=2}^{n} (a_k - a_1)\, V(a_2, \dots, a_{n+1}) = \prod_{1 \leqslant i < j \leqslant n+1} (a_j - a_i)$$

cqfd

7.2 Déterminant circulant (droit)

Soient $n > 1$ et $(a_1, \dots, a_n) \in \mathsf{K}^n$; alors

$$\boxed{C(a_1, \dots, a_n) = \begin{vmatrix} a_1 & a_2 & \cdots & a_n \\ a_n & a_1 & \cdots & a_{n-1} \\ \cdots & \cdots & \cdots & \cdots \\ a_2 & a_3 & \cdots & a_1 \end{vmatrix} = \prod_{k=1}^{n} \Big(\sum_{s=1}^{n} a_s \zeta^{k(s-1)} \Big) \text{ où } \zeta = \exp(\frac{2i\pi}{n})}$$

7.3 Déterminant de Cauchy

Soient $n > 1$, $(a_1, \dots, a_n) \in \mathsf{K}^n$ et $(b_1, \dots, b_n) \in \mathsf{K}^n$ tels que, pour tout i et j de $[\![1, n]\!]$, on ait $a_i + b_j \neq 0$; alors

$$\boxed{\begin{vmatrix} (a_1 + b_1)^{-1} & (a_1 + b_2)^{-1} & \cdots & (a_1 + b_n)^{-1} \\ (a_1 + b_1)^{-1} & (a_1 + b_2)^{-1} & \cdots & (a_1 + b_n)^{-1} \\ \cdots & \cdots & \cdots & \cdots \\ (a_n + b_1)^{-1} & (a_n + b_2)^{-1} & \cdots & (a_n + b_n)^{-1} \end{vmatrix} = \frac{\displaystyle\prod_{1 \leqslant i < j \leqslant n} (a_j - a_i)(b_j - b_i)}{\displaystyle\prod_{(i,j) \in [\![1,n]\!]^2} (a_i + b_j)}}$$

Application au déterminant de la matrice de Hilbert $H_n = \left[(i+j)^{-1}\right]$

$$\boxed{\det H_n = \frac{\left[1! \times 2! \times \cdots \times (n-1)!\right]^3 n!}{(n+1)! \times (n+2)! \times \cdots \times (2n)!}}$$

8 Résolution des systèmes linéaires

8.1 Quelques notations

Soient n et p deux entiers au moins égaux à 1, et $\mathcal{E} = (E_1, \dots, E_n)$ la base canonique de $\mathcal{M}_{n,1}(\mathsf{K})$. À toute matrice $M = [a_{i,j}] = (C_1, \dots, C_p) \in \mathcal{M}_{n,p}(\mathsf{K})$, on associe l'application linéaire u de K^p vers K^n telle que :

$$\mathcal{M}\mathrm{at}_{\mathcal{B},\mathcal{C}}(u) = M$$

Pour $j \in [\![1, p]\!]$, on note $\mathbf{c}_j = (a_{1,j}, \dots, a_{n,j}) = {}^tC_j \in \mathsf{K}^n$, $\mathbf{b} = (b_1, \dots, b_n) = {}^tB \in \mathsf{K}^n$ et $\mathbf{x} = (x_1, \dots, x_p) = {}^tX \in \mathsf{K}^p$ divers vecteurs.
Tout système linéaire $(\mathcal{L})$ de n équations à p inconnues peut s'écrire de manière

– analytique : $\begin{cases} a_{1,1}x_1 + \cdots + a_{1,p}x_p & = & b_1 \\ a_{2,1}x_1 + \cdots + a_{2,p}x_p & = & b_2 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{n,1}x_1 + \cdots + a_{n,p}x_p & = & b_n \end{cases}$

– vectorielle : $\sum_{j=1}^{p} x_j \mathbf{c}_j = \mathbf{b}$ ou encore $\sum_{j=1}^{p} x_j C_j = B$

– matricielle : $MX = B$;

– fonctionnelle : $u(\mathbf{x}) = \mathbf{b}$.

Les rangs de M, de u, de la famille de vecteurs $(\mathbf{b}_1, \dots, \mathbf{b}_p)$ ou de la famille $(C_1, \dots, C_p)$ sont égaux ; cet entier est noté r et appelé *rang du système linéaire* $(\mathcal{L})$.

8.2 Cas des systèmes de Cramer

Définition 8.1 (Systèmes de Cramer).
Un système linéaire $(\mathcal{L})$ de n équations à n inconnues est appelé *système de Cramer* si, et seulement si, l'une des conditions équivalentes suivantes est vérifiée

(i) $n = p = r$;
(ii) $M \in \mathcal{GL}_n(\mathsf{K})$;
(iii) u est inversible.

Théorème 8.1 (Formules de Cramer).
Avec les notations précédentes, l'unique solution d'un système de Cramer est donnée par

$$\forall j \in [\![1, n]\!],\ x_j = \frac{\det M_j}{\det M}$$

où M_j est la matrice déduite de M en remplaçant le vecteur colonne C_j par le second membre B.

Démonstration. $X = {}^t(x_1, \dots, x_p)$ est solution de $(\mathcal{L})$ si, et seulement si, $\sum_{j=1}^{p} x_j C_j = B$. On a donc

$$\begin{aligned}
\det M_j &= \det(C_1, \dots, C_{j-1}, B, C_{j+1}, \dots, C_n) \\
&= \det(C_1, \dots, C_{j-1}, \sum_{k=1}^{p} x_k C_k, C_{j+1}, \dots, C_n) \\
&= \sum_{k=1}^{p} x_k \det(C_1, \dots, C_{j-1}, C_k, C_{j+1}, \dots, C_n) \qquad (\det \text{ est } n\text{-linéaire}) \\
&= x_j \det(C_1, \dots, C_{j-1}, C_j, C_{j+1}, \dots, C_n) \qquad (\det \text{ est alterné}) \\
&= x_j \det M
\end{aligned}$$

Puisque $\det M \neq 0$, on a le résultat annoncé. cqfd

8.3 Cas des systèmes homogènes

Un système linéaire est dit *homogène* si le second membre **b** est nul ; il admet toujours au moins une solution : la solution nulle.
L'ensemble des solutions d'un système linéaire homogène est un sous-espace vectoriel de K^p de dimension $p - r$, c'est $\ker u$; il suffit d'en exhiber une base.

A) Cas particulier : $p = n + 1$, $r = n$

Les solutions d'un système linéaire homogène de n équations à $n+1$ inconnues et de rang maximum n constituent une droite vectorielle ; il suffit donc d'exhiber une solution non nulle.
Appelons M_j la matrice carrée d'ordre n déduite de M par suppression de la j^{e} colonne et $\tilde{M}$ la matrice carrée d'ordre $n+1$ obtenue en ajoutant à M la ligne $(b_1, \dots, b_{n+1})$; $\det M_j$ est le mineur de $\tilde{M}$ relatif à b_j et le développement de $\tilde{M}$ suivant sa dernière ligne donne

$$\det \tilde{M} = \sum_{j=1}^{n+1} b_j (-1)^{n+1+j} \det M_j$$

Au lieu de compléter M par la ligne $(b_j)_j$, complétons-la par sa i^{e} ligne $(a_{i,j})_j$; dans ce cas, $\tilde{M}$ possède deux lignes identiques et l'égalité précédente devient

$$\forall i \in [\![1, n]\!],\ 0 = \sum_{j=1}^{n+1} a_{i,j} (-1)^{n+1+j} \det M_j = (-1)^{n+1} \sum_{j=1}^{n+1} a_{i,j} (-1)^j \det M_j$$

ce qui signifie que $\big((-1)^j \det M_j\big)_j$ est une solution du système, solution non nulle puisque M est de rang n.Cette solution constitue donc une base de la droite vectorielle des solutions.

B) Intersection de deux plans de K^3

Considérons le système homogène

$$(\mathcal{H}) : \begin{cases} (\mathcal{P}_1) : u_1 x + v_1 y + w_1 z = 0 \\ (\mathcal{P}_2) : u_2 x + v_2 y + w_2 z = 0 \end{cases}$$

avec $(u_i, v_i, w_i) \neq \mathbf{0}$ pour $i = 1$ et $i = 2$. On pose

$$d_1 = \begin{vmatrix} v_1 & w_1 \\ v_2 & w_2 \end{vmatrix},\ d_2 = \begin{vmatrix} w_1 & u_1 \\ w_2 & u_2 \end{vmatrix} = -\begin{vmatrix} u_1 & w_1 \\ u_2 & w_2 \end{vmatrix},\ d_3 = \begin{vmatrix} u_1 & v_1 \\ u_2 & v_2 \end{vmatrix},$$

$$\Delta_1 = \begin{vmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_1 & v_1 & w_1 \end{vmatrix},\ \Delta_2 = \begin{vmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_2 & v_2 & w_2 \end{vmatrix}$$

En développant Δ_1 et Δ_2 suivant leur dernière ligne, on trouve $\Delta_1 = 0 = u_1 d_1 + v_1 d_2 + w_1 d_3$ et $\Delta_2 = 0 = u_2 d_1 + v_2 d_2 + w_2 d_3$.
Si les deux plans $(\mathcal{P}_1)$ et $(\mathcal{P}_2)$ sont identiques, alors $(d_1, d_2, d_3) = \mathbf{0}$. Sinon, (d_1, d_2, d_2) dirige la droite $\mathcal{P}_1 \cap \mathcal{P}_2$.
Dans R^3 euclidien, le vecteur (d_1, d_2, d_3) s'interprète comme le produit vectoriel des vecteurs (u_1, v_1, w_1) et (u_2, v_2, w_2) normaux respectivement à $(\mathcal{P}_1)$ et $(\mathcal{P}_2)$.

9 Déterminant et rang

Le rang d'une matrice $M = [a_{i,j}] = (C_1, \dots, C_p) = {}^t(L_1, \dots, L_n) \in \mathcal{M}_{n,p}(\mathsf{K})$ est le rang de ses vecteurs colonnes $(C_j)_{j \in [\![1,p]\!]}$ ou celui de ses vecteurs lignes $(L_i)_{i \in [\![1,n]\!]}$ car le rang d'une matrice est égal à celui de sa transposée ; on a donc :

$$\operatorname{rg} M \leqslant \inf(n, p)$$

Définition 9.1 (Matrice extraite).
Si I est une partie non vide de $[\![1, n]\!]$ et J une partie non vide de $[\![1, p]\!]$, on appelle *matrice extraite de M associée* à I et J, la matrice $R = [a_{i,j}]$ où $i \in I$ et $j \in J$.

Lemme 9.1. *Le rang d'une matrice extraite de M est inférieur ou égal au rang de M.*

Démonstration. Soit R une matrice extraite de M associée à I et J. Considérons la matrice Q extraite de M et associée à $[\![1, n]\!]$ et J ; les vecteurs colonnes $(C_j)_{j \in J}$ de Q constituent une sous-famille des vecteurs colonnes de M, donc $\operatorname{rg} Q \leqslant \operatorname{rg} M$.
De même, les vecteurs lignes $(L_i)_{i \in I}$ de R constituent une sous-famille des vecteurs lignes de Q, d'où $\operatorname{rg} R \leqslant \operatorname{rg} Q$, et le résultat. cqfd

Théorème 9.2 (Caractérisation du rang d'une matrice).
Le rang d'une matrice non nulle est l'ordre maximal des matrices carrés inversibles extraites.

Démonstration. Soit $M \in \mathcal{M}_{n,p}(\mathsf{K})$ une matrice non nulle.
L'ensemble des ordres des matrices carrées inversibles extraites de M n'est pas vide (il contient 1 puisque M n'est pas la matrice nulle), et est majoré par $\operatorname{rg} M$ d'après le lemme. On note r son plus grand élément ; on a donc $1 \leqslant r \leqslant \operatorname{rg} M$.
De la famille $(C_1, \dots, C_p)$ des vecteurs colonnes de M, on peut extraire une sous-famille libre $(C_j)_{j \in J}$ de cardinal $\operatorname{rg} M$; on note Q la matrice extraite de M associée à $[\![1, n]\!]$ et J, et $rgM = \operatorname{rg} Q$. Des vecteurs lignes $(L_1, \dots, L_n)$ de Q, on peut encore extraire une sous-famille libre $(L_i)_{i \in I}$ de cardinal $\operatorname{rg} Q$; on note R la matrice extraite de M et associée à I et J et $\operatorname{rg} R = \operatorname{rg} Q$.
R est une matrice carrée de rang maximum ($\#I = \#J = \operatorname{rg} R = \operatorname{rg} Q = \operatorname{rg} M$), donc une matrice inversible. En conséquence, $r \geqslant \operatorname{rg} M$.
Finalement r est égal au rang de M. cqfd

Corollaire (Caractérisation des familles libres).
Si $\mathcal{F} = (\mathbf{c}_1, \dots, \mathbf{c}_p)$ est une famille de p vecteurs d'un K-espace vectoriel E de dimension finie n, si $M = \mathcal{M}\mathrm{at}_{\mathcal{B}}(\mathcal{F}) \in \mathcal{M}_{n,p}(\mathsf{K})$ est la matrice des composantes de $\mathcal{F}$ relatives à une base $\mathcal{B}$ de E, $\mathcal{F}$ est une famille libre si, et seulement si, il existe une matrice carrée d'ordre p, extraite de M et de déterminant non nul.

CHAPITRE

3

Réduction des endomorphismes et des matrices

Sommaire

1 Sous-espaces vectoriels stables

1.1 Généralités

Définition 1.1 (Sous-espace vectoriel stable).
Soient E un K-espace vectoriel et u un endomorphisme de E ; un sous-espace vectoriel F de E est dit *stable pour* u si, et seulement si, $u(F)$ est inclus dans F.

$$\boxed{F \text{ est stable pour } u \iff \forall \mathbf{x},\ \mathbf{x} \in F \implies u(\mathbf{x}) \in F}$$

Proposition 1.1 (Stabilité et famille génératrice).
F est un sous-espace vectoriel stable pour u si, et seulement si, l'image par u d'une famille génératrice de F est contenue dans F.

Démonstration. La condition nécessaire est évidente : les éléments d'une famille génératrice de F sont des éléments particuliers de F.
Si $(\mathbf{f}_1, \dots, \mathbf{f}_p)$ engendre F, alors $(u(\mathbf{f}_1), \dots, u(\mathbf{f}_p))$ engendre $u(F)$, grâce à la linéarité de u :

$$\mathbf{x} = \sum_{j=1}^{p} \lambda_j \mathbf{f}_j \in F \implies u(\mathbf{x}) = \sum_{j=1}^{p} \lambda_j u(\mathbf{f}_j) \in u(F)$$

ce qui donne la condition suffisante cqfd

Proposition 1.2 (Application induite sur un sous-espace vectoriel stable).
Si u est un endomorphisme de E et F un sous-espace vectoriel stable pour u, l'application $v : \mathbf{x} \in F \mapsto u(\mathbf{x})$ induite par u sur F, est un endomorphisme de F.

Démonstration. L'application v est à valeurs dans F et est linéaire puisque u l'est. cqfd

Théorème 1.3 (Stabilité de l'image et du noyau).
Si u et v sont deux endomorphismes de E qui commutent, $\operatorname{Im} u$ *et* $\ker u$ *sont stables par v.*

Démonstration. Soit $\mathbf{y} = u(\mathbf{x})$ un vecteur quelconque de $\operatorname{Im} u$, son image par v, $v(\mathbf{y}) = v\big(u(\mathbf{x})\big) = u\big(v(\mathbf{x})\big)$ reste dans $\operatorname{Im} u$.
Si $\mathbf{x}$ est élément de $\ker u$, alors $u\big(v(\mathbf{x})\big) = v\big(u(\mathbf{x})\big) = v(\mathbf{0}) = \mathbf{0}$, et $v(\mathbf{x})$ reste dans $\ker u$. cqfd

1.2 Cas de la dimension finie

E est un K-espace vectoriel de dimension finie n ; rappelons que si F est un sous-espace vectoriel de dimension p, toute base de E dont les p premiers vecteurs constituent une base de F, est appelée *base de E adaptée* à F. Le théorème de la base incomplète montre l'existence de telle base.

Théorème 1.4 (Caractérisation des sous-espaces vectoriels stables).
Soient F est un sous-espace vectoriel de E, $\mathcal{B}$ une base de E adaptée à F et u un endomorphisme de E ; alors F est stable pour u si, et seulement si, la matrice de u relativement à $\mathcal{B}$ est triangulaire supérieure par blocs, soit

$$\mathcal{M}\mathrm{at}_{\mathcal{B}}(u) = \begin{pmatrix} A & \vdots & B \\ \cdots & \cdots & \cdots \\ 0 & \vdots & C \end{pmatrix}$$

Démonstration. Soit $\mathcal{B} = (\mathbf{e}_i, \dots, \mathbf{e}_p, \mathbf{e}_{p+1}, \dots, \mathbf{e}_n)$ une base de E telle que les p premiers vecteurs $(\mathbf{e}_1, \dots, \mathbf{e}_p)$ constituent une base de F ; F est stable par u si, et seulement si, pour tout $j \in [\![1, p]\!]$, $u(\mathbf{e}_j)$ est dans F. Or, $u(\mathbf{e}_j) = \sum_{i=1}^{n} a_{i,j} \mathbf{e}_i$ appartient à F si, et seulement si, $a_{i,j} = 0$ pour $i > p$, *i.e.* si, et seulement si, les p premières colonnes de la matrice de u ont des 0 à partir de la ligne $(p+1)$. cqfd

1.3 Généralisation

Si $E = \bigoplus_{i=1}^{p} F_i$, et si $\mathcal{B}_j$ est une base de F_j, $\mathcal{B} = \bigcup_{j=1}^{p} \mathcal{B}_j$ est une base de E *adaptée* à cette décomposition en somme directe.

Théorème 1.5. *Soient E est un K-espace vectoriel de dimension finie, $F_1, \dots, F_p$ des sous-espaces vectoriels dont E est la somme directe, $\mathcal{B}$ une base de E adaptée à cette décomposition et u un endomorphisme de E ; alors u stabilise les sous-espaces F_j si, et seulement si, la matrice de u relativement à $\mathcal{B}$ est diagonale par blocs,* i.e.

$$\mathcal{M}\mathrm{at}_{\mathcal{B}}(u) = \begin{pmatrix} A_1 & \vdots & 0 & \vdots & \cdots & \vdots & 0 \\ 0 & \vdots & A_2 & \vdots & \cdots & \vdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \vdots & 0 & \vdots & \cdots & \vdots & A_p \end{pmatrix}$$

En appelant u_j l'endomorphisme induit par u sur le sous-espace stable F_j, A_j est la matrice de u_j relativement à la base $\mathcal{B}_|$ et

$$\det u = \det A_1 \times \cdots \times \det A_p = \det u_1 \times \cdots \times \det u_p$$

Si $\mathcal{D}$ est une droite vectorielle, tout endomorphisme de $\mathcal{D}$ est une homothétie ; ainsi, si $\mathcal{D}$ est une droite vectorielle de E, stable pour u, l'endomorphisme de $\mathcal{D}$ induit par u est une homothétie de $\mathcal{D}$ ce qui donne le

Théorème 1.6 (Endomorphisme de matrice diagonale).
Soient E un K-espace vectoriel de dimension finie n, $\mathcal{B} = (\mathbf{e}_1, \dots, \mathbf{e}_n)$ une base de E et u un endomorphisme de E ; alors, la matrice de u dans $\mathcal{B}$ est diagonale si, et seulement si, pour tout $j \in [\![1, n]\!]$, la restriction u_j de u à $\mathsf{K}\mathbf{e}_j$ est une homothétie.

1.4 Drapeau

Définition 1.2 (Drapeau).
Si E est un K-espace vectoriel de dimension finie $n \geqslant 1$, un *drapeau* est une suite $(E_1, \dots, E_n)$ de sous-espaces vectoriels de E, croissante pour l'inclusion et telle que $\dim E_k = k$.

Soit $\mathcal{B} = (\mathbf{e}_1, \dots, \mathbf{e}_n)$ est une base de E ; on définit $E_k = \oplus_{i=1}^{k} \mathsf{K}\mathbf{e}_i$ le sous-espace vectoriel engendré par les k premiers vecteurs de $\mathcal{B}$; la suite $(E_1, \dots, E_n)$ est un drapeau de E, c'est le *drapeau associé* à $\mathcal{B}$.
Réciproquement, à tout drapeau $(E_1, \dots, E_n)$ de E, on associe une base $\mathcal{B} = (\mathbf{e}_1, \dots, \mathbf{e}_n)$ de E telle que $E_k = \oplus_{i=1}^{k} \mathsf{K}\mathbf{e}_i$; c'est la *base adaptée* au drapeau $(E_1, \dots, E_n)$.

Théorème 1.7 (Endomorphisme de matrice triangulaire).
Soient E un K-espace vectoriel de dimension finie n, $(E_1, \dots, E_n)$ un drapeau de E, $\mathcal{B}$ une base adaptée à ce drapeau et u un endomorphisme de E ; alors, u stabilise les E_k si, et seulement si, la matrice de u relativement à $\mathcal{B}$ est une matrice triangulaire supérieure.

Démonstration. Pour tout $k \in [\![1, n]\!]$, l'image de E_k est contenue dans E_k si, et seulement si, $u(\mathbf{e}_k)$ est dans E_k. Si $(a_{i,j})_i$ sont les composantes de $u(\mathbf{e}_j)$ dans $\mathcal{B}$, $u(\mathbf{e}_k)$ appartient à E_k si, et seulement si, $a_{i,k}$ est nul pour $i > k$. cqfd

2 Polynômes d'un endomorphisme

2.1 Puissance d'un endomorphisme

Définition 2.1 (Puissance d'un endomorphisme).

Si u est un endomorphisme d’un K-espace vectoriel E, on définit, par récurrence sur $k \in \mathsf{N}$, les puissances k^{e} u^k de u en posant

$$u^0 = I_E \text{ et } \forall k \in \mathsf{N},\ u^{k+1} = u^k \circ u$$

Proposition 2.1. *Pour tout entier naturel p et q, on a $u^p \circ u^q = u^q \circ u^p = u^{p+q}$.*

Démonstration. Montrons que $u^p \circ u^q = u^{p+q}$ en effectuant une récurrence sur q. La formule est vraie pour $q = 0$ et

$$u^p \circ u^{q+1} = u^p \circ (u^q \circ u) = (u^p \circ u^q) \circ u = u^{p+q} \circ u = u^{p+q+1}$$

La seconde égalité se montre en échangeant le rôle de p et q. cqfd

Remarque. Les endomorphismes u^p et u^q commutent pour tout entier naturel p et q.

Proposition 2.2 (Noyaux itérés).
Si u est un endomorphisme d’un K-espace vectoriel E,

(i) *la suite $(\ker u^p)_p$ est croissante pour l’inclusion ;*

(ii) *si l’ensemble $\{p \in \mathsf{N} \ /\ \ker u^p = \ker u^{p+1}\}$ n’est pas vide, il possède un plus petit élément noté i_u, appelé* indice *de u ; dans ces conditions*

$$\forall q \in \mathsf{N},\ \ker u^{i_u+q} = \ker u^{i_u} \text{ et } \{\mathbf{0}\} \subsetneq \ker u \subsetneq \cdots \subsetneq \ker u^{i_u}$$

(iii) *si E est de dimension finie $n \geqslant 1$, l’existence de i_u est assurée et $i_u \leqslant n$.*

Démonstration.

(i) L’implication $u^p(\mathbf{x}) = \mathbf{0} \implies \mathbf{0} = u\big(u^p(\mathbf{x})\big) = u^{p+1}(\mathbf{x})$ vraie pour tout $p \in \mathsf{N}$ et tout $\mathbf{x} \in E$, montre l’inclusion $\ker u^p \subset \ker u^{p+1}$.

(ii) Remarquons que $\ker u^p = \ker u^{p+1} \implies \ker u^{p+1} = \ker u^{p+2}$; il suffit de démontrer l’inclusion $\ker u^{p+2} \subset \ker u^{p+1}$. Or, $\mathbf{x} \in \ker u^{p+2}$ s’écrit $\mathbf{0} = u^{p+2}(\mathbf{x}) = u^{p+1}\big(u(\mathbf{x})\big)$, soit $u(\mathbf{x}) \in \ker u^{p+1}$; comme $\ker u^{p+1} = \ker u^p$, on obtient $\mathbf{0} = u^p\big(u(\mathbf{x})\big) = u^{p+1}(\mathbf{x})$ et $\mathbf{x} \in \ker u^{p+1}$. Par récurrence sur q, on montre que $\ker u^p = \ker u^{p+q}$.
Si l’ensemble $\{p \in \mathsf{N} \ /\ \ker u^p = \ker u^{p+1}\}$ n’est pas vide, il contient un plus petit élément noté i_u. L’égalité $\ker u^{i_u} = \ker u^{i_u+1}$ montre que pour tout $q \in \mathsf{N}$, $\ker u^{i_u} = \ker u^{i_u+q}$, et puisque i_u est le premier entier qui vérifie l’égalité $\ker u^p = \ker u^{p+1}$, la suite $(\ker u^p)_{p\in[\![0,i_u]\!]}$ est strictement croissante pour l’inclusion.

(iii) Si la suite $(\ker u^p)_p$ est strictement croissante pour l’inclusion, la suite $(\dim \ker u^p)_p$ des dimensions est strictement croissante ; comme ces dimensions sont majorées par n, la dimension de E, il y a contradiction ; l’existence de i_u est donc assurée et $i_u \leqslant n$.

cqfd

Remarques.
Un endomorphisme nilpotent u possède un indice, il est donné par

$$i_u = \inf\{p \in \mathsf{N} \ /\ \ker u^p = E\} = \inf\{p \in \mathsf{N} \ /\ u^p = 0\}$$

La dérivation D de $\mathsf{K}[X]$ n’a pas d’indice ; la suite des noyaux itérés $(\ker \mathrm{D}^p)_p = (\mathsf{K}_{p-1}[X])_p$ n’est pas stationnaire.

2.2 Polynômes d’un endomorphisme ou d’une matrice

Définition 2.2 (Polynôme d’un endomorphisme).
Si u est un endomorphisme d’un K-espace vectoriel E et $P = a_0 + a_1 X + \cdots + a_p X^p$ un polynôme à coefficients dans K, on pose

$$P(u) = a_0 I_E + a_1 u + \cdots + a_p u^p$$

$P(u)$ est un *polynôme de l’endomorphisme* u.

Définition 2.3 (Polynôme d'une matrice).
De même, si M est une matrice carrée d'ordre n, on pose

$$P(M) = a_0 I_n + a_1 M + \cdots + a_p M^p$$

$P(M)$ est un *polynôme de la matrice M*.

Remarque. Rappelons qu'un polynôme P peut s'écrire $P = \sum_i a_i X^i$ où la suite $(a_i)_i$ est nulle à partir d'un certain rang ; dans ces conditions, $P(u) = \sum_i a_i u^i$ et $P(M) = \sum_i a_i M^i$.

Proposition 2.3 (Règles de calcul, cas des endomorphismes).
L'application $\varphi_u : P \in \mathsf{K}[X] \mapsto P(u) \in \mathcal{L}(E)$ est un morphisme d'algèbre ; en particulier

$$\forall u \in \mathcal{L}(E),\ \forall (P, Q) \in \mathsf{K}[X]^2, \quad \big(PQ\big)(u) = P(u) \circ Q(u)$$

Démonstration.
$\varphi_u(\lambda P + \mu Q) = \sum_i (\lambda a_i + \mu b_i) u^i = \lambda \sum_i a_i u^i + \mu \sum_i b_i u^i = \lambda \varphi_u(P) + \mu \varphi_u(Q)$
$\varphi_u(1) = I_E$
Quant à l'égalité $\varphi_u(PQ) = \varphi_u(P) \circ \varphi_u(Q)$, on commence par la démontrer pour $Q = X^k$:

$$\varphi_u(PX^k) = \sum_i a_i u^{i+k} = \big(\sum_i a_i u^i\big) \circ u^k = \varphi_u(P) \circ \varphi_u(X^k)$$

puis dans le cas général $Q = \sum_k b_k X^k$ en utilisant la linéarité de φ_u :

$$\begin{aligned}\varphi_u(PQ) &= \varphi_u\Big(\sum_k b_k P X^k\Big) = \sum_k b_k \varphi_u(PX^k) = \sum_k b_k \varphi_u(P) \circ \varphi_u(X^k)\\ &= \varphi_u(P) \circ \big(\sum_k b_k \varphi_u(X^k)\big) = \varphi_u(P) \circ \varphi_u(Q)\end{aligned}$$

cqfd

Proposition 2.4 (Règles de calcul, cas des matrices).
L'application $\varphi_M : P \in \mathsf{K}[X] \mapsto P(M) \in \mathcal{M}_n(\mathsf{K})$ est un morphisme d'algèbre ; en particulier

$$\forall M \in \mathcal{M}_n(\mathsf{K}),\ \forall (P, Q) \in \mathsf{K}[X]^2, \quad \big(PQ\big)(M) = P(M)Q(M)$$

Démonstration. Même démonstration que précédemment. cqfd

Puisque u^p commute avec u^q, les endomorphismes u et $P(u)$ commutent, ainsi que $P(u)$ et $Q(u)$, ce qui donne la

Proposition 2.5 (u-stabilité de $\ker P(u)$ et de $\operatorname{Im} P(u)$).
Pour tout polynôme P à coefficients dans K et tout endomorphisme u de E, $\operatorname{Im} P(u)$ et $\ker P(u)$ sont stables par u.

3 Valeurs propres, vecteurs propres d'un endomorphisme

E est un K-espace vectoriel de dimension finie ou non et u un endomorphisme de E.

3.1 Droite stable par un endomorphisme

Si u est un endomorphisme de E et $\mathcal{D}$ une droite (vectorielle) stable pour u, $u|_{\mathcal{D}}$ est un endomorphisme de $\mathcal{D}$, donc une homothétie de $\mathcal{D}$; il existe donc un unique scalaire λ, dépendant de $\mathcal{D}$, tel que

$$u|_{\mathcal{D}} = \lambda I_{\mathcal{D}} \quad \textit{i.e.} \quad \forall \mathbf{x} \in \mathcal{D},\ u(\mathbf{x}) = \lambda \mathbf{x}$$

3.2 Vecteur propre

Définitions 3.1 (Vecteur propre, valeur propre associée).
Si u est un endomorphisme de E, tout vecteur $\mathbf{e}$ non nul qui dirige une droite vectorielle stable pour u est appelé un *vecteur propre* de u.
Le rapport de l'homothétie induite par u sur $\mathcal{D} = \mathsf{K}\mathbf{e}$ est appelé *valeur propre* associée au vecteur propre $\mathbf{e}$.

$$\boxed{\mathbf{e} \neq \mathbf{0} \text{ est un vecteur propre pour } u \iff \mathcal{D} = \mathsf{K}\mathbf{e} \text{ est stable par } u \iff \exists! \lambda \in \mathsf{K},\ u(\mathbf{e}) = \lambda\mathbf{e}}$$

Remarques.
La valeur propre associée à un vecteur propre est unique.
Un vecteur non nul $\mathbf{e}$ est un vecteur propre si, et seulement si, la famille $\big(\mathbf{e}, u(\mathbf{e})\big)$ est une famille liée.
Si $\mathbf{e}$ est un vecteur propre pour u, tous les vecteurs non nuls de la droite stable $\mathsf{K}\mathbf{e}$ sont des vecteurs propres pour u et sont associés à la même valeur propre. Remarquons donc que, si $\mathbf{e}$ est un vecteur propre de u associé à la valeur propre λ, pour tout $\alpha \in \mathsf{K}^*$, $\alpha\mathbf{e}$ est encore un vecteur propre de u associée à la même valeur propre λ.

Exemples 3.1.
Tous les vecteurs non nuls de E sont des vecteurs propres pour I_E associés à la valeur propre 1 (resp. des vecteures propres pour $0_{\mathcal{L}(E)}$, associés à la valeur propre 0).
Les rotations vectorielles planes d'angle non nul *modulo* π, n'admettent aucun vecteur propre. Les vecteurs propres des rotations vectorielles de l'espace d'angle non nul *modulo* π, sont les vecteurs non nuls de leur axe et sont associés à la valeur propre 1, ce sont des vecteurs invariants.

3.3 Valeur propre

Définitions 3.2 (Valeur propre, spectre).
Soit u est un endomorphisme de E ; on appelle *valeur propre* de u, tout scalaire λ tel qu'il existe un vecteur $\mathbf{x}$ non nul de E vérifiant $u(\mathbf{x}) = \lambda\mathbf{x}$.
L'ensemble des valeurs propres d'un endomorphisme u est appelé le *spectre* de u et noté $\operatorname{sp} u$.

$$\boxed{\lambda \in \operatorname{sp} u \iff \exists \mathbf{x} \in E,\ \mathbf{x} \neq \mathbf{0}_E \text{ et } u(\mathbf{x}) = \lambda\mathbf{x}}$$

Remarque. Le vecteur $\mathbf{x}$ de la définition précédente est un vecteur propre de u ; il est dit associé à la valeur propre λ.

Exemples 3.2.
$\operatorname{sp} I_E = \{1\}$ et $\operatorname{sp} 0_{\mathcal{L}(E)} = \{0\}$.
Le spectre d'une rotation plane d'angle non nul *modulo* π est vide ; celui d'une rotation de l'espace d'angle non nul *modulo* π est $\{1\}$.

Proposition 3.1 (Caractérisation des valeurs propres).
Si u un endomorphisme de E, alors

$$\lambda \in \operatorname{sp} u \iff \ker(u - \lambda I_E) \neq \{\mathbf{0}_E\} \iff u - \lambda I_E \textit{ non injective}$$

En particulier,

$$0 \in \operatorname{sp} u \iff u \textit{ non injective}$$

Si u est un automorphisme de E, $\operatorname{sp}(u^{-1}) = \{\lambda^{-1} \ /\ \lambda \in \operatorname{sp} u\}$.

Démonstration. Si λ est une valeur propre de u, il existe un vecteur (propre) non nul $\mathbf{x} \in E$ tel que $u(\mathbf{x}) = \lambda\mathbf{x}$, égalité que l'on peut encore écrire $\mathbf{0} = u(\mathbf{x}) - \lambda\mathbf{x} = (u - \lambda I_E)(\mathbf{x})$; ceci donne les équivalences annoncées.

Si u est un automorphisme de E, toutes ses valeurs propres sont non nulles ; pour toute valeur propre λ et tout vecteur propre $\mathbf{x}$ associé, on a :

$$u(\mathbf{x}) = \lambda\mathbf{x} \iff \mathbf{x} = u^{-1}(\lambda\mathbf{x}) = \lambda u^{-1}(\mathbf{x}) \iff \frac{1}{\lambda}\mathbf{x} = u^{-1}(\mathbf{x}) \tag{3.1}$$

Ainsi λ est valeur propre de u si, et seulement si, λ^{-1} et valeur propre de u^{-1} et $\mathrm{sp}(u^{-1}) = \{\lambda^{-1} \ / \ \lambda \in \mathrm{sp}\, u\}$. cqfd

Proposition 3.2 (Caractérisation de valeurs propres en dimension finie).
Si u est un endomorphisme d'un K-espace vectoriel de dimension finie *n, alors*

$$\lambda \in \mathrm{sp}\, u \iff u - \lambda I_E \textit{ non injective} \iff u - \lambda I_E \textit{ non surjective}$$
$$\mathrm{rg}(u - \lambda I_E) \leqslant n - 1 \iff \det(u - \lambda I_E) = 0$$

et aussi

$$\lambda \notin \mathrm{sp}\, u \iff (u - \lambda I_E) \textit{ est inversible}$$

Démonstration. C'est l'application des célèbres caractérisations des automorphismes d'un K-espace vectoriel de dimension *finie* : si v est un endomorphisme de E, alors

$$v \text{ est inversible} \iff v \text{ est injectif} \iff v \text{ est surjectif} \iff \mathrm{rg}\, u = n \iff \det v \neq 0$$

Ces caractérisations sont appliquées à $v = u - \lambda I_E$. cqfd

3.4 Sous-espace propre

Définition 3.3 (Sous-espace propre).
Si u est un endomorphisme de E et λ une valeur propre de u, le sous-espace vectoriel

$$\boxed{E_\lambda(u) = \ker(u - \lambda I_E) = \{\mathbf{x} \in E \ / \ u(\mathbf{x}) = \lambda\mathbf{x}\}}$$

est appelé *sous-espace vectoriel propre* de u associé à λ.

Remarques.
$E_\lambda(u)$ est constitué de $\mathbf{0}_E$ et des vecteurs propres de u associés à λ. Si $\mathbf{e}$ est un vecteur propre de u associé à λ, la droite $\mathsf{K}\mathbf{e}$ est contenue dans $E_\lambda(u)$.
$E_0(u) = \ker u$ est le noyau de u.
$E_1(u) = \ker(u - I_E) = \{\mathbf{x} \in E \ / \ u(\mathbf{x}) = \mathbf{x}\}$ est le sous-espace des vecteurs de E invariants par u.

Exemples 3.3. Voici les éléments propres de quelques endomorphismes :

- homothétie $h = \lambda I_E$: $\mathrm{sp}\, h = \{\lambda\}$, $E_\lambda(h) = E$;
- projecteur p : $\mathrm{sp}\, p = \{0, 1\}$, $E_0(p) = \ker p$ et $E_1(p) = \mathrm{Im}\, p$;
- symétrie s : $\mathrm{sp}\, s = \{-1, 1\}$, $E_{-1}(s) = F_1$ et $E_1(s) = F_2$;
- affinité a : $\mathrm{sp}\, a = \{1, \lambda\}$, $E_1(a) = F_1$ et $E_\lambda(a) = F_2$.

Proposition 3.3 (Stabilité des sous-espaces vectoriels propres).
Si les endomorphismes u et v commutent, tout sous-espace vectoriel propre relativement à u est stable par v et réciproquement.

Démonstration. Puisque u et v commutent, $u - \lambda I_E$ et v commutent, et donc, $\ker(u - \lambda I_E) = E_\lambda(u)$ est stable par v. De même $\ker(v - \lambda I_E) = E_\lambda(v)$ est stable par u. cqfd

Théorème 3.4 (Somme directe de sous-espaces propres).
La somme d'une famille finie de sous-espaces propres associés à des valeurs propres distinctes deux à deux, est directe.

Démonstration. Démonstration par récurrence sur le nombre k de sous-espaces vectoriels propres.
$E_{\lambda_1}(u) \cap E_{\lambda_2}(u) = \{\mathbf{0}\}$ car à tout vecteur propre est associé une seule valeur propre. La propriété est vraie pour $k = 2$.
Soient $(\lambda_1, \dots, \lambda_k, \lambda_{k+1})$ $(k+1)$ valeurs propres distinctes de u et montrons que $E_{\lambda_{k+1}}(u) \cap \sum_{i=1}^{k} E_{\lambda_i}(u) = \{\mathbf{0}_E\}$.
Soit $\mathbf{x}_{k+1} = \sum_{i=1}^{k} \mathbf{x}_i \in E_{\lambda_{k+1}}(u) \cap \sum_{i=1}^{k} E_{\lambda_i}(u)$; alors

$$\begin{aligned} u(\mathbf{x}_{k+1}) &= \lambda_{k+1}\mathbf{x}_{k+1} = \lambda_{k+1} \sum_{i=1}^{k} \mathbf{x}_i \\ &= u\Big(\sum_{i=1}^{k} \mathbf{x}_i\Big) = \sum_{i=1}^{k} u(\mathbf{x}_i) = \sum_{i=1}^{k} \lambda_i \mathbf{x}_i \\ \text{donc } \mathbf{0} &= \sum_{i=1}^{k} (\lambda_{k+1} - \lambda_i)\mathbf{x}_i \end{aligned}$$

ce qui montre que $(\lambda_{k+1} - \lambda_i)\mathbf{x}_i = \mathbf{0}$ puisque la somme $\sum_{i=1}^{k} E_{\lambda_i}(u)$ est directe (la propriété est vraie au rang k, puisque les valeurs propres sont distinctes deux à deux) et $\mathbf{x}_i = \mathbf{0}$. Ainsi $\mathbf{x}_{k+1} = \mathbf{0}$ et la propriété est vraie au rang $k + 1$.
Le théorème de récurrence montre que la propriété est vraie pour tout k. cqfd

Corollaire (Liberté d'une famille de vecteurs propres).
Toute famille de vecteurs propres, associés à des valeurs distinctes deux à deux, est une famille libre.

3.5 Éléments propres et polynôme d'endomorphisme

Théorème 3.5 (Éléments propres d'un polynôme d'endomorphisme).
Si $\mathbf{x}$ est un vecteur propre de u associé à λ, pour tout polynôme P, $\mathbf{x}$ est un vecteur propre de $P(u)$ associé à $P(\lambda)$.

Démonstration. $\mathbf{x}$ est non nul et

$$u(\mathbf{x}) = \lambda\mathbf{x} \implies u^k(\mathbf{x}) = \lambda^k\mathbf{x} \text{ (par récurrence sur } k)$$

Si $P = \sum_k a_k X^k$, on a :

$$P(u)(\mathbf{x}) = \Big(\sum_k a_k u^k\Big)(\mathbf{x}) = \sum_k a_k u^k(\mathbf{x}) = \sum_k a_k \lambda^k \mathbf{x} = P(\lambda)\mathbf{x}$$

cqfd

Corollaire (Valeur propre et polynôme annulateur).
Si P est un polynôme annulateur pour u, les valeurs propres de u sont des *racines de P.*

Démonstration. Si λ est une valeur propre de u, $P(\lambda)$ est une valeur propre de $P(u) = 0_{\mathcal{L}(E)}$. Or, 0 est la seule valeur propre de $0_{\mathcal{L}(E)}$, d'où $P(\lambda) = 0$ et λ est racine de P. cqfd

3.6 Éléments propres et automorphismes intérieurs

Proposition 3.6. *Si $a \in \mathcal{GL}(E)$ est un automorphisme de E, l'application $\varphi_a : u \mapsto aua^{-1}$ est un morphisme bijectif de l'algèbre $\mathcal{L}(E)$.*

Démonstration. La notation $\circ$ pour la composition des endomorphismes a été remplacée par une notation multiplicative, *i.e.* par une absence de notation. Il nous faut vérifier un certain nombre de propriétés :

– $\varphi_a(u + v) = a(u + v)a^{-1} = aua^{-1} + ava^{-1} = \varphi_a(u) + \varphi_a(v)$;

– $\varphi_a(\lambda u) = a(\lambda u)a^{-1} = \lambda aua^{-1} = \lambda\varphi_a(u)$;
– $\varphi_a(u \circ v) = auva^{-1} = (aua^{-1})(ava^{-1}) = \varphi_a(u) \circ \varphi_a(v)$;
– $\varphi_a(I_E) = aI_E\, a^{-1} = I_E$;
– $\varphi_a(u) = aua^{-1} = v \iff u = a^{-1}va = \varphi_{a^{-1}}(v)$, soit $(\varphi_a)^{-1} = \varphi_{a^{-1}}$.

cqfd

Définition 3.4 (Automorphismes intérieurs).
Les automorphismes de $\mathcal{L}(E)$ de la forme $u \mapsto aua^{-1}$ sont appelés *automorphismes intérieurs* de $\mathcal{L}(E)$.

Théorème 3.7. *Si u est un endomorphisme et a un automorphisme de E, alors*

$$\operatorname{sp} u = \operatorname{sp}(aua^{-1}) \text{ et } E_\lambda(aua^{-1}) = a\big(E_\lambda(u)\big)$$

Démonstration. Considérons λ une valeur propre de u et $\mathbf{x}$ un vecteur propre associé ; posons $\mathbf{y} = a(\mathbf{x})$, ce qui revient à $\mathbf{x} = a^{-1}(\mathbf{y})$. Ainsi

$$u(\mathbf{x}) = \lambda\mathbf{x} \iff u\big(a^{-1}(\mathbf{y})\big) = \lambda a^{-1}(\mathbf{y}) = a^{-1}(\lambda\mathbf{y}) \iff aua^{-1}(\mathbf{y}) = \lambda\mathbf{y}$$

ce qui montre que λ est une valeur propre de u et $\mathbf{x}$ un vecteur propre associé, si, et seulement si, λ est une valeur propre de aua^{-1} et $\mathbf{y} = a(\mathbf{x})$ un vecteur propre associé. Ainsi

$$\operatorname{sp} u = \operatorname{sp}(aua^{-1}) \quad \text{et} \quad a\big(E_\lambda(u)\big) = E_\lambda(aua^{-1})$$

cqfd

4 Valeurs propres, vecteurs propres d'une matrice carrée

Dans ce paragraphe, M est une matrice carrée d'ordre $n \geqslant 1$, à coefficients dans K.

4.1 Éléments propres d'une matrice carrée

Définition 4.1 (Vecteur propre d'une matrice carrée).
Un élément non nul X de $\mathcal{M}_{n,1}(\mathsf{K})$ est appelé *vecteur propre* de M s'il existe un scalaire λ tel que $MX = \lambda X$. Ce scalaire est unique ; on l'appelle valeur propre associée à X.

Définition 4.2 (Valeur propre d'une matrice carrée).
Un scalaire λ est une *valeur propre* de M si $M - \lambda I_n$ n'est pas inversible ; tout élément non nul de $\ker(M - \lambda I_n)$ est appelé vecteur propre associé à λ.

Définition 4.3 (Sous-espace vectoriel propre d'une matrice carrée).
Si λ est une valeur propre de M, le sous-espace vectoriel $\ker(M - \lambda I_n) = \{X \in \mathcal{M}_{n,1}(\mathsf{K}) \ / \ MX = \lambda X\}$ est appelé *sous-espace vectoriel propre* associé à λ ; il est noté $E_\lambda(M)$.

Définition 4.4 (Spectre d'une matrice carrée).
L'ensemble des valeurs propres de M est le *spectre* de M ; il est noté $\operatorname{sp}_{\mathsf{K}} M$ ou $\operatorname{sp} M$ si le corps n'est pas ambigu.

$$\operatorname{sp}_{\mathsf{K}} M = \{\lambda \in \mathsf{K} \ / \ M - \lambda I_n \notin \mathcal{GL}_n(\mathsf{K})\}$$

Remarque. Rappelons les équivalences pour une matrice carrée

$$\begin{aligned} \lambda \in \operatorname{sp}_{\mathsf{K}} M &\iff M - \lambda I_n \notin \mathcal{GL}_n(\mathsf{K}) \iff \det(M - \lambda I_n) = 0 \\ &\iff \ker(M - \lambda I_n) \neq \{0\} \iff \operatorname{rg}(M - \lambda I_n) \leqslant n - 1 \end{aligned}$$

4.2 Lien avec les endomorphismes

À la matrice M, on associe l'endomorphisme u_M de $\mathcal{M}_{n,1}(\mathsf{K})$ par la formule

$$u_M : X \in \mathcal{M}_{n,1}(\mathsf{K}) \mapsto MX$$

ce qui donne la

Proposition 4.1 (Éléments propres d'une matrice et de son endomorphisme).
Les éléments propres de la matrice M sont les élément propres de l'endomorphisme u_M associé.

Soient E un K-espace vectoriel de dimension finie n, $\mathcal{B}$ une base de E, u un endomorphisme de E et M sa matrice relativement à $\mathcal{B}$. Si $\mathbf{x}$, vecteur de E, est de matrice X relativement à $\mathcal{B}$, $u(\mathbf{x})$ est de matrice MX relativement à $\mathcal{B}$; d'autre part, $\mathbf{x}$ et X sont simultanément nuls ou non, et

$$u(\mathbf{x}) = \lambda\mathbf{x} \iff MX = \lambda X$$

ce qui montre le

Théorème 4.2 (Éléments propres d'un endomorphisme et de sa matrice).
Si u est un endomorphisme de E et $M = \mathcal{M}\mathrm{at}_{\mathcal{B}}(u)$ sa matrice relativement à une base $\mathcal{B}$, alors :

(i) *u et M ont le même spectre ;*
(ii) *$\mathbf{x}$ est vecteur propre de u associé à λ si, et seulement si, sa matrice $X = \mathcal{M}\mathrm{at}_{\mathcal{B}}(\mathbf{x})$ relativement à $\mathcal{B}$ est vecteur propre de M associé à λ.*

4.3 Cas de deux matrices semblables

Soient A et B deux matrices semblables d'ordre n, et P une matrice inversible telle que $B = P^{-1}AP$; à A est associé l'endomorphisme u_A de $\mathcal{M}_{n,1}(\mathsf{K})$. La matrice P peut s'interpréter comme la matrice de passage de la base naturelle (canonique) $\mathcal{E}$ de $\mathcal{M}_{n,1}(\mathsf{K})$ à une base $\mathcal{B}$ constituée des vecteurs colonnes de P. Relativement à cette base, la matrice de u_A est B, d'où le

Théorème 4.3 (Matrices semblables et endomorphisme).
Deux matrices sont semblables si, et seulement si, elles représentent le même endomorphisme (relativement à deux bases).

Corollaire (Spectre de deux matrices semblables).
Deux matrices semblables ont le même spectre.

Démonstration. C'est le spectre de l'unique endomorphisme qu'elle représente. cqfd

4.4 Cas des matrices réelles

Une matrice à coefficients réels est aussi une matrice à coefficients complexes. La notion de valeur propre dépend du corps envisagé et on se gardera de confondre $\mathrm{sp}_{\mathsf{R}}\, M$ et $\mathrm{sp}_{\mathsf{C}}\, M$; mais on a le

Théorème 4.4 (Spectres réel et complexe d'une matrice réelle).
Si M est une matrice à coefficients réels, $\mathrm{sp}_{\mathsf{R}}\, M$ est contenue dans $\mathrm{sp}_{\mathsf{C}}\, M$; en général, les deux spectres sont distincts.

Démonstration. Si λ est un élément de $\mathrm{sp}_{\mathsf{R}}\, M$ et X un vecteur propre associé, alors $MX = \lambda X$, X est non nul et appartient à $\mathcal{M}_{1,n}(\mathsf{R})$ donc à $\mathcal{M}_{1,n}(\mathsf{C})$; ainsi λ appartient à $\mathrm{sp}_{\mathsf{C}}\, M$.
Considérons la rotation plane d'angle $\pi/2$; sa matrice, dans une base orthonormale directe, est $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ et $\mathrm{sp}_{\mathsf{R}}\, A$ est vide puisque cette rotation n'admet aucune droite stable. Par contre, i est une valeur propre complexe de A, puisque $A - iI_2 = \begin{pmatrix} -i & -1 \\ 1 & -i \end{pmatrix}$ n'est pas inversible. cqfd

Proposition 4.5 (Valeur propre, vecteur propre, conjugaison).
Soit M une matrice carrée d'ordre n et à coefficients complexes ; si X est un vecteur propre de M associé à la valeur propre λ, $\overline{X}$ est un vecteur propre de $\overline{M}$ associé à la valeur propre $\overline{\lambda}$, et

$$\mathrm{sp}_{\mathbb{C}}(\overline{M}) = \{\overline{\lambda} \ / \ \lambda \in \mathrm{sp}\, M\}, \qquad E_{\overline{\lambda}}(\overline{M}) = \{\overline{X} \ / \ X \in E_\lambda(M)\}$$

Démonstration. Rappelons la relation $\overline{A\,B} = \overline{A}\ \overline{B}$ pour $A \in \mathcal{M}_{n,p}(\mathbb{C})$ et $B \in \mathcal{M}_{p,q}(\mathbb{C})$. Ainsi on a l'équivalence

$$MX = \lambda X \iff \overline{M}\ \overline{X} = \overline{\lambda}\ \overline{X}$$

ce qui donne les résultats annoncés. cqfd

Proposition 4.6 (Valeur propre complexe non réelle d'une matrice réelle).
Soit M une matrice carrée d'ordre $n \geqslant 2$ et à coefficients réels *; si λ est une valeur propre complexe* non réelle *de M et X un vecteur propre à coefficients complexes associé à λ, alors :*

(i) *$\overline{\lambda}$ est une valeur propre de M et $\overline{X}$ un vecteur propre associé à $\overline{\lambda}$;*
(ii) *le plan $(\mathcal{P})$ orienté par $(\Im\mathrm{m}\, X, \Re\mathrm{e}\, X)$ est un plan stable pour M, et M induit sur $(\mathcal{P})$ la similitude de rapport $|\lambda|$ et d'angle* $\arg\lambda$.

Démonstration.

(i) M est une matrice réelle, alors $\overline{M} = M$, et

$$MX = \lambda X \implies \overline{M}\ \overline{X} = M\overline{X} = \overline{\lambda}\ \overline{X}$$

(ii) Puisque $\lambda = a + ib$ n'est pas réel, $\overline{\lambda} \neq \lambda$ et la famille $\mathcal{B} = (X, \overline{X})$ est libre (vecteurs propres associés à deux valeurs propres distinctes). Les formules

$$\Re\mathrm{e}\, X = \frac{1}{2}(X + \overline{X}), \qquad \Im\mathrm{m}\, X = \frac{1}{2i}(X - \overline{X})$$
$$X = \Re\mathrm{e}\, X + i\,\Im\mathrm{m}\, X, \qquad \overline{X} = \Re\mathrm{e}\, X - i\,\Im\mathrm{m}\, X$$

montrent que la famille $\mathcal{C} = (\Im\mathrm{m}\, X, \Re\mathrm{e}\, X)$ est aussi une famille libre (sur $\mathbb{C}$, donc sur $\mathbb{R}$). Séparant les parties réelle et imaginaire, on obtient

$$\begin{aligned} MX &= M\,\Re\mathrm{e}\, X + iM\,\Im\mathrm{m}\, X \\ &= \lambda X = (a+ib)(\Re\mathrm{e}\, X + i\,\Im\mathrm{m}\, X) \\ &= (a\,\Re\mathrm{e}\, X - b\,\Im\mathrm{m}\, X) + i(b\,\Re\mathrm{e}\, X + a\,\Im\mathrm{m}\, X) \end{aligned}$$

En appelant s l'endomorphisme induit sur le plan $(\mathcal{P})$ par M, $\mathcal{M}\mathrm{at}_{\mathcal{C}}(s) = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$, ce qui est la matrice de la similitude de rapport (complexe) $a + ib = \lambda$, *i.e.* le produit de l'homothétie de rapport $|\lambda|$ et de la rotation d'angle $\arg\lambda$.

cqfd

5 Polynôme caractéristique

5.1 Définitions

A) Polynôme caractéristique d'une matrice carrée

Si M est une matrice carrée d'ordre $n \geqslant 1$ à coefficients dans $\mathbb{K}$, $M - XI_n$ est une matrice à coefficients dans $\mathbb{K}[X]$ sous-anneau du corps $\mathbb{K}(X)$ des fractions rationnelles. Le déterminant de $M - XI_n$ est polynomial en les coefficients de cette matrice, il est donc élément de $\mathbb{K}[X]$, ce qui permet la

Définition 5.1 (Polynôme caractéristique d'une matrice carrée).
Si $M = [a_{i,j}]$ est une matrice carrée d'ordre $n \geqslant 1$, le *polynôme caractéristique* de M, que l'on note χ_M, est le déterminant

$$\chi_M(X) = \det(M - XI_n) = \begin{vmatrix} a_{1,1} - X & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} - X & \cdots & a_{2,n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} - X \end{vmatrix}$$

Remarque. X est ici une indéterminée ; en cas d'indigestion, remplacer X par x ; le polynôme devient une fonction polynôme ; servir chaud.

Exemples 5.1. Voici quelques polynômes caractéristiques de matrices :

- matrice nulle : $\chi_0 = (-X)^n$;
- matrice unité d'ordre n : $\chi_{I_n} = (1 - X)^n$;
- matrice diagonale $D = \mathrm{Diag}(\lambda_1, \dots, \lambda_n)$: $\chi_D = \prod_{i=1}^n (\lambda_i - X)$;
- matrice triangulaire $T = [t_{i,j}]$: $\chi_T = \prod_{i=1}^n (t_{i,i} - X)$.

Proposition 5.1 (Polynôme caractéristique et transposée).
Une matrice et sa transposée ont le même polynôme caractéristique.

Démonstration. Rappelons qu'une matrice et sa transposée ont le même déterminant ; d'où

$${}^tM - XI_n = {}^t(M - XI_n) \implies \det({}^tM - XI_n) = \det\big({}^t(M - XI_n)\big) = \det(M - XI_n)$$

ce qui donne le résultat. cqfd

Proposition 5.2 (Polynôme caractéristique et matrices semblables).
Deux matrices semblables ont le même polynôme caractéristique.

Démonstration. Rappelons que deux matrices semblables ont le même déterminant. Si A et B sont deux matrices semblables, et P une matrice inversible telle que $B = P^{-1}AP$, alors

$$B = P^{-1}AP \implies B - XI_n = P^{-1}AP - XI_n = P^{-1}(A - XI_n)P$$

les matrices $B - XI_n$ et $A - XI_n$ sont semblables, d'où

$$\det(B - XI_n) = \det\Big(P^{-1}(A - XI_n)P\Big) = \det(A - XI_n)$$

ce qui donne le résultat. cqfd

B) Polynôme caractéristique d'un endomorphisme

Soient u un endomorphisme de E, $\mathcal{B}$ et $\mathcal{C}$ deux bases de E ; les matrices $\mathcal{M}\mathrm{at}_{\mathcal{B}}(u)$ et $\mathcal{M}\mathrm{at}_{\mathcal{C}}(u)$ sont des matrices semblables ; elles admettent donc le même polynôme caractéristique, ce qui permet la

Définition 5.2 (Polynôme caractéristique d'un endomorphisme).
Si u est un endomorphisme d'un K-espace vectoriel E de dimension finie $n \geqslant 1$, le polynôme caractéristique de la matrice qui représente u dans une base donnée $\mathcal{B}$ de E, est indépendant du choix de cette base ; on l'appelle *polynôme caractéristique* de l'endomorphisme u ; il est noté χ_u.

$$\boxed{\chi_u = \det\big(\mathcal{M}\mathrm{at}_{\mathcal{B}}(u) - XI_n\big) \quad \text{où } \mathcal{B} \text{ est une base de } E}$$

Exemple 5.2. Le polynôme caractéristique de l'application identique est $\chi_{I_E} = \chi_{I_n} = (1 - X)^n$ et celui de l'application nulle $\chi_{0_{\mathcal{L}(E)}} = \chi_{0_{\mathcal{M}_n(\mathsf{K})}} = (-X)^n$.

5.2 Propriétés du polynôme caractéristique

Nous indiquerons, dans ce paragraphe, les propriétés du polynôme caractéristique d'une matrice ; le lecteur, dans sa grande bonté, effectuera lui-même les modifications qu'il jugera utiles pour exprimer les propriétés du polynôme caractéristique d'un endomorphisme.
Quelle est l'utilité du polynôme caractéristique ? Déterminer les valeurs propres d'une matrice, ou d'un endomorphisme.

Théorème 5.3 (Polynôme caractéristique et valeurs propres).
Les valeurs propres d'une matrice sont les racines sur K *de son polynôme caractéristique.*

Démonstration. Évidente, car $\lambda \in \operatorname{sp}_{\mathsf{K}} M \iff M - \lambda I_{,} \notin \mathcal{GL}_n(\mathsf{K}) \iff 0 = \det(M - \lambda I_n) = \chi_M(\lambda)$ et $\lambda \in \mathsf{K}$.
cqfd

Corollaire (Spectre d'une matrice complexe).
Toute matrice à coefficients complexes admet, au moins, une valeur propre (complexe).

Démonstration. Tout polynôme à coefficients complexes admet une racine complexe, c'est le théorème de D'Alembert.
cqfd

Proposition 5.4 (Coefficients du polynôme caractéristique).
χ_M *est un polynôme de degré n et*

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

Démonstration. L'égalité $\chi_M(0) = \det M$ donne le coefficient constant.
Appelons $a_{i,j}$ le terme général de M et $\tilde{a}_{i,j}$ celui de $M - XI_n$. Si s est une permutation de $\mathfrak{S}_n$ distincte de l'identité, $\prod_{i=1}^n \tilde{a}_{i,s(i)}$ est un polynôme de degré au plus $(n-2)$, ce qui donne

$$\begin{aligned}
\chi_M &= \begin{vmatrix} a_{1,1} - X & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} - X & \cdots & a_{2,n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} - X \end{vmatrix} \\
&= \prod_{i=1}^n (a_{i,i} - X) + \underbrace{\sum_{s \in \mathfrak{S}_n, s \neq e} \varepsilon(s) \prod_{i=1}^n \tilde{a}_{i,s(i)}}_{\text{polynôme de degré} \leqslant n-2} \\
&= (-X)^n + \Big(\sum_{i=1}^n a_{i,i}\Big)(-X)^{n-1} + \underbrace{\cdots + \cdots + \cdots}_{\text{polynôme de degré} \leqslant n-2} \\
&= (-X)^n + \operatorname{tr} M(-X)^{n-1} + \underbrace{\cdots + \cdots + \cdots}_{\text{polynôme de degré} \leqslant n-2}
\end{aligned}$$

Quant au coefficient de $-X$, c'est une autre histoire . . . cqfd

Remarques. Le calcul de χ_M est facile en petite dimension :

- pour $n = 2$, $\chi_M = X^2 - (\operatorname{tr} M)X + \det M$;
- pour $n = 3$, $\chi_M = -X^3 + (\operatorname{tr} M)X^2 - \operatorname{tr}(\operatorname{Com} M)X + \det M$.

Corollaire (Nombre de valeurs propres distinctes d'une matrice carrée).
Le nombre de valeurs propres d'une matrice carrée est au plus égal à son ordre.

Démonstration. Le nombre de valeurs propres distinctes d'une matrice carrée d'ordre n est égal au nombre de racines distinctes du polynôme caractéristique, *i.e.* d'un polynôme de degré n ; ce nombre est compris entre 0 et n.
cqfd

Corollaire (Cas des polynômes caractéristiques scindés).
Si le polynôme caractéristique de M est scindé sur K, avec $\chi_M = \prod_{i=1}^{n}(\lambda_i - X)$, alors

$$\operatorname{sp} M = \{\lambda_1, \dots, \lambda_n\},\ \operatorname{tr} M = \sum_{i=1}^{n} \lambda_i,\ \det M = \prod_{i=1}^{n} \lambda_i$$

Démonstration. Il suffit de développer $\prod_{i=1}^{n}(\lambda_i - X)$ et d'identifier ce polynôme avec celui de la formule précédente. cqfd

Corollaire (Cas des matrices réelles).
Tout matrice à coefficients réels et d'ordre impair admet, au moins, une valeur propre réelle.

Démonstration. Si n est impair, χ_M est de degré impair et $\chi_M(x) \underset{\pm\infty}{\sim} (-x)^n$; le théorème des valeurs intermédiaires assure l'existence d'une racine réelle pour χ_M, et $\operatorname{sp}_{\mathsf{R}} M$ n'est pas vide. cqfd

Théorème 5.5 (Cas d'une matrice triangulaire par blocs).
Le polynôme caractéristique d'une matrice triangulaire par blocs se calcule aisément :

$$M = \begin{pmatrix} A & B \\ 0 & C \end{pmatrix},\ \textit{avec } A \in \mathcal{M}_p(\mathsf{K}) \textit{ et } C \in \mathcal{M}_q(\mathsf{K}) \implies \chi_M = \chi_A \chi_C$$

Démonstration. $M - XI_n = \begin{pmatrix} A - XI_p & B \\ 0 & C - XI_q \end{pmatrix}$ et le calcul du déterminant d'une matrice triangulaire par blocs donne le résultat. cqfd

5.3 Multiplicité d'une valeur propre

Définition 5.3 (Multiplicité d'une valeur propre).
L'ordre de multiplicité d'une racine λ de χ_M est appelé *multiplicité de la valeur propre* λ de M ; elle est noté $m(\lambda)$.

Remarque. Si χ_M est scindé et si $\lambda_1, \dots, \lambda_p$ sont les valeurs propres deux à deux distinctes de M, le polynôme caractéristique de M s'écrit

$$\chi_M = \prod_{j=1}^{p} (\lambda_j - X)^{m(\lambda_j)} = \prod_{\lambda \in \operatorname{sp} M} (\lambda - X)^{m(\lambda)}$$

et la somme des multiplicités des valeurs propres de M est égale à son ordre.

Théorème 5.6 (Dimension d'un sous-espace vectoriel propre).
Soient λ une valeur propre de M, $m(\lambda)$ sa multiplicité et $q(\lambda)$ la dimension du sous-espace propre associé ; alors

$$\boxed{1 \leqslant q(\lambda) = \dim E_\lambda(M) = n - \operatorname{rg}(M - \lambda I_n) \leqslant m(\lambda)}$$

Démonstration. Appelons u l'endomorphisme de $\mathcal{M}_{n,1}(\mathsf{K})$ associé à M. Pour $\lambda \in \operatorname{sp} u = \operatorname{sp} M$, on a

- $E_u(\lambda) = \ker(u - \lambda I_E) \neq \{\mathbf{0}\}$, donc $q(\lambda) \geqslant 1$;
- $\dim E_\lambda(u) = \dim \ker(u - \lambda I_E) = n - \operatorname{rg}(u - \lambda I_E)$;
- $E_\lambda(u)$ est un sous-espace stable par u ; dans une base adaptée, la matrice de u est triangulaire (supérieure) par blocs, soit
$$N = \begin{pmatrix} \lambda I_{q(\lambda)} & B \\ 0 & C \end{pmatrix}$$
donc $\chi_N = \chi_u = \chi_M = (\lambda - X)^{p(\lambda)} \chi_C$; ceci montre que $q(\lambda) \leqslant m(\lambda)$.

cqfd

Corollaire (Cas des valeurs propres simples).
Le sous-espace propre associé à une valeur propre simple est une droite vectorielle.

Démonstration. Les inégalités $1 \leqslant q(\lambda) \leqslant m(\lambda) = 1$ donnent le résultat. cqfd

5.4 Polynôme caractéristique et polynôme annulateur

Théorème 5.7 (Cayley-Hamilton).
Le polynôme caractéristique d'un endomorphisme ou d'une matrice est un polynôme annulateur de cet endomorphisme ou de cette matrice, soit :

$$\chi_M(M) = 0_{\mathcal{M}_n(\mathsf{K})} \qquad et \qquad \chi_u(u) = 0_{\mathcal{L}(E)}$$

Démonstration. Hors programme cqfd

6 Diagonalisation

On suppose ici que E est un K-espace vectoriel vectoriel de dimension finie $n \geqslant 1$ et u un endomorphisme de E.

6.1 Endomorphisme diagonalisable

Définition 6.1. Un endomorphisme de E est dit *diagonalisable* si E est la somme directe de ses sous-espaces propres.

$$\boxed{u \text{ est diagonalisable} \iff E = \bigoplus_{\lambda \in \operatorname{sp} u} E_\lambda(u)}$$

Remarque. La somme des sous-espaces propres de E est *toujours* directe, mais pas nécessairement égale à E.

Exemples 6.1. Les homothéties, les projecteurs, les symétries et les affinités sont diagonalisables.

6.2 Caractérisation des endomorphismes diagonalisables

Théorème 6.1. *Si u est un endomorphisme d'un K-espace vectoriel E de dimension finie, alors les propositions suivantes sont équivalentes :*

(i) *u est diagonalisable ;*
(ii) *E est la somme directe de sous-espaces stables sur lesquels u induit une homothétie ;*
(iii) *il existe une base de E constituée de vecteurs propres de u ;*
(iv) *il existe une base de E sur laquelle la matrice de u est diagonale.*

Démonstration.

$(i) \implies (ii)$ Pour toute valeur propre λ de u, $E_\lambda(u)$ est un sous-espace vectoriel stable sur lequel u induit une homothétie (de rapport λ).

$(ii) \implies (iii)$ Si F est un sous-espace stable sur lequel u induit une homothétie (de rapport k), tous les vecteurs non nuls de F sont des vecteurs propres de u (associés à k). Une base de E adaptée à la décomposition de E en somme directe est constituée de vecteurs propres de u.

$(iii) \implies (iv)$ Si $\mathcal{B} = (\mathbf{e}_1, \dots, \mathbf{e}_n)$ est une base constituée de vecteurs propres de u, la matrice de u relativement à $\mathcal{B}$ est diagonale $\operatorname{Diag}(\lambda_1, \dots, \lambda_n)$ avec λ_i valeur propre de u associée au vecteur propre $\mathbf{e}_i$ ($u(\mathbf{e}_i) = \lambda_i \mathbf{e}_i$).

$(iv) \implies (i)$ Quitte à réordonner la base de E sur laquelle la matrice de u est diagonale, on peut supposer que cette matrice est diagonale par blocs : $M = \operatorname{Diag}(\lambda_j I_{m(\lambda_j)})_{\lambda_j \in [\![1,p]\!]}$. Le polynôme caractéristique de u est $\chi_u = \chi_M = \prod_{i=j}^{p} (\lambda_j - X)^{m(\lambda_j)}$, les λ_j sont les valeurs propres de u et $E_{\lambda_j}(u)$ est de dimension $m(\lambda_j)$ de vecteures de la base de E ; ainsi E est somme directe des $E_{\lambda_j}(u)$ puisque $\sum_j \dim E_{\lambda_j}(u) = \sum_j m(\lambda_j) = n$.

cqfd

6.3 Endomorphisme diagonalisable, dimension des sous-espaces propres

Théorème 6.2. *Soit u un endomorphisme d'un* $\mathbb{K}$*-espace vectoriel de dimension finie ; les propositions suivantes sont équivalentes :*

(i) *u est diagonalisable ;*
(ii) *la somme des dimensions des sous-espaces propres est égale à la dimension de E ;*
(iii) *le polynôme caractéristique de u est scindé, et, pour toute valeur propre de u, la multiplicité est égale à la dimension du sous-espace propre associé.*

$$\boxed{\begin{array}{c} u \text{ est diagonalisable} \iff \displaystyle\sum_{\lambda\in\operatorname{sp} u} \dim E_\lambda(u) = n \\ \iff \chi_u \text{ est scindé et } \forall\lambda \in \operatorname{sp} u,\ m(\lambda) = \dim E_\lambda(u) \end{array}}$$

Démonstration.

$(i) \iff (ii)$ La somme des sous-espaces vectoriels propres est directe ; cette somme est égale à E si, et seulement si, sa dimension est la dimension de E ; ceci donne le résultat, puisque la dimension d'une somme directe est égale à la somme des dimensions de chacun des facteurs.

$(i) \implies (iii)$ Puisque u est diagonalisable, il existe une base de E sur laquelle la matrice de u est diagonale $\operatorname{Diag}(\lambda_1, \dots, \lambda_n)$; ainsi $\chi_u = \prod_{i=1}^n (\lambda_i - X)$ est un polynôme scindé.
On sait que pour toute valeur propre λ, $q(\lambda) = \dim E_\lambda(u) \leqslant m(\lambda)$; d'autre part

$$n = \sum_{\lambda\in\operatorname{sp} u} m(\lambda) = \sum_{\lambda\in\operatorname{sp} u} \dim E_\lambda(u)$$

ce qui implique que $0 = \sum_{\lambda\in\operatorname{sp} u}\big(m(\lambda) - q(\lambda)\big)$ et chacun des termes de la somme est positif ; tous les termes de cette somme sont donc nuls.

$(iii) \implies (ii)$ Puisque χ_u est scindé, $\sum_{\lambda\in\operatorname{sp} u} m(\lambda) = n$; ceci implique que $\sum_{\lambda\in\operatorname{sp} u} \dim E_\lambda(u) = n$ et u est diagonalisable.
cqfd

Théorème 6.3 (Cas des valeurs propres simples).
Tout endomorphisme d'un $\mathbb{K}$*-espace vectoriel de dimension finie dont le polynôme caractéristique est scindé et a toutes ses racines* simples*, est diagonalisable et ses sous-espaces propres sont des droites vectorielles.*

Démonstration. Dans ce cas, $m(\lambda) = \dim E_\lambda(u) = 1$ pour toute valeur propre λ de u. cqfd

Corollaire. *Tout endomorphisme d'un* $\mathbb{K}$*-espace vectoriel de dimension finie n dont le polynôme caractéristique admet n racines distinctes est diagonalisable et ses sous-espaces propres sont des droites vectorielles.*

Démonstration. Dans ce cas, les racines du polynôme caractéristique sont simples. cqfd

6.4 Endomorphisme diagonalisable et polynôme annulateur

Théorème 6.4. *Si u est un endomorphisme d'un* $\mathbb{K}$*-espace vectoriel de dimension finie, u est diagonalisable si, et seulement si, u annule un polynôme scindé dont toutes les racines sont simples.*

Démonstration.

$\boxed{\implies}$ Appelons $\lambda_1, \dots, \lambda_p$ les valeurs propres deux à deux *distinctes* de u. Dans une base adaptée à la décomposition $E = \oplus_{j=1}^p E_{\lambda_j}(u)$, la matrice de u est diagonale par blocs $M = \operatorname{Diag}(\lambda_j I_{m(\lambda_j)})_j$. Le polynôme $P = \prod_{j=1}^p (X-\lambda_j)$ est annulateur de M, car, par blocs,

$$P(M) = \operatorname{Diag}(P(\lambda_j) I_{m(\lambda_j)})_{j\in[\![1,p]\!]} = 0 \qquad \text{car } \lambda_j \text{ est racine de } P \text{ pour tout } j \in [\![1, p]\!]$$

Ainsi le polynôme P est scindé, à racines simples et annulateur de u.

$\Longleftarrow$ La démonstration est hors programme ; elle est basée sur la propriété suivante : si λ et μ sont deux scalaires distincts, alors
$$\ker\big((u-\lambda I_E)\circ(u-\mu I_E)\big)=\ker(u-\lambda I_E)\oplus\ker(u-\mu I_E)$$
Un raisonnement par récurrence sur le nombre de racines de $P=\prod_{j=1}^{p}(X-\lambda_j)$ montre que :
$$\ker P(u)=\ker\prod_{j=1}^{p}(u-\lambda_j I_E)=\bigoplus_{j=1}^{p}\ker(u-\lambda_j I_E)$$
Si P est annulateur de u, $\ker P(u)=E$ et u est diagonalisable. cqfd

6.5 Endomorphisme diagonalisable et sous-espace stable

Théorème 6.5. *Si u est un endomorphisme diagonalisable de E et F un sous-espace stable pour u, l'endomorphisme v, induit par u sur F, est diagonalisable.*

Démonstration. Soit P un polynôme annulateur de u, scindé et à racines simples. La restriction de $P(u)$ à F est $P(v)$, et donc, P est annulateur de v. Ainsi v est diagonalisable. cqfd

6.6 Cas des matrices

Définition 6.2 (Matrice diagonalisable).
Une matrice carrée d'ordre n à coefficients dans K est *diagonalisable* si l'endomorphisme u_M de $\mathcal{M}_{n,1}(\mathsf{K})$ associé est diagonalisable.

Théorème 6.6 (Caractérisation des matrices diagonalisables).
Pour une matrice carrée M d'ordre n et à coefficients dans K, les propriétés suivantes sont équivalentes :

- (i) *M est diagonalisable ;*
- (ii) *M est la matrice d'un endomorphisme diagonalisable ;*
- (iii) *M est semblable à une matrice diagonale ;*
- (iv) *la somme des dimensions des sous-espaces propres de M est n ;*
- (v) *χ_M est scindé sur K, et, pour toute valeur propre (de $\mathrm{sp}_{\mathsf{K}}\,M$), la multiplicité est égale à la dimension du sous-espace propre associé ;*
- (vi) *il existe une matrice inversible $P\in\mathcal{GL}_n(\mathsf{K})$ telle que $P^{-1}MP=\mathrm{Diag}(\lambda_1,\dots,\lambda_n)$ soit diagonale, P étant la matrice de changement de base, de la base canonique à une base constituée de vecteurs propres de M, le j^{e} étant relatif à la valeur propre λ_j.*

Démonstration. Les cinq premières assertions sont la traduction au cas de u_M des théorèmes qui caractérisent les endomorphismes diagonalisables.
La sixième assertion est l'explication pratique : recherche d'une base de $\mathcal{M}_{n,1}(\mathsf{K})$ constituée de vecteurs propres de M et utilisation de cette base pour construire une matrice inversible P, la matrice de passage de la base canonique à la base de vecteurs propres, qui diagonalise M. cqfd

CHAPITRE

4

Analyse hilbertienne

Sommaire

Le but de chapitre est de généraliser l'espace ordinaire et son produit scalaire

- à la dimension n : espace euclidien ;
- à la dimension infinie : espace préhilbertien réel ;
- aux nombres complexes : espace préhilbertien complexe, espace hermitien.

La généralisation portera essentiellement sur les notions d'orthogonalité et de projection orthogonale.

1 Produit scalaire sur un espace vectoriel réel

Dans cette section, E désigne un R-espace vectoriel de dimension finie ou infinie, sauf avis contraire.

Définition 1.1 (Produit scalaire réel).
On appelle *produit scalaire réel* sur E, toute forme bilinéaire symétrique et définie positive, *i.e.* toute application $\varphi : E \times E \to \mathsf{R}$ telle que

(i) $\forall \mathbf{x} \in E,\ \varphi_{\mathbf{x}} : \mathbf{y} \mapsto \varphi(\mathbf{x}, \mathbf{y})$ est linéaire — *linéarité à droite* ;
(ii) $\forall (\mathbf{x}, \mathbf{y}) \in E^2,\ \varphi(\mathbf{x}, \mathbf{y}) = \varphi(\mathbf{y}, \mathbf{x})$ — *symétrie* ;
(iii) $\forall \mathbf{x} \in E,\ \mathbf{x} \neq \mathbf{0} \implies \varphi(\mathbf{x}, \mathbf{x}) > 0$ — *définie positive*.

Définitions 1.2 (Espace préhilbertien réel, espace euclidien).
E muni du produit scalaire φ est appelé un *espace préhilbertien réel*. Si E est un espace de dimension finie, (E, φ) est un *espace euclidien*.
Le produit scalaire scalaire $\varphi(\mathbf{x}, \mathbf{y})$ de deux vecteurs est noté $\langle \mathbf{x} \mid \mathbf{y} \rangle$, ou encore $\mathbf{x} \cdot \mathbf{y}$, $<\mathbf{x}, \mathbf{y}>$, $(\mathbf{x} \,|\, \mathbf{y})$ ...

Remarques.
La linéarité à droite et la symétrie impliquent la linéarité à gauche.
Si l'un des vecteurs est nul, le produit scalaire est nul.
Le caractère « défini positif » du produit scalaire peut s'établir en montrant que

$$\forall \mathbf{x} \in E,\ \langle \mathbf{x} \mid \mathbf{x} \rangle \geqslant 0 \text{ et } \langle \mathbf{x} \mid \mathbf{x} \rangle = 0 \implies \mathbf{x} = \mathbf{0}$$

Si F est un R-sous-espace de E, tout produit scalaire réel sur E induit un produit scalaire réel sur F.

Exemples 1.1.

(i) Produit scalaire canonique sur R^n : il est défini par

$$\forall \mathbf{x} = (x_1, \dots, x_n),\ \forall \mathbf{y} = (y_1, \dots, y_n), \quad \langle \mathbf{x} \mid \mathbf{y} \rangle = \sum_{k=1}^{n} x_k y_k$$

(ii) Produit scalaire canonique sur $\mathcal{M}_{n,1}(\mathsf{R})$: il est défini par

$$\forall (X, Y) \in \big(\mathcal{M}_{n,1}(\mathsf{R})\big)^2, \quad \langle X \mid Y \rangle = {}^t X Y = \sum_{k=1}^{n} x_k y_k$$

(iii) Produit scalaire canonique sur $\mathcal{M}_{n,p}(\mathsf{R})$: il est défini par

$$\forall (A, B) \in \big(\mathcal{M}_{n,p}(\mathsf{R})\big)^2, \quad \langle A \mid B \rangle = \operatorname{tr}({}^t A B) = \sum_{i,j} a_{i,j} b_{i,j}$$

(iv) Produit scalaire sur l'espace des fonctions continues sur le segment $[a, b]$ et à valeurs réelles :

$$\forall (f, g) \in \big(\mathcal{C}([a, b], \mathsf{R})\big)^2, \quad \langle f \mid g \rangle = \int_a^b f(t) g(t)\, dt$$

(v) Produit scalaire sur l'espace des fonctions continues, 2π-périodiques sur R et à valeurs réelles :

$$\forall (f,g) \in \left(\mathcal{C}_{2\pi}(\mathsf{R})\right)^2, \quad \langle f \mid g\rangle = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)g(t)\,dt$$

(vi) Produit scalaire sur l'espace des fonctions continues et de carré intégrable sur l'intervalle I, et à valeurs réelles :

$$\forall (f,g) \in \left(\mathcal{L}^2(I,\mathsf{R})\right)^2, \quad \langle f \mid g\rangle = \int_I f(t)g(t)\,dt$$

(vii) Produits scalaires sur l'espace des polynômes à coefficients réels ; si I est un intervalle et w une fonction continue sur I et à valeurs réelles (strictement) positives, telle que, pour tout entier n, $t \mapsto t^n w(t)$ soit intégrable sur I, l'application

$$(P,Q) \in \left(\mathsf{R}[X]\right)^2 \mapsto \langle P \mid Q\rangle = \int_I P(t)Q(t)w(t)\,dt$$

est un produit scalaire. Les cas classiques sont

$$I = [-1,1],\ w(t) = 1 \text{ et } \langle P \mid Q\rangle = \int_{-1}^{1} P(t)Q(t)\,dt$$
$$I =]-1,1[,\ w(t) = \frac{1}{\sqrt{1-t^2}} \text{ et } \langle P \mid Q\rangle = \int_{-1}^{1} P(t)Q(t)\,\frac{dt}{\sqrt{1-t^2}}$$
$$I = [0,+\infty[,\ w(t) = e^{-t} \text{ et } \langle P \mid Q\rangle = \int_{0}^{+\infty} P(t)Q(t)e^{-t}\,dt$$
$$I =]-\infty,+\infty[,\ w(t) = e^{-t^2} \text{ et } \langle P \mid Q\rangle = \int_{-\infty}^{+\infty} P(t)Q(t)e^{-t^2}\,dt$$

$(viii)$ Produit scalaire sur l'espace des suites réelles de carré sommable $\ell^2_{\mathsf{N}}(\mathsf{R})$, *i.e.* les suites réelles $(a_n)_n$ telles que la série $\sum_n a_n^2$ soit convergente.

$$\forall (\mathbf{a},\mathbf{b}) \in \left(\ell^2_{\mathsf{N}}(\mathsf{R})\right)^2, \quad \langle \mathbf{a} \mid \mathbf{b}\rangle = \sum_{k=0}^{\infty} a_k b_k$$

2 Norme et distance associée à un produit scalaire réel

E désigne un espace préhilbertien réel dont le produit scalaire est noté $\langle\ |\ \rangle$.

Définitions 2.1 (Norme et distance associée).
La norme associée au produit scalaire $\langle\ |\ \rangle$ est définie par :

$$\boxed{\forall \mathbf{x} \in E, \quad \|\mathbf{x}\| = \sqrt{\langle \mathbf{x} \mid \mathbf{x}\rangle}}$$

La distance associée au produit scalaire est définie par :

$$\boxed{\forall (\mathbf{x},\mathbf{y}) \in E^2, \quad \mathrm{d}(\mathbf{x},\mathbf{y}) = \|\mathbf{y}-\mathbf{x}\| = \sqrt{\langle (\mathbf{y}-\mathbf{x}) \mid (\mathbf{y}-\mathbf{x})\rangle}}$$

Dans ce cas réel, la norme et la distance sont qualifiées d'*euclidiennes*.

Proposition 2.1. $\mathbf{x} \mapsto \|\mathbf{x}\| = \sqrt{\langle \mathbf{x} \mid \mathbf{x}\rangle}$ *est une application de* E *sur* $[0,+\infty[$ *qui vérifie*

(i) $\forall \mathbf{x} \in E,\ \|\mathbf{x}\| = 0 \iff \mathbf{x} = \mathbf{0}$ axiome de séparation ;

(ii) $\forall(\lambda, \mathbf{x}) \in \mathsf{R} \times E,\ \|\lambda\mathbf{x}\| = |\lambda|\,\|\mathbf{x}\|$ axiome d'homogénéité.

Démonstration.

(i) $0 = \|\mathbf{x}\|^2 = \langle \mathbf{x} \mid \mathbf{x}\rangle \iff \mathbf{x} = \mathbf{0}$;
(ii) $\|\lambda\mathbf{x}\| = \sqrt{\langle \lambda\mathbf{x} \mid \lambda\mathbf{x}\rangle} = \sqrt{\lambda^2\langle \mathbf{x} \mid \mathbf{x}\rangle} = |\lambda|\langle \mathbf{x} \mid \mathbf{x}\rangle$

cqfd

L'inégalité triangulaire sera démontrée à la fin de cette section.

2.1 Expression du produit scalaire en fonction de la norme

Proposition 2.2 (Règles de calcul).
Voici trois relations pour $\mathbf{x}$ *et* $\mathbf{y}$ *éléments de* E *:*

(i) $\|\mathbf{x}+\mathbf{y}\|^2 = \|\mathbf{x}\|^2 + 2\langle \mathbf{x} \mid \mathbf{y}\rangle + \|\mathbf{y}\|^2$ *;*
(ii) $\|\mathbf{x}+\mathbf{y}\|^2 + \|\mathbf{x}-\mathbf{y}\|^2 = 2\|\mathbf{x}\|^2 + 2\|\mathbf{y}\|^2$ *égalité du parallélogramme ;*
(iii) $\langle \mathbf{x} \mid \mathbf{y}\rangle = \frac{1}{2}\big(\|\mathbf{x}+\mathbf{y}\|^2 - \|\mathbf{x}\|^2 - \|\mathbf{y}\|^2\big) = \frac{1}{4}\big(\|\mathbf{x}+\mathbf{y}\|^2 - \|\mathbf{x}-\mathbf{y}\|^2\big)$
expression du produit scalaire réel à l'aide de la norme.

Démonstration.

(i) Utilisons la linéarité à droite et à gauche, et la symétrie du produit scalaire

$$\begin{aligned}\|\mathbf{x}+\mathbf{y}\|^2 &= \langle \mathbf{x}+\mathbf{y} \mid \mathbf{x}+\mathbf{y}\rangle = \langle \mathbf{x} \mid \mathbf{x}\rangle + \langle \mathbf{x} \mid \mathbf{y}\rangle + \langle \mathbf{y} \mid \mathbf{x}\rangle + \langle \mathbf{y} \mid \mathbf{y}\rangle \\ &= \|\mathbf{x}\|^2 + 2\langle \mathbf{x} \mid \mathbf{y}\rangle + \|\mathbf{y}\|^2\end{aligned}$$

(ii) En changeant $\mathbf{y}$ en $-\mathbf{y}$, on obtient

$$\|\mathbf{x}-\mathbf{y}\|^2 = \|\mathbf{x}\|^2 - 2\langle \mathbf{x} \mid \mathbf{y}\rangle + \|\mathbf{y}\|^2$$

Il suffit d'additionner les deux formules pour obtenir le résultat annoncé.

cqfd

La deuxième égalité s'interprète par le

Corollaire (Égalité du parallélogramme).
La somme des carrés des longueurs des côtés d'un parallélogramme est égale au double de la somme des carrés des longueurs des diagonales.

Remarque. L'égalité du parallélogramme caractérise les normes euclidiennes, *i.e.* les normes qui sont associées à un produit scalaire (réel).

2.2 Inégalité de Cauchy-Schwarz

Théorème 2.3 (Inégalité de Cauchy-Schwarz).
Pour $\mathbf{x}$ *et* $\mathbf{y}$ *de* E*, on a*

$$\boxed{\forall(\mathbf{x}, \mathbf{y}) \in E^2,\ |\langle \mathbf{x} \mid \mathbf{y}\rangle| \leqslant \|\mathbf{x}\|\,\|\mathbf{y}\|}$$

L'égalité a lieu si, et seulement si, la famille $(\mathbf{x}, \mathbf{y})$ *est liée.*

Démonstration. Si $\|\mathbf{x}\| = 0$, $\mathbf{x}$ est le vecteur nul, l'inégalité, qui devient une égalité dans ce cas, est vérifiée, et la famille $(\mathbf{x} = \mathbf{0}, \mathbf{y})$ est une famille liée.
Si $\|\mathbf{x}\| \neq 0$, on pose, pour $\lambda \in \mathsf{R}$, $T(\lambda) = \|\lambda\mathbf{x}+\mathbf{y}\|^2 = \lambda^2\|\mathbf{x}\|^2 + 2\lambda\langle\mathbf{x} \mid \mathbf{y}\rangle + \|\mathbf{y}\|^2$. $T(\lambda)$ est un trinôme du second degré, que l'on écrit sous sa forme canonique

$$0 \leqslant T(\lambda) = \|\mathbf{x}\|^2\Big(\lambda + \frac{\langle\mathbf{x} \mid \mathbf{y}\rangle}{\|\mathbf{x}\|^2}\Big)^2 + \frac{\|\mathbf{x}\|^2\|\mathbf{y}\|^2 - \langle\mathbf{x} \mid \mathbf{y}\rangle^2}{\|\mathbf{x}\|^2} \tag{2.1}$$

En donnant la valeur particulière $\lambda_0 = -\frac{\langle\mathbf{x}|\mathbf{y}\rangle}{\|\mathbf{x}\|^2}$, on obtient l'inégalité annoncée.
Dans le cas de l'égalité de Cauchy-Schwarz, on a

$$T(\lambda) = \|\mathbf{x}\|^2\Big(\lambda + \frac{\langle\mathbf{x} \mid \mathbf{y}\rangle}{\|\mathbf{x}\|^2}\Big)^2 \tag{2.2}$$

Donnant à λ la valeur particulière $\lambda_0 = -\frac{\langle\mathbf{x}|\mathbf{y}\rangle}{\|\mathbf{x}\|^2}$, on obtient $0 = T(\lambda_0) = \|\lambda_0\mathbf{x}+\mathbf{y}\|^2$, soit $\lambda_0\mathbf{x}+\mathbf{y} = \mathbf{0}$ et la famille $(\mathbf{x}, \mathbf{y})$ est une famille liée.
Réciproquement, si la famille $(\mathbf{x}, \mathbf{y})$ est une famille liée, par exemple $\mathbf{y} = \mu\mathbf{x}$, alors

$$|\langle\mathbf{x} \mid \mathbf{y}\rangle| = |\langle\mathbf{x} \mid \mu\mathbf{x}\rangle| = |\mu|\langle\mathbf{x} \mid \mathbf{x}\rangle = |\mu|\,\|\mathbf{x}\|^2 = \|\mathbf{x}\|\,\|\mu\mathbf{x}\| = \|\mathbf{x}\|\,\|\mathbf{y}\|$$

cqfd

Exemples 2.1. Voici quelques exemples d'application de l'inégalité de Schwarz :

(i) cas de R^n :

$$|\langle\mathbf{x} \mid \mathbf{y}\rangle| = \Big|\sum_{k=1}^n x_k y_k\Big| \leqslant \|\mathbf{x}\|\,\|\mathbf{y}\| = \Big(\sum_{k=1}^n x_k^2\Big)^{\frac12}\Big(\sum_{k=1}^n y_k^2\Big)^{\frac12}$$

(ii) cas de $\mathcal{M}_{n,1}(\mathsf{R})$:

$$|\langle X \mid Y\rangle| = |{}^tXY| = \Big|\sum_{k=1}^n x_k y_k\Big| \leqslant \big({}^tXX\big)^{\frac12}\big({}^tYY\big)^{\frac12} = \Big(\sum_{k=1}^n x_k^2\Big)^{\frac12}\Big(\sum_{k=1}^n y_k^2\Big)^{\frac12}$$

(iii) cas de $\mathcal{M}_{n,p}(\mathsf{R})$:

$$|\langle A \mid B\rangle| = |\mathrm{tr}({}^tAB)| = \Big|\sum_{i,j} a_{i,j}b_{i,j}\Big| \leqslant \big(\mathrm{tr}({}^tAA)\big)^{\frac12}\big(\mathrm{tr}({}^tBB)\big)^{\frac12} = \Big(\sum_{i,j} a_{i,j}^2\Big)^{\frac12}\Big(\sum_{i,j} b_{i,j}^2\Big)^{\frac12}$$

(iv) cas de $\mathcal{C}([a,b],\mathsf{R})$:

$$|\langle f \mid g\rangle| = \Big|\int_a^b f(t)g(t)\,dt\Big| \leqslant \Big(\int_a^b \big(f(t)\big)^2\,dt\Big)^{\frac12}\Big(\int_a^b \big(g(t)\big)^2\,dt\Big)^{\frac12}$$

(v) cas de $\mathcal{L}^2(I,\mathsf{R})$:

$$|\langle f \mid g\rangle| = \Big|\int_I f(t)g(t)\,dt\Big| \leqslant \Big(\int_I \big(f(t)\big)^2\,dt\Big)^{\frac12}\Big(\int_I \big(g(t)\big)^2\,dt\Big)^{\frac12}$$

(vi) cas de $\mathsf{R}[X]$:

$$|\langle P \mid Q\rangle| = \Big|\int_I P(t)Q(t)w(t)\,dt\Big| \leqslant \Big(\int_I \big(P(t)\big)^2 w(t)\,dt\Big)^{\frac12}\Big(\int_I \big(Q(t)\big)^2 w(t)\,dt\Big)^{\frac12}$$

Par exemple :

$$\Big|\int_{-\infty}^{+\infty} P(t)Q(t)e^{-t^2}\,dt\Big| \leqslant \Big(\int_{-\infty}^{+\infty} \big(P(t)\big)^2 e^{-t^2}\,dt\Big)^{\frac12}\Big(\int_{-\infty}^{+\infty} \big(Q(t)\big)^2 e^{-t^2}\,dt\Big)^{\frac12}$$

(vii) cas de $\ell^2_{\mathsf{N}}(\mathsf{R})$:

$$|\langle \mathbf{a} \mid \mathbf{b}\rangle| = \Big|\sum_{k=0}^{\infty} a_k b_k\Big| \leqslant \Big(\sum_{k=0}^{\infty} a_k^2\Big)^{\frac{1}{2}} \Big(\sum_{k=0}^{\infty} b_k^2\Big)^{\frac{1}{2}}$$

L'inégalité de Schwarz montre que, pour deux vecteurs non nuls $\mathbf{x}$ et $\mathbf{y}$ de E, le quotient $\frac{\langle \mathbf{x}|\mathbf{y}\rangle}{\|\mathbf{x}\|\,\|\mathbf{y}\|}$ est un réel de $[-1, 1]$, réel que l'on écrit $\cos\theta$, pour un θ unique du segment $[0, \pi]$; ce qui donne la

Définition 2.2 (Écart angulaire entre deux vecteurs réels).
Si $\mathbf{x}$ et $\mathbf{y}$ sont deux vecteurs non nuls d'un espace préhilbertien réel, il existe un unique $\theta \in [0, \pi]$ tel que

$$\langle \mathbf{x} \mid \mathbf{y}\rangle = \|\mathbf{x}\|\,\|\mathbf{y}\| \cos\theta$$

θ est appelé *l'angle (non orienté) entre* $\mathbf{x}$ *et* $\mathbf{y}$; cet angle est défini à π près.

2.3 Inégalité de Minkowski, ou inégalité triangulaire

Proposition 2.4 (Inégalité de Minkowski).
On a l'inégalité, dite de MinKowski,

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

Démonstration. Développons $\|\mathbf{x}+\mathbf{y}\|^2$ et utilisons l'inégalité de Schwarz :

$$\|\mathbf{x}+\mathbf{y}\|^2 = \|\mathbf{x}\|^2 + 2\langle \mathbf{x} \mid \mathbf{y}\rangle + \|\mathbf{y}\|^2 \leqslant \|\mathbf{x}\|^2 + 2\|\mathbf{x}\|\,\|\mathbf{y}\| + \|\mathbf{y}\|^2 = \big(\|\mathbf{x}\| + \|\mathbf{y}\|\big)^2$$

cqfd

Corollaire. $\mathbf{x} \mapsto \|\mathbf{x}\| = \sqrt{\langle \mathbf{x} \mid \mathbf{x}\rangle}$ *est une norme sur* E.

3 Produit scalaire sur un espace vectoriel complexe

Dans cette section, E désigne un C-espace vectoriel de dimension finie ou infinie, sauf avis contraire.

Remarque. Sur R, l'égalité $x_1^2 + x_2^2 = 0$ est équivalente à $x_1 = 0$ *et* $x_2 = 0$. Sur C, la situation est différente ; on a :

$$0 = z_1^2 + z_2^2 = (z_1 + iz_2)(z_1 - iz_2) \iff z_1 + iz_2 = 0 \textbf{ ou } z_1 - iz_2 = 0$$

tandis que

$$0 = |z_1|^2 + |z_2|^2 = z_1\overline{z_1} + z_2\overline{z_2} \iff z_1 = 0 \textbf{ et } z_2 = 0$$

Définition 3.1 (Produit scalaire complexe ou hermitien).
On appelle *produit scalaire complexe* ou *produit scalaire hermitien* sur E, toute forme *sesquilinéaire à symétrie hermitienne* et définie positive, *i.e.* toute application $\varphi : E \times E \to \mathsf{C}$ telle que

- (i) $\forall \mathbf{x} \in E,\ \varphi_{\mathbf{x}} : \mathbf{x} \mapsto \varphi(\mathbf{x}, \mathbf{y})$ est linéaire — *linéarité à droite* ;
- (ii) $\forall(\mathbf{x}, \mathbf{y}) \in E^2,\ \varphi(\mathbf{x}, \mathbf{y}) = \overline{\varphi(\mathbf{y}, \mathbf{x})}$ — *symétrie hermitienne* ;
- (iii) $\forall \mathbf{x} \in E,\ \mathbf{x} \neq \mathbf{0} \implies \varphi(\mathbf{x}, \mathbf{x}) > 0$ — *définie positive.*

Définitions 3.2 (Espace préhilbertien complexe, espace hermitien).
E muni du produit scalaire φ est appelé *espace préhilbertien complexe*. Si E est un espace de dimension finie, (E, φ) est un *espace hermitien*.
Le produit scalaire $\varphi(\mathbf{x}, \mathbf{y})$ de deux vecteurs est noté $\langle \mathbf{x} \mid \mathbf{y}\rangle$, ou encore $\mathbf{x} \cdot \mathbf{y}$, $<\mathbf{x}, \mathbf{y}>$, $(\mathbf{x} \,|\, \mathbf{y})$, ...

Remarques.
La linéarité à droite et la symétrie hermitienne impliquent la *semi-linéarité* à gauche, *i.e.* pour tous nombres complexes λ_1 et λ_2, pour tous vecteurs $\mathbf{x}_1$, $\mathbf{x}_2$ et $\mathbf{y}$ de E,

$$\begin{aligned} \langle \lambda_1 \mathbf{x}_1 + \lambda_2 \mathbf{x}_2 \mid \mathbf{y} \rangle &= \overline{\langle \mathbf{y} \mid \lambda_1 \mathbf{x}_1 + \lambda_2 \mathbf{x}_2 \rangle} && \text{symétrie hermitienne} \\ &= \overline{\lambda_1 \langle \mathbf{y} \mid \mathbf{x}_1 \rangle + \lambda_2 \langle \mathbf{y} \mid \mathbf{x}_2 \rangle} && \text{linéarité à droite} \\ &= \overline{\lambda_1} \langle \mathbf{x}_1 \mid \mathbf{y} \rangle + \overline{\lambda_2} \langle \mathbf{x}_2 \mid \mathbf{y} \rangle && \text{symétrie hermitienne} \end{aligned}$$

Si l'un des vecteurs est nul, le produit scalaire est nul.
Le caractère « défini positif » du produit scalaire peut s'établir en montrant que

$$\forall \mathbf{x} \in E,\ \langle \mathbf{x} \mid \mathbf{x} \rangle \geqslant 0 \text{ et } \langle \mathbf{x} \mid \mathbf{x} \rangle = 0 \implies \mathbf{x} = \mathbf{0}$$

Si F est un sous-espace vectoriel de E, tout produit scalaire hermitien sur E induit un produit scalaire hermitien sur F.

Exemples 3.1. Reprenons les mêmes exemples que dans le cas réel, arrangés à la sauce complexe par l'utilisation de la conjugaison de la première variable.

(i) Produit scalaire canonique sur C^n : il est défini par

$$\forall \mathbf{x} = (x_1, \ldots, x_n),\ \forall \mathbf{y} = (y_1, \ldots, y_n), \quad \langle \mathbf{x} \mid \mathbf{y} \rangle = \sum_{k=1}^{n} \overline{x_k} y_k$$

(ii) Produit scalaire canonique sur $\mathcal{M}_{n,1}(\mathsf{C})$: il est défini par

$$\forall (X, Y) \in \mathcal{M}_{n,1}(\mathsf{C})^2, \quad \langle X \mid Y \rangle = {}^t\overline{X} Y = \sum_{k=1}^{n} \overline{x_k} y_k$$

(iii) Produit scalaire canonique sur $\mathcal{M}_{n,p}(\mathsf{C})$: il est défini par

$$\forall (A, B) \in \big(\mathcal{M}_{n,p}(\mathsf{C})\big)^2, \quad \langle A \mid B \rangle = \operatorname{tr}({}^t\overline{A} B) = \sum_{i,j} \overline{a_{i,j}} b_{i,j}$$

(iv) Produit scalaire sur l'espace des fonctions continues sur le segment $[a, b]$ et à valeurs complexes :

$$\forall (f, g) \in \big(\mathcal{C}([a, b], \mathsf{C})\big)^2, \quad \langle f \mid g \rangle = \int_a^b \overline{f(t)} g(t)\, dt$$

(v) Produit scalaire sur l'espace des fonctions continues, 2π-périodiques sur R et à valeurs complexes :

$$\forall (f, g) \in \big(\mathcal{C}_{2\pi}\big)^2, \quad \langle f \mid g \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} \overline{f(t)} g(t)\, dt$$

(vi) Produit scalaire sur l'espace des fonctions continues, de carré intégrable sur l'intervalle I et à valeurs complexes :

$$\forall (f, g) \in \big(\mathcal{L}^2(I)\big)^2, \quad \langle f \mid g \rangle = \int_I \overline{f(t)} g(t)\, dt$$

(vii) Produits scalaires sur l'espace des polynômes à coefficients complexes ; si I est un intervalle et w une fonction continue sur I et à valeurs réelles (strictement) positives, telle que, pour tout entier n, l'application $t \mapsto t^n w(t)$ soit intégrable sur I, l'application

$$(P, Q) \in \big(\mathsf{C}[X]\big)^2 \mapsto \langle P \mid Q \rangle = \int_I \overline{P(t)} Q(t)\, w(t)\, dt$$

est un produit scalaire. Les cas classiques sont

$$I = [-1,1],\ w(t) = 1 \text{ et } \langle P \mid Q\rangle = \int_{-1}^{1} \overline{P(t)}Q(t)\,dt$$
$$I =]-1,1[,\ w(t) = \frac{1}{\sqrt{1-t^2}} \text{ et } \langle P \mid Q\rangle = \int_{-1}^{1} \overline{P(t)}Q(t)\,\frac{dt}{\sqrt{1-t^2}}$$
$$I = [0,+\infty[,\ w(t) = e^{-t} \text{ et } \langle P \mid Q\rangle = \int_{0}^{+\infty} \overline{P(t)}Q(t)\,e^{-t}\,dt$$
$$I =]-\infty,+\infty[,\ w(t) = e^{-t^2} \text{ et } \langle P \mid Q\rangle = \int_{-\infty}^{+\infty} \overline{P(t)}Q(t)\,e^{-t^2}\,dt$$

$(viii)$ Produit scalaire sur l'espace des suites complexes de carré sommable $\ell^2_{\mathbb{N}}(\mathbb{C})$, *i.e.* les suites complexes $(a_n)_n$ telles que la série $\sum_n |a_n|^2$ soit convergente :

$$\forall(\mathbf{a},\mathbf{b}) \in \left(\ell^2_{\mathbb{N}}(\mathbb{C})\right)^2, \quad \langle \mathbf{a} \mid \mathbf{b}\rangle = \sum_{k=0}^{\infty} \overline{a_k} b_k$$

4 Norme et distance associées à un produit scalaire complexe

E désigne un espace préhilbertien complexe dont le produit scalaire est noté $\langle\ \mid\ \rangle$.

Définitions 4.1 (Norme et distance hermitienne).
Les définitions sont identiques au cas réel. La norme associée au produit scalaire hermitien $\langle\ \mid\ \rangle$ est définie par :

$$\boxed{\forall \mathbf{x} \in E, \quad \|\mathbf{x}\| = \sqrt{\langle \mathbf{x} \mid \mathbf{x}\rangle}}$$

La distance associée au produit scalaire hermitien est définie par :

$$\boxed{\forall(\mathbf{x},\mathbf{y}) \in E^2, \quad \mathrm{d}(\mathbf{x},\mathbf{y}) = \|\mathbf{y}-\mathbf{x}\| = \sqrt{\langle \mathbf{y}-\mathbf{x} \mid \mathbf{y}-\mathbf{x}\rangle}}$$

Dans ce cas complexe, on donne le qualificatif d'*hermitienne* à la norme et à la distance associées au produit scalaire (hermitien).

Proposition 4.1. $\mathbf{x} \mapsto \|\mathbf{x}\| = \sqrt{\langle \mathbf{x} \mid \mathbf{x}\rangle}$ *est une application de* E *sur* $[0,+\infty[$ *qui vérifie*

(i) $\forall \mathbf{x} \in E,\ \|\mathbf{x}\| = 0 \iff \mathbf{x} = \mathbf{0}$ — axiome de séparation ;
(ii) $\forall(\lambda,\mathbf{x}) \in \mathbb{C}\times E,\ \|\lambda\mathbf{x}\| = |\lambda|\,\|\mathbf{x}\|$ — axiome d'homogénéité.

Démonstration.

(i) $0 = \|\mathbf{x}\|^2 = \langle \mathbf{x} \mid \mathbf{x}\rangle \iff \mathbf{x} = \mathbf{0}$;
(ii) $\|\lambda\mathbf{x}\| = \sqrt{\langle \lambda\mathbf{x} \mid \lambda\mathbf{x}\rangle} = \sqrt{|\lambda|^2\langle \mathbf{x} \mid \mathbf{x}\rangle} = |\lambda|\langle \mathbf{x} \mid \mathbf{x}\rangle$

cqfd

L'inégalité triangulaire sera démontrée à la fin de cette section.

4.1 Expression du produit scalaire en fonction de la norme

Proposition 4.2 (Règles de calcul).
Voici des relations pour $\mathbf{x}$ *et* $\mathbf{y}$ *éléments de* E *:*

(i) $\|\mathbf{x}+\mathbf{y}\|^2 = \|\mathbf{x}\|^2 + 2\,\Re\mathrm{e}\langle \mathbf{x} \mid \mathbf{y}\rangle + \|\mathbf{y}\|^2$ *;*

(ii) $\|\mathbf{x}+\mathbf{y}\|^2 + \|\mathbf{x}-\mathbf{y}\|^2 = 2\|\mathbf{x}\|^2 + 2\|\mathbf{y}\|^2$ *égalité du parallélogramme ;*

(iii) $\Re\mathrm{e}\langle \mathbf{x} \mid \mathbf{y}\rangle = \frac{1}{2}\big(\|\mathbf{x}+\mathbf{y}\|^2 - \|\mathbf{x}\|^2 - \|\mathbf{y}\|^2\big) = \frac{1}{4}\big(\|\mathbf{x}+\mathbf{y}\|^2 - \|\mathbf{x}-\mathbf{y}\|^2\big)$

$\Im\mathrm{m}\langle \mathbf{x} \mid \mathbf{y}\rangle = \frac{1}{2}\big(\|\mathbf{x}-i\mathbf{y}\|^2 - \|\mathbf{x}\|^2 - \|\mathbf{y}\|^2\big) = \frac{1}{4}\big(\|\mathbf{x}-i\mathbf{y}\|^2 - \|\mathbf{x}+i\mathbf{y}\|^2\big)$

et $\langle \mathbf{x} \mid \mathbf{y}\rangle = \frac{1}{4}\big(\|\mathbf{x}+\mathbf{y}\|^2 - \|\mathbf{x}-\mathbf{y}\|^2\big) + \frac{i}{4}\big(\|\mathbf{x}-i\mathbf{y}\|^2 - \|\mathbf{x}+i\mathbf{y}\|^2\big)$

expression du produit scalaire complexe à l'aide de la norme.

Démonstration.

(i) Utilisons la linéarité à droite, la semi-linéarité à gauche et la symétrie hermitienne du produit scalaire

$$\begin{aligned}\|\mathbf{x}+\mathbf{y}\|^2 &= \langle \mathbf{x}+\mathbf{y} \mid \mathbf{x}+\mathbf{y}\rangle = \langle \mathbf{x} \mid \mathbf{x}\rangle + \langle \mathbf{x} \mid \mathbf{y}\rangle + \langle \mathbf{y} \mid \mathbf{x}\rangle + \langle \mathbf{y} \mid \mathbf{y}\rangle \\ &= \|\mathbf{x}\|^2 + \langle \mathbf{x} \mid \mathbf{y}\rangle + \overline{\langle \mathbf{x} \mid \mathbf{y}\rangle} + \|\mathbf{y}\|^2 \\ &= \|\mathbf{x}\|^2 + 2\,\Re\mathrm{e}\langle \mathbf{x} \mid \mathbf{y}\rangle + \|\mathbf{y}\|^2\end{aligned}$$

(ii) En changeant $\mathbf{y}$ en $-\mathbf{y}$, on obtient

$$\|\mathbf{x}-\mathbf{y}\|^2 = \|\mathbf{x}\|^2 - 2\,\Re\mathrm{e}\langle \mathbf{x} \mid \mathbf{y}\rangle + \|\mathbf{y}\|^2$$

Il suffit d'additionner les deux formules pour obtenir le résultat annoncé.

(iii) Changeons $\mathbf{y}$ en $-i\mathbf{y}$; on obtient, en utilisant $\Re\mathrm{e}(-iz) = \Im\mathrm{m}\, z$,

$$\|\mathbf{x}-i\mathbf{y}\|^2 = \|\mathbf{x}\|^2 + 2\,\Re\mathrm{e}(-i\langle \mathbf{x} \mid \mathbf{y}\rangle) + \|\mathbf{y}\|^2 = \|\mathbf{x}\|^2 + 2\,\Im\mathrm{m}\langle \mathbf{x} \mid \mathbf{y}\rangle + \|\mathbf{y}\|^2$$

Par addition, on obtient les identités demandées.

cqfd

La deuxième égalité s'interprète toujours par le

Corollaire (Égalité du parallélogramme).
La somme des carrés des longueurs des côtés d'un parallélogramme est égale au double de la somme des carrés des longueurs des diagonales.

Remarque. L'égalité du parallélogramme caractérise les normes hermitiennes, *i.e.* les normes associées à un produit scalaire complexe (ou hermitien).

4.2 Inégalité de Cauchy-Schwarz

Théorème 4.3 (Inégalité de Cauchy-Schwarz).
Pour $\mathbf{x}$ *et* $\mathbf{y}$ *de* E, *on a*

$$\boxed{\forall(\mathbf{x},\mathbf{y}) \in E^2,\ |\langle \mathbf{x} \mid \mathbf{y}\rangle| \leqslant \|\mathbf{x}\|\,\|\mathbf{y}\|}$$

L'égalité a lieu si, et seulement si, la famille $(\mathbf{x},\mathbf{y})$ *est liée.*

Démonstration. Si $\|\mathbf{x}\| = 0 \ldots$ voir le cas réel.
Si $\|\mathbf{x}\| \neq 0$, on pose $T(\lambda) = \|\lambda\mathbf{x}+\mathbf{y}\|^2$ pour $\lambda \in \mathbb{C}$. On développe et on trouve :

$$\begin{aligned}0 \leqslant T(\lambda) = \|\lambda\mathbf{x}+\mathbf{y}\|^2 &= \langle \lambda\mathbf{x}+\mathbf{y} \mid \lambda\mathbf{x}+\mathbf{y}\rangle = \lambda\overline{\lambda}\|\mathbf{x}\|^2 + \overline{\lambda}\langle \mathbf{x} \mid \mathbf{y}\rangle + \lambda\overline{\langle \mathbf{x} \mid \mathbf{y}\rangle} + \|\mathbf{y}\|^2 \\ &= \|\mathbf{x}\|^2\Big(\overline{\lambda} + \frac{\overline{\langle \mathbf{x} \mid \mathbf{y}\rangle}}{\|\mathbf{x}\|^2}\Big)\Big(\lambda + \frac{\langle \mathbf{x} \mid \mathbf{y}\rangle}{\|\mathbf{x}\|^2}\Big) + \|\mathbf{y}\|^2 - \frac{\langle \mathbf{x} \mid \mathbf{y}\rangle\overline{\langle \mathbf{x} \mid \mathbf{y}\rangle}}{\|\mathbf{x}\|^2} \\ &= \|\mathbf{x}\|^2\Big|\lambda + \frac{\langle \mathbf{x} \mid \mathbf{y}\rangle}{\|\mathbf{x}\|^2}\Big|^2 + \frac{\|\mathbf{x}\|^2\|\mathbf{y}\|^2 - |\langle \mathbf{x} \mid \mathbf{y}\rangle|^2}{\|\mathbf{x}\|^2}\end{aligned}$$

$T(\lambda)$ est un trinôme du second degré en la variable *complexe* λ, que l'on a écrit sous sa forme canonique. En donnant la valeur particulière $\lambda_0 = -\frac{\langle \mathbf{x} \mid \mathbf{y}\rangle}{\|\mathbf{x}\|^2}$, on obtient l'inégalité annoncée.
Le reste de la démonstration se traite comme dans le cas réel. cqfd

Exemples 4.1. Toujours les mêmes exemples d'application de l'inégalité de Cauchy-Schwarz ; il suffit d'ajouter une pincée de condiment « conjugaison » sur la première variable et le plat est prêt.

(i) cas de $\mathbb{C}^n$:

$$|\langle \mathbf{x} \mid \mathbf{y}\rangle| = \Big|\sum_{k=1}^{n} \overline{x_k} y_k\Big| \leqslant \|\mathbf{x}\|\,\|\mathbf{y}\| = \Big(\sum_{k=1}^{n}|x_k|^2\Big)^{\frac12}\Big(\sum_{k=1}^{n}|y_k|^2\Big)^{\frac12}$$

(ii) cas de $\mathcal{M}_{n,1}(\mathbb{C})$:

$$|\langle X \mid Y\rangle| = |{}^t\overline{X}Y| = Big|\sum_{k=1}^{n} \overline{x_k} y_k Big| \leqslant \big({}^t\overline{X}X\big)^{\frac12}\big({}^t\overline{Y}Y\big)^{\frac12} = \Big(\sum_{k=1}^{n}|x_k|^2\Big)^{\frac12}\Big(\sum_{k=1}^{n}|y_k|^2\Big)^{\frac12}$$

(iii) cas de $\mathcal{M}_{n,p}(\mathbb{C})$:

$$|\langle A \mid B\rangle| = |\mathrm{tr}({}^t\overline{A}B)| = \Big|\sum_{i,j} a_{i,j} b_{i,j}\Big| \leqslant \big(\mathrm{tr}({}^t\overline{A}A)\big)^{\frac12}\big(\mathrm{tr}({}^t\overline{B}B)\big)^{\frac12} = \Big(\sum_{i,j}|a_{i,j}|^2\Big)^{\frac12}\Big(\sum_{i,j}|b_{i,j}|^2\Big)^{\frac12}$$

(iv) cas de $\mathcal{C}([a,b],\mathbb{C})$:

$$|\langle f \mid g\rangle| = \Big|\int_a^b \overline{f(t)} g(t)\,dt\Big| \leqslant \Big(\int_a^b \big|f(t)\big|^2\,dt\Big)^{\frac12}\Big(\int_a^b \big|g(t)\big|^2\,dt\Big)^{\frac12}$$

(v) cas de $\mathcal{L}^2(I)$:

$$|\langle f \mid g\rangle| = \Big|\int_I \overline{f(t)} g(t)\,dt\Big| \leqslant \Big(\int_I \big|f(t)\big|^2\,dt\Big)^{\frac12}\Big(\int_I \big|g(t)\big|^2\,dt\Big)^{\frac12}$$

(vi) cas de $\mathbb{C}[X]$:

$$|\langle P \mid Q\rangle| = \Big|\int_I \overline{P(t)} Q(t) w(t)\,dt\Big| \leqslant \Big(\int_I \big|P(t)\big|^2 w(t)\,dt\Big)^{\frac12}\Big(\int_I \big|Q(t)\big|^2 w(t)\,dt\Big)^{\frac12}$$

Par exemple :

$$\Big|\int_{-1}^{1} \overline{P(t)} Q(t)\,\frac{dt}{\sqrt{1-t^2}}\Big| \leqslant \Big(\int_{-1}^{1} \big|P(t)\big|^2\,\frac{dt}{\sqrt{1-t^2}}\Big)^{\frac12}\Big(\int_{-1}^{1} \big|Q(t)\big|^2\,\frac{dt}{\sqrt{1-t^2}}\Big)^{\frac12}$$

(vii) cas de $\ell^2_{\mathbb{N}}(\mathbb{C})$:

$$|\langle \mathbf{a} \mid \mathbf{b}\rangle| = \Big|\sum_{k=0}^{\infty} \overline{a_k} b_k\Big| \leqslant \Big(\sum_{k=0}^{\infty}|a_k|^2\Big)^{\frac12}\Big(\sum_{k=0}^{\infty}|b_k|^2\Big)^{\frac12}$$

L'inégalité de Schwarz montre que, pour deux vecteurs non nuls $\mathbf{x}$ et $\mathbf{y}$ de E, le quotient $\frac{\langle \mathbf{x}|\mathbf{y}\rangle}{\|\mathbf{x}\|\,\|\mathbf{y}\|}$ est un nombre *complexe* de module inférieur ou égal à 1 ; on ne peut plus parler d'écart angulaire entre deux vecteurs d'un espace vectoriel complexe.

4.3 Inégalité de Minkowski, ou inégalité triangulaire

Proposition 4.4. *Pour* $\mathbf{x}$ *et* $\mathbf{y}$ *de* E, *on a*

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

Démonstration. Démonstration identique au cas réel :

$$\|\mathbf{x}+\mathbf{y}\|^2 = \|\mathbf{x}\|^2 + 2\,\Re\mathrm{e}\langle \mathbf{x} \mid \mathbf{y}\rangle + \|\mathbf{y}\|^2 \leqslant \|\mathbf{x}\|^2 + 2\|\mathbf{x}\|\,\|\mathbf{y}\| + \|\mathbf{y}\|^2 = \big(\|\mathbf{x}\| + \|\mathbf{y}\|\big)^2$$

cqfd

Corollaire. $\mathbf{x} \mapsto \|\mathbf{x}\| = \sqrt{\langle \mathbf{x} \mid \mathbf{x}\rangle}$ *est une norme sur* E.

5 Orthogonalité

E désigne un espace préhilbertien réel ou complexe dont le produit scalaire est noté $\langle\ |\ \rangle$.

Définition 5.1 (Vecteur unitaire).
Un vecteur *unitaire* est un vecteur $\mathbf{x}$ de norme 1, *i.e.* vérifiant $\langle \mathbf{x} \mid \mathbf{x} \rangle = 1$.

Définition 5.2 (Orthogonalité).
Deux vecteurs sont *orthogonaux* si leur produit scalaire est nul ; la relation d'orthogonalité est notée $\perp$.

$$\boxed{\mathbf{x} \perp \mathbf{y} \iff \langle \mathbf{x} \mid \mathbf{y} \rangle = 0}$$

Définition 5.3 (Famille orthogonale).
Une famille de vecteurs $(\mathbf{x}_k)_{k\in\Lambda}$ est une *famille orthogonale* si les vecteurs de cette famille sont orthogonaux deux à deux, *i.e.*

$$\forall (k,l) \in \Lambda^2, \quad k \neq l \implies \langle \mathbf{x}_k \mid \mathbf{x}_l \rangle = 0 \tag{5.1}$$

Définition 5.4 (Famille orthonormale).
Une famille orthogonale de vecteurs unitaires est appelée une famille *orthonormale*, *i.e.*

$$\forall (k,l) \in \Lambda^2, \quad \langle \mathbf{x}_k \mid \mathbf{x}_l \rangle = \delta_{k,l} = \begin{cases} 0 & \text{si } k \neq l \\ 1 & \text{si } k = l \end{cases} \tag{5.2}$$

Remarque. Si $(\mathbf{v}_k)_k$ est une famille orthogonale de vecteurs *non nuls*, la famille $(\frac{1}{\|\mathbf{v}_k\|}\mathbf{v}_k)_k$ est orthogonale.

Exemples 5.1.

(i) Les bases naturelles (canoniques) de R^n, C^n, $\mathcal{M}_{n,1}(\mathsf{R})$, $\mathcal{M}_{n,1}(\mathsf{C})$, $\mathcal{M}_{n,p}(\mathsf{R})$ et $\mathcal{M}_{n,p}(\mathsf{C})$ sont des familles orthonormales pour le produit scalaire naturel (canonique) des espaces considérés.

(ii) La famille $\{e_k : t \mapsto \exp(ikt) \ /\ k \in \mathsf{Z}\}$ est une famille orthonormale de $\mathcal{C}_{2\pi}$.
La famille $\{1\} \cup \{t \mapsto \cos kt \ /\ k \in \mathsf{N}^*\} \cup \{t \mapsto \sin kt \ /\ k \in \mathsf{N}^*\}$ est une famille orthogonale de $\mathcal{C}_{2\pi}$; la famille orthonormale associée est $\{1\} \cup \{t \mapsto \sqrt{2}\cos kt \ /\ k \in \mathsf{N}^*\} \cup \{t \mapsto \sqrt{2}\sin kt \ /\ k \in \mathsf{N}^*\}$.

(iii) Les divers produits scalaires sur $\mathsf{R}[X]$ et $\mathsf{C}[X]$ donnent des familles orthogonales de polynômes, encore appelées familles de polynômes orthogonaux :

- les polynômes de Legendre $P_n(t) = \dfrac{d^n}{dt^n}\big((t^2-1)^n\big)$ constituent une famille de polynômes orthogonaux pour le produit scalaire $\langle P \mid Q \rangle = \displaystyle\int_{-1}^{1} \overline{P(t)}\,Q(t)\,dt$.
- les polynômes de Tchebychev $T_n(t) = \cos(n \arccos t)$ constituent une famille de polynômes orthogonaux pour le produit scalaire $\langle P \mid Q \rangle = \displaystyle\int_{-1}^{1} \overline{P(t)}\,Q(t)\,\frac{dt}{\sqrt{1-t^2}}$.
- les polynômes de Laguerre $L_n(t) = \dfrac{d^n}{dt^n}(t^n e^{-t})$ constituent une famille de polynômes orthogonaux pour le produit scalaire $\langle P \mid Q \rangle = \displaystyle\int_{0}^{+\infty} \overline{P(t)}\,Q(t)\,e^{-t}\,dt$.
- les polynômes d'Hermite $H_n(t) = e^{t^2}\dfrac{d^n}{dt^n}(e^{-t^2})$ constituent une famille de polynômes orthogonaux pour le produit scalaire $\langle P \mid Q \rangle = \displaystyle\int_{-\infty}^{+\infty} \overline{P(t)}\,Q(t)\,e^{-t^2}\,dt$.

5.1 Relation de Pythagore

Voici tout d'abord, deux règles de calcul :

Lemme 5.1.

(i) $\langle \sum_{k=1}^{p} \lambda_k \mathbf{x}_k \mid \sum_{l=1}^{q} \mu_l \mathbf{y}_l \rangle = \sum_{k=1}^{p} \sum_{l=1}^{q} \overline{\lambda_k} \mu_l \langle \mathbf{x}_k \mid \mathbf{y}_l \rangle$

(ii) $\left\| \sum_{k=1}^{p} \lambda_k \mathbf{x}_k \right\|^2 = \sum_{k=1}^{p} \sum_{l=1}^{p} \overline{\lambda_k} \lambda_l \langle \mathbf{x}_k \mid \mathbf{x}_l \rangle$

Démonstration. Y'a qu'à développer en utilisant la linéarité à droite et la semi-linéarité à gauche du produit scalaire.

(i) $\langle \sum_{k=1}^{p} \lambda_k \mathbf{x}_k \mid \sum_{l=1}^{q} \mu_l \mathbf{y}_l \rangle = \sum_{k=1}^{p} \overline{\lambda_k} \langle \mathbf{x}_k \mid \sum_{l=1}^{q} \mu_l \mathbf{y}_l \rangle = \sum_{k=1}^{p} \sum_{l=1}^{q} \overline{\lambda_k} \mu_l \langle \mathbf{x}_k \mid \mathbf{y}_l \rangle$

(ii) Attention de ne pas oublier d'utiliser deux indices !

$$\left\| \sum_{k=1}^{p} \lambda_k \mathbf{x}_k \right\|^2 = \langle \sum_{k=1}^{p} \lambda_k \mathbf{x}_k \mid \sum_{l=1}^{p} \mu_l \mathbf{x}_l \rangle = \sum_{k=1}^{p} \sum_{l=1}^{p} \overline{\lambda_k} \lambda_l \langle \mathbf{x}_k \mid \mathbf{x}_l \rangle$$

cqfd

Remarque. En particulier, dans un espace préhilbertien réel :

$$\|\mathbf{x} + \mathbf{y} + \mathbf{z}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 + \|\mathbf{z}\|^2 + 2\langle \mathbf{x} \mid \mathbf{y} \rangle + 2\langle \mathbf{y} \mid \mathbf{z} \rangle + 2\langle \mathbf{z} \mid \mathbf{x} \rangle$$

Théorème 5.2 (de Pythagore).

(i) *Si* $\mathbf{u}$ *et* $\mathbf{v}$ *sont deux vecteurs* orthogonaux*, on a*

$$\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 \tag{5.3}$$

La réciproque est vraie dans un espace préhilbertien réel.

(ii) *Si la famille* $(\mathbf{v}_k)_{1 \leqslant k \leqslant p}$ *est une famille* orthogonale*, alors*

$$\left\| \sum_{k=1}^{p} \mathbf{v}_k \right\|^2 = \sum_{k=1}^{p} \|\mathbf{v}_k\|^2 \tag{5.4}$$

Démonstration.

(i) Cas réel : $\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + 2\langle \mathbf{u} \mid \mathbf{v} \rangle + \|\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$ si, et seulement si, $\langle \mathbf{u} \mid \mathbf{v} \rangle = 0$.
Cas complexe : $\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + 2\,\Re\mathrm{e} \langle \mathbf{u} \mid \mathbf{v} \rangle + \|\mathbf{v}\|^2$ et $\langle \mathbf{u} \mid \mathbf{v} \rangle = 0$ implique $\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$.

(ii) Puisque les vecteurs sont orthogonaux deux à deux, $\langle \mathbf{v}_k \mid \mathbf{v}_l \rangle = 0$ pour $k \neq l$ et

$$\left\| \sum_{k=1}^{p} \mathbf{v}_k \right\|^2 = \sum_{k,l} \langle \mathbf{v}_k \mid \mathbf{v}_l \rangle = \sum_{k=1}^{p} \|\mathbf{v}_k\|^2 + \sum_{\substack{k,l=1 \\ k \neq l}}^{p} \langle \mathbf{v}_k \mid \mathbf{v}_l \rangle = \sum_{k=1}^{p} \|\mathbf{v}_k\|^2$$

cqfd

Théorème 5.3 (Famille orthogonale et famille libre).
Toute famille orthogonale de vecteurs non nuls est une famille libre.

Démonstration. Soit $(\mathbf{v}_k)_k$ une famille orthogonale de vecteurs non nuls ; pour toute combinaison linéaire nulle $\sum_{k=1}^{p} \lambda_k \mathbf{v}_k = \mathbf{0}$, on a, en utilisant la formule de Pythagore,

$$0 = \left\| \sum_{k=1}^{p} \lambda_k \mathbf{v}_k \right\|^2 = \sum_{k=1}^{p} \|\lambda_k \mathbf{v}_k\|^2 = \sum_{k=1}^{p} |\lambda|^2 \|\mathbf{v}_k\|^2$$

ce qui montre que $|\lambda_k|^2 \|\mathbf{v}_k\|^2 = 0$ pour tout $k \in [\![1, p]\!]$, et $\lambda_k = 0$ puisque les vecteurs $\mathbf{v}_k$ sont tous non nuls. cqfd

Corollaire. *Toute famille orthonormale est une famille libre.*

5.2 Procédé d'orthonormalisation de Schmidt

Comment construire une famille orthonormale à partir d'une famille libre ? C'est l'objet du procédé d'orthonormalisation d'Erhard Schmidt.

Théorème 5.4. *Si* $(\mathbf{f}_k)_{k \geqslant 1}$ *est une suite libre d'un espace préhilbertien réel ou complexe, il existe une unique suite* orthonormale $(\mathbf{u}_k)_{k \geqslant 1}$ *telle que, pour tout entier* $p \geqslant 1$,

(i) *les espaces engendrés par les familles* $(\mathbf{f}_1, \dots, \mathbf{f}_p)$ *et* $(\mathbf{u}_1, \dots, \mathbf{u}_p)$ *sont identiques ;*

(ii) $\langle \mathbf{f}_p \mid \mathbf{u}_p \rangle > 0$.

Démonstration. Par récurrence sur p.
La propriété est vraie pour $p = 1$. Posons $\mathbf{u}_1 = \lambda \mathbf{f}_1$; puisque $0 < \langle \mathbf{f}_1 \mid \mathbf{u}_1 \rangle = \lambda \langle \mathbf{f}_1 \mid \mathbf{f}_1 \rangle$, le scalaire λ est > 0, $1 = \|\mathbf{u}_1\| = \lambda \|\mathbf{f}_1\|$ ce qui détermine λ. Ainsi $\mathbf{u}_1 = \frac{1}{\|\mathbf{f}_1\|}\mathbf{f}_1$.
La propriété est héréditaire.
Commençons par analyser la situation. Puisque les sous-espaces engendrés par $(\mathbf{f}_1, \dots, \mathbf{f}_p)$ et $(\mathbf{u}_1, \dots, \mathbf{u}_p)$ sont identiques, on peut poser

$$\mathbf{u}_{p+1} = \lambda \mathbf{f}_{p+1} + \sum_{k=1}^{p} \alpha_k \mathbf{u}_k \tag{5.5}$$

L'orthogonalité de $\mathbf{u}_k$ avec $\mathbf{u}_{p+1}$ pour tout $k \in [\![1, p]\!]$ montre que

$$\begin{aligned} 0 = \langle \mathbf{u}_k \mid \mathbf{u}_{p+1} \rangle &= \lambda \langle \mathbf{u}_k \mid \mathbf{f}_{p+1} \rangle + \sum_{j=1}^{p} \alpha_j \langle \mathbf{u}_k \mid \mathbf{u}_j \rangle \\ &= \lambda \langle \mathbf{u}_k \mid \mathbf{f}_{p+1} \rangle + \sum_{j=1}^{p} \alpha_j \delta_{k,j} = \lambda \langle \mathbf{u}_k \mid \mathbf{f}_{p+1} \rangle + \alpha_k \end{aligned} \tag{5.6}$$

ce qui détermine α_k : $\alpha_k = -\lambda \langle \mathbf{u}_k \mid \mathbf{f}_{p+1} \rangle$ et

$$\mathbf{u}_{p+1} = \lambda \mathbf{f}_{p+1} - \sum_{k=1}^{p} \lambda \langle \mathbf{u}_k \mid \mathbf{f}_{p+1} \rangle \mathbf{u}_k = \lambda \mathbf{v}_{p+1}$$

où $\mathbf{v}_{p+1} = \mathbf{f}_{p+1} - \sum_{k=1}^{p} \langle \mathbf{u}_k \mid \mathbf{f}_{p+1} \rangle \mathbf{u}_k$. D'autre part,

$$0 < \langle \mathbf{f}_{p+1} \mid \mathbf{u}_{p+1} \rangle = \langle \mathbf{v}_{p+1} + \sum_{k=1}^{p} \langle \mathbf{u}_k \mid \mathbf{f}_{p+1} \rangle \mathbf{u}_k \mid \mathbf{u}_{p+1} \rangle = \langle \mathbf{v}_{p+1} \mid \mathbf{u}_{p+1} \rangle = \lambda \|\mathbf{v}_{p+1}\|^2$$

Ainsi le scalaire λ est > 0, $\lambda = \frac{1}{\|\mathbf{v}_{p+1}\|}$ puisque $\mathbf{u}_{p+1}$ est unitaire, et le vecteur $\mathbf{u}_{p+1}$ est unique.
Reste à montrer, c'est la synthèse, que ce vecteur convient. En effet, $\mathbf{v}_{p+1} = \mathbf{f}_{p+1} - \sum_{k=1}^{p} \langle \mathbf{u}_k \mid \mathbf{f}_{p+1} \rangle \mathbf{u}_k$ n'est pas nul, puisque la famille $(\mathbf{u}_1, \dots, \mathbf{u}_p, \mathbf{f}_{p+1})$ est libre, et est orthogonal, par construction, à $\mathbf{u}_1, \dots, \mathbf{u}_p$. cqfd

Algorithme de calcul de l'orthonormalisée

Le calcul effectif de l'orthonormalisée $(\mathbf{u}_k)_k$ d'une suite libre $(\mathbf{f}_k)_k$ s'effectue à l'aide de l'algorithme suivant :

$$\mathbf{u}_1 = \frac{1}{\|\mathbf{f}_1\|}\mathbf{f}_1$$
$$\forall p \geqslant 1, \quad \mathbf{v}_{p+1} = \mathbf{f}_{p+1} - \sum_{k=1}^{p} \langle \mathbf{u}_k \mid \mathbf{f}_{p+1} \rangle \mathbf{u}_k \text{ et } \mathbf{u}_{p+1} = \frac{1}{\|\mathbf{v}_{p+1}\|}\mathbf{v}_{p+1}$$

ou bien à l'aide de celui-ci, qui permet, en général, des calculs plus simples :

(i) calcul de la suite orthogonale $(\mathbf{v}_k)_k$:

$$\mathbf{v}_1 = \mathbf{f}_1 \text{ et } \forall p \geqslant 1, \quad \mathbf{v}_{p+1} = \mathbf{f}_{p+1} - \sum_{k=1}^{p} \frac{\langle \mathbf{v}_k \mid \mathbf{f}_{p+1} \rangle}{\|\mathbf{v}_k\|^2}\mathbf{v}_k$$

(ii) orthonormalisation de la suite $(\mathbf{v}_k)_k$:

$$\forall k \geqslant 1, \quad \mathbf{u}_k = \frac{1}{\|\mathbf{v}_k\|}\mathbf{v}_k$$

5.3 Base orthonormale d'un sous-espace vectoriel de dimension finie

Théorème 5.5 (Existence de base orthonormale).
Tout sous-espace vectoriel de dimension finie d'un espace préhilbertien réel ou complexe, admet une base orthonormale.

Démonstration. L'orthonormalisée d'une base de F répond à la question. cqfd

Corollaire. *Les espaces euclidien et hermitien admettent des bases orthonormales.*

Dans un sous-espace vectoriel de dimension finie rapporté à une base orthonormale $(\mathbf{u}_1, \dots, \mathbf{u}_p)$, les expressions des coordonnées d'un vecteur, de sa norme, du produit scalaire et de la distance de deux vecteurs, sont particulièrement simples :

(i) expression des coordonnées de $\mathbf{x}$: $\mathbf{x} = \sum_{k=1}^{p} \langle \mathbf{u}_k \mid \mathbf{x} \rangle \mathbf{u}_k = \sum_{k=1}^{p} x_k \mathbf{u}_k$

(ii) expression de la norme : $\|\mathbf{x}\|^2 = \sum_{k=1}^{p} |x_k|^2 = \sum_{k=1}^{p} |\langle \mathbf{u}_k \mid \mathbf{x} \rangle|^2$

(iii) expression du produit scalaire : $\langle \mathbf{x} \mid \mathbf{y} \rangle = \sum_{k=1}^{p} \overline{x_k} y_k = \sum_{k=1}^{p} \overline{\langle \mathbf{u}_k \mid \mathbf{x} \rangle} \langle \mathbf{u}_k \mid \mathbf{y} \rangle$

(iv) expression de la distance : $\mathrm{d}(\mathbf{x}, \mathbf{y})^2 = \sum_{k=1}^{p} |y_k - x_k|^2$

6 Projection orthogonale

E désigne encore et toujours un espace préhilbertien réel ou complexe muni d'un produit scalaire noté $\langle \mid \rangle$; le corps des scalaires est noté K.

6.1 Orthogonal d'une partie

Définition 6.1 (Orthogonal d'une partie).
L'*orthogonal* d'une partie non vide A de E est l'ensemble des vecteurs de E qui sont orthogonaux à tous les vecteurs de A : on le note $A^{\perp}$.

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

Proposition 6.1 (Propriétés de l'orthogonal).
L'orthogonal d'une partie possède les propriétés suivantes :

(i) *l'orthogonal de $\mathbf{0}_E$ est E et l'orthogonal de E est $\{\mathbf{0}_E\}$;*
(ii) *si $\mathbf{a}$ est un vecteur non nul, l'orthogonal de $\mathbf{a}$ est un hyperplan de E ;*
(iii) *pour toute partie A non vide, $A^{\perp}$ est un sous-espace vectoriel de E ;*
(iv) *si F est un sous-espace de E, $F \cap F^{\perp}$ est réduit à $\{\mathbf{0}\}$;*
(v) *si F est un sous-espace vectoriel engendré par la famille $(\mathbf{f}_1, \ldots, \mathbf{f}_p)$, l'orthogonal de F est caractérisé par :*

$$\mathbf{x} \in F^{\perp} \iff \forall k \in [\![1, p]\!], \quad \langle \mathbf{f}_k \mid \mathbf{x} \rangle = 0 \iff \mathbf{x} \in \bigcap_{k=1}^{p} \{\mathbf{f}_k\}^{\perp} \tag{6.1}$$

Démonstration.

(i) Pour tout $\mathbf{x}$ de E, $\langle \mathbf{x} \mid \mathbf{0} \rangle = 0$, soit $\mathbf{x} \perp \mathbf{0}$. Ainsi $E \subset \{\mathbf{0}\}^{\perp}$ et $E = \{\mathbf{0}\}^{\perp}$.
Si $\mathbf{x}$ est orthogonal à E, $\mathbf{x}$ est, en particulier, orthogonal à lui-même ; $\mathbf{x}$ est donc nul. Ainsi $E^{\perp} \subset \{\mathbf{0}\}$ et $E^{\perp} = \{\mathbf{0}\}$.
(ii) Si $\mathbf{a}$ n'est pas nul, $\{\mathbf{a}\}^{\perp} = \{\mathbf{x} \ / \ \langle \mathbf{a} \mid \mathbf{x} \rangle = 0\} = \ker\{\mathbf{x} \mapsto \langle \mathbf{a} \mid \mathbf{x} \rangle\}$ est le noyau d'une forme linéaire non nulle ; c'est donc un hyperplan.
(iii) $A^{\perp}$ est l'intersection de tous les hyperplans $\{\mathbf{a}\}^{\perp}$ où $\mathbf{a}$ décrit A, c'est donc un sous-espace vectoriel.
(iv) $\mathbf{x} \in F \cap F^{\perp} \implies \langle \mathbf{x} \mid \mathbf{x} \rangle = 0$, soit $\mathbf{x} = \mathbf{0}$.
(v) Les éléments de F sont des combinaisons linéaires des vecteurs $\mathbf{f}_k$, et

$$\begin{aligned} \mathbf{x} \in F^{\perp} &\iff \forall (\lambda_1, \ldots, \lambda_p) \in \mathsf{K}^p, \quad 0 = \langle \mathbf{x} \mid \sum_{k=1}^{p} \lambda_k \mathbf{f}_k \rangle = \sum_{k=1}^{p} \lambda_k \langle \mathbf{x} \mid \mathbf{f}_k \rangle \\ &\iff \forall k \in [\![1, p]\!], \quad \langle \mathbf{x} \mid \mathbf{f}_k \rangle = 0 \iff \mathbf{x} \in \bigcap_{k=1}^{p} \{\mathbf{f}_k\}^{\perp} \end{aligned}$$

cqfd

6.2 Supplémentaires orthogonaux, projecteurs orthogonaux

Définition 6.2 (Sous-espaces orthogonaux).
Deux sous-espaces vectoriels F et G sont *orthogonaux* si tous les vecteurs de F sont orthogonaux à tous les vecteurs de G, *i.e.* si $F \subset G^{\perp}$ ou bien si tous les vecteurs de G sont orthogonaux à tous les vecteurs de F, *i.e.* $G \subset F^{\perp}$.

Théorème 6.2 (Caractérisation des supplémentaires orthogonaux).
Si F et G sont deux sous-espaces vectoriels supplémentaires *dans E, les propositions suivantes sont équivalentes :*

(i) *F et G sont orthogonaux ;*
(ii) *$F^{\perp} = G$;*
(iii) *$G^{\perp} = F$.*

Démonstration. (ii) et (iii) donnent (i).
Tout vecteur $\mathbf{x}$ de $F^{\perp}$ se décompose suivant $F \oplus G$ en $\mathbf{x} = \mathbf{y} + \mathbf{z}$; or $\langle \mathbf{x} \mid \mathbf{y} \rangle = 0$ ($\mathbf{x} \in F^{\perp}$ et $\mathbf{y} \in F$) et $\langle \mathbf{z} \mid \mathbf{y} \rangle = 0$ ($\mathbf{z} \in G \subset F^{\perp}$ et $\mathbf{y} \in F$), ce qui impose $\langle \mathbf{y} \mid \mathbf{y} \rangle = 0$ donc $\mathbf{y} = \mathbf{0}$, et $\mathbf{x} = \mathbf{z} \in G$. Chère lectrice, cher lecteur, vous venez de démontrer que (i) implique (ii).
Vous démontrerez de même que (i) implique (iii). cqfd

Dans ce cas, on dira que F et G sont *supplémentaires orthogonaux* dans E, et les projecteurs associés sont qualifiés de *projecteurs orthogonaux*.

Exemples 6.1.
L'hyperplan $\mathcal{H}$ d'équation $x_1 + x_2 + x_3 + x_4 = 0$ et la droite vectorielle dirigée par ${}^t(1, 1, 1, 1)$ sont supplémentaires orthogonaux dans R^4 muni du produit scalaire canonique.
Plus généralement, l'hyperplan $\mathcal{H}$ d'équation $\sum_{j=1}^{n} a_j x_j = 0$ et la droite $\mathcal{D} = \mathsf{R}\,{}^t(a_1, \ldots, a_n)$ sont supplémentaires orthogonaux dans R^n muni de son produit scalaire naturel (canonique).
Dans l'espace hermitien C^n muni de son produit scalaire naturel (canonique), l'hyperplan $\mathcal{H}$ d'équation $\sum_{j=1}^{n} a_j x_j = 0$ et la droite $\mathcal{D} = \mathsf{C}\,{}^t(\overline{a_1}, \ldots, \overline{a_n})$ sont supplémentaires orthogonaux.

Généralisation. Si $E = F_1 \oplus \cdots \oplus F_p$ et si les sous-espaces F_k sont orthogonaux deux à deux, la somme des sous-espaces F_k est une *somme directe orthogonale* de E.

6.3 Projection orthogonale sur un sous-espace vectoriel de dimension finie

Théorème 6.3 (Théorème de la projection).
Si F est un sous-espace vectoriel de dimension finie p d'un espace préhilbertien réel ou complexe E, alors

(i) *pour tout vecteur $\mathbf{x}$ de E, il existe un unique vecteur de F noté $p_F(\mathbf{x})$ tel que $\mathbf{x} - p_F(\mathbf{x})$ soit orthogonal à F ;*

(ii) *si $(\mathbf{u}_1, \ldots, \mathbf{u}_p)$ est une base* orthonormale *de F, on a*

$$p_F(\mathbf{x}) = \sum_{k=1}^{p} \langle \mathbf{u}_k \mid \mathbf{x} \rangle \mathbf{u}_k \tag{6.2}$$

(iii) *E est somme directe orthogonale de F et $F^{\perp}$, et p_F est le projecteur orthogonal d'image F,* i.e. *le projecteur sur F parallèlement à $F^{\perp}$.*

Démonstration.
(i) et *ii.* Appelons $(\mathbf{u}_1, \ldots, \mathbf{u}_p)$ une base orthonormale de F et considérons un vecteur $\mathbf{y} = \sum_{k=1}^{p} y_k \mathbf{u}_k$ de F. Alors

$$\begin{aligned}
\mathbf{x} - \mathbf{y} \in F^{\perp} = (\mathsf{K}\mathbf{u}_1 \oplus \cdots \oplus \mathsf{K}\mathbf{u}_p)^{\perp} &\iff \forall j \in [\![1, p]\!],\ (\mathbf{x} - \mathbf{y}) \perp \mathbf{u}_j \\
&\iff \forall j,\ 0 = \langle \mathbf{u}_j \mid \mathbf{x} - \mathbf{y} \rangle = \langle \mathbf{u}_j \mid \mathbf{x} - \sum_{k=1}^{p} y_k \mathbf{u}_k \rangle = \langle \mathbf{u}_j \mid \mathbf{x} \rangle - \sum_{k=1}^{p} y_k \delta_{k,j} = \langle \mathbf{u}_j \mid \mathbf{x} \rangle - y_j \\
&\iff \mathbf{y} = \sum_{k=1}^{p} \langle \mathbf{u}_k \mid \mathbf{x} \rangle \mathbf{u}_k = p_F(\mathbf{x})
\end{aligned}$$

(iii) Tout vecteur $\mathbf{x}$ de E se décompose, de manière unique, en un élément $p_F(\mathbf{x})$ de F et un élément $\mathbf{x} - p_F(\mathbf{x})$ de $F^{\perp}$; ainsi $E = F \oplus F^{\perp}$.
p_F est le projecteur sur F, parallèlement à $F^{\perp}$; c'est donc le projecteur orthogonal de E sur F.
cqfd

Remarque. Avec les mêmes notations, $\mathbf{x} - p_F(\mathbf{x})$ est le projeté orthogonal de $\mathbf{x}$ sur $F^{\perp}$, et $p_{F^{\perp}} = I_E - p_F$.

Corollaire (Existence de supplémentaire orthogonal).
Tout sous-espace vectoriel de dimension finie de E admet un supplémentaire orthogonal dans E.

Remarque. Si F n'est pas de dimension finie, la somme directe $F \oplus F^{\perp}$ peut être différente de E.

Projection orthogonale sur une droite, sur un hyperplan

La projection orthogonale sur la droite (vectorielle) $\mathcal{D}$ dirigée par $\mathbf{a}$, est donnée par :

$$p_{\mathcal{D}} : \mathbf{x} \in E \mapsto \frac{\langle \mathbf{a} \mid \mathbf{x} \rangle}{\|\mathbf{a}\|^2}\mathbf{a} \tag{6.3}$$

La projection orthogonale sur l'hyperplan $\mathcal{H}$ orthogonal à $\mathbf{a}$, est donnée par :

$$p_{\mathcal{H}} = I_E - p_{\mathcal{D}} : \mathbf{x} \in E \mapsto \mathbf{x} - \frac{\langle \mathbf{a} \mid \mathbf{x} \rangle}{\|\mathbf{a}\|^2}\mathbf{a} \tag{6.4}$$

La réflexion (ou symétrie orthogonale) $r_{\mathcal{H}}$ par rapport à l'hyperplan $\mathcal{H}$, est donnée par

$$r_{\mathcal{H}} = 2p_{\mathcal{H}} - I_E : \mathbf{x} \in E \mapsto 2\frac{\langle \mathbf{a} \mid \mathbf{x} \rangle}{\|\mathbf{a}\|^2}\mathbf{a} - \mathbf{x} \tag{6.5}$$

Interprétation géométrique de l'orthonormalisation de Schmidt

En notant F_p le sous-espace engendré par la famille $(\mathbf{f}_1, \ldots, \mathbf{f}_p)$, ou engendré par la famille orthonormale $(\mathbf{u}_1, \ldots, \mathbf{u}_p)$, on peut écrire

$$\mathbf{v}_{p+1} = \mathbf{f}_{p+1} - \sum_{k=1}^{p} \langle \mathbf{u}_k \mid \mathbf{f}_{p+1} \rangle \mathbf{u}_k = \mathbf{f}_{p+1} - p_{F_p}(\mathbf{f}_{p+1}) = p_{F_p^{\perp}}(\mathbf{f}_{p+1}) \tag{6.6}$$

et $\mathbf{v}_{p+1}$ s'interprète comme la projection orthogonale de $\mathbf{f}_{p+1}$ sur $F_p^{\perp}$.

6.4 Distance d'un vecteur à un sous-espace de dimension finie

Définition 6.3 (Distance à un sous-espace).
Si F est un sous-espace vectoriel de E et $\mathbf{x}$ un vecteur de E, on pose :

$$\mathrm{d}(\mathbf{x}, F) = \inf_{\mathbf{y} \in F} \mathrm{d}(\mathbf{x}, \mathbf{y}) = \inf_{\mathbf{y} \in F} \|\mathbf{y} - \mathbf{x}\| \tag{6.7}$$

Ce nombre existe puisqu'il est la borne inférieure d'une partie non vide de $[0, +\infty[$; il est appelé *distance de* $\mathbf{x}$ *à* F.

Théorème 6.4 (Expression de la distance à un sous-espace).
Si F est un sous-espace vectoriel de dimension finie *de E,*

(i) *l'application* $\mathbf{y} \mapsto \mathrm{d}(\mathbf{x}, \mathbf{y}) = \|\mathbf{y} - \mathbf{x}\|$ *admet un minimum global strict sur F, atteint en* $p_F(\mathbf{x})$ *;*

(ii) $\mathrm{d}(\mathbf{x}, F) = \mathrm{d}\big(\mathbf{x}, p_F(\mathbf{x})\big)$ *et* $\mathrm{d}^2(\mathbf{x}, F) = \|\mathbf{x}\|^2 - \|p_V(\mathbf{x})\|^2 = \langle \mathbf{x} - p_F(\mathbf{x}) \mid \mathbf{x} \rangle$.

Démonstration.

(i) Considérons un vecteur $\mathbf{y}$ de F distinct de $p_F(\mathbf{x})$; $\mathbf{x} - p_F(\mathbf{x})$ est orthogonal à F, donc en particulier à $\mathbf{y} - p_F(\mathbf{x})$ et le théorème de Pythagore donne

$$\mathrm{d}^2(\mathbf{x}, \mathbf{y}) = \mathrm{d}^2(\mathbf{x}, p_F(\mathbf{x})) + \mathrm{d}^2(p_F(\mathbf{x}), \mathbf{y}) > \mathrm{d}^2(\mathbf{x}, p_F(\mathbf{x})) \tag{6.8}$$

(ii) La question précédente montre que $\{\mathrm{d}(\mathbf{x}, \mathbf{y}) \ / \ \mathbf{y} \in F\}$ possède un unique plus petit élément à savoir $\mathrm{d}(\mathbf{x}, p_F(\mathbf{x}))$; on peut écrire :

$$\begin{aligned}
\mathrm{d}^2(\mathbf{x}, F) &= \|\mathbf{x} - p_F(\mathbf{x})\|^2 = \|\mathbf{x}\|^2 - \|p_F(\mathbf{x})\|^2 \text{ relation de Pythagore} \\
&= \langle \mathbf{x} - p_F(\mathbf{x}) \mid \mathbf{x} - p_F(\mathbf{x}) \rangle = \langle \mathbf{x} - p_F(\mathbf{x}) \mid \mathbf{x} \rangle \text{ car } p_F(\mathbf{x}) \perp (\mathbf{x} - p_F(\mathbf{x}))
\end{aligned}$$

cqfd

Théorème 6.5 (Projection et application lipschitzienne).
p_F est une application lipschitzienne de rapport 1, i.e.

$$\forall(\mathbf{x},\mathbf{y})\in E^2,\quad \|p_F(\mathbf{y})-p_F(\mathbf{x})\|\leqslant\|\mathbf{y}-\mathbf{x}\| \tag{6.9}$$

Démonstration. $\|p_F(\mathbf{x})\|^2=\|\mathbf{x}\|^2-\mathrm{d}^2(\mathbf{x},F)\leqslant\|\mathbf{x}\|^2$ et la linéarité de p_F donne le résultat. cqfd

Remarque. Les propriétés de ce paragraphe se généralisent à un sous-espace vectoriel F qui admet un supplémentaire orthogonal dans E.

6.5 Inégalité de Bessel

Théorème 6.6 (Inégalité de Bessel).
Si $(u_1,\dots,u_p)$ est une famille orthonormale *de E et* $\mathbf{x}$ *un vecteur de E, alors :*

$$\sum_{k=1}^{p}|\langle\mathbf{u}_k\mid\mathbf{x}\rangle|^2\leqslant\|\mathbf{x}\|^2 \tag{6.10}$$

Si $(\mathbf{u}_k)_{k\in\mathsf{N}}$ est une suite libre de E et $\mathbf{x}$ *un vecteur de E, alors :*

$$\big(\langle\mathbf{u}_k\mid\mathbf{x}\rangle\big)_k\in\ell^2_{\mathsf{N}}(\mathsf{C})\quad et\quad \sum_{k=0}^{+\infty}|\langle\mathbf{u}_k\mid\mathbf{x}\rangle|^2\leqslant\|\mathbf{x}\|^2 \tag{6.11}$$

Démonstration. Notons F_p le sous-espace vectoriel engendré par $(\mathbf{u}_1,\dots,\mathbf{u}_p)$; la projection de $\mathbf{x}$ sur F_p donne

$$\|p_{F_p}(\mathbf{x})\|^2=\Big\|\sum_{k=1}^{p}\langle\mathbf{u}_k\mid\mathbf{x}\rangle\,\mathbf{u}_k\Big\|^2=\sum_{k=1}^{p}|\langle\mathbf{u}_k\mid\mathbf{x}\rangle|^2\leqslant\|\mathbf{x}\|^2$$

L'inégalité précédente montre que les sommes partielles de la série de terme général $\sum|\langle\mathbf{u}_k\mid\mathbf{x}\rangle|^2$ sont majorées par $\|\mathbf{x}\|^2$; la série est donc convergente (elle est à termes positifs) et sa somme est majorée par $\|\mathbf{x}\|^2$. cqfd

6.6 Séries de Fourier, le retour

La famille $(e_k:t\mapsto e^{ikt})_{k\in\mathsf{Z}}$ est une famille orthonormale pour le produit scalaire hermitien de l'espace $\mathcal{C}_{2\pi}$ des fonctions continues 2π-périodiques $\langle f\mid g\rangle=\frac{1}{2\pi}\int_{-\pi}^{\pi}\overline{f(t)}g(t)\,dt$
Les coefficients exponentiels de Fourier $c_k(f)$ sont donnés par $c_k(f)=\langle e_k\mid f\rangle$.
L'espace $\mathcal{T}_n$ des polynômes trigonométriques de degré au plus n admet la famille des e_k pour $k\in[\![-n,n]\!]$ pour base orthonormale.
La somme partielle de Fourier $S_n(f)=\sum_{k=-n}^{n}c_k(f)e_k=\sum_{k=-n}^{n}\langle e_k\mid f\rangle e_k$ s'interprète comme la projection orthogonale de f sur $\mathcal{T}_n$; $S_n(f)$ est donc le polynôme trigonométrique de degré au plus n de meilleure approximation, et

(*i*) $\|S_n(f)\|^2=\Big\|\sum_{k=-n}^{n}c_k(f)e_k\Big\|^2=\sum_{k=-n}^{n}|c_k(f)|^2$ relation de Pythagore ;

(*ii*) $\|f-S_n(f)\|^2=\|f\|^2-\|S_n(f)\|^2$ relation de Pythagore ;

(*iii*) $\|S_n(f)\|^2=\sum_{k=-n}^{n}|c_k|^2\leqslant\|f\|^2=\frac{1}{2\pi}\int_{-\pi}^{\pi}|f|^2$ inégalité de Bessel.

Puisque $(S_n(f))_n$ tend vers f pour la norme de la convergence en moyenne quadratique, la suite $(\|f - S_n(f)\|^2)_n$, tend vers 0, soit

$$\|f - S_n(f)\|^2 = \|f\|^2 - \|S_n(f)\|^2 \xrightarrow[n]{} 0$$

ce qui donne l'égalité de Bessel-Parseval :

$$\lim_n \|S_n(f)\|^2 = \|f\|^2 \text{ soit } |c_0|^2 + \sum_{k=1}^{\infty}(|c_k|^2 + |c_{-k}|^2) = \frac{1}{2\pi}\int_{-\pi}^{\pi} |f|^2$$

CHAPITRE

5

Espace euclidien

Sommaire

On note E un R-espace vectoriel de *dimension finie n*, muni d'un produit scalaire (réel) qui est noté $\langle \mid \rangle$. Rappelons qu'un C-espace vectoriel de dimension finie et muni d'un produit scalaire (complexe), est appelé espace hermitien.

1 Résumé des épisodes précédents

1.1 Produit scalaire

A) Son expression dans une base quelconque

Soit $\mathcal{B} = (\mathbf{e}_1, \dots, \mathbf{e}_n)$ une base *quelconque* de E. On note $(x_1, \dots, x_n)$ (resp. $(y_1, \dots, y_n)$) les composantes de $\mathbf{x}$ (resp. $\mathbf{y}$) relativement à la base $\mathcal{B}$. Ainsi,

$$\mathbf{x} = \sum_{j=1}^{n} x_j\, \mathbf{e}_j \text{ et } \mathbf{y} = \sum_{j=1}^{n} y_j\, \mathbf{e}_j \tag{1.1}$$

ce qui donne

$$\langle \mathbf{x} \mid \mathbf{y} \rangle = \langle \sum_{i=1}^{n} x_i \mathbf{e}_i \mid \sum_{j=1}^{n} x_j \mathbf{e}_j \rangle = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i y_j \langle \mathbf{e}_i \mid \mathbf{e}_j \rangle = \sum_{i,j} g_{i,j}\, x_i y_j \tag{1.2}$$

avec $g_{i,j} = \langle \mathbf{e}_i \mid \mathbf{e}_j \rangle$. Dans le cas complexe, on a

$$\langle \mathbf{x} \mid \mathbf{y} \rangle = \sum_{i=1}^{n} \sum_{j=1}^{n} \overline{x_i} y_j \langle \mathbf{e}_i \mid \mathbf{e}_j \rangle = \sum_{i,j} g_{i,j}\, \overline{x_i} y_j$$

B) Expression matricielle du produit scalaire

Appelons G la matrice de terme général $g_{i,j} = \langle \mathbf{e}_i \mid \mathbf{e}_j \rangle$; G est une matrice carrée d'ordre n, à coefficients réels et symétrique. L'expression du produit scalaire s'écrit à l'aide de G :

$$\langle \mathbf{x} \mid \mathbf{y} \rangle = \sum_{i,j} g_{i,j}\, x_i y_j = {}^t X G Y \tag{1.3}$$

où $X = \mathcal{M}\mathrm{at}_{\mathcal{B}}(\mathbf{x})$ (resp. $Y = \mathcal{M}\mathrm{at}_{\mathcal{B}}(\mathbf{y})$) est la matrice-colonne des composantes de $\mathbf{x}$ (resp. $\mathbf{y}$) relativement à $\mathcal{B}$.

C) Effet d'un changement de base

Soient $\mathcal{B}' = (\mathbf{e}'_1, \dots, \mathbf{e}'_n)$ une autre base de E et P la matrice de passage de la base $\mathcal{B}$ à la base $\mathcal{B}'$, *i.e.* la matrice $P = \mathcal{M}\mathrm{at}_{\mathcal{B}}(\mathcal{B}')$ dont la j^{e} colonne est la colonne des composantes du vecteur $\mathbf{e}'_j$ relativement à $\mathcal{B}$. Ainsi,

$$\mathcal{M}\mathrm{at}_{\mathcal{B}}(\mathbf{x}) = \mathcal{M}\mathrm{at}_{\mathcal{B}}(\mathcal{B}')\mathcal{M}\mathrm{at}_{\mathcal{B}'}(\mathbf{x}), \text{ soit } X = P\, X'$$

et le produit scalaire devient :

$$\langle \mathbf{x} \mid \mathbf{y} \rangle = {}^t X P Y = {}^t(P X') G (P Y) = {}^t X' ({}^t P G P) Y' = {}^t X' G' Y' \tag{1.4}$$

La matrice G' du produit scalaire relativement à la base $\mathcal{B}'$ s'écrit :

$$G' = {}^t P G P$$

D) Cas d'une base orthonormale

Si $\mathcal{U} = (u_1, \ldots, u_n)$ est une base *orthonormale* de E, alors $\langle \mathbf{u}_i \mid \mathbf{u}_j \rangle = \delta_{i,j}$ et la matrice du produit scalaire relativement à la base *orthonormale* $\mathcal{U}$ est la matrice unité I_n.
Si $\mathcal{V} = (v_1, \ldots, v_n)$ est une base *orthogonale* de E, alors $\langle \mathbf{v}_i \mid \mathbf{v}_j \rangle = \delta_{i,j} \|\mathbf{v}_i\|^2$; la matrice du produit scalaire relativement à la base *orthogonale* $\mathcal{V}$ est la matrice diagonale $\mathrm{Diag}(\|\mathbf{v}_1\|^2, \ldots, \|\mathbf{v}_n\|^2)$.

1.2 Base orthonormale

A) Leur existence

Tout espace euclidien (resp. hermitien) possède, au moins, une base orthonormale : l'orthonormalisation d'une base (quelconque) de E donne le résultat.

B) Diverses expressions dans une base orthonormale

Considérons une base orthonormale $\mathcal{U} = (\varepsilon_1, \ldots, \varepsilon_n)$ de E. La composante x_j de $\mathbf{x}$ suivant ε_j est $x_j = \langle \varepsilon_j \mid \mathbf{x} \rangle = \langle \mathbf{x} \mid \varepsilon_j \rangle$ et

$$\mathbf{x} = \sum_{j=1}^{n} x_j \varepsilon_j = \sum_{j=1}^{n} \langle \varepsilon_j \mid \mathbf{x} \rangle \varepsilon_j$$

$$\langle \mathbf{x} \mid \mathbf{y} \rangle = \sum_{j,k=1}^{n} x_j y_k \langle \varepsilon_j \mid \varepsilon_k \rangle = \sum_{j,k=1}^{n} x_j y_k \delta_{j,k} = \sum_{k=1}^{n} x_k y_k$$

$$\|\mathbf{x}\|^2 = \langle \mathbf{x} \mid \mathbf{x} \rangle = \sum_{k=1}^{n} x_k^2$$

$$\mathrm{d}^2(\mathbf{x}, \mathbf{y}) = \|\mathbf{y} - \mathbf{x}\|^2 = \sum_{k=1}^{n} (y_k - x_k)^2$$

Dans le cas complexe, le lecteur mettra les barres de module et de conjugaison là où il faut !

C) Isomorphisme de E sur $\mathcal{M}_{n,1}(\mathsf{R})$

La donnée d'une base orthonormale $\mathcal{U} = (\varepsilon_1, \ldots, \varepsilon_n)$ de E permet de construire un isomorphisme de sur $\mathcal{M}_{n,1}(\mathsf{R})$ muni de son produit scalaire naturel (canonique) :

$$\mathbf{x} \mapsto \mathcal{M}\mathrm{at}_{\mathcal{U}}(\mathbf{x}) = {}^t(x_1, \ldots, x_n) = X$$

D) Matrice d'un endomorphisme relativement à une base orthonormale

Soient u un endomorphisme de E et $A = [a_{i,j}] = \mathcal{M}\mathrm{at}_{\mathcal{U}}(u)$ sa matrice relativement à une base orthonormale $\mathcal{U}$ de E ; le terme $a_{i,j}$, i^{e} composante du vecteur $u(\varepsilon_j)$ relativement à la base $\mathcal{U}$, se calcule par :

$$a_{i,j} = \langle \varepsilon_i \mid u(\varepsilon_j) \rangle$$

E) Caractérisation de la matrice d'un produit scalaire (réel)

Théorème 1.1. *Soient E un espace euclidien de dimension n, $\mathcal{B} = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ une base (quelconque) de E et G une matrice symétrique réelle d'ordre n ; alors $\langle \mathbf{x} \mid \mathbf{y} \rangle = {}^tXGY$ définit un produit scalaire sur E si, et seulement si, il existe une matrice inversible $P \in \mathcal{GL}_n(\mathsf{R})$ telle que $G = {}^tPP$*

Démonstration.

$\boxed{\Longrightarrow}$ G est la matrice du produit scalaire $\langle \mid \rangle$ relativement à $\mathcal{B}$. Si $\mathcal{U}$ est une base orthonormale de E, et si $P = \mathcal{M}\mathrm{at}_{\mathcal{B}}(\mathcal{U})$ est la matrice de passage de $\mathcal{B}$ à $\mathcal{U}$, $G' = {}^tPGP$ est la matrice du produit scalaire relativement à $\mathcal{U}$, donc $G' = I_n$ et $G = {}^tP^{-1}P^{-1}$.

$\boxed{\Longleftarrow}$ $(\mathbf{x}, \mathbf{y}) \mapsto {}^tX\,{}^tPP\,Y = {}^t(PX)\,PY$ est une forme bilinéaire symétrique définie positive sur E, démonstration à rédiger.
cqfd

1.3 Supplémentaires orthogonaux

A) Leur existence

Tout sous-espace vectoriel F de E admet un supplémentaire orthogonal $F^\perp$, puisque F est de dimension finie. D'autre part $\left(F^\perp\right)^\perp = F$.

B) Leur dimension

$\dim F^\perp = \dim E - \dim F$

C) Leurs équations

Si $(\mathbf{f}_1, \dots, \mathbf{f}_p)$ est une base de F, ou plus généralement une famille génératrice de F, alors

$$\mathbf{x} \in F^\perp = \mathcal{V}\mathrm{ect}\{\mathbf{f}_1, \dots, \mathbf{f}_p\}^\perp \iff \forall j \in [\![1, p]\!],\ \langle \mathbf{f_j} \mid \mathbf{x} \rangle = 0 \tag{1.5}$$

D) Base orthonormale adaptée à un sous-espace vectoriel

Puisque $E = F \oplus F^\perp$, on peut construire une base orthonormale de E, en réunissant une base orthonormale de F et une base orthonormale de $F^\perp$; une telle base est appelée *base orthonormale adaptée* à F.

Théorème 1.2 (de la base orthonormale incomplète).
Toute famille orthonormale $(\mathbf{u}_1, \dots, \mathbf{u}_p)$ de E peut être complétée en une base orthonormale $(\mathbf{u}_1, \dots, \mathbf{u}_p, \mathbf{u}_{p+1}, \dots, \mathbf{u}_n)$ de E.

Démonstration. Il suffit d'appliquer la remarque précédente au sous-espace vectoriel F engendré par $(\mathbf{u}_1, \dots, \mathbf{u}_p)$.
cqfd

2 Adjoint d'un endomorphisme

2.1 Isomorphisme naturel (canonique) de E sur son dual

Théorème 2.1 (Isomorphisme naturel de E sur son dual).
L'application $\Phi : \mathbf{x} \in E \mapsto \langle \mathbf{x} \mid \cdot \rangle$ réalise un isomorphisme de E sur son dual E^* ; on le qualifie de naturel *ou de* canonique.

Démonstration.
Pour tout $\mathbf{x} \in E$, $\Phi(\mathbf{x}) = \langle \mathbf{x} \mid \cdot \rangle : \mathbf{y} \mapsto \langle \mathbf{x} \mid \mathbf{y} \rangle$ est une application linéaire de E vers R, *i.e.* une forme linéaire sur E, *i.e.* un élément de E^*.
Φ est une application linéaire car, pour tout $\mathbf{x}_1$ et $\mathbf{x}_2$ dans E, pour tout λ_1 et λ_2 dans R, on a l'égalité des applications $\langle \lambda_1\mathbf{x}_1 + \lambda_2\mathbf{x}_2 \mid \cdot \rangle$ et $\lambda_1\langle \mathbf{x}_1 \mid \cdot \rangle + \lambda_2\langle \mathbf{x}_2 \mid \cdot \rangle$. En effet, pour tout $\mathbf{y}$ dans E,

$$\langle \lambda_1\mathbf{x}_1 + \lambda\mathbf{x}_2 \mid \mathbf{y} \rangle = \lambda_1\langle \mathbf{x}_1 \mid \mathbf{y} \rangle + \lambda_2\langle \mathbf{x}_2 \mid \mathbf{y} \rangle = \left[\Phi(\lambda_1\mathbf{x}_1 + \lambda_2\mathbf{x}_2)\right](\mathbf{y})$$

$\mathbf{z} \in \ker\big(\mathbf{x} \mapsto \langle \mathbf{x} \mid \cdot \rangle\big)$ si, et seulement si, $\langle \mathbf{z} \mid \cdot \rangle$ est la forme linéaire nulle, *i.e.* $\forall \mathbf{y} \in E,\ \langle \mathbf{z} \mid \mathbf{y} \rangle = 0$, soit $\mathbf{z} = \mathbf{0}$.
L'application linéaire Φ est donc injective.
E et E^* sont deux espaces vectoriels de même dimension finie ; l'application linéaire injective Φ est donc un isomorphisme de E sur E^*. cqfd

Théorème 2.2 (de représentation de Riesz).
À toute forme linéaire φ sur E correspond un unique vecteur $\mathbf{a}$ *de E tel que*

$$\forall \mathbf{y} \in E,\ \varphi(\mathbf{y}) = \langle \mathbf{a} \mid \mathbf{y} \rangle$$

Démonstration. Utilisons l'isomorphisme naturel Φ et posons $\mathbf{a} = \Phi^{-1}(\varphi)$. Ainsi $\varphi = \Phi(\mathbf{a}) = \langle \mathbf{a} \mid \cdot \rangle$. cqfd

Exemple 2.1.
À toute forme linéaire φ sur $\mathcal{M}_{n,p}(\mathsf{R})$ correspond une une unique matrice $A \in \mathcal{M}_{n,p}(\mathsf{R})$ qui « représent » φ pour le produit scalaire naturel (canonique) de $\mathcal{M}_{n,p}(\mathsf{R})$, soit

$$\forall M \in \mathcal{M}_{n,p}(\mathsf{R}),\ \varphi(M) = \operatorname{tr}({}^t AM)$$

Produit vectoriel en dimension 3

Soit E un espace vectoriel euclidien de dimension 3, orienté par la donnée d'une base orthonormale $\mathcal{B} = (\mathbf{i}, \mathbf{j}, \mathbf{k})$. Le produit mixte

$$[\mathbf{x}, \mathbf{y}, \mathbf{z}] = \det(\mathbf{x}, \mathbf{y}, \mathbf{z})$$

est indépendant de la base orthonormale directe $\mathcal{B}$ choisie. Pour $\mathbf{x}$ et $\mathbf{y}$ fixés dans E, l'application $\mathbf{z} \mapsto [\mathbf{x}, \mathbf{y}, \mathbf{z}]$ est une forme linéaire que l'on peut représenter à l'aide d'un vecteur $\mathbf{w}$que l'on note $\mathbf{x} \wedge \mathbf{y}$, et l'on a :

$$\forall \mathbf{z} \in E,\ \langle \mathbf{x} \wedge \mathbf{y} \mid \mathbf{z} \rangle = [\mathbf{x}, \mathbf{y}, \mathbf{z}] = \det(\mathbf{x}, \mathbf{y}, \mathbf{z})$$

À l'aide de cette définition, on retrouve les propriétés habituelles du produit vectoriel. Pouvez-vous les démontrer ?

2.2 Endomorphisme associé à une forme bilinéaire

Théorème 2.3. *À toute forme bilinéaire Ψ sur E est associée un unique endomorphisme u de E tel que*

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

u est appelé endomorphisme naturel *(canonique)* associé à la forme bilinéaire Ψ.

Démonstration.
$\mathbf{y} \mapsto \Psi(\mathbf{x}, \mathbf{y})$ est une forme linéaire sur E. Le théorème de représentation de Riesz montre l'existence d'un unique vecteur de E, que l'on note $u(\mathbf{x})$, tel que

$$\forall \mathbf{y} \in E,\ \Psi(\mathbf{x}, \mathbf{y}) = \langle u(\mathbf{x}) \mid \mathbf{y} \rangle \tag{2.1}$$

L'application u est linéaire, car pour tout vecteur $\mathbf{x}_1$, $\mathbf{x}_2$ et $\mathbf{y}$, pour tout réel λ_1 et λ_2, on a

$$\begin{aligned}
\langle u(\lambda_1 \mathbf{x}_1 + \lambda_2 \mathbf{x}_2) \mid \mathbf{y} \rangle &= \Psi(\lambda_1 \mathbf{x}_1 + \lambda_2 \mathbf{x}_2, \mathbf{y}) = \lambda_1 \Psi(\mathbf{x}_1, \mathbf{y}) + \lambda_2 \Psi(\mathbf{x}_2, \mathbf{y}) \\
&= \lambda_1 \langle u(\mathbf{x}_1) \mid \mathbf{y} \rangle + \lambda_2 \langle u(\mathbf{x}_2) \mid \mathbf{y} \rangle \\
&= \langle \lambda_1 u(\mathbf{x}_1) + \lambda_2 u(\mathbf{x}_2) \mid \mathbf{y} \rangle
\end{aligned}$$

L'unicité de la représentation de Riesz montre l'égalité entre $u(\lambda_1 \mathbf{x}_1 + \lambda_2 \mathbf{x}_2)$ et $\lambda_1 u(\mathbf{x}_1) + \lambda_2 u(\mathbf{x}_2)$.
cqfd

2.3 Adjoint d'un endomorphisme

Théorème 2.4 (Existence et unicité de l'adjoint).
À tout endomorphisme u de E est associé un unique endomomorphisme de E noté u^* tel que

$$\boxed{\forall(\mathbf{x},\mathbf{y})\in E^2,\ \langle u^*(\mathbf{x})\mid\mathbf{y}\rangle=\langle\mathbf{x}\mid u(\mathbf{y})\rangle}$$

L'endomorphisme u^* est appelé l'adjoint de u.

Démonstration. u^* est l'endomorphisme naturel (canonique) associé à la forme bilinéaire symétrique

$$\Psi:\mathbf{x}\in E\mapsto\langle\mathbf{x}\mid u(\mathbf{y})\rangle \tag{2.2}$$

Cher lecteur, la vérification de la bilinéarité de Ψ est laissée à votre initiative. cqfd

Théorème 2.5 (Matrice de l'adjoint dans une base orthonormale).
*Si $\mathcal{B}$ est une base **orthonormale** de E, alors*

$$\mathcal{M}\mathrm{at}_{\mathcal{B}}(u^*)={}^t\mathcal{M}\mathrm{at}_{\mathcal{B}}(u)$$

Démonstration. Appelons $(\mathbf{e}_1,\dots,\mathbf{e}_n)$ les vecteurs de la base orthonormale $\mathcal{B}$ de E ; alors

$$\big(\mathcal{M}\mathrm{at}_{\mathcal{B}}(u^*)\big)_{i,j}=\langle\mathbf{e}_i\mid u^*(\mathbf{e}_j)\rangle=\langle e_j\mid u(\mathbf{e}_i)\rangle=\big(\mathcal{M}\mathrm{at}_{\mathcal{B}}(u)\big)_{j,i}$$

ce qui donne l'égalité annoncée. cqfd

Proposition 2.6 (Règles de calcul).
Si u et v sont des endomorphismes de E et λ un réel, alors

(i) $(u+v)^*=u^*+v^*$; $(\lambda u)^*=\lambda\,u^*$; $(u^*)^*=u$; *ainsi* $u\mapsto u^*$ *est un automorphisme involutif de* $\mathcal{L}(E)$;
(ii) $(u\circ v)^*=v^*\circ u^*$; $(I_E)^*=I_E$;
(iii) $u\in\mathcal{GL}(E)\iff u^*\in\mathcal{GL}(E)$ *et, dans ce cas,* $(u^*)^{-1}=(u^{-1})^*$;
(iv) $\mathrm{tr}(u^*)=\mathrm{tr}\,u$ *et* $\det(u^*)=\det u$.

Démonstration. Si $\mathbf{x}$ et $\mathbf{y}$ sont des vecteurs de E, alors

(i) $\langle(u+v)^*(\mathbf{x})\mid\mathbf{y}\rangle=\langle\mathbf{x}\mid(u+v)(\mathbf{y})\rangle=\langle\mathbf{x}\mid u(\mathbf{y})\rangle+\langle\mathbf{x}\mid v(\mathbf{y})\rangle=\langle u^*(\mathbf{x})\mid\mathbf{y}\rangle+\langle v^*(\mathbf{x})\mid\mathbf{y}\rangle=\langle u^*(\mathbf{x})+v^*(\mathbf{x})\mid\mathbf{y}\rangle=\langle(u^*+v^*)(\mathbf{x})\mid\mathbf{y}\rangle$; l'unicité de la représentation de Riesz montre l'égalité entre $(u+v)^*(\mathbf{x})$ et $(u^*+v^*)(\mathbf{x})$, ceci pour tout $\mathbf{x}$, donc l'égalité des applications $(u+v)^*$ et u^*+v^*.
$\langle(\lambda u)^*(\mathbf{x})\mid\mathbf{y}\rangle=\langle\mathbf{x}\mid\big(\lambda u\big)(\mathbf{y})\rangle=\lambda\langle\mathbf{x}\mid u(\mathbf{y})\rangle=\lambda\langle u^*(\mathbf{x})\mid\mathbf{y}\rangle=\langle\lambda u^*(\mathbf{x})\mid\mathbf{y}\rangle$; l'unicité de la représentation de Riesz montre l'égalité entre $(\lambda u)^*(\mathbf{x})$ et $\lambda u^*(\mathbf{x})$, ceci pour tout $\mathbf{x}$, donc l'égalité des applications $(\lambda u)^*$ et λu^*.
$\langle(u^*)^*(\mathbf{x})\mid\mathbf{y}\rangle=\langle\mathbf{x}\mid u^*(\mathbf{y})\rangle=\langle u(\mathbf{x})\mid\mathbf{y}\rangle$; et, d'après l'unicité de la repré ..., l'égalité des applications $(u^*)^*$ et u est démontrée.
Ainsi, $u\mapsto u^*$ est une application linéaire qui est son propre inverse ; $u\mapsto u^*$ est une involution.

(ii) $\langle\big(u\circ v\big)^*(\mathbf{x})\mid\mathbf{y}\rangle=\langle\mathbf{x}\mid u\big(v(\mathbf{y})\big)\rangle=\langle u^*(\mathbf{x})\mid v(\mathbf{y})\rangle=\langle v^*\circ u^*(\mathbf{x})\mid\mathbf{y}\rangle$; d'où l'égalité des applications $(u\circ v)^*$ et $v^*\circ u^*$.
$\langle\big(I_E^*\big)(\mathbf{x})\mid\mathbf{y}\rangle=\langle\mathbf{x}\mid I_E(\mathbf{y})\rangle=\langle\mathbf{x}\mid\mathbf{y}\rangle=\langle I_E(\mathbf{x})\mid\mathbf{y}\rangle$ et $(I_E)^*=I_E$.

(iii) $u\circ v=v\circ u=I_E\implies v^*\circ u^*=u^*\circ v^*=(I_E)^*=I_E$; ceci montre l'implication $u\in\mathcal{GL}(E)\implies u^*\in\mathcal{GL}(E)$ et l'égalité de $v^*=(u^{-1})^*$ avec $(u^*)^{-1}$. L'égalité $(u^*)^*=u$ montre la réciproque.

(iv) Si $\mathcal{B}$ est une base *orthonormale* de E, la matrice de u^* relativement à $\mathcal{B}$ est la transposée de la matrice de u relativement à la même base ; ces matrices ont la même trace et le même déterminant ; u et u^* ont donc même trace et même déterminant.

cqfd

2.4 Adjoint et stabilité

Théorème 2.7. *Si u est un endomorphisme de E, alors,*

(i) $\ker u^* = \big(\operatorname{Im} u\big)^{\perp}$, $\operatorname{Im} u^* = \big(\ker u\big)^{\perp}$ *et* $\operatorname{rg} u^* = \operatorname{rg} u$ *;*
(ii) *soit F un sous-espace vectoriel de E ; F est stable par u si, et seulement si, $F^{\perp}$ est stable par u^*.*

Démonstration.

(i) $\mathbf{x} \in \ker u^* \iff u^*(\mathbf{x}) = \mathbf{0} \iff \forall \mathbf{y} \in E,\ 0 = \langle u^*(\mathbf{x}) \mid \mathbf{y}\rangle = \langle \mathbf{x} \mid u(\mathbf{y})\rangle \iff \mathbf{x} \in \big(\operatorname{Im} u\big)^{\perp}$
On applique l'égalité précédente à u^* et on trouve : $\big(\operatorname{Im} u^*\big)^{\perp} = \ker(u^*)^* = \ker u$; en prenant les orthogonaux, on obtient l'égalité annocée.
$\operatorname{rg} u^* = \dim\big(\operatorname{Im} u^*\big) = \dim\big(\ker u^{\perp}\big) = \dim E - \dim(\ker u) = \operatorname{rg} u$;

(ii) si F est stable par u, alors, pour tout $\mathbf{y} \in F$, $u(\mathbf{y}) \in F$. Si $\mathbf{x} \in F^{\perp}$, alors, pour tout $\mathbf{y} \in F$, $0 = \langle \mathbf{x} \mid u(\mathbf{y})\rangle = \langle u^*(\mathbf{x}) \mid \mathbf{y}\rangle$, ce qui montre que $u^*(\mathbf{x}) \in F^{\perp}$; ainsi, $F^{\perp}$ est stable par u^*.
Si $F^{\perp}$ est stable par u^*, $(F^{\perp})^{\perp} = F$ est stable par $(u^*)^* = u$.

cqfd

3 Endomorphisme auto-adjoint

3.1 Généralités

Définitions 3.1 (Endomorphisme auto-adjoint ou symétrique, anti-symétrique).
Si u est un endomorphisme d'un espace euclidien E, u est dit *auto-adjoint* ou *symétrique* si

$$\forall (\mathbf{x}, \mathbf{y}) \in E^2,\ \langle u(\mathbf{x}) \mid \mathbf{y}\rangle = \langle \mathbf{x} \mid u(\mathbf{y})\rangle$$

u est dit *anti-symétrique* si

$$\forall (\mathbf{x}, \mathbf{y}) \in E^2,\ \langle u(\mathbf{x}) \mid \mathbf{y}\rangle = -\langle \mathbf{x} \mid u(\mathbf{y})\rangle$$

On note $\mathcal{S}(E)$ l'ensemble des endomorphismes auto-adjoints ou symétriques et $\mathcal{A}(E)$ l'ensemble des endomorphismes antisymétriques.

Théorème 3.1 (Caractérisation des endomorphismes auto-adjoints).
Si u est un endomorphisme d'un espace euclidien E, les assertions suivantes sont équivalentes :

(i) *u est auto-adjoint ;*
(ii) *$u = u^*$;*
(iii) *pour toute base* orthonormale *$\mathcal{B}$ de E, la matrice de u relative à $\mathcal{B}$ est symétrique ;*
(iv) *il existe une base* orthonormale *$\mathcal{B}$ de E telle que la matrice de u relative à $\mathcal{B}$ soit symétrique ;*

Démonstration. u est auto-adjoint $\iff \forall (\mathbf{x}, \mathbf{y}) \in E^2,\ \langle u(\mathbf{x}) \mid \mathbf{y}\rangle = \langle \mathbf{x} \mid u(\mathbf{y})\rangle = \langle u^*(\mathbf{x}) \mid \mathbf{y}\rangle \iff u = u^*$, ce qui montre l'équivalence des deux premières assertions.

$(ii) \implies (iii)$ Les matrices de u et u^*, relatives à $\mathcal{B}$, sont égales puisque $u = u^*$; d'autre part, $\mathcal{M}\mathrm{at}_{\mathcal{B}}(u^*) = {}^t\mathcal{M}\mathrm{at}_{\mathcal{B}}(u)$ puisque $\mathcal{B}$ est une base orthonormale.

$(iii) \implies (iv)$ Qui peut le plus, peut le moins.

$(iv) \implies (ii)$ Les égalités $\mathcal{M}\mathrm{at}_{\mathcal{B}}(u) = {}^t\mathcal{M}\mathrm{at}_{\mathcal{B}}(u) = \mathcal{M}\mathrm{at}_{\mathcal{B}}(u^*)$ montrent que $u = u^*$. cqfd

Endomorphisme symétrique et matrice symétrique réelle

Si S est une matrice symétrique réelle d'ordre n, l'endomorphisme u_S associé à S, est un endomorphisme auto-adjoint (symétrique) de $\mathcal{M}_{n,1}(\mathsf{R})$ muni de son produit scalaire naturel (canonique). En effet :

$$\langle u_S(X) \mid Y\rangle = {}^t(SX)Y = {}^tX{}^tSY = {}^tXSY = \langle X \mid SY\rangle = \langle X \mid u_S(Y)\rangle \tag{3.1}$$

Soient E un espace euclidien et $\mathcal{B}$ une base orthonormale de E ; à toute matrice symétrique réelle d'ordre n, on associe l'endomorphisme u_S de E, de matrice S relativement à $\mathcal{B}$; u_S est un endomorphisme auto-adjoint (symétrique) de E car

$$\langle u_S(\mathbf{x}) \mid \mathbf{y}\rangle = {}^t(SX)Y = {}^tXSY = \langle \mathbf{x} \mid u_S(\mathbf{y})\rangle \tag{3.2}$$

Tout ceci donne la

Proposition 3.2. *Si $\mathcal{B}$ est une base orthonormale d'un espace euclidien E, l'application $u \mapsto \mathcal{M}\mathrm{at}_{\mathcal{B}}(u)$ réalise un isomorphisme du* R*-espace vectoriel $\mathcal{S}(E)$ des endomorphismes auto-adjoints (symétriques), sur le* R*-espace vectoriel $\mathcal{S}_n(\mathsf{R})$ des matrices symétriques réelles d'ordre n, et*

$$\dim \mathcal{S}(E) = \dim \mathcal{S}_n(\mathsf{R}) = \frac{n(n+1)}{2}$$

Théorème 3.3 (Caractérisation des endomorphismes antisymétriques).
Si u est un endomorphisme d'un espace euclidien E, les assertions suivantes sont équivalentes :

(i) *u est antisymétrique ;*
(ii) *$u^* = -u$;*
(iii) *pour toute base* orthonormale *$\mathcal{B}$ de E, la matrice de u relative à $\mathcal{B}$ est antisymétrique ;*
(iv) *il existe une base* orthonormale *$\mathcal{B}$ de E telle que la matrice de u relative à $\mathcal{B}$ soit antisymétrique ;*

Démonstration. La démonstration ressemble comme une goutte d'eau à la démonstration précédente ; elle est laissée aux soins de la lectrice ou du lecteur. cqfd

Proposition 3.4. *Si $\mathcal{B}$ est une base orthonormale d'un espace euclidien E, l'application $u \mapsto \mathcal{M}\mathrm{at}_{\mathcal{B}}(u)$ réalise un isomorphisme du* R*-espace vectoriel $\mathcal{A}(E)$ des endomorphismes antisymétriques, sur le* R*-espace vectoriel $\mathcal{A}_n(\mathsf{R})$ des matrices antisymétriques réelles d'ordre n, et*

$$\dim \mathcal{A}(E) = \dim \mathcal{A}_n(\mathsf{R}) = \frac{n(n-1)}{2}$$

3.2 Projecteur orthogonal

Définition 3.2 (Projecteur orthogonal).
Si F est un sous-espace vectoriel d'un espace euclidien E, le projecteur p_F d'image F, parallèlement à $F^\perp$, est appelé *projecteur orthogonal* de E sur F.

Proposition 3.5 (Expression analytique d'un projecteur orthogonal).
Si $(\mathbf{u}_1, \dots, \mathbf{u}_r)$ est une base orthonormale *de F, on a*

$$\forall \mathbf{x} \in E,\ p_F(\mathbf{x}) = \sum_{j=1}^{r} \langle \mathbf{u}_j \mid \mathbf{x}\rangle\, \mathbf{u}_j$$

Cas d'une droite $\mathcal{D} = \mathsf{R}\mathbf{a}$: $p_{\mathcal{D}} : \mathbf{x} \mapsto \dfrac{\langle \mathbf{a} \mid \mathbf{x}\rangle}{\|\mathbf{a}\|^2}\mathbf{a}$

Cas d'un hyperplan $\mathcal{H} = \{\mathbf{a}\}^\perp$: $p_{\mathcal{H}} = I_E - p_{\mathcal{D}} : \mathbf{x} \mapsto \mathbf{x} - \dfrac{\langle \mathbf{a} \mid \mathbf{x}\rangle}{\|\mathbf{a}\|^2}\mathbf{a}$

Proposition 3.6 (Expression matricielle d'un projecteur orthogonal).
Si $\mathcal{B}$ est une base orthonormale de E, si $(\mathbf{u}_1, \dots, \mathbf{u}_r)$ est une base orthonormale *de F et si $U_j = \mathcal{M}\mathrm{at}_{\mathcal{B}}(\mathbf{u}_j)$ est le vecteur-colonne des composantes de $\mathbf{u}_j$ relativement à $\mathcal{B}$, alors*

$$\mathcal{M}\mathrm{at}_{\mathcal{B}}(p_F) = \sum_{j=1}^{r} U_j\,{}^tU_j$$

Démonstration. Appelons $P_F = \mathcal{M}\mathrm{at}_{\mathcal{B}}(p_F)$ la matrice de p_F relative à $\mathcal{B}$. L'égalité $p_F(\mathbf{x}) = \sum_{j=1}^{r} \langle \mathbf{u}_j \mid \mathbf{x}\rangle\, \mathbf{u}_j$ s'écrit matriciellement $P_F X = \sum_{j=1}^{r} \big({}^tU_j X\big) U_j$. Or, ${}^tU_j X$ est un nombre réel ou plutôt une matrice de taille 1×1 qui commute avec toute matrice. Ainsi

$$PX = \sum_{j=1}^{r} U_j\big({}^tU_j X\big) = \sum_{j=1}^{r} \big(U_j\,{}^tU_j\big) X = \Big(\sum_{j=1}^{r} U_j\,{}^tU_j\Big) X$$

cqfd

Théorème 3.7 (Caractérisation des projecteurs orthogonaux parmi les projecteurs).
Si p est un projecteur d'un espace euclidien E, p est un projecteur orthogonal si, et seulement si, p est auto-adjoint (symétrique).

Démonstration. Tout projecteur p est la projection de E sur $\mathrm{Im}\, p$, parallèlement à $\ker p$.

$\boxed{\Longrightarrow}$ Soit p le projecteur orthogonal sur $\mathrm{Im}\, p$ parallèlement à $(\mathrm{Im}\, p)^{\perp}$; dans une base orthonormale $\mathcal{B}$ adaptée à la somme directe orthogonale $E = \mathrm{Im}\, p \oplus (\mathrm{Im}\, p)^{\perp}$, la matrice de p relative à $\mathcal{B}$ est la matrice diagonale par blocs $\mathrm{Diag}(I_r, 0_{n-r})$ où r est le rang de p. Cette matrice est symétrique réelle, donc p est auto-adjoint.

$\boxed{\Longleftarrow}$ Si p est auto-adjoint, $\ker p = \ker p^* = (\mathrm{Im}\, p)^{\perp}$ et p est le projecteur de E sur $\mathrm{Im}\, p$ et parallèlement à $(\mathrm{Im}\, p)^{\perp}$; p est un projecteur orthogonal. cqfd

Proposition 3.8 (Interprétation matricielle).
Soient p un endomorphisme d'un espace euclidien E, $\mathcal{B}$ une base orthonormale *de E et $M = \mathcal{M}\mathrm{at}_{\mathcal{B}}(p)$ la matrice de p relative à $\mathcal{B}$; alors, p est un projecteur orthogonal si, et seulement si, $M \in \mathcal{S}_n(\mathbb{R})$ et $M^2 = M$.*

3.3 Symétrie orthogonale

Définition 3.3 (Symétrie orthogonale).
Si F est un sous-espace vectoriel d'un espace euclidien E, la symétrie par rapport à F et parallèlement à $F^{\perp}$ est appelée *symétrie orthogonale* par rapport à F.

Proposition 3.9 (Symétrie et projecteur orthogonaux).
Soient F un sous-espace vectoriel d'un espace euclidien, s_F la symétrie orthogonale par rapport à F et p_F le projecteur orthogonal d'image F ; alors :

$$s_F + I_E = 2p_F$$

Cas d'une droite $\mathcal{D} = \mathbb{R}\mathbf{a}$: $s_{\mathcal{D}} = 2p_{\mathcal{D}} - I_E : \mathbf{x} \mapsto 2\dfrac{\langle \mathbf{a} \mid \mathbf{x}\rangle}{\|\mathbf{a}\|^2}\mathbf{a} - \mathbf{x}$

Cas d'un hyperplan $\mathcal{H} = \{\mathbf{a}\}^{\perp}$: $s_{\mathcal{H}} = 2p_{\mathcal{H}} - I_E = -s_{\mathcal{D}} : \mathbf{x} \mapsto \mathbf{x} - 2\dfrac{\langle \mathbf{a} \mid \mathbf{x}\rangle}{\|\mathbf{a}\|^2}\mathbf{a}$

Démonstration. Faire un dessin et se rappeler des propriétés de la diagonale d'un parallélogramme, qui est un losange dans notre cas. cqfd

Théorème 3.10 (Caractérisation des syméties orthogonales parmi les symétries).
Si s est une symétrie d'un espace euclidien E, s est une symétrie orthogonale si, et seulement si, s est auto-adjoint (symétrique).

Démonstration. Toue symétrie, *i.e.* tout endomorphisme s de E vérifiant $s \circ s = I_E$, est la symétrie par rapport à $\ker(s - I_E)$ (espace des invariants) parallèlement à $\ker(s + I_E)$ qui est identique à $\mathrm{Im}(s - I_E)$ dans ce cas.

$\boxed{\Longrightarrow}$ Soit s la symétrie orthogonale par rapport à F ; dans une base orthonormale $\mathcal{B}$ adaptée à la somme directe orthogonale $E = F \oplus F^{\perp}$, la matrice de s relative à $\mathcal{B}$ est la matrice diagonale par blocs $\mathrm{Diag}(I_r, -I_{n-r})$ où r est la dimension de F. Cette matrice est symétrique réelle, donc s est auto-adjoint.

$\boxed{\Longleftarrow}$ Si s est auto-adjoint, alors $\ker(s + I_E) = \ker(s^* + I_E) = \ker(s + I_E)^* = \big(\mathrm{Im}(s + I_E)\big)^{\perp} = \big(\ker(s - I_E)\big)^{\perp}$; s est la symétrie orthogonale par rapport à $\ker(s - I_E)$. cqfd

3.4 Endomorphisme symétrique positif, défini positif

Définition 3.4 (Endomorphisme symétrique positif, défini positif).
Soit u un endomorphisme auto-adjoint (symétrique) d'un espace euclidien E ; u est dit *positif* si

$$\forall \mathbf{x} \in E,\ \langle u(\mathbf{x}) \mid \mathbf{x} \rangle \geqslant 0$$

u est dit *défini positif* si

$$\forall \mathbf{x} \in E \setminus \{\mathbf{0}\},\ \langle u(\mathbf{x}) \mid \mathbf{x} \rangle > 0$$

Proposition 3.11 (Endomorphisme symétrique défini positif et automorphisme).
Tout endomorphisme symétrique défini positif est un automorphisme de E.

Démonstration. Le noyau d'un endomorphisme u symétrique défini positif est réduit à $\{\mathbf{0}\}$ car $\mathbf{x} \neq \mathbf{0} \implies \langle u(\mathbf{x}) \mid \mathbf{x} \rangle > 0$, donc $u(\mathbf{x}) \neq \mathbf{0}$. En dimension finie, ceci montre que u est application linéaire inversible, *i.e.* un automorphisme. cqfd

Définition 3.5 (Matrice symétrique positive, définie positive).
Soit $A \in \mathcal{S}_n(\mathsf{R})$ une matrice symétrique réelle d'ordre n ;
A est dite *positive* si

$$\forall X \in \mathcal{M}_{n,1}(\mathsf{R}),\ {}^t X A X \geqslant 0$$

A est dite *définie positive* si

$$\forall X \in \mathcal{M}_{n,1}(\mathsf{R}) \setminus \{0\},\ {}^t X A X > 0$$

Proposition 3.12 (Matrice symétrique positive et endomorphisme associé).
Si $A \in \mathcal{S}_n(\mathsf{R})$ est une matrice symétrique et $u_A : X \mapsto AX$ l'endomrphisme auto-adjoint (symétrique) de $\mathcal{M}_{n,1}(\mathsf{R})$ muni de son produit scalaire naturel, alors

(i) *A est positive $\iff$ u_A est positif ;*
(ii) *A est définie positive $\iff$ u_A est défini positif.*

Démonstration. Remarquons que $\langle u_A(X) \mid X \rangle = {}^t(XA)X = {}^t XAX$, puisque A est symétrique. cqfd

Proposition 3.13 (Cas des matrices diagonales).
Si $D = \mathrm{Diag}(\lambda_1, \dots, \lambda_n)$ est une matrice diagonale à coefficients réels, alors

(i) *D est positive $\iff \forall i \in [\![1, n]\!],\ \lambda_i \geqslant 0$*
(ii) *D est définie positive $\iff \forall i \in [\![1, n]\!],\ \lambda_i > 0$*

Démonstration. Soit $X = {}^t(x_1, \dots, x_n)$; alors ${}^t XDX = \sum_{i=1}^{n} \lambda_i x_i^2$ et

$$\forall X \in \mathcal{M}_{n,1}(\mathsf{R}),\ {}^t XDX = \sum_{i=1}^{n} \lambda_i x_i^2 \geqslant 0 \iff \forall i \in [\![1, n]\!],\ \lambda_i \geqslant 0$$

$$\forall X \in \mathcal{M}_{n,1}(\mathsf{R}) \setminus \{0\},\ {}^t XDX = \sum_{i=1}^{n} \lambda_i x_i^2 > 0 \iff \forall i \in [\![1, n]\!],\ \lambda_i > 0$$

cqfd

Proposition 3.14 (Cas des adjoints).
Si u est un endomorphisme d'un espace euclidien E, alors

(i) *$u^* \circ u$ et $u \circ u^*$ sont des endomorphismes symétriques positifs,* $\ker(u^* \circ u) = \ker u$ *et* $\ker(u \circ u^*) = \ker u^*$ *;*
(ii) *$u \circ u^*$ (resp. $u^* \circ u$) est défini positif si, et seulement si, u est inversible.*

Démonstration.

(i) $\forall \mathbf{x} \in E$, $\langle u^* \circ u(\mathbf{x}) \mid \mathbf{x}\rangle = \langle u(\mathbf{x}) \mid u(\mathbf{x})\rangle = \|u(\mathbf{x})\|^2 \geqslant 0$ et $\langle u \circ u^*(\mathbf{x}) \mid \mathbf{x}\rangle = \langle u^*(\mathbf{x}) \mid u^*(\mathbf{x})\rangle = \|u^*(\mathbf{x})\|^2 \geqslant 0$; ainsi, $u^* \circ u$ et $u \circ u^*$ sont positifs.
$\mathbf{x} \in \ker(u \circ u^*) \iff u^* \circ u(\mathbf{x}) = \mathbf{0} \implies 0 = \langle u^* \circ u(\mathbf{x}) \mid \mathbf{x}\rangle = \|u(\mathbf{x})\|^2 \iff u(\mathbf{x}) = \mathbf{0} \iff \mathbf{x} \in \ker u$, ce qui montre l'inclusion $\ker(u^* \circ u) \subset \ker u$; puisque $\ker u \subset \ker(u^* \circ u)$, l'égalité annoncée est démontrée par double inclusion.
Changeons u en u^* dans l'égalité précédente : $\ker u^* = \ker\big((u^*)^* \circ u^*\big) = \ker(u \circ u^*)$.
(ii) Si $u^* \circ u$ est défini positif, $u^* \circ u$ est inversible et $\{\mathbf{0}\} = \ker(u^* \circ u) = \ker u$, ce qui montre que u est inversible.
Réciproquement, si u est inversible, u^* est inversible, et, par composition, $u^* \circ u$ aussi.
La démonstration est identique dans l'autre cas.

cqfd

Proposition 3.15 (Traduction matricielle).
Si $A \in \mathcal{M}_n(\mathsf{R})$ est une matrice carrée, à coefficients réels et d'ordre n, alors

(i) *${}^tAA$ et $A{}^tA$ sont des matrices symétriques positives,* $\ker({}^tAA) = \ker A$ *et* $\ker(A{}^tA) = \ker {}^tA$ *;*
(ii) *${}^tAA$ (resp. $A{}^tA$) est définie positive si, et seulement si* $\det A \neq 0$.

Démonstration. Appliquons la proposition précédente à l'endomormorphisme u_A de $\mathcal{M}_{n,1}(\mathsf{R})$ associé à A, en utilisant les relations $(u_A)^* = u_{{}^tA}$, $(u_A)^* \circ u_A = u_{{}^tA} \circ u_A = u_{{}^tAA}$, ... cqfd

4 Automorphismes orthogonaux

E est soit un espace préhilbertien réel, soit un espace euclidien de dimension n ; son produit scalaire est noté $\langle \mid \rangle$.

4.1 Généralités

Lemme 4.1 (Conservation de la norme, conservation du produit scalaire).
Si E est un espace préhibertien réel et u une application de E vers E, les propriétés suivantes sont équivalentes :

(i) *l'application u est linéaire et conseve la norme,* i.e. $\forall \mathbf{x} \in E$, $\|u(\mathbf{x})\| = \|\mathbf{x}\|$ *;*
(ii) *l'application u conserve le produit scalaire,* i.e. $\forall (\mathbf{x}, \mathbf{y}) \in E$, $\langle u(\mathbf{x}) \mid u(\mathbf{y})\rangle = \langle \mathbf{x} \mid \mathbf{y}\rangle$.

Démonstration.

$(i) \implies (ii)$ De l'identité $4\langle \mathbf{u} \mid \mathbf{v}\rangle = \|\mathbf{u}+\mathbf{v}\|^2 - \|\mathbf{u}-\mathbf{v}\|^2$, on tire que :

$$\begin{aligned} 4\langle u(\mathbf{x}) \mid u(\mathbf{y})\rangle &= \|u(\mathbf{x}) + u(\mathbf{y})\|^2 - \|u(\mathbf{x}) - u(\mathbf{y})\|^2 \\ &= \|u(\mathbf{x}+\mathbf{y})\|^2 - \|u(\mathbf{x}-\mathbf{y})\|^2 \qquad u \text{ est linéaire} \\ &= \|\mathbf{x}+\mathbf{y}\|^2 - \|\mathbf{x}-\mathbf{y}\|^2 \qquad u \text{ conserve la norme} \\ &= 4\langle \mathbf{x} \mid \mathbf{y}\rangle \end{aligned}$$

$(ii) \implies (i)$ Si u conserve le produit scalaire, u conserve la norme : faire $\mathbf{y} = \mathbf{x}$.
Montrer la linéarité de u, c'est montrer que $\forall (\mathbf{x}, \mathbf{y}, \lambda) \in E^2 \times \mathsf{R}$, le vecteur $u(\lambda\mathbf{x}+\mathbf{y}) - \lambda u(\mathbf{x}) - u(\mathbf{y})$ est le vecteur nul, *i.e.* de norme nulle. De l'identité

$$\|\mathbf{u}+\mathbf{v}+\mathbf{w}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 + \|\mathbf{w}\|^2 + 2\langle \mathbf{u} \mid \mathbf{v}\rangle + 2\langle \mathbf{v} \mid \mathbf{w}\rangle + 2\langle \mathbf{w} \mid \mathbf{u}\rangle$$

on tire :

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

cqfd

Lemme 4.2 (Conservation de la norme et injectivité).
Soit u un endomorphisme d'un espace préhilbertien réel E ; si u conserve la norme, alors u est une injection.

Démonstration. Il suffit de montrer que le noyau de u est réduit à $\{\mathbf{0}\}$: $u(\mathbf{x})=\mathbf{0} \iff \|u(\mathbf{x})\|=0 \iff \|\mathbf{x}\|=0$ (u conserve la norme) $\iff \mathbf{x}=\mathbf{0}$. cqfd

Corollaire. *Soit u un endomorphisme d'un espace euclidien E ; si u conserve la norme, alors u est une bijection.*

Démonstration. Tout endomorphisme injectif dans un espace vectoriel de dimension finie est bijectif, donc un automorphisme. cqfd

Définitions 4.1 (Automorphisme orthogonal).
Soit u un endomorphisme d'un espace euclidien E ; on dit que u est un *automorphisme orthogonal* si u conserve la norme.
L'ensemble des automorphismes orthogonaux de E est noté $\mathcal{O}(E)$. $\mathcal{O}(\mathsf{R}^n)=\mathcal{O}_n(\mathsf{R})$ désigne l'ensemble des automorphismes orthogonaux de R^n muni de son produit scalaire naturel (canonique).

$$\boxed{u\in\mathcal{O}(E) \iff u\in\mathcal{L}(E) \text{ et } \forall\mathbf{x}\in E,\ \|u(\mathbf{x})\|=\|\mathbf{x}\|}$$

Remarques.
Le lemme précédent montre que les endomorphismes d'un espace euclidien qui conservent la norme sont des automorphismes.
Si u conserve le produit scalaire d'un espace euclidien, u est nécessairement linéaire et conserve la norme ; u est donc un automorphisme orthogonal.

Théorème 4.3 (Caractérisation des automorphismes orthogonaux).
Si u est un endomorphisme d'un espace euclidien E, les assertions suivantes sont équivalentes :

(i) $u\in\mathcal{O}(E)$;
(ii) *u conserve le produit scalaire et donc l'orthogonalité et les angles ;*
(iii) $u^*\circ u=I_E$;
(iv) $u\circ u^*=I_E$;
(v) $u\in\mathcal{GL}(E)$ *et* $u^{-1}=u^*$;
(vi) *l'image par u de toute base orthonormale de E est une base orthonormale de E ;*
(vii) *l'image par u d'une base orthonormale de E est une base orthonormale de E.*

Démonstration.

$(i) \implies (ii)$ Le lemme 1 montre la conservation de la norme ;
$\mathbf{x}\perp\mathbf{y} \iff \langle\mathbf{x}\mid\mathbf{y}\rangle=0 \implies \langle u(\mathbf{x})\mid u(\mathbf{y})\rangle=\langle\mathbf{x}\mid\mathbf{y}\rangle=0$, puisque u conserve le produit scalaire $\iff u(\mathbf{x})\perp u(\mathbf{y})$

$$\cos\big(u(\mathbf{x}),u(\mathbf{y})\big)=\frac{\langle u(\mathbf{x})\mid u(\mathbf{y})\rangle}{\|u(\mathbf{x})\|\,\|\mathbf{y}\|}=\frac{\langle\mathbf{x}\mid\mathbf{y}\rangle}{\|\mathbf{x}\|\,\|\mathbf{y}\|}=\cos(\mathbf{x},\mathbf{y})$$

$(ii) \iff (iii)$ $\forall(\mathbf{x},\mathbf{y})\in E^2,\ \langle u^*\circ u(\mathbf{x})\mid\mathbf{y}\rangle=\langle u(\mathbf{x})\mid u(\mathbf{y})\rangle=\langle\mathbf{x}\mid\mathbf{y}\rangle \iff u^*\circ u=I_E$

$(iii) \iff (v)$ $I_E = u^* \circ u \implies 1 = \det I_E = \det u^* \det u$; ainsi $\det u \neq 0$, u est inversible et $u^{-1} = u^*$. La réciproque est évidente.

$(iv) \iff (v)$ Même démonstration.

$(ii) \implies (vi)$ Si $\mathcal{B} = (\mathbf{e}_1, \dots, \mathbf{e}_n)$ est une base orthonormale de E, alors, puisque u conserve le produit scalaire, $\langle u(\mathbf{e}_i) \mid u(\mathbf{e}_j)\rangle = \langle e_i \mid e_j\rangle = \delta i, j$. La famille $\big(u(e_1), \dots, u(e_n)\big)$ est une famille orthonormale de E, maximale, donc une base orthonormale de E.

$(vi) \implies (vii)$ Un espace euclidien admet une base orthonormale.

$(vii) \implies (i)$ Appelons $\mathcal{B} = (\mathbf{e}_1, \dots, \mathbf{e}_n)$ est une base orthonormale de E. Si $\mathbf{x} = \sum_{j=1}^n x_j \mathbf{e}_j$ est un vecteur de E, alors $u(\mathbf{x}) = \sum_{j=1}^n x_j u(\mathbf{e}_j)$ et

$$\|\mathbf{x}\|^2 = \langle \sum_{j=1}^n x_j \mathbf{e}_j \mid \sum_{k=1}^n x_k \mathbf{e}_k \rangle = \sum_{j,k} x_j x_k \langle \mathbf{e}_j \mid \mathbf{e}_k\rangle = \sum_{j=1}^n x_j^2$$
$$\|u(\mathbf{x})\|^2 = \langle \sum_{j=1}^n x_j u(\mathbf{e}_j) \mid \sum_{k=1}^n x_k u(\mathbf{e}_k)\rangle = \sum_{j,k} x_j x_k \langle u(\mathbf{e}_j) \mid u(\mathbf{e}_k)\rangle = \sum_{j=1}^n x_j^2$$

puisque $\mathcal{B}$ et $u(\mathcal{B})$ sont deux bases orthonormales ; ainsi u conserve la norme. cqfd

Théorème 4.4 (Traduction matricielle).
Soient E un espace euclidien, $\mathcal{B}$ une base orthonormale de E, u un endomorphisme de E et A la matrice de u relative à $\mathcal{B}$; Les assertions suivantes sont équivalentes :

(i) $u \in \mathcal{O}(E)$;
(ii) ${}^tAA = I_n$;
(iii) $A{}^tA = I_n$;
(iv) $A \in \mathcal{GL}_n(\mathsf{R})$ *et* $A^{-1} = {}^tA$.

Démonstration. La traduction matricielle de $u^* \circ u = I_E$ est $\mathcal{M}\mathrm{at}_{\mathcal{B}}(u^* \circ u) = \mathcal{M}\mathrm{at}_{\mathcal{B}}(I_E)$, ce qui devient ${}^tAA = I_n$; ainsi $i. \iff ii.$. Les autres équivalences traduisent « $u \circ u^* = I_E$ » et « $u \in \mathcal{GL}(E)$, $u^{-1} = u^*$ ». cqfd

4.2 Exemples

A) Symétrie orthogonale

Rappelons que s est une symétrie orthogonale si, et seulement si, $s \circ s = I_E$ et $s = s^*$; ainsi $s = s^{-1} = s^*$ et toute symétrie orthogonale est un automorphisme orthogonal. Un autre argument aurait pu être employé : les symétries orthogonales conservent la norme.

B) Réflexion

Lemme 4.5. *Si* $\mathbf{a}$ *et* $\mathbf{b}$ *sont deux vecteurs d'un espace préhilbertien E, alors* $\mathbf{a} + \mathbf{b}$ *et* $\mathbf{a} - \mathbf{b}$ *sont orthognaux si, et seulement si,* $\mathbf{a}$ *et* $\mathbf{b}$ *ont même longueur.*

Démonstration. L'égalité $\langle \mathbf{a} + \mathbf{b} \mid \mathbf{a} - \mathbf{b}\rangle = \|\mathbf{a}\|^2 - \|\mathbf{b}\|^2$ donne l'équivalence. cqfd

Corollaire (Parallélogamme et losange).
Les diagonales d'un parallélogramme sont orthogonales si, et seulement si, ce parallélogramme est un losange.

Corollaire (Bissectrice(s) de deux vecteurs).
Si $\mathbf{a}$ *et* $\mathbf{b}$ *ne sont pas deux vecteurs colinéaires de même sens,* $\mathbf{a} + \mathbf{b}$ *dirige la bissectrice des vecteurs* $\mathbf{a}$ *et* $\mathbf{b}$ *si, et seulement si,* $\mathbf{a}$ *et* $\mathbf{b}$ *sont de même longueur.*

Démonstration. Si $\mathbf{a}$ et $\mathbf{b}$ ne sont pas des vecteurs colinéaires de même sens, la quantité $\|\mathbf{a}\|\,\|\mathbf{b}\| - \langle \mathbf{a} \mid \mathbf{b}\rangle$ est stictement positive et

$$\cos(\mathbf{a}+\mathbf{b},\mathbf{a}) - \cos(\mathbf{a}+\mathbf{b},\mathbf{b}) = \frac{\langle \mathbf{a}+\mathbf{b} \mid \mathbf{a}\rangle}{\|\mathbf{a}+\mathbf{b}\|\,\|\mathbf{a}\|} - \frac{\langle \mathbf{a}+\mathbf{b} \mid \mathbf{b}\rangle}{\|\mathbf{a}+\mathbf{b}\|\,\|\mathbf{b}\|} = \frac{\big(\|\mathbf{a}\|\,\|\mathbf{b}\| - \langle \mathbf{a} \mid \mathbf{b}\rangle\big)\big(\|\mathbf{a}\| - \|\mathbf{b}\|\big)}{\|\mathbf{a}+\mathbf{b}\|\,\|\mathbf{a}\|\,\|\mathbf{b}\|}$$

Les angles$(\mathbf{a}+\mathbf{b}, vca)$ et $(\mathbf{a}+\mathbf{b},\mathbf{b})$ ont le même cosinus si, et seulement si, les vecteurs $\mathbf{a}$ et $\mathbf{b}$ sont de même longueur. cqfd

Définition 4.2 (Réflexion).
Étant donnée deux vecteurs $\mathbf{a}$ et $\mathbf{b}$ de *même norme*, on appelle *réflexion* de $\mathbf{a}$ sur $\mathbf{b}$, la symétrie orthogonale par rapport à l'hyperplan médiateur de $\mathbf{a}$ et $\mathbf{b}$, *i.e.* l'hyperplan orthogonal à $\mathbf{e} = \mathbf{a} - \mathbf{b}$.
Cette réflexion est notée $s_{\mathbf{a},\mathbf{b}}$ et

$$s_{\mathbf{a},\mathbf{b}} : \mathbf{x} \in E \mapsto \mathbf{x} - 2\frac{\langle \mathbf{e} \mid \mathbf{x}\rangle}{\|\mathbf{e}\|^2}\mathbf{e} \qquad \text{avec } \mathbf{e} = \mathbf{a} - \mathbf{b}$$

Remarques.
Dauns une base orthogonale adaptée à la decomposition $E = \mathsf{R}\mathbf{e} \oplus \{\mathbf{e}\}^\perp$, la matrice de $s_{\mathbf{a},\mathbf{b}}$ est par blocs $\mathrm{Diag}(-1, I_{n-1})$ et donc $\det s_{\mathbf{a},\mathbf{b}} = -1$.
$s_{\mathbf{a},\mathbf{b}}$ et $s_{\mathbf{a},-\mathbf{b}}$ sont deux réflexions qui envoient la droite $\mathsf{R}\mathbf{a}$ sur la droite $\mathsf{R}\mathbf{b}$.

4.3 Groupe orthogonal

Théorème 4.6 (Groupes orthogonal et spécial orthogonal de E).
Si E est un espace euclidien, alors :

(i) $\mathcal{O}(E)$ *est un sous-groupe de* $\mathcal{GL}(E)$ *appelé* groupe orthogonal *de E ;*
(ii) *si* $u \in \mathcal{O}(E)$, *alors* $\det u \in \{-1, 1\}$ *; la réciproque est* fausse *;*
(iii) $\mathcal{SO}(E) = \{u \in \mathcal{O}(E) \ /\ \det u = 1\}$ *est un sous-groupe de* $\mathcal{O}(E)$ *appelé* groupe spécial orthogonal *de E, ou encore* groupe des rotations vectorielles *de E ; il est encore noté* $\mathcal{O}^+(E)$ *;*
(iv) $\mathcal{O}^-(E) = \mathcal{O}(E) \setminus \mathcal{SO}(E) = \{u \in \mathcal{O}(E) \ /\ \det u = -1\}$ *n'est pas un groupe.*

Démonstration.

(i) $\mathcal{O}(E) \subset \mathcal{GL}(E)$ et $I_E \in \mathcal{O}(E)$;
stabilité de la composition : si u et v appartiennent à $\mathcal{O}(E)$, alors $\forall \mathbf{x} \in E$, $\|v\big(u(\mathbf{x})\big)\| = \|u(\mathbf{x})\| = \|\mathbf{x}\|$;
stabilité de la prise d'inverse : si $u \in \mathcal{O}(E)$, $\|u^{-1}(\mathbf{x})\| = \|u\big(u^{-1}(\mathbf{x})\big)\| = \|\mathbf{x}\|$ et $u^{-1} \in \mathcal{O}(E)$.
Ainsi, $\mathcal{O}(E)$ est un sous-groupe de $\mathcal{GL}(E)$ muni de la composition des applications.

(ii) $u \in \mathcal{O}(E) \iff u^* \circ u = I_E \boxed{\implies} \det(u^* \circ u) \det I_E = 1 = \det u^* \det u = (\det u)^2$ et donc $\det u = \pm 1$
Si $A = \left(\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right)$, alors ${}^tAA = \left(\begin{smallmatrix} 1 & 1 \\ 1 & 2 \end{smallmatrix}\right) \neq I_2$, ce qui montre que la réciproque est fausse.

(iii) $u \mapsto \det u$ est un morphisme du groupe $\big(\mathcal{O}(E), \circ\big)$ sur le groupe $\big(\{-1, 1\}, \times\big)$; le noyau $\mathcal{SO}(E)$ de ce morphisme est un sous-groupe de $\mathcal{O}(E)$;

(iv) $\mathcal{O}^-(E)$ ne peut être un groupe, car la composition des applications n'est pas *stable*. Si s est une réflexion de E, $u \mapsto u \circ s$ est une bijection de $\mathcal{O}(E)$ (c'est une involution) qui envoie $\mathcal{O}^-(E)$ sur $\mathcal{O}^+(E) = \mathcal{SO}(E)$.

cqfd

Théorème 4.7 (Éléments propres d'un automorphisme orthogonal).
Si u est un automorphisme orthogonal d'un espace euclidien, alors

(i) $\mathrm{sp}\, u \subset \{-1, 1\}$, *les sous-espaces vectoriels propres sont orthogonaux ;*
(ii) *u est diagonalisable si, et seulement si, u est une symétrie (vectorielle) orthogonale ;*
(iii) *si F est un sous-espace vectoriel stable pour u, alors $F^\perp$ est aussi un sous-espace vectoriel stable pour u ; u induit sur F et $F^\perp$ des automorphismes orthogonaux ; en particulier, $u(F) = F$ et $u(F^\perp) = F^\perp$;*

(iv) $\big(\ker(u - I_E)\big)^\perp = \operatorname{Im}(u - I_E)$ *et* $\big(\ker(u + I_E)\big)^\perp = \operatorname{Im}(u + I_E)$.

Démonstration.

(i) Si λ est une valeur propre de u et $\mathbf{x}$ un vecteur propre associé, alors $u(\mathbf{x}) = \lambda\mathbf{x}$ et, puisque u conserve la norme, $\|\mathbf{x}\| = \|u(\mathbf{x})\| = |\lambda|\,\|\mathbf{x}\|$; ainsi $|\lambda| = 1$.
Si $\mathbf{x}$ est un vecteur propre associé à la valeur propre 1 (vecteur invariant) et $\mathbf{y}$ un vecteur propre associé à la valeur propre -1, alors, puisque u conserve le produit scalaire, $\langle \mathbf{x} \mid \mathbf{y}\rangle = \langle u(\mathbf{x}) \mid u(\mathbf{y})\rangle = \langle \mathbf{x} \mid -\mathbf{y}\rangle = -\langle \mathbf{x} \mid \mathbf{y}\rangle$, ce qui montre que $\langle \mathbf{x} \mid \mathbf{y}\rangle = 0$. Ainsi $E_1(u) \perp E_{-1}(u)$.

(ii) u est diagonalisable si, et seulement si, $E = E_1(u) \oplus E_{-1}(u)$, *i.e.* si, et seulement si, u est la symétrie par rapport à $E_1(u)$ parallèlement à $E_{-1}(u)$, soit la symétrie orthogonale par rapport à $E_1(u)$

(iii) L'endomorphisme $u|_F$ induit par u sur le sous-espace vectoriel stable F, conserve la norme ; c'est donc un automorphisme orthogonal de F, en particulier une bijection et $u(F) = F$.
Soit $\mathbf{x} \in F^\perp$; poour tout $\mathbf{y} \in F$, il existe $\mathbf{y}' \in F$ tel que $\mathbf{y} = u(\mathbf{y}')$; alors

$$\begin{aligned}
\langle u(\mathbf{x}) \mid \mathbf{y}\rangle &= \langle u(\mathbf{x}) \mid u(\mathbf{y}')\rangle \\
&= \langle \mathbf{x} \mid \mathbf{y}'\rangle && u \text{ conserve le produit scalaire} \\
&= 0 && \mathbf{x} \in F^\perp \text{ et } \mathbf{y}' \in F
\end{aligned}$$

ce qui montre que $u(\mathbf{x}) \in F^\perp$, *i.e.* la stabilité de $F^\perp$. u induit donc sur $F^\perp$ un automorphisme orthogonal et $u(F^\perp) = F^\perp$.

(iv) Deux temps pour la démonstration : d'abord l'inclusion $\operatorname{Im}(u - I_E) \subset \big(\ker(u - I_E)\big)^\perp$, puis l'égalité des dimensions.
$\mathbf{x} \in \operatorname{Im}(u - I_E) \iff \exists \mathbf{x}' \in E,\ \mathbf{x} = (u - I_E)(\mathbf{x}') = u(\mathbf{x}') - \mathbf{x}'$ et $\mathbf{y} \in \ker(u - I_E) \iff u(\mathbf{y}) = \mathbf{y}$ et

$$\begin{aligned}
\langle \mathbf{x} \mid \mathbf{y}\rangle = \langle u(\mathbf{x}') - \mathbf{x}' \mid \mathbf{y}\rangle &= \langle u(\mathbf{x}') \mid \mathbf{y}\rangle - \langle \mathbf{x}' \mid \mathbf{y}\rangle \\
&= \langle u(\mathbf{x}') \mid u(\mathbf{y})\rangle - \langle \mathbf{x}' \mid \mathbf{y}\rangle && \text{car } \mathbf{y} = u(\mathbf{y}) \\
&= 0 && u \text{ conserve le produit scalaire}
\end{aligned}$$

Ainsi $\forall \mathbf{x} \in \operatorname{Im}(u - I_E)$, $\forall \mathbf{x} \in \ker(u - I_E)$, $\mathbf{x} \perp \mathbf{y}$, *i.e.* les sous-espaces $\operatorname{Im}(u - I_E)$ et $\ker(u - I_E)$ sont orthogonaux et $\operatorname{Im}(u - I_E) \subset \ker(u - I_E)$.
Les dimensions sont égales car :

$$\dim \operatorname{Im}(u - I_E) = \dim E - \dim \ker(u - I_E) = \dim\big(\ker(u - I_E)\big)^\perp$$

La démonstration est identique pour l'égalité de $\operatorname{Im}(u + I_E)$ avec $\big(\ker(u + I_E)\big)^\perp$.
Ainsi, $\ker(u - I_E)$ et $\operatorname{Im}(u - I_E)$ sont supplémentaires orthogonaux, ainsi que $\ker(u + I_E)$ et $\operatorname{Im}(u + I_E)$.

cqfd

5 Matrices orthogonales

5.1 Généralités

Définition 5.1 (Matrice orthogonale).
Une matrice A carrée d'ordre n à coefficients réels est dite *orthogonale* si ${}^tAA = I_n$.
L'ensemble des matrices orthogonales est notée $\mathcal{O}(n)$.

Exemple 5.1. Les matrices de pemutation sont des matrices orthogonales.

Théorème 5.1 (Caractérisation des matrices orthogonales).
Si E est un espace euclidien de dimension n et $A \in \mathcal{M}_n(\mathbb{R})$, alors les propriétés suivantes sont équivalentes :

(i) *$A \in \mathcal{O}(n)$;*
(ii) *${}^tAA = I_n$;*
(iii) *$A{}^tA = I_n$;*
(iv) *$A \in \mathcal{GL}_n(\mathsf{R})$ et $A^{-1} = {}^tA$;*
(v) *A est la matrice relative à une base* orthonormale *d'un automorphisme orthogonal de E ;*
(vi) *les colonnes de A constituent une base orthonormale de* R^n *pour le produit scalaire naturel ;*
(vii) *les lignes de A constituent une base orthonormale de* R^n *pour le produit scalaire naturel ;*
$(viii)$ *A est la matrice de passage d'une base* orthonormale *de E à une base* orthonormale *de E.*

Démonstration.
L'égalité $BA = I_n$ implique $\det B \det A = 1$; ainsi $\det A \neq 0$, A est inversible et $A^{-1} = B$. Cette remarque montre l'équivalence de (ii) et de (iii) avec (iv).

$(ii) \iff (v)$ Déjà vu lors de la caractérisation des automorphismes orthogonaux.

$(ii) \iff (vi)$ Notons $(C_1, \dots, C_n)$ les colonnes de A et remarquons que $\left(I_n\right)_{i,j} = \delta_{i,j}$. Ainsi

$$\left({}^tA\,A\right)_{i,j} = \left(\begin{pmatrix} {}^tC_i & \cdots\cdots \\ & \\ & \end{pmatrix}\begin{pmatrix} & C_j & \\ & \vdots & \\ & \vdots & \end{pmatrix}\right)_{i,j} = {}^tC_i\, C_j = \langle C_i \mid C_j\rangle$$

ce qui montre que

$${}^tA\,A = I_n \iff \forall (i,j) \in [\![1,n]\!]^2,\ \langle C_i \mid C_j\rangle = \delta_{i,j}$$

$(iii) \iff (vii)$ Notons $(L_1, \dots, L_n)$ les colonnes de A. Ainsi

$$\left({}^tA\,A\right)_{i,j} = \left(\begin{pmatrix} L_i & \cdots\cdots \\ & \\ & \end{pmatrix}\begin{pmatrix} & {}^tL_j & \\ & \vdots & \\ & \vdots & \end{pmatrix}\right)_{i,j} = L_i{}^tL_j = \langle L_i \mid L_j\rangle$$

ce qui montre que

$${}^tA\,A = I_n \iff \forall (i,j) \in [\![1,n]\!]^2,\ \langle L_i \mid L_j\rangle = \delta_{i,j}$$

$(ii) \iff (viii)$ Soient $\mathcal{B}$ une base *orthonomale* de E, $\mathcal{B}' = (\mathbf{e}'_1, \dots, \mathbf{e}'_n)$ une autre base, a priori quelconque, de E et A la matrice de passage de la base $\mathcal{B}$ à la base $\mathcal{B}'$; ainsi la j^{e} colonne C_j de A est la matrice $\mathcal{M}\text{at}_{\mathcal{B}}(\mathbf{e}'_j)$ de $\mathbf{e}'_j$ relative à $\mathcal{B}$. Puisque $\mathcal{B}$ est une base orthonoormale, on a :

$$\langle \mathbf{e}'_i \mid \mathbf{e}'_j\rangle = {}^t\mathcal{M}\text{at}_{\mathcal{B}}(\mathbf{e}'_i)\,\mathcal{M}\text{at}_{\mathcal{B}}(\mathbf{e}'_j) = {}^tA_i\,A_j = \left({}^tA\,A\right)_{i,j}$$

ce qui donne l'équivalence. cqfd

5.2 Goupe orthogonal d'ordre n

Théorème 5.2 (Groupe orthogonal et spécial orthogonal d'ordre n).

(i) *$\mathcal{O}(n)$ est un sous-groupe de $\mathcal{GL}_n(\mathsf{R})$ appelé* groupe orthogonal d'ordre *n ;*
(ii) *si $A \in \mathcal{O}(n)$, alors* $\det A \in \{-1, 1\}$ *; la réciproque est fausse ;*
(iii) *$\mathcal{SO}(n) = \{A \in \mathcal{O}(n) \ /\ \det A = 1\}$ est un sous-groupe de $\mathcal{O}(n)$ appelé* groupe spécial orthogonal d'ordre *n, ou encore* groupe des rotations vectorielles d'ordre *n ; il est encore noté $\mathcal{O}^+(n)$;*
(iv) *$\mathcal{O}^-(n) = \mathcal{O}(n) \setminus \mathcal{SO}(n) = \{A \in \mathcal{O}(n) \ /\ \det u = -1\}$ n'est pas un groupe.*

Démonstration.

C'est la traduction matricielle du théorème correspondant des propriétés du groupe orthogonal $\mathcal{O}(E)$, où E est un espace euclidien de dimension n rapporté à une base orthonormale.

Pourtant, dans sa grande bonté, le scribe de service va en donner une démonstration spécifique.

(i) $\mathcal{O}(n) \subset \mathcal{GL}_n(\mathsf{R})$ et $I_n \in \mathcal{O}(n)$;
stabilité de la multiplication : si A et B appartiennent à $\mathcal{O}(n)$, alors
${}^t(AB)AB = {}^tB{}^tA\,AB = {}^tBI_nB = I_n$;
stabilité de la prise d'inverse : si $A \in \mathcal{O}(n)$, alors
$I_n = ({}^tA\,A)^{-1} = A^{-1}({}^tA)^{-1} = A^{-1}\,{}^tA^{-1}$ et $A^{-1} \in \mathcal{O}(n)$.
Ainsi $\mathcal{O}(n)$ est un sous-groupe de $\mathcal{GL}_n(\mathsf{R})$ muni du produit des matrices.

(ii) ${}^tA\,A = I_n \implies 1 = \det I_n = \det({}^tA\,A) = \det{}^tA\,\det A = (\det A)^2$, et donc $\det A \in \{-1, 1\}$.

(iii) $A \mapsto \det A$ est un morphisme du groupe $\big(\mathcal{O}(n), \times\big)$ sur le groupe $\big(\{-1, 1\}, \times\big)$; le noyau $\mathcal{SO}(n)$ de ce morphisme est un sous-groupe de $\mathcal{O}(n)$;

(iv) $\mathcal{O}^-(n)$ ne peut être un groupe car la multiplication n'est pas stable : le produit de deux matrices orthogonales de déterminant -1 est une matrice orthogonale de déterminant $+1$. Si $S = \mathrm{Diag}(-1, I_{n-1})$ (matrice diagonale par blocs), $A \mapsto SA$ est une bijection de $\mathcal{O}(n)$ (c'est une involution) qui envoie $\mathcal{O}^-(n)$ sur $\mathcal{O}^+(n) = \mathcal{SO}(n)$.

cqfd

Théorème 5.3 (Éléments propres d'une matrice orthogonale).

(i) *Toute propriété vraie pour les automorphismes orthogonaux, est encore vraie pour les matrices orthogonales ; en particulier, le spectre d'une matrice orthogonale est inclus dans* $\{-1, 1\}$ *;*

(ii) *si* $A \in \mathcal{O}(n) \cap \mathcal{SO}(n)$, *A est la matrice d'une symétrie orthogonale relativement à une base orthonormale ;*

(iii) *si n est* impair *et* $A \in \mathcal{SO}(n) = \mathcal{O}^+(n)$, *alors 1 est valeur propre de A ;*

(iv) *si* $A \in \mathcal{O}^-(n)$, *alors* -1 *est valeur propre de A ;*

Démonstration.

(i) Toute matrice orthogonale est la matrice d'un automorphisme orthogonal relativement à une base orthonormale. Toute les propriétés vraies pour les automorphismes orthogonaux, sont donc vraies pour les matrices orthogonales ; en particulier le spectre d'une matrice orthogonale est inclus dans $\{-1, 1\}$.
Voici une démonstration directe de cette propriété. Si λ est une valeur propre de A et X un vecteur propre associé, $AX = \lambda X$ et

$$\lambda^2\|X\|^2 = \|AX\|^2 = {}^t(AX)AX = {}^tX{}^tAAX = {}^tXX = \|X\|^2$$

et $\lambda^2 = 1$.

(ii) Considérons l'endomorphisme u de E de matrice A relativement à une base orthonormale $\mathcal{B}$; alors :

$$ {}^tA\,A = I_n \iff u^* \circ u = I_E \iff u \text{ est un automorphisme orthogonal ;}$$
$$ {}^tA = A \iff u^* = u \iff u \text{ est auto-adjoint ;}$$

Puisque qu'une symétrie (vectorielle) orthogonale est un automorphisme orthogonal et auto-adjoint, le résultat annoncée est démontré.

(iii) Montrer que 1 est valeur propre de A, c'est montrer la nullité de $\det(A - I_n)$. Puisque $A \in \mathcal{SO}(n)$, $A - I_n = A - {}^tA\,A = A(I_n - {}^tA)$ et $\det A = 1$. Ainsi,

$$\det(A - I_n) = \det A\,\det(I_n - {}^tA) = 1 \times \det\big({}^t(I_n - A)\big) = \det(I_n - A) = (-1)^n\det(A - I_n)$$

Puisque n est impair, $\det(A - I_n) = -\det(A - I_n)$ et $\det(A - I_n) = 0$.

(iv) Si $A \in \mathcal{O}^-(n)$, en utilisant la même transformation que ci-dessus, on obtient :

$$\det(A + I_n) = \det(A + A{}^tA) = \det A\,\det(I_n + {}^tA) = -1 \times \det\big({}^t(I_n + A)\big) = -\det(A + I_n)$$

Ainsi, $\det(A + I_n) = 0$ et -1 est valeur propre de A.

cqfd

5.3 Le groupe orthogonal d'ordre 1

$\mathcal{O}(1) = \{1, -1\}$, $\mathcal{SO}(1) = \mathcal{O}^+(1) = \{1\}$ et $\mathcal{O}^-(1) = \{-1\}$.
Notons $\mathcal{D}$ un espace euclidien de dimension 1, *i.e.* une droite vectorielle euclidienne ; alors
$\mathcal{O}(\mathcal{D}) = \{I_\mathcal{D}, -I_\mathcal{D}\}$, $\mathcal{SO}(\mathcal{D}) = \mathcal{O}^+(\mathcal{D}) = \{I_\mathcal{D}\}$ et $\mathcal{O}^-(\mathcal{D}) = \{-I_\mathcal{D}\}$.

5.4 Le groupe orthogonal d'ordre 2

$\mathcal{SO}(2) = \mathcal{O}^+(2) = \left\{ \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix},\ \theta \in \mathsf{R} \right\}$
$\mathcal{O}^-(2) = \left\{ \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix},\ \theta \in \mathsf{R} \right\}$
Notons $\mathcal{P}$ un espace euclidien de dimension 2, *i.e.* un plan vectoriel euclidien ; alors
$\mathcal{SO}(\mathcal{P}) = \mathcal{O}^+(\mathcal{P}) =$ { rotations vectorielles planes } et
$\mathcal{O}^-(\mathcal{P}) =$ { symétries vectorielles planes et orthogonales }
$=$ { réflexions de $\mathbf{i} = (1, 0)$ sur $\mathbf{u}_\theta = (\cos\theta, \sin\theta)$, $\theta \in \mathsf{R}$ }.

5.5 Le groupe orthogonal d'ordre 3

$A \in \mathcal{SO}(3) = \mathcal{O}^+(3)$ si, et seulement si, A est une matrice orthogonale dont la troisième colonne est le produit vectoriel de ses deux premières, si, et seulement si, il existe $P \in \mathcal{O}^+(3)$, matrice de passage de la base canonique de $\mathcal{M}_{3,1}(\mathsf{R})$ (identifié à R^3) à une base orthonormale *directe*, et $\theta \in \mathsf{R}$ tel que

$$P^{-1}AP = {}^tPAP = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$A \in \mathcal{O}^-(3)$ si et seulement si $-A \in \mathcal{O}^+(3)$, si, et seulement si, A est une matrice orthogonale dont la troisième colonne est l'opposé du produit vectoriel de ses deux premières, si, et seulement si, il existe $P \in \mathcal{O}^+(3)$, matrice de passage de la base canonique de $\mathcal{M}_{3,1}(\mathsf{R})$ (identifié à R^3) à une base orthonormale *directe*, et $\theta \in \mathsf{R}$ tel que

$$P^{-1}AP = {}^tPAP = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

Notons $\mathcal{E}$ un espace euclidien de dimension 3, *i.e.* l'espace vectoriel euclidien habituel de la physique (newtonienne) et u un automorphisme orthogonal de $\mathcal{E}$ distinct de $\pm I_\mathcal{E}$.
$u \in \mathcal{SO}(\mathcal{E}) = \mathcal{O}^+(\mathcal{E})$ si, et seulement si, $\mathcal{E} = \ker(u - I_\mathcal{E}) \oplus \big(\ker(u - I_\mathcal{E})\big)^\perp = \mathsf{R}\mathbf{w} \oplus \mathcal{P}$ avec $u|_{\mathsf{R}\mathbf{w}} = I_{\mathsf{R}\mathbf{w}}$ et $u|_\mathcal{P}$ une rotation vectorielle du plan $\mathcal{P}$; ainsi, u est une rotation axiale.
$u \in \mathcal{O}^-(\mathcal{E})$ si, et seulement si, $\mathcal{E} = \ker(u + I_\mathcal{E}) \oplus \big(\ker(u + I_\mathcal{E})\big)^\perp = \mathsf{R}\mathbf{w} \oplus \mathcal{P}$ avec $u|_{\mathsf{R}\mathbf{w}} = -I_{\mathsf{R}\mathbf{w}}$ et $u|_\mathcal{P}$ une rotation vectorielle du plan $\mathcal{P}$; ainsi, u est le composé (commutatif) d'une rotation axiale et d'une symétrie orthogonale par rapport au plan orthogonal à l'axe de la rotation, *i.e.* d'une réflexion qui laisse globalement invariant l'axe de la rotation.
Remarques.
Si $u \in \mathcal{SO}(\mathcal{E}) = \mathcal{O}^+(\mathcal{E})$, $\operatorname{tr} u = 1 + 2\cos\theta$, ce qui détermine $\cos\theta$ (exactement) et $\sin\theta$ (au signe près) ; si $\mathbf{u}$ est un vecteur orthogonal à l'axe $\mathsf{R}\mathbf{w}$ de la rotation u, l'égalité $\mathbf{u} \wedge u(\mathbf{u}) = \sin\theta\mathbf{w}$ donne le signe de θ.
Si $u \in \mathcal{O}^-(\mathcal{E})$, $\operatorname{tr} u = -1 + 2\cos\theta$, ce qui détermine $\cos\theta$ (exactement) et $\sin\theta$ (au signe près) ; si $\mathbf{u}$ est un vecteur orthogonal à l'axe $\mathsf{R}\mathbf{w}$ de l'anti-rotation u, l'égalité $\mathbf{u} \wedge u(\mathbf{u}) = \sin\theta\mathbf{w}$ donne le signe de θ.
Si $\mathbf{u}$ est un vecteur *unitaire* et θ un nombre réel, la rotation $r_{\mathbf{u},\theta}$ d'axe $\mathsf{R}\mathbf{u}$ et d'angle θ est donnée par la formule due à Euler :

$$r_{\mathbf{u},\theta} : \mathbf{x} \mapsto \cos\theta\,\mathbf{x} + (1 - \cos\theta)\langle\mathbf{u} \mid \mathbf{x}\rangle\mathbf{u} + \sin\theta(\mathbf{u} \wedge \mathbf{x}) \tag{5.1}$$

6 Réduction des endomorphismes auto-adjoints

E est un espace euclidien de dimension n.

6.1 Réalité des valeurs propres et orthogonalité des sous-espaces vectoriels propres d'un endomorphisme auto-adjoint (symétrique)

Théorème 6.1. *Si u est un endomorphisme auto-adjoint (symétrique) d'un espace euclidien E, alors*

(i) *le polynôme caractéristique de u est scindé sur R ;*

(ii) *les sous-espaces vectoriels propres de u associés à des valeurs propres distinctes sont orthogonaux.*

Démonstration.

(i) Soient $\mathcal{B}$ une base orthonormale de E et $A = \mathcal{M}\text{at}_{\mathcal{B}}(u)$ la matrice de u relative à $\mathcal{B}$. Puisque u est auto-adjoint, A est une matrice symétrique réelle et $\chi_A = \chi_u$.
χ_A, polynôme réel, est scindé sur C ; si $\lambda \in \text{sp}_{\mathsf{C}}(A)$ est une valeur propre complexe de A et $X \in \mathcal{M}_{n,1}(\mathsf{C})$ un vecteur propre associé, $\overline{\lambda} \in \text{sp}_{\mathsf{C}}(A)$ est une valeur propre complexe de A et $\overline{X} \in \mathcal{M}_{n,1}(\mathsf{C})$ un vecteur propre associé, car

$$\begin{aligned} AX = \lambda X \implies \overline{AX} &= \overline{\lambda X} = \overline{\lambda}\,\overline{X} \\ &= \overline{A}\,\overline{X} = A\overline{X} \end{aligned}$$

Ainsi

$$\begin{aligned} {}^t\overline{X}AX &= {}^t\overline{X}(AX) = {}^t\overline{X}(\lambda X) = \lambda {}^t\overline{X}X = \lambda\|X\|^2 \\ &= ({}^t\overline{X}A)X = {}^t({}^tA\overline{X})X = {}^t(A\overline{X})X = {}^t(\overline{\lambda}\,\overline{X})X = \overline{\lambda}{}^t\overline{X}X = \overline{\lambda}\|X\|^2 \end{aligned}$$

et, puisque $X \neq 0$, $\|X\| > 0$ et $\lambda = \overline{\lambda}$, ce qui montre que toutes les valeurs propres complexes de A, donc de u, sont en fait réelles et $\chi_A = \chi_u$ est un poynôme scindé sur R.

(ii) Si λ et μ sont deux valeurs propres distinctes de u, et $\mathbf{x}$ et $\mathbf{y}$ deux vecteurs propres associés, alors

$$\begin{gathered} \langle \mathbf{x} \mid u(\mathbf{y})\rangle = \langle \mathbf{x} \mid \mu\mathbf{y}\rangle = \mu\langle \mathbf{x} \mid \mathbf{y}\rangle \\ = \\ \langle u(\mathbf{x}) \mid \mathbf{y}\rangle = \langle \lambda\mathbf{x} \mid \mathbf{y}\rangle = \lambda\langle \mathbf{x} \mid \mathbf{y}\rangle \end{gathered}$$

Puisque $\lambda \neq \mu$, $\langle \mathbf{x} \mid \mathbf{y}\rangle = 0$ et donc :

$$\lambda \neq \mu \implies \forall(\mathbf{x},\mathbf{y}) \in E_\lambda(u) \times E_\mu(u),\ \mathbf{x} \perp \mathbf{y} \iff E_\lambda(u) \perp E_\mu(u)$$

cqfd

6.2 Diagonalisation des endomorphismes auto-adjoints (symétriques)

Théorème 6.2 (fondamental).
Tout endomorphisme auto-adjoint u d'un espace euclidien E est diagonalisable ; plus précisément, E admet une base orthonormale constituée de vecteurs propres de u.

Démonstration. Par récurrence sur la dimension n de E.
Initalisation : $n = 1$. Dans ce cas E est une droite vectorielle et u est une homothétie ; si $\mathbf{a}$ est un vecteur non nul de E, $\|\mathbf{a}\|^{-1}\mathbf{a}$ est une base orthonormale qui convient.
Hérédité : $n \implies n+1$. Supposons l'existence d'une base orthonormale constituée de vecteurs propres pour tout endomorphisme auto-adjoint d'un espace euclidien de dimension n. Soient E un espace euclidien de dimension

$n+1$ et u un endomorphisme auto-adjoint de E. Puisque u est scindé sur R, prenons λ une valeur propre (réelle) de u et $\mathbf{a}$ un vecteur propre (non nul) associé. La droite $\mathcal{D} = \mathbb{R}\mathbf{a}$ est stable par u ; l'hyperplan $\mathcal{H} = \mathcal{D}^{\perp}$ est stable par $u^* = u$. Ainsi, u induit sur $\mathcal{H}$ un endomorphisme auto-adjoint et, puisque $\mathcal{H}$ est un R-espace vectoriel de dimension n, l'hypothèse de récurrence montre l'existence d'une base orthonormale de $\mathcal{H}$ constituée de vecteurs propres de u. En complétant cette base par $\|\mathbf{a}\|^{-1}\mathbf{a}$, on obtient le résultat. cqfd

Théorème 6.3 (Traduction matricielle).
Toute matrice symétrique réelle A est diagonalisable (sur R) ; plus précisément, il existe une matrice orthogonale $P \in \mathcal{O}(n)$ et une matrice diagonale D telles que :

$$D = P^{-1}AP = {}^tPAP = \mathrm{Diag}(\lambda_1, \dots, \lambda_n)$$

Les colonnes de P constituent une base orthonormale de vecteurs propres de A et λ_j est la valeur propre associé à la j^e colonne.

Démonstration. L'endomorphisme $u_A : X \mapsto AX$ associé à A est un endomorphisme auto-adjoint de $\mathcal{M}_{n,1}(\mathbb{R})$ muni de son produit scalaire naturel ; il existe donc une base orthonormale $\mathcal{B}$ de $\mathcal{M}_{n,1}(\mathbb{R})$, constituée de vecteurs propres de u_A, *i.e.* de vecteurs propres de A. La matrice de passage P de la base canonique, qui est orthonormale, à $\mathcal{B}$ est une matrice orthogonale et donne le résultat. cqfd

6.3 Applications

Proposition 6.4 (Spectre d'un endomorphisme positif, défini positif).
Soit u un endomorphisme auto-adjoint d'un espace euclidien E ; alors

(i) *u est positif,* i.e. $\forall \mathbf{x} \in E,\ \langle u(\mathbf{x}) \mid \mathbf{x} \rangle \geqslant 0 \iff \mathrm{sp}\, u \subset \mathbb{R}_+ = [0, +\infty[$;

(ii) *u est défini positif,* i.e. $\forall \mathbf{x} \in E \setminus \{\mathbf{0}\},\ \langle u(\mathbf{x}) \mid \mathbf{x} \rangle > 0 \iff \mathrm{sp}\, u \subset \mathbb{R}_+^* =]0, +\infty[$;

Démonstration. Soit $\mathcal{B} = (\mathbf{e}_1, \dots, \mathbf{e}_n)$ une base orthonormale de E constituée de vecteurs propres de E ; on appelle λ_j la valeur propre associée au vecteur propre $\mathbf{e}_j$ et $(x_1, \dots, x_n)$ les composantes de $\mathbf{x}$ relatives à $\mathcal{B}$. Ainsi, pour tout $\mathbf{x} \in E$,

$$u(\mathbf{x}) = u\Big(\sum_{j=1}^{n} x_j \mathbf{e}_j\Big) = \sum_{j=1}^{n} x_j u(\mathbf{e}_j) = \sum_{j=1}^{n} x_j \lambda_j \mathbf{e}_j$$

et

$$\langle u(\mathbf{x}) \mid \mathbf{x} \rangle = \langle \sum_{j=1}^{n} x_j \lambda_j \mathbf{e}_j \mid \sum_{k=1}^{n} x_k \mathbf{e}_k \rangle = \sum_{j,k} \lambda_j x_j x_k \langle \mathbf{e}_j \mid \mathbf{e}_k \rangle = \sum_{j=1}^{n} \lambda_j x_j^2$$

On obtient donc

$$\forall \mathbf{x} \in E,\ 0 \leqslant \langle u(\mathbf{x}) \mid \mathbf{x} \rangle = \sum_{j=1}^{n} \lambda_j x_j^2 \iff \forall j \in [\![1, n]\!],\ \lambda_j \geqslant 0 \iff \mathrm{sp}\, u \subset \mathbb{R}_+$$

$$\forall \mathbf{x} \in E \setminus \{\mathbf{0}\},\ 0 < \langle u(\mathbf{x}) \mid \mathbf{x} \rangle = \sum_{j=1}^{n} \lambda_j x_j^2 \iff \forall j \in [\![1, n]\!],\ \lambda_j > 0 \iff \mathrm{sp}\, u \subset \mathbb{R}_+^*$$

cqfd

Proposition 6.5 (Spectre de ${}^tA\,A$).
Si $A \in \mathcal{M}_{p,n}(\mathbb{R})$, alors ${}^tA\,A \in \mathcal{S}_n(\mathbb{R})$ et $\mathrm{sp}({}^tAA) \subset \mathbb{R}_+$.

Démonstration. ${}^tA\,A$ est une matrice carrée d'ordre n, symétrique réelle ; l'endomorphisme de $\mathcal{M}_{n,1}(\mathbb{R})$ associé est positif, car

$$\forall X \in \mathcal{M}_{n,1}(\mathbb{R}),\ \langle u_A(X) \mid X \rangle = {}^tX({}^tA\,A)X = {}^t(AX)AX = \|AX\|^2 \geqslant 0$$

cqfd

Proposition 6.6 (Caractérisation de la matrice d'un produit scalaire).
Soient φ une forme bilinéaire symétrique sur un espace euclidien E et $G = [\varphi(\mathbf{e}_i, \mathbf{e}_j)]$ la matrice de φ dans une base (quelconque) de E ; alors

$$\varphi \textit{ est un produit scalaire sur } E \iff \operatorname{sp} G \subset \mathsf{R}_+^* =]0, +\infty[$$

Démonstration.

$$\begin{aligned}
\varphi \text{ est un produit scalaire sur } E &\iff \forall \mathbf{x} \in E \setminus \{\mathbf{0}\},\ \varphi(\mathbf{x}, \mathbf{x}) > 0 \\
&\iff \forall X \in \mathcal{M}_{n,1}(\mathsf{R}) \setminus \{0\},\ {}^tXGX > 0 \\
&\iff \text{l'endomorphisme } u_G \text{ associé à } G \text{ est défini positif} \\
&\iff \operatorname{sp} u_G = \operatorname{sp} G \subset \mathsf{R}_+^* =]0, +\infty[
\end{aligned}$$

cqfd

7 Réduction des formes bilinéaires symétriques

7.1 Réduction des formes bilinéaires symétriques

Soient E un espace euclidien de dimension n, $\mathcal{B} = (\mathbf{e}_1, \dots, \mathbf{e}_n)$ une base orthonormale de E (la base de référence), φ une forme bilinéaire symétrique sur E et u_φ l'endomorphisme auto-adjoint de E associé :

$$\forall (\mathbf{x}, \mathbf{y}) \in E^2,\ \varphi(\mathbf{x}, \mathbf{y}) = \langle u_\varphi(\mathbf{x}) \mid \mathbf{y} \rangle = \langle \mathbf{x} \mid u_\varphi(\mathbf{y}) \rangle$$

Notons $A = [a_{i,j}] = [\varphi(\mathbf{e}_i, \mathbf{e}_j)]$ la matrice de φ relative à $\mathcal{B}$, $X = {}^t(x_1, \dots, x_n) = \mathcal{M}\text{at}_{\mathcal{B}}(\mathbf{x})$ et $Y = {}^t(y_1, \dots, y_n) = \mathcal{M}\text{at}_{\mathcal{B}}(\mathbf{y})$; ainsi

$$\forall (\mathbf{x}, \mathbf{y}) \in E^2,\ \varphi(\mathbf{x}, \mathbf{y}) = {}^tXAY = \sum_{i,j=1}^{n} a_{i,j} x_i y_j$$

Puisque $\mathcal{B}$ est une base orthonormale, les égalités

$$a_{i,j} = \varphi(\mathbf{e}_i, \mathbf{e}_j) = \langle u_\varphi(\mathbf{e}_i) \mid \mathbf{e}_j \rangle$$

montrent que $a_{i,j}$ est la composante suivant $\mathbf{e}_j$ de $u_\varphi(\mathbf{e}_i)$; ainsi A est aussi la matrice de u_φ relative à $\mathcal{B}$.
Relativement à une base orthonormale $\mathcal{B}'$ constituée de vecteurs propres de u_φ, la matrice de u_φ relative à $\mathcal{B}'$ est diagonale ; en notant P la matrice (orthogonale) de passage de $\mathcal{B}$ à $\mathcal{B}'$, on a :

$$\mathcal{M}\text{at}_{\mathcal{B}'}(u_\varphi) = \operatorname{Diag}(\lambda_1, \dots, \lambda_n) = P^{-1}AP = {}^tPAP$$

En posant $X = PX'$ et $Y = PY'$, il vient

$$\varphi(\mathbf{x}, \mathbf{y}) = {}^t(PX')A(PY') = {}^tX'({}^tPAP)Y' = {}^tX'DY' = \sum_{j=i}^{n} \lambda_j x'_j y'_j$$

Nous venons de *réduire* la forme bilinéaire symétrique φ dans une base orthonormale de E et

$$\varphi(\mathbf{x}, \mathbf{x}) = \sum_{j=1}^{n} \lambda_j {x'_j}^2 \qquad \text{avec } X' = {}^t(x'_1, \dots, x'_n) = \mathcal{M}\text{at}_{\mathcal{B}'}(\mathbf{x})$$

7.2 Réduction des formes quadratiques sur R^n

Définition 7.1 (Forme quadratique à n variables).
Une *forme quadratique* q à n variables, ou sur R^n, est un polynôme homogène de degré 2 de ces n variables ; cette forme quadratique s'écrit :

$$q(x_1, x_2, \ldots, x_n) = \sum_{i,j=1}^{n} a_{i,j} x_i, x_j = \sum_{i=1}^{n} a_{i,i}\, {x_i}^2 + 2 \sum_{1 \leqslant i < j \leqslant n} a_{i,j}\, x_i x_j$$

en prenant la convention $a_{j,i} = a_{i,j}$.

Écriture matricielle d'une forme quadratique

La matrice $A = [a_{i,j}]$, symétrique, réelle, d'ordre n, est dite associée à la forme quadratique q et on a :

$$q(x_1, x_2, \ldots, x_n) = \sum_{i,j=1}^{n} a_{i,j}\, x_i x_j = {}^t XAX$$

Forme bilinéaire symétrique associée

Puisque A est une matrice symétrique réelle, il existe une matrice orthogonale $P \in \mathcal{O}(n)$ telle que $P^{-1}AP = {}^tPAP = \mathrm{Diag}(\lambda_1, \ldots, \lambda_n)$. En posant $X = PX'$, il vient :

$$q(x_1, \ldots, x_n) = {}^t(PX')A(PX') = {}^tX'({}^tPAP)X' = {}^tX'\mathrm{Diag}(\lambda_1, \ldots, \lambda_n)X' = \sum_{j=i}^{n} \lambda_j {x'_j}^2$$

Nous venons de *réduire* la forme quadratique q en *somme de carrés* dans une base orhononormale de $\mathcal{M}_{n,1}(\mathsf{R})$ (que l'on peut identifier à R^n), base orthonormale constituée de vecteurs propres de A.

TABLE DES MATIÈRES