Table des matières

Chapitre 1

Produit scalaire et espaces préhilbertiens

1. Définitions

1.1. Le cas réel.

DÉFINITION 1.1. Soit E un $\mathbf{R}$-espace vectoriel. Un **produit scalaire** dans E est une application $\langle\cdot,\cdot\rangle\colon E\times E\to\mathbf{R}$ qui est :
(l'image de $x, y \in E$ sera notée $\langle x,y\rangle$)

— Symétrique : pour tous $x, y \in E$,

$$\langle x,y\rangle = \langle y,x\rangle \qquad \forall\, x,y\in E.$$

— **Linéaire à droite**. Autrement dit, si on fixe $x\in E$ (l'argument de gauche), alors l'application $y\mapsto\langle x,y\rangle$ est linéaire : pour tous $y, y'\in E$ et $\alpha\in\mathbf{R}$,

$$\langle x, y+\alpha y'\rangle = \langle x,y\rangle + \alpha\langle x,y'\rangle.$$

— **Définie positive** : pour tout $x\in E$, si $x\neq 0$ alors

$$\langle x,x\rangle > 0.$$

On définit la **norme** associée au produit scalaire par :

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

(pour $x\in E$)

EXAMPLE. (i) Produit scalaire standard dans $\mathbf{R}^n$:

$$\langle X,Y\rangle = {}^tXY = \sum_1^n x_iy_i, \qquad \|X\| = \sqrt{\sum_1^n x_i^2}.$$

(ii) $E = C([a,b],R)$,

$$\langle f,g\rangle = \int_a^b f(t)g(t)dt.$$

EXERCICE 1.2. Montrer que ce sont des produit scalaire.

REMARQUE 1.3. Soit $\langle\cdot,\cdot\rangle$ un produit scalaire sur un $\mathbf{R}$-e.v. E.

(i) Puisqu'il est linéaire à droite et symétrique,le produit scalaire est également linéaire à gauche. Ainsi, le produit scalaire est **bilinéaire** (linéaire à droite et à gauche).

(ii) Si $x=0$ alors $\langle x,x\rangle=0$. Ainsi, $\langle x,x\rangle\geq 0$ pour **tout** $x\in E$ (nul ou non).

(iii) On a :

$$\|x+y\|^2 = \|x\|^2 + 2\langle x,y\rangle + \|y\|^2, \qquad \|\alpha x\|^2 = \alpha^2\|x\|^2$$

On observe que si $\langle\cdot,\cdot\rangle$ est un produit scalaire alors

$$x=0 \quad\Longleftrightarrow\quad \forall y\ \langle x,y\rangle = 0 \quad\Longleftrightarrow\quad \|x\|=0.$$

THÉORÈME 1.4 (Inégalité de Cauchy-Schwartz). *Soit $\langle\cdot,\cdot\rangle$ un produit scalaire sur un $\mathbf{R}$-e.v., $\|\cdot\|$ la norme associée. Alors $|\langle x,y\rangle| \leq \|x\|\|y\|$, et on a égalité si et seulement si x et y sont colinéaires.*

DÉMONSTRATION. Si $x = 0$: pas de problème. On suppose donc que $x \neq 0$, et on étudie $\langle tx+y, tx+y\rangle$. □

COROLLAIRE 1.5. *Soit $\langle\cdot,\cdot\rangle$ un produit scalaire sur un $\mathbf{R}$-e.v., $\|\cdot\|$ la norme associée. Alors $\|\cdot\|$ vérifie l'inégalité du triangle, et c'est bien une norme. Elle induit donc une* **distance euclidienne** *dans E, définie par $d(x,y) = \|x-y\|$.*

Le produit scalaire est déterminé par la norme associée (identité de polarisation) :

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

1.2. Le cas complexe. La même définition ne marche pas pour un $\mathbf{C}$-e.v. En effet, si $\langle\cdot,\cdot\rangle$ est symétrique, linéaire à droite et définie positive, on a pour $x \neq 0$:

$$\langle x,x\rangle > 0, \qquad \langle ix,ix\rangle > 0,$$
$$\langle ix,ix\rangle = i^2\langle x,x\rangle = -\langle x,x\rangle,$$

ce qui est absurde.

Il faudra « sacrifier, » ou plutôt modifier, l'une des trois propriétés : « symétrique » sera remplacé par une nouvelle notion : « hermitien »

DÉFINITION 1.6. Soit E un $\mathbf{C}$-espace vectoriel. Un **produit scalaire** dans E est une application $\langle\cdot,\cdot\rangle\colon E\times E \to \mathbf{C}$ qui est :
(l'image de $x,y\in E$ sera notée $\langle x,y\rangle$)

— **Hermitienne** : pour tous $x,y\in E$,

$$\langle x,y\rangle = \overline{\langle y,x\rangle}.$$

Ici $\overline{\langle y,x\rangle}$ est le conjugué complexe : $\overline{a+ib} = a-ib$ pour $a,b\in\mathbf{R}$.
En particulier, $\langle x,x\rangle = \overline{\langle x,x\rangle}$, d'où $\langle x,x\rangle\in\mathbf{R}$.

— Linéaire à droite. Autrement dit, si on fixe $x\in E$ (l'argument de gauche), alors l'application $y\mapsto\langle x,y\rangle$ est linéaire : pour tous $y,y'\in E$ et $\alpha\in\mathbf{C}$,

$$\langle x, y+\alpha y'\rangle = \langle x,y\rangle + \alpha\langle x,y'\rangle.$$

— Définie positive : pour tout $x\in E$, si $x\neq 0$ alors

$$\langle x,x\rangle > 0.$$

On définit la **norme** associée au produit scalaire par :

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

(pour $x\in E$)

EXAMPLE. Le produit scalaire standard sur $\mathbf{C}^n$ est :

$$\langle X,Y\rangle = {}^t\overline{X}Y = \sum_{i=1}^{n}\overline{x_i}y_i.$$

EXERCICE 1.7. Montrer que c'est un produit scalaire.

REMARQUE 1.8. Soit $\langle\cdot,\cdot\rangle$ un produit scalaire sur un $\mathbf{C}$-e.v. E.

(i) Puisqu'il est linéaire à droite et hermitien,le produit scalaire est **semi-linéaire à gauche** :

$$\langle x + \alpha x', y\rangle = \langle x, y\rangle + \overline{\alpha}\langle x', y\rangle.$$

Ainsi, le produit scalaire est **sesquilinéaire** (= une-fois-et-demi-linéaire).

(ii) Si $x = 0$ alors $\langle x, x\rangle = 0$. Ainsi, $\langle x, x\rangle \geq 0$ pour **tout** $x \in E$ (nul ou non).

(iii) On a :

$$\|x + y\|^2 = \|x\|^2 + 2\Re\big(\langle x, y\rangle\big) + \|y\|^2, \qquad \|\alpha x\|^2 = |\alpha|^2\|x\|^2$$

Théorème 1.9 (Inégalité de Cauchy-Schwartz). *Même énoncé pour les* $\mathbf{C}$*-e.v. que pour les* $\mathbf{R}$*-e.v.*

Corollaire 1.10. *Soit* $\langle\cdot,\cdot\rangle$ *un produit scalaire sur un* $\mathbf{C}$*-e.v.,* $\|\cdot\|$ *la norme associée. Alors* $\|\cdot\|$ *vérifie l'inégalité du triangle, et c'est bien une norme.*

On a

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

Or

$$\Im\langle x, y\rangle = \Re\bar{i}\langle x, y\rangle = \Re\langle yx, y\rangle$$

D'où l'identité de polarisation pour les produits scalaires complexes :

$$\langle x, y\rangle = \frac{1}{4}\big(\|x + y\|^2 - \|x - y\|^2 + i\|ix + y\|^2 - i\|ix - y\|^2\big).$$

Désormais, K est le corps réel $\mathbf{R}$ or le corps complexe $\mathbf{C}$. Ainsi, on traitera du **cas réel** et du **cas complexe** en parallèle.

Définition 1.11. Un espace **préhilbertien** est un K-e.v. E muni d'un produit scalaire ($K \in \{\mathbf{R}, \mathbf{C}\}$). On le notera par une paire $\big(E, \langle\cdot,\cdot\rangle\big)$ ou par E seul, considérant que la donnée de E comprend la donnée du produit scalaire.

Un espace **euclidien** est un espace préhilbertien réel de dimension fini.

Un espace **hermitien** est un espace préhilbertien complexe de dimension fini.

On remarque que si E est un espace préhilbertien et $F \subseteq E$ est un sous-espace vectoriel, alors F est lui aussi préhilbertien, une fois muni de la restriction de E à F du produit scalaire.

Lemme 1.12 (Identité du parallélogramme). *Soit E un espace préhilbertien. Alors la norme associée au produit scalaire vérifie l'identité du parallélogramme :*

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

Exercice 1.13. En déduire que la norme $\|x\|_1 = |x_1| + |x_2|$ sur $\mathbf{R}^2$ (ou sur $\mathbf{C}^2$) n'est pas associée à un produit scalaire.

Remarque 1.14. La réciproque est vraie aussi, mais plus difficile à démontrer : toute norme vérifiant l'identité du parallélogramme est associée à un produit scalaire (obtenu par la formule de polarisation).

2. Vecteurs orthogonaux

Partout, E est un espace préhilbertien : un K-e.v. muni d'un produit scalaire, où $K = \mathbf{R}$ or $K = \mathbf{C}$.

DÉFINITION 1.15. (i) Deux vecteurs $x, y \in E$ sont **orthogonaux** si $\langle x, y\rangle = 0$, noté $x \perp y$.

(ii) Si $A \subseteq E$ et $x \in E$, on dit que x est orthogonal à A ($x \perp A$) si $x \perp y$ pour tout $y \in A$.

(iii) Si $A, B \subseteq E$, on dit que A est orthogonal à B ($A \perp B$) si $x \perp y$ pour tout $x \in A$ et $y \in B$, ce qui revient à $x \perp B$ pour tout $x \in A$.

On observe que $x \perp y$ ssi $y \perp x$, et $A \perp B$ ssi $B \perp A$.

LEMME 1.16. *On a*

$$\begin{aligned} x \perp A \quad &\iff x \perp \mathrm{Vect}(A) \\ A \perp B \quad &\iff \mathrm{Vect}(A) \perp \mathrm{Vect}(B). \end{aligned}$$

DÉFINITION 1.17. Soit $A \subseteq E$. L'**orthogonal** de A est l'ensemble de vecteurs de E orthogonaux à A :

$$A^{\perp} = \{x \in E : x \perp A\} = \{x \in E : \forall y \in A \text{ on a } \langle x, y\rangle = 0\}.$$

EXERCICE 1.18. (i) Montrer que $A \perp A^{\perp}$.

(ii) Montrer que $A^{\perp}$ est un sous-espace vectoriel de E.

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

(iv) Montrer que $\mathrm{Vect}(A) \subseteq (A^{\perp})^{\perp}$.

DÉFINITION 1.19. Une famille de vecteurs de E est dite **orthogonale** si les vecteurs de cette famille sont deux à deux orthogonaux.

Une famille de vecteurs de E est dite **orthonormée** si elle est orthogonales et tous ses vecteurs sont de norme 1.

D'une famille orthogonale $(e_1, e_2, ...)$ sans vecteurs nuls on peut facilement passer à une famille orthonormée en normalisant : $e'_i = \frac{e_i}{\|e_i\|}$.

On verra que tout espace préhilbertien **de dimension finie** admet une base orthonormée. C'est faux en dimension infinie : l'espace des fonctions continues $C([a,b], R)$ n'admet pas de base orthogonale au sens algébrique.

LEMME 1.20. *Toute famille orthogonale sans vecteur nul est libre.*
Toute famille orthonormée est libre.

EXAMPLE. Dans $\mathbf{R}^n$ ou $\mathbf{C}^n$ (muni, sauf mention contraire, du produit scalaire standard) $\langle x, y\rangle = \sum x_i y_i$, la base standard $(e_1, e_2, \ldots)$ est orthonormée.

Exemple. 1. Dans l'espace des fonctions continues $C([0, 2\pi], R)$ avec le produit scalaire $\langle f, g\rangle = \int_0^{2\pi} f(t)g(t)dt$ la famille $(1, \cos nx, \sin nx)_{n\geq 1}$ est orthogonale et la famille $(\frac{1}{\sqrt{2\pi}}, \frac{1}{\sqrt{\pi}}\cos nx, \frac{1}{\sqrt{\pi}}\sin nx)_{n\geq 1}$ est orthonormée.

3. Calculs dans une base orthogonale/orthonormée

Soit $(e_1, ..., e_n)$ une base **orthogonale**, et $x = \sum x_i e_i \in E$.

On a le "Théorème de Pythagore" :

$$\|x\|^2 = \sum |x_i|^2 \|e_i\|^2.$$

Plus généralement, si $y = \sum y_i e_i$ on a :

$$\langle x, y\rangle = \sum \overline{x_i} y_i \|e_i\|^2.$$

(Si $K = \mathbf{R}$, ceci revient à $\sum x_i y_i \|e_i\|^2$, sans la conjugaison complexe.)

En développant $\langle e_i, x\rangle$, on obtient :

$$x_i = \frac{\langle e_i, x\rangle}{\|e_i\|^2},$$

que l'on peut substituer dans les formules précédentes, par exemple :

$$\|x\|^2 = \sum |x_i|^2 \|e_i\|^2 = \sum \frac{|\langle e_i, x\rangle|^2}{\|e_i\|^2}$$

Dans le cas ou la base est **orthonormée**, ces formules prennent une forme plus simple :

$$x_i = \langle e_i, x\rangle,$$
$$\|x\|^2 = \sum |x_i|^2,$$
$$\langle x, y\rangle = \sum \overline{x_i} y_i.$$

Autrement dit, dans une base orthonormée, un produit scalaire revient au produit scalaire standard.

4. Orthogonalisation de Gram-Schmidt

LEMME 1.21. *Soit $(u_1, \ldots, u_k)$ une famille orthogonale dans E et $v \in E$. Nous posons :*

$$u = v - \sum_{i=1}^{k} \frac{\langle u_i, v\rangle}{\|u_i\|^2} u_i.$$

Alors

(i) $u \perp u_1, \ldots, u_k$

(ii) $\mathrm{Vect}(u_1, \ldots, u_k, v) = \mathrm{Vect}(u_1, \ldots, u_k, u)$.

(iii) $u = 0$ *ssi* $v \in \mathrm{Vect}(u_1, \ldots, u_k)$.

THÉORÈME 1.22 (Orthogonalisation de Gram-Schmidt). *Soit $(v_1, ..., v_m)$ une famille libre dans E. On construit une suite de vecteurs $u_1, \ldots, u_m$ par :*

$$u_1 = v_1$$
$$u_{k+1} = v_{k+1} - \sum_{i=1}^{k} \frac{\langle u_i, v_{k+1}\rangle}{\|u_i\|^2} u_i.$$

(i) *La famille $(u_1, \ldots, u_m)$ ainsi construite est une base orthogonale pour* $\mathrm{Vect}(v_1, \ldots, v_m)$.

(ii) *Si on pose $e_i = \frac{u_i}{\|u_i\|}$, alors $(e_1, \ldots, e_m)$ est une base orthonormée pour* $\mathrm{Vect}(v_1, \ldots, v_m)$.

Variante : on peut normaliser $e_i = \frac{u_i}{\|u_i\|}$ pendant la construction et utiliser la formule

$$u_{k+1} = v_{k+1} - \sum_{i=1}^{k} \langle e_i, v_{k+1}\rangle e_i.$$

Bien que cette formule semble plus simple, le vecteurs normalisés e_i auront souvent une forme plus compliquée et le plus souvent on n'y gagne rien.

Si la famille $(v_1, \dots, v_m)$ n'est pas libre, on peut toujours appliquer le procédé de Gram-Schmidt. Pour chaque v_k qui est engendré par les vecteurs qui lui précèdent, on aura $u_k = 0$. Dans ce cas on peut enlever v_k de la famille et continuer avec le vecteur suivant.

COROLLAIRE 1.23. *Tout espace préhilbertien de dimension finie (= espace euclidien ou hermitien) admet un base orthonormée.*

5. Projection orthogonale

DÉFINITION 1.24. Une application linéaire $p\colon E \to E$ est un **projecteur** si $p(x) = x$ pour tout $x \in \operatorname{img} p$. Autrement dit, p est un projecteur si $p \circ p = p$ (on écrira plutôt : $p^2 = p$).

C'est un **projecteur orthogonal** si de surcroît on a $\operatorname{img} p \perp \ker p$.

LEMME 1.25. *Soit $p\colon E \to E$ un projecteur. Alors $E = \operatorname{img} p \oplus \ker p$.*

D'ailleurs, si $x = x_1 + x_2 \in E$ avec $x_1 \in \operatorname{img} p$ et $x_2 \in \ker p$, alors $p(x) = x_1$. Ainsi, si on connaît $\operatorname{img} p$ et $\ker p$, on connaît p.

DÉMONSTRATION. Si $x \in \operatorname{img} p \cap \ker p$ alors $x = p(x) = 0$.

Et pour $x \in E$ on pose $x_1 = p(x) \in \operatorname{img} E$ et $x_2 = x - x_1$. Alors $p(x_2) = 0$, donc $x_2 \in \ker p$, d'où $E = \operatorname{img} p \oplus \ker p$.

Pour la deuxième partie : on a $p(x) = x_1$. □

LEMME 1.26. *Soit $p\colon E \to E$ un projecteur orthogonal. Alors $\ker p = (\operatorname{img} p)^\perp$.*

Ainsi, si on connaît $\operatorname{img} p$ alors on connaît p.

DÉMONSTRATION. Par définition $\ker p \subseteq (\operatorname{img} p)^\perp$. Réciproquement, soit $x \in (\operatorname{img} p)^\perp$, $x_1 = p(x)$, $x_2 = x - p(x)$. Alors $x_1 \in \operatorname{img} x$ et $x_2 \in \ker p$. Donc $x \perp x_1$, d'où :

$$0 = \langle x, x_1 \rangle = \langle x_2, x_1 \rangle + \|x_1\|^2 = \|x_1\|^2$$

Donc $x_1 = 0$ et $x = x_2 \in \ker p$. □

DÉFINITION 1.27. Soit $F \subseteq E$ un s.e.v. L'unique projecteur orthogonal d'image F, **si un tel existe**, sera noté p_F. Il s'appelle **le projecteur orthogonal sur** F.

LEMME 1.28. *Si p est un projecteur alors $p' = \mathrm{id}_E - p$ l'est aussi, et $\ker p = \operatorname{img} p'$, $\operatorname{img} p = \ker p'$.*
Si p est un projecteur orthogonal alors p' l'est aussi.

QUESTION. Soit $F \subseteq E$ un s.e.v. Le projecteur orthogonal p_F existe-t-il ?

Si p_F existe on a $F = \operatorname{img} p_F$, $\ker p_F = F^\perp$, d'où $E = F \oplus F^\perp$. Si on pose $p' = \mathrm{id}_E - p_F$ alors $\operatorname{img} p' = \ker p_F = F^\perp$, i.e. $p' = p_{F^\perp}$:

$$\mathrm{id}_E = p_F + p_{F^\perp}.$$

On obtient :

$$(F^\perp)^\perp = \ker p_{F^\perp} = \operatorname{img}(\mathrm{id}_E - p_{F^\perp}) = \operatorname{img} p_F = F.$$

EXERCICE 1.29. Montrer les réciproques : si $E = F \oplus F^\perp$ (comment pourrait-ce être faux ?) alors p_F existe. Si $F = (F^\perp)^\perp$, alors $E = F \oplus F^\perp$.

LEMME 1.30. *Si $\dim F < \infty$ le projecteur orthogonal sur F existe. Plus précisément, si $\mathcal{B} = (e_1, \dots, e_n)$ est une base orthogonale de F alors pour tout $x \in E$:*

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

Si c'est une base orthonormée :

$$p_F(x) = \sum_{i=1}^{n} \langle e_i, x\rangle e_i.$$

DÉMONSTRATION. Posons, pour une base orthogonale $(e_1, \ldots, e_n)$ de F :

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

Alors : C'est linéaire car le produit scalaire est linéaire à droite. $\operatorname{img} p \subseteq F$ par construction, et si $x \in F$ alors $p(x) = x$ par une formule vue plus tôt. D'où : $\operatorname{img} p = F$ et p est un projecteur. Finalement, si $p(x) = 0$ on obtient $x \perp e_1, \ldots, e_n$. Donc $\ker p \perp F = \operatorname{img} p$, et p est bien un projecteur orthogonal d'image F : c'est bien p_F. $\square$

LEMME 1.31. *Soit $F \subseteq E$ un s.e.v. et $p = p_F$ le projecteur orthogonal sur F. Soit $x \in E$. Alors $y = p_F(x)$ est l'unique vecteur de F tel que $x - y \perp F$.*

On remarque que

$$x - y \perp F \iff \forall\, z \in F : \langle x - y, z\rangle = 0 \iff \forall\, z \in F : \langle x, z\rangle = \langle y, z\rangle.$$

Lien avec Gram-Schmidt : soit $(v_1, ..., v_n)$ une famille libre dans E. Pour $k = 1 \ldots n$ ponsons $F_k = \mathrm{Vect}(v_1, \ldots, v_k)$. Alors

$$\begin{aligned} u_1 &= v_1 \\ u_k &= v_k - P_{F_{k-1}}(v_k) = P_{F_{k-1}^{\perp}}(v_k) \end{aligned}$$

EXAMPLE. $E = \mathbf{R}^3$ muni du produit scalaire canonique.

$$v_1 = (1, 1, 0), \quad v_2 = (1, 0, 1/\sqrt{2}), \quad F = \mathrm{Vect}(v_1, v_2)$$

Calculer explicitement $P_F(e_1) = P_F(1, 0, 0)$.

On applique Gram-Schmidt pour obtenir une base orthogonale de F :

$$u_1 = (1, 1, 0), \quad u_2 = \left(\frac{1}{2}, -\frac{1}{2}, \frac{1}{\sqrt{2}}\right),$$

Alors :

$$P_F(e_1) = \frac{\langle e_1, u_1\rangle}{\|u_1\|^2} u_1 + \frac{\langle e_1, u_2\rangle}{\|u_2^2} u_2 = \frac{1}{2} u_1 + \frac{1}{2} u_2 = \left(\frac{3}{4}, \frac{1}{4}, \frac{1}{2\sqrt{2}}\right)$$

EXAMPLE. $E = \mathbf{R}^3$ muni du produit scalaire canonique,

$$F = \big\{(x, y, z) : x - y - \sqrt{2} z = 0\big\}.$$

Calculer explicitement $P_F(e_1)$.

Au lieu de trouver une base orthonormée pour F observons que $F = w^{\perp} = \mathrm{Vect}(w)^{\perp}$ où $w = (1, -1, -\sqrt{2})$. Autrement dit, $F^{\perp} = \mathrm{Vect}(w)$. Alors,

$$P_{F^{\perp}}(e_1) = \frac{\langle e_1, w\rangle}{\|w\|^2}\} w = \frac{1}{4} w = \left(\frac{1}{4}, -\frac{1}{4}, -\frac{1}{2\sqrt{2}}\right).$$

Or, puisque $P_F(e_1) + P_{F^{\perp}}(e_1) = \mathrm{id}(e_1) = e_1$:

$$P_F(e_1) = e_1 - P_{F^{\perp}}(e_1) = \left(\frac{3}{4}, \frac{1}{4}, \frac{1}{2\sqrt{2}}\right)$$

Au fait dans les deux exemples c'est le même F donnée de deux manières différentes.

6. Distance à un sous-espace

LEMME 1.32. *Soit F est un sous-espace de dimension finie et $x \in E$. Le point $y = P_F(x)$ est le point de F le plus proche de x. En d'autres termes, c'est un minimum global strict de la fonction $z \mapsto \|x - z\|$ sur F.*

En particulier, $d(x, F) = d(x, P_F(x)) = \|P_{F^\perp}(x)\|$.

DÉMONSTRATION. En effet, soit $z \in F$, $z \neq y$. Alors $y - z \in F$ d'où $x - y \perp y - z$ et $\|x - z\|^2 = \|x - y\|^2 + \|y - z\|^2 > \|x - y\|^2$. □

7. Bijection entre E et E^* en dimension finie

DÉFINITION 1.33. Soit E un K-e.v. Une **forme linéaire** sur E est une application linéaire $\varphi\colon E \to K$. L'ensemble de toutes les formes linéaires forme un espace vectoriel noté E^*, appelé le **dual** de E.

THÉORÈME 1.34. *Soit E un espace euclidien. Pour chaque $x \in E$ définissons*

$$\varphi_x\colon E \to \mathbf{R}, \qquad \varphi_x(y) = \langle x, y \rangle.$$

Alors $\varphi_x \in E^*$, et l'application $x \mapsto \varphi_x$ est un isomorphisme d'espaces vectoriels entre E et E^*.

DÉMONSTRATION. C'est facilement linéaire, de noyau nul. On pourrait montrer que $\dim E = \dim E^*$ et déduire que $x \mapsto \varphi_x$ est surjective, mais on préfère donner une construction plus explicite, qui est d'ailleurs proche de l'argument que l'on utiliserait (sous des hypothèses supplémentaires) en dimension infinie.

Soit $\varphi \in E^*$. Si $\varphi = 0$ alors $\varphi = \varphi_0$ et c'est bon. Supposons donc que $\varphi \neq 0$ et posons $F = \ker \varphi$. Alors $\dim \operatorname{img} \varphi = 1$ d'où $\dim F = \dim E - 1$. Or en dimension finie on a $E = F \oplus F^\perp$, donc $\dim F^\perp = 1$. Choisissons $x_0 \in F^\perp$ non nul, et posons

$$x = \frac{\varphi(x_0)}{\|x_0\|^2} x_0$$

Alors $\varphi = \varphi_{x_0}$. En effet, si $y \in E$, on décompose $y = y_1 + y_2$ avec $y_1 \in F$ et $y_2 \in F^\perp$, si bien que $y_2 = \alpha x_0$.

$$\varphi_{x_0}(y) = \varphi_{x_0}(y_2) = \frac{\varphi(x_0)\langle x_0, \alpha x_0 \rangle}{\|x_0\|^2} = \varphi(\alpha x_0) = \varphi(y_2) = \varphi(y).$$

□

REMARQUE 1.35. Pour $\varphi \in E^*$ on définit sa norme par :

$$\|\varphi\| = \sup_{y : \|y\| \leq 1} |\varphi(y)|$$

Alors la bijection donnée polus haut préserve la norme : $\|\varphi_x\| = \|x\|$.

REMARQUE 1.36. Cas complexe : la seule différence est que $x \mapsto \varphi_x$ est semi-linéaire : $\varphi_{x+\alpha x'} = \varphi_x + \overline{\alpha} \varphi_{x'}$

Chapitre 2

Endomorphismes d'espaces euclidiens

1. L'endomorphisme adjoint

Soit E un espace euclidien ($K = \mathbf{R}$) (ou hermitien ($K = \mathbf{C}$) et soit $\mathrm{End}(E)$ l'espace des **endomorphismes** de E (c'est-à-dire, des applications linéaires $E \to E$).

Pour un endomorphisme f nous allons s'intéresser aux deux applications $E \times E \to K$:

$$(x, y) \mapsto \langle x, f(y)\rangle, \qquad (x, y) \mapsto \langle f(x), y\rangle$$

DÉFINITION 2.1. Soit $f \in \mathrm{End}(E)$. On appelle **adjoint** de f un endomorphisme $f^* \in \mathrm{End}(E)$ tel que pour tout $x, y \in E$:

$$\langle f^*x, y\rangle = \langle x, fy\rangle.$$

Fixons-nous une base orthonormée $\mathcal{B} = (e_1, \ldots, e_n)$. Si X et Y sont les coordonnées de x et y dans cette base, on a déjà observé que :

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

Si $A = \mathrm{Mat}_{\mathcal{B}}\, f$, on obtient :

$$\langle x, f(y)\rangle = {}^t\overline{X}AY$$
$$\langle f(x), y\rangle = {}^t\overline{AX}Y = {}^t\overline{X}\,{}^t\overline{A}Y.$$

PROPOSITION 2.2. (i) *Pour tout $f \in \mathrm{End}(E)$, il existe un unique endomorphisme adjoint f^*.*

(ii) *Si $\mathcal{B}$ est une base orthonormée, f^* est l'unique endomorphisme tel que* $\mathrm{Mat}_{\mathcal{B}}(f^*) = {}^t\mathrm{Mat}_{\mathcal{B}}(\mathbf{f})$.

LEMME 2.3. (i) *L'application $f \to f^*$ est semi-linéaire (linéaire si $K = \mathbf{R}$) :*

$$(f + g)^* = f^* + g^*, \qquad (\alpha f)^* = \overline{\alpha} f^*.$$

(ii) *On a*

$$(fg)^* = g^*f^*, \qquad (f^*)^* = f.$$

(iii) *Si f est inversible, f^* l'est aussi et*

$$(f^*)^{-1} = (f^{-1})^*.$$

DÉFINITION 2.4. Un endomorphisme f est dit **auto-adjoint** si $f = f^*$. Si $K = \mathbf{R}$, f est auto-adjoint si et seulement si sa matrice dans un base orthonormée est symétrique. Ainsi, un tel f s'appelle parfois **symétrique**

EXAMPLE. Toute projection orthogonale P_F vérifie $P_F = P_F^2 = P_F^*$. Elle est en particulier auto-adjointe.

En effet, si $(e_1, \ldots, e_m)$ est une base orthonormée de F et $(e_{m+1}, \ldots, e_n)$ une base orthonormée de $F^\perp$ alors $\mathcal{B} = (e_1, \ldots, e_n)$ est une base orthonormée de E et $A = \mathrm{Mat}_{\mathcal{B}}\, P_F = \begin{pmatrix} I_m & 0 \\ 0 & 0 \end{pmatrix}$. Ainsi, $A = A^2 = {}^tA$ d'où le résultat.

Exercice 2.5. La réciproque est vraie aussi : si $f \in \mathrm{End}(E)$ vérifie $f = f^* = f^2$ alors f est égal à P_F pour un certain s.e.v. $F \subseteq E$. Comment trouver F ?

2. Endomorphismes orthogonaux, matrices orthogonales

Définition 2.6. Un endomorphisme $f \in \mathrm{End}(E)$ est dit **orthogonal** si $f^{-1} = f^*$.
Une matrice $O \in M_n(\mathbf{R})$ est dites **orthogonale** si $O^{-1} = {}^tO$.

Un endomorphisme f est orthogonal ssi sa matrice dans une (toute) base orthonormée est orthogonale.

Lemme 2.7. *Soit $O \in M_n(\mathbf{R})$ une matrice orthogonale. Alors toutes les valeurs propres (complexes) de O sont de module 1.*
Ainsi, le déterminant de O est de module 1.

Lemme 2.8. *Soit $U \in M_n(\mathbf{C})$ tel que $U^{-1} = {}^t\overline{U}$ (une telle matrice est appelée* **unitaire***). Alors toutes les valeurs propres de U sont de module 1.*
Ainsi, le déterminant de U est de module 1.

Démonstration. Supposons que $X \in \mathbf{C}^n \setminus \{0\}$, $UX = \lambda X$, et posons $\|X\|^2 = {}^t\overline{X}X = \sum |x_i|^2 > 0$. Alors :

$$\|X\|^2 = {}^t\overline{X}X = {}^t\overline{X}({}^t\overline{U}U)X = {}^t\overline{(UX)}UX = (\bar{\lambda}{}^tX)(\lambda X) = \lambda\bar{\lambda}\|X\|^2,$$

d'où $|\lambda|^2 = \lambda\bar{\lambda} = 1$. □

Lemme 2.9. *Soit $\mathcal{B} = (e_1, \ldots, e_n)$ une base orthonormée, $\mathcal{B}' = (e'_1, \ldots, e'_n)$ une base quelconque, P la matrice de passage. Alors P est orthogonale ssi $\mathcal{B}'$ est orthonormée.*

Lemme 2.10. *Une matrice $A \in M_n(\mathbf{R})$ est orthogonale ssi ses colonnes (lignes) forment une base orthonormée de $\mathbf{R}^n$ par rapport au produit scalaire standard.*

3. Théorème spectral : diagonalisation des endomorphismes et matrices symétriques

Définition 2.11. Un sous-espace $F \subseteq E$ est dit **stable** par $f \in \mathrm{End}(E)$ si $f(F) \subseteq F$.

Si $F \subseteq E$ est un sous-espace stable par f, la restriction $f\restriction_F$ est un endomorphisme de F : $f\restriction_F \in \mathrm{End}(F)$.

Lemme 2.12. (i) *Si F est stable par f alors $F^\perp$ est stable par f^*.*
(ii) $\ker f^* = (\mathrm{img}\, f)^\perp$ *et* $\mathrm{img}\, f^* = (\ker f^*)^\perp$.

Corollaire 2.13. *Si f est auto-adjoint, alors $\ker f = (\mathrm{img}\, f)^\perp$, et $E = \ker f \oplus \mathrm{img}\, f$. De plus, l'orthogonal d'un sous-espace stable par f est stable par f, i.e., si $f(F) \subseteq F$ alors $f(F^\perp) \subseteq F^\perp$.*

Lemme 2.14. *Soit f un endomorphisme auto-adjoint. Alors f admet une valeur propre réelle.*

Lemme 2.15. *Soit $A \in M_n(\mathbf{R})$ symétrique. Alors toutes les valeurs propres (complexes) de A sont réelles.*

Lemme 2.16. *Soit $A \in M_n(\mathbf{C})$ telle que ${}^t\overline{A} = A$ (une telle matrice est dite* **hermitienne***). Alors toutes les valeurs propres de A sont réelles.*

Théorème 2.17 (Théorème spectral). *Soit f un endomorphisme auto-adjoint. Alors f est diagonalisable dans une base orthonormée. Autrement dit, il existe une base orthonormée qui consiste en des vecteurs propres de f.*

DÉMONSTRATION. L'endomorphisme f admets au moins une valeur propre réelle λ, avec vecteur propre e_1 que l'on peut supposer normalisé : $\|e_1\| = 1$. Puisque $fe_1 = \lambda e_1$, l'espace $\mathrm{Vect}(e_1)$ est stable par f. Puisque $f = f^*$, l'orthogonal $F := e_1^\perp = \mathrm{Vect}(e_1)$ est lui aussi stable par f. Ainsi, f se restreint en un endomorphisme $g = f\restriction_F : F \to F$, qui est lui aussi auto-adjoint. Or, $\dim F = \dim E - 1$, et par l'hypothèse de récurrence F admet une base orthonormée $(e_2, \ldots, e_n)$ qui consiste en des vecteurs propres de g, qui sont aussi des vecteurs propres de f. On conclut que $(e_1, e_2, \ldots, e_n)$ est une base orthonormée de E qui consiste en des vecteurs propres de f. □

COROLLAIRE 2.18. *Soit f un endomorphisme auto-adjoint. Soit $(\lambda_1, \ldots, \lambda_k)$ les valeurs propres de f, et soit $E_i = \ker(f - \lambda_i \,\mathrm{Id})$ le sous-espace propre correspondant à λ_i. Alors*

(i) *Les sous-espaces propres de f sont deux à deux orthogonaux, i.e., $E_i \perp E_j$ pour $i \neq j$. En plus, $E_1 \oplus E_2 \oplus \ldots \oplus E_k = E$.*

(ii) **Décomposition spectrale :** $f = \sum_{i=1}^k \lambda_i P_{E_i}$.

EXERCICE 2.19. Montrer directement, sans passer par le théorème spectral, que si $f \in \mathrm{End}(E)$ est auto-adjoint et $v, w \in E$ sont des vecteurs propres correspondant à des valeurs propres distinctes, alors $v \perp w$.

COROLLAIRE 2.20. *Soit $A \in M_n(\mathbf{R})$ une matrice symétrique. Alors A est diagonalisable par une matrice orthogonale : il existe une matrice orthogonale O telle que $O^{-1}AO = {}^tOAO$ est diagonale.*

On observe que la réciproque est facile : si O est orthogonale et $O^{-1}AO$ est diagonale, A est symétrique.

THÉORÈME 2.21. *Soient $f_1, f_2, \ldots \in \mathrm{End}(E)$ des endomorphismes auto-adjoints qui en plus* **commutent entre eux** *: $f_i f_j = f_j f_i$ pour tout i, j. Alors ils sont simultanément diagonalisables dans une base orthonormée. Autrement dit, il existe une base orthonormée $\mathcal{B}$ dans laquelle toutes les matrices $M_{\mathcal{B}}(f_i)$ sont diagonales, i.e., les membres de $\mathcal{B}$ sont des vecteurs propres de chacun des f_i.*

Une famille de matrices symétriques qui commutent entre elles sont simultanément diagonalisables par une matrice orthogonale.

DÉMONSTRATION. On démontre pour le cas de deux endomorphismes auto-adjoints f, g tels que $fg = gf$. Soit λ une valeur propre de f et $E_\lambda \subseteq E$ le sous-espace propre correspondant. Alors pour tout $x \in E_\lambda$ on a :

$$f(gx) = g(fx) = g(\lambda x) = \lambda(gx),$$

i.e., $gx \in E_\lambda$. Ainsi E_λ est stable par g. D'après le théorème spectral appliqué à la restriction de g à E_λ, l'espace propre E_λ admet une base $\mathcal{B}_\lambda$ orthonormée qui consiste en des vecteurs propres de g.

Posons $\mathcal{B} = \mathcal{B}_{\lambda_1} \cup \ldots \cup \mathcal{B}_{\lambda_k}$ où $\lambda_1, \ldots, \lambda_k$ sont toutes les valeurs propres de f. Puisque les espaces E_{λ_i} sont deux-à-deux orthogonaux, $\mathcal{B}$ est une famille orthonormée. Puisque $E = E_{\lambda_1} \oplus \ldots \oplus E_{\lambda_k}$, $\mathcal{B}$ est une base. Et finalement, les membres de $\mathcal{B}$ sont des vecteurs propres de f et de g à la fois.

Le cas général de m endomorphismes auto-adjoints qui commutent entre eux se démontre de la même manière, par récurrence. □

4. Endomorphismes positifs

Si f est un endomorphisme symétrique, la forme bilinéaire $\varphi_f = \langle x, fy \rangle$ est symétrique (et coïncide avec $\langle fx, y \rangle$). On l'appelle la **forme bilinéaire associée** à f.

Un endomorphisme symétrique f est dit **positif** (respectivement, **défini positif**) si la forme associée $\langle x, fy \rangle$ est positive (respectivement, définie positive). Une matrice symétrique A est dite **positive** (respectivement, **définie positive**) si ${}^tXAX \geq 0$ pour tout $X \in \mathbf{R}^n$ (respectivement, ${}^tXAX > 0$ pour tout X non-nul).

Corollaire 2.22. *Un endomorphisme symétrique est positif (respectivement, défini positive) si et seulement si toutes ses valeurs propres sont positives (respectivement, strictement positives).*

Théorème 2.23 (Critère de Sylvester pour une matrice définie positive). *Soit $A = (a_{ij})_{i,j=1}^n \in M_n(\mathbf{R})$ une matrice symétrique. Pour $k \leq n$ soit $A_k = (a_{ij})_{i,j=1}^k$ le bloc $k \times k$ supérieur à gauche de A. Alors A est définie positive si et seulement si* $\det A_k > 0$ *pour $k = 1, \ldots, n$.*

Démonstration. D'abord, si A est définie positive alors $\det A > 0$: en effet, toutes ses valeurs propres sont > 0. Aussi, si A est définie positive alors A_k l'est aussi, si bien que $\det A_k > 0$ pour chaque k. Ceci montre que le condition est nécessaire. Supposons maintenant que A est symétrique et $\det A_k > 0$ pour chaque k. D'après l'hypothèse de récurrence, A_{n-1} est définie positive, et on a :

$$A = \begin{pmatrix} A_{n-1} & B \\ {}^tB & c \end{pmatrix}$$

où $B \in \mathbf{R}^{n-1}$ est un vecteur colonne et $c \in \mathbf{R}$. Soit $X_1, \ldots, X_{n-1}$ les colonnes de A_{n-1}. Comme A_{n-1} est inversible, on peut écrire :

$$B = \lambda_1 X_1 + \ldots + \lambda_{n-1} X_{n-1}.$$

Posons :

$$P = I_n + \begin{pmatrix} & -\lambda_1 \\ 0 & \vdots \\ & -\lambda_{n-1} \\ & 0 \end{pmatrix}$$

Et

$$A' = {}^tPAP = {}^tP \begin{pmatrix} A_{n-1} & 0 \\ {}^tB & ? \end{pmatrix} = \begin{pmatrix} A_{n-1} & 0 \\ 0 & d \end{pmatrix}$$

Puisque P est inversible on a $\det P^2 > 0$ d'où $\det A' > 0$. Or $\det A' = d \cdot \det A_{n-1}$, si bien que $d > 0$. On en déduit que A' est définie positive, donc A aussi [DÉTAILS À COMPLÉTER]. $\square$

Corollaire 2.24 (Racine carré d'une matrice positive). *Soit f un endomorphisme positif, Π_i le projecteur spectral associé à la valeur propre λ_i, $i = 1, ..., k$. On a $f = \sum \lambda_i \Pi_i$. Posons $g = \sum \sqrt{\lambda_i}\Pi_i$. Alors g est symétrique positif et $g^2 = f$. On montre facilement qu'une telle racine carré positive $\sqrt{f} = g$ est unique (en considérant la décomposition spectrale de g).*

5. Décomposition polaire

Théorème. Soit f un endomorphisme inversible. Il existe un unique endomorphisme orthogonal U et un unique endomorphisme défini positif S tels que $f = US$.

Construction : On pose $S = \sqrt{f^* f}$; S est défini positif. On définit U par $U = fS^{-1}$ et on vérifie que U est orthogonal : $U^*U = S^{-1} f^* f S^{-1} = S^{-1} S^2 S^{-1} = \mathrm{Id}$.

6. Transformations orthogonales

Un endomorphisme inversible s'appelle également une **transformation**. Ainsi, tout endomorphisme orthogonal est une transformation. Autrement dit, un endomorphisme orthogonal et une transformation orthogonale sont la même chose.

Nous avons vu que le déterminant d'une transformation orthogonale est ± 1.

Une transformation orthogonale de déterminant 1 est dite **directe**. Une transformation orthogonale de déterminant -1 est dite **indirecte**.

Proposition 2.25. *Si f, g sont des endomorphismes orthogonaux alors f^{-1} et fg le sont aussi. En particulier, l'ensemble des endomorphismes orthogonaux, noté $O(E)$, est un groupe.*

Pareil pour les matrices. Le groupe des matrices orthogonales est $O_n(\mathbf{R}) = \{A \in M_n(\mathbf{R}) : {}^tAA = \mathrm{Id}\}$.

Proposition 2.26. *Pour $f \in \mathrm{End}(E)$, les propriétés suivantes sont équivalentes.*

(i) *f est orthogonal.*

(ii) *f préserve la norme : $\|fx\| = \|x\|$ pour tout x.*

(iii) *f préserve le produit scalaire : $\langle fx, fy\rangle = \langle x, y\rangle$ pour tout x, y.*

(iv) *f transforme toute base orthonormée en base orthonormée.*

(v) *f transforme une base orthonormée en base orthonormée.*

Démonstration. (i) $\Longrightarrow$ (ii). $\|fx\|^2 = \langle fx, fx\rangle = \langle x, f^*fx\rangle = \langle x, f^{-1}fx\rangle = \langle x, x\rangle = \|x\|^2$.

(ii) $\Longrightarrow$ (iii). Puisque le produit scalaire est détermine par la norme (identité de polarisation).

(iii) $\Longrightarrow$ (iv). Soit $\mathcal{B} = (e_1, \ldots, e_n)$ une base orthonormée. Alors $\langle fe_i, fe_j\rangle = \langle e_i, e_j\rangle = \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$ (le delta de Kronecker). Donc, $(fe_1, \ldots, fe_n)$ est aussi une famille orthonormée. Elle est donc libre, de taille $n = \dim E$, c'est donc une base.

(iv) $\Longrightarrow$ (v). Clair.

(v) $\Longrightarrow$ (i). Soit $\mathcal{B} = (e_1, \ldots, e_n)$ une base orthonormée telle que $(fe_1, \ldots, fe_n)$ est aussi une base orthonormée, et montrons que $f^*fx = x$ pour tout x. En effet on a $x = \sum_{i=1}^n \langle x, e_i\rangle e_i$ d'où $fx = \sum_{i=1}^n \langle x, e_i\rangle fe_i$. Puisque $(fe_1, \ldots, fe_n)$ est une base orthonormée : $fx = \sum_{i=1}^n \langle fx, fe_i\rangle fe_i = \sum_{i=1}^n \langle f^*fx, e_i\rangle fe_i$. Il en découle que $\langle x, e_i\rangle = \langle f^*fx, e_i\rangle$ pour tout $i = 1, \ldots, n$. Par conséquent :

$$x = \sum_{i=1}^n \langle x, e_i\rangle e_i = \sum_{i=1}^n \langle f^*fx, e_i\rangle e_i = f^*fx.$$

Donc $f^* = f^{-1}$. □

Une application quelconque $f\colon E \to E$ est une **isométrie** si :

$$\forall x, y \in E : \quad d(x, y) = d\big(f(x), f(y)\big),$$

i.e.,

$$\forall x, y \in E : \quad \|x - y\| = \|f(x) - f(y)\|.$$

Lorsque f est linéaire, ceci est encore équivalent à :

$$\forall x, y \in E : \quad \|x - y\| = \|f(x - y)\|,$$

ce qui revient à :

$$\forall x \in E : \quad \|x\| = \|f(x)\|.$$

Autrement dit, pour $f \in \mathrm{End}(E)$: f est une isométrie ssi f est orthogonal.

Lemme 2.27. *Soit E un espace euclidien. Une isométrie $f\colon E \to E$ qui fixe 0 est linéaire. C'est donc un endomorphisme orthogonal de E.*

Démonstration. Montrons d'abord que $\|fx\| = \|x\|$. En effet, $\|fx\| = d(fx, 0) = d(fx, f0) = d(x, 0) = \|x\|$.

Ainsi, pour tout $x, y \in E$:

$$\begin{aligned}\langle fx, fy\rangle &= \frac{1}{2}\left[\|fx\|^2 + \|fy\|^2 - \|d(fx, fy)\|^2\right]\\ &= \frac{1}{2}\left[\|x\|^2 + \|y\|^2 - \|d(x, y)\|^2\right]\\ &= \langle x, y\rangle.\end{aligned}$$

Soit maintenant $\mathcal{B} = (e_1, \dots, e_n)$ une base orthonormée de E, et posons $\mathcal{B}' = \big(f(e_1), \dots, f(e_n)\big)$. D'après le point précédent, $\mathcal{B}'$ est aussi une base orthonormée. On se rappelle que dans une base orthonormée on a pour tout x :

$$x = \sum_i \langle e_i, x\rangle e_i.$$

En d'autres mots, $x = \sum x_i e_i$ avec $x_i = \langle e_i, x\rangle$. Ainsi :

$$f\left(\sum x_i e_i\right) = fx = \sum_i \langle f(e_i), fx\rangle f(e_i) = \sum_i \langle e_i, x\rangle f(e_i) = \sum_i x_i f(e_i).$$

Ainsi f est linéaire. Il suit que f est orthogonal. □

7. Symétrie par rapport à un sous-espace. Réflexions.

Soit F un sous-espace de E et $E = F \oplus F^\perp$ la décomposition orthogonale. Pour $x \in E$ on écrit $x = x_1 + x_2$ avec $x_1 \in F$ et $x_2 \in F^\perp$. La **symétrie** s_F par rapport à F est définie par $s_F(x) = x_1 - x_2$.

On a :

$$\begin{gathered} s_F = P_F - P_{F^\perp} = \operatorname{id} - 2P_{F^\perp} = 2P_F - \operatorname{id}\\ s_F^* = s_F\\ s_F^2 = \operatorname{id}. \end{gathered}$$

Autrement dit, une symétrie s_F est orthogonale et auto-adjointe.

Exercice 2.28. Réciproquement, tout endomorphisme f qui est à la fois auto-adjoint et orthogonal (i.e., tel que $f = f^*$ et $f^2 = \operatorname{id}$) est une symétrie. Comment trouver l'espace F à partir de f ?

Une symétrie par rapport à un hyperplan s'appelle **réflexion**. Si F est un hyperplan et e le vecteur unitaire orthogonal à F, la réflexion par rapport au F s'écrit $s_F(x) = x - 2\langle x, e\rangle e$.

Plus généralement, si z est un vecteur quelconque et $F = z^\perp$ alors $s_F(x) = x - 2\frac{\langle x, z\rangle}{\|z\|^2} z$.

Matrice d'une réflexion :

$$\begin{pmatrix} \operatorname{Id}_{n-1} & \\ & -1 \end{pmatrix}$$

Le déterminant d'une réflexion est donc toujours -1.

Théorème 2.29. *Soit $f \in \operatorname{End}(E)$ orthogonal, $r = \operatorname{rg}(f - \operatorname{Id})$. Alors f s'exprime comme produit d'au plus r réflexions. En particulier, f s'exprime comme produit d'au plus $n = \dim E$ réflexions.*

Démonstration. Par récurrence sur r.

Supposons que c'est vrai pour $r' < r$ et montrons pour r. Si $r = 0$: $f = \mathrm{Id}$ est la composition de zéro réflexions. Sinon, $f \neq \mathrm{Id}$, et il existe y tel que $fy \neq y$. Soit $z = fy - y \neq 0$ et $F = z^\perp$. Alors :

$\|z\|^2 = 2\|y\|^2 - 2\langle y, fy\rangle$.

Pour tout x : $s_F(x) = x - 2\frac{\langle x,z\rangle}{\|z\|^2} z = x - \frac{\langle x,z\rangle}{\|y\|^2 - \langle y,fy\rangle} z$.

D'où $s_F(y) = f(y)$ et $s_F(z) = -z$.

Si $fx = x$ alors $x \perp z$ et $s_F(x) = x$.

On applique l'hypothèse de récurrence à $s_F f$. □

8. Transformations orthogonales en petite dimension

Dimension 1. $O(1) = \{\pm 1\}$.

Dimension 2. Soit $A \in O(2)$:

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

Alors $a^2 + b^2 = c^2 + d^2 = 0$, si bien qu'il existe des angles θ et μ telles que :

$$a = \cos\theta, \quad b = \sin\theta \quad c = \sin\mu \quad d = \cos\mu$$
$$\det A = ad - bc = \cos\theta\cos\mu - sin\theta\sin\mu = \cos(\theta + \mu).$$

Deux cas :

Transformation orthogonale directe. C'est que $\mu = -\theta$ et

$$A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

C'est la matrice de rotation d'angle θ.

Transformation orthogonale indirecte. Si $\det A = -1$, c'est que $\mu = \pi - \theta$ et

$$A = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}$$

C'est une réflexion. Par rapport à quel droite ?

Exercice 2.30. Une matrice de rotation est la composition de combien de réflexions ? Comment faire pour les trouver concrètement ?

Dimension 3.

Lemme 2.31. *Soit E un espace euclidien et $g \in O(E)$ un endomorphisme orthogonal. Si $F \subseteq E$ est un sous-espace stable par g, alors $F^\perp$ est stable par g.*

Maintenant soit $g \in O(E)$ avec $\dim E = 3$. Alors g admets au moins une valeur propre (réelle).

Le plus souvent, g aura une **unique** valeur propre réelle λ. Puisque $|\lambda| = 1$, on a $\lambda = \pm 1$. On choisit e_3 tel que $g(e_3) = \lambda e_3$ et $\|e_3\| = 1$. Soit (e_1, e_2) une base orthonormée de $e_3^\perp$. Alors $\mathcal{B} = (e_1, e_2, e_3)$ est une base orthonormée et $\mathrm{Vect}(e_3)$ et $\mathrm{Vect}(e_1, e_2)$ sont stables par g :

$$\mathrm{Mat}_{\mathcal{B}}\, g = \begin{pmatrix} R & 0 \\ 0 & \pm\lambda \end{pmatrix}$$

La matrice R est orthogonale et non-diagonalisable, donc une rotation :

$$\mathrm{Mat}_{\mathcal{B}}\, g = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & \pm 1 \end{pmatrix}$$

Les même deux cas :
Transformation orthogonale directe. $\lambda = 1$, et g est une rotation autour d'un axe.

$$\mathrm{Mat}_{\mathcal{B}}\, g = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Comment trouver l'axe ? Comme trouver l'angle de rotation ?

Transformation orthogonale indirecte. $\lambda = -1$, et g est une rotation-réflexion : c'est la composition d'une rotation autour d'un axe et d'une réflexion par rapport au plan orthogonal à cet axe. De surcroît, ces deux opérations commutent :

$$\begin{aligned}\mathrm{Mat}_{\mathcal{B}}\, g &= \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & -1 \end{pmatrix} \\ &= \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}\end{aligned}$$

Comment trouver l'axe ? Le plan ? L'angle de rotation ?

9. Réduction des endomorphismes orthogonaux

Lemme 2.32. *Soit $A \in M_n(\mathbf{R})$ quelconque. Il existe un sous espace $F \subseteq \mathbf{R}^n$, $1 \leq \dim F \leq 2$, qui est stable par A.*

Démonstration. Si A admet une valeur propre réelle λ, il existe $X \in \mathbf{R}^n \setminus \{0\}$ tel que $AX = \lambda X$, donc $F = \mathrm{Vect}(X)$ est stable par A, de dimension 1.

Sinon, soit $\lambda \in \mathbf{C} \setminus \mathbf{R}$ une valeur propre, et soit $X \in \mathbf{C}^n$ un vecteur propre, $AX = \lambda X$, d'où $A\overline{X} = \bar{\lambda}\overline{X}$. Puisque $\lambda \neq \bar{\lambda}$, X et $\overline{X}$ ne sont pas colinéaire, on peu donc poser $Y = X + \overline{X} \in \mathbf{R}^n \setminus \{0\}$. An a $AY = \lambda X + \bar{\lambda}\overline{X}$ et $A^2 Y = \lambda^2 X + \bar{\lambda}^2 \overline{X} = (\lambda + \bar{\lambda})AY - \lambda\bar{\lambda}Y$, où $\lambda + \bar{\lambda}, \lambda\bar{\lambda} \in \mathbf{R}$. Donc $F = \mathrm{Vect}(Y, AY) \subseteq \mathbf{R}^n$ est stable par A. □

Proposition 2.33. *Soit U un endomorphisme orthogonal.*

(i) *Toute valeur propre de U est de module 1.*

(ii) *E se décompose en somme orthogonale des sous-espaces stables par U de dimension 1 ou 2.*

Démonstration. (i) On a $U(F) \subseteq F$ et puisque U est inversible, $\dim U(F) = \dim F$, d'où $F = U(F)$. Ainsi, F est stable par U^{-1}, i.e., par U^*. Maintenant, si $x \in F^\perp$ et $y \in F$ alors $U^* y \in F$ et $\langle Ux, y\rangle = \langle x, U^* y\rangle = 0$, donc $Ux \in F^\perp$ également.

(ii) Déjà vu.

(iii) D'après le Lemme précédent, E admet un sous-espace vectoriel $F \subseteq E$, tel que F est stable par U et $\dim F \leq 2$. Puisque $F^{\perp}$ est aussi stable par E on peut procéder par récurrence. □

10. Espace euclidiens orientés

QUESTION. Comment trouver le signe de l'angle entre deux vecteurs x et y dans le plan ?

Soit E un espace euclidien. Nous disons que deux bases orthonormées $\mathcal{B}$ et $\mathcal{B}'$ définissent **la même orientation** sur E si la matrice de passage $\mathcal{B} \to \mathcal{B}'$ est directe (i.e., de déterminant 1, sachant déjà qu'elle est orthogonale).

EXEMPLES DANS LE PLAN AU TABLEAU.

Ceci sépare les base orthonormées de E en deux classes d'équivalence.

Une **orientation** sur E est donnée par l'une de ces classes d'équivalence, c.à.d. par une base orthonormée **à transformation orthogonale directe près**. Un espace euclidien admet donc deux orientations possibles.

Pour $\mathbf{R}^n$ il y a toujours une orientation canonique, à savoir, celle donnée par la base canonique.

Si $k < n = \dim E$, alors toute famille orthogonale $(e_1, \ldots, e_k)$ peut être complété en une base orthonormée qui définit une orientation désirée. En effet, on complète en une base o.n. quelconque, et si elle ne donne pas la bonne orientation, on remplace e_n par $-e_n$.

10.1. Plan euclidien orienté. Calcul de l'aire d'un parallélogramme, par le déterminant. Invariance par une transformation orthogonale directe.

Angle signé.

10.2. Espace euclidien orienté de dim 3. Volume d'un parallélépipède, par le déterminant. Invariance par une transformation orthogonale directe.

Notons-le $\mathrm{vol}(x, y, z)$. Fixant x, y, c'est une forme linéaire en z. D'après un théorème précédent, il existe un unique vecteur $x \wedge y \in E$ tel que pour tout z on ait

$$\langle x \wedge y, z \rangle = \mathrm{vol}(x, y, z).$$

On appelle $x \wedge y$ le **produit vectoriel** de x et y.

EXERCICE 2.34. Vérifier que dans $\mathbf{R}^3$ avec l'orientation canonique on obtient la formule habituelle.

On appelle $\langle x \wedge y, z \rangle$ le **produit mixte** de x, y, z

Chapitre 3

Géométrie affine

1. Généralités

On peut donner deux définitions d'un espace affine. Une « concrète » avec laquelle il est plus facile de travailler, une plus abstraite, et d'une manière meilleure (plus générale). Toujours, E est un $\mathbf{R}$-espace vectoriel de dimension finie.

Définition 3.1. Une partie $\mathcal{F} \subseteq E$ est une **partie (ou sous-espace) affine** de E s'il existe $u \in E$ et un s.e.v. $F \subseteq E$ tel que

$$\mathcal{F} = F + u = \{y + u : y \in F\}.$$

En particulier, une partie affine de E n'est jamais vide.

Si $\mathcal{F} = F + u$ c'est que forcément $x \in \mathcal{F}$ et $F = \mathcal{F} - u = \{v - u : v \in \mathcal{F}\}$. Si $u' \in \mathcal{F}$ est un autre point, alors $u' - u \in F$. Donc $F \pm (u' - u) = F$, et on a :

$$F = F - (u' - u) = \mathcal{F} - u - (u' - u) = \mathcal{F} - u'$$
$$\mathcal{F} = F + u = F + (u' - u) + u = F + u'.$$

Autrement dit, $\mathcal{F} = F + u$ pour **tout** $u \in \mathcal{F}$, et $F = \mathcal{F} - u$ ne dépend pas du choix de u.

On appelle F l'**espace directeur** de $\mathcal{F}$. Pour $u, v \in \mathcal{F}$ on note aussi $\overrightarrow{uv} = v - u$. C'est le vecteur qui mène du point u de $\mathcal{F}$ au point v :

$$u + \overrightarrow{uv} = v.$$

Example. Point. Droite affine. Etc...

Tout espace vectoriel est un sous-espace affine de lui-même.

Un sous-espace affine dirigée par un hyperplan vectoriel s'appelle un **hyperplan affine**.

Lemme 3.2. *Supposons que F est un un hyperplan vectoriel. C'est donc le noyau d'une forme linéaire non nulle $\varphi \colon E \to \mathbf{R}$. Alors pour tout $c \in \mathbf{R}$, l'ensemble*

$$\mathcal{F}_c = \{u \in E : \varphi(u) = c\}$$

est une partie affine dirigée par F. Réciproquement, toute partie affine dirigée par F est de cette forme.

Example. Équation d'une droite affine dans le plan, d'un plan affine en dim 3.

Si E est un espace euclidien, on peux aussi prendre un vecteur z normal à $\mathcal{F}$. L'équation prends alors la forme :

$$\langle u, z \rangle = c.$$

Cette approche a le désavantage d'exiger que l'espace affine $\mathcal{F}$ soit une partie d'un espace vectoriel. Voici la définition générale, d'un espace affine « indépendant » :

Définition 3.3. Soit E un $\mathbf{R}$-espace vectoriel de dimension finie.

Un **espace affine** dirigé par E est un ensemble non-vide $\mathcal{E}$ muni d'une application $\mathcal{E} \times \mathcal{E} \to E$ qui associe au couple de points $A, B \in \mathcal{E}$ un vecteur $\overrightarrow{AB}$, et qui vérifie

(i) Pour tout $O \in \mathcal{E}$ l'application $A \mapsto \overrightarrow{OA}$ est une bijection de $\mathcal{E}$ sur E.

(ii) Pour tous $A, B, C \in \mathcal{E}$ on a la *relation de Chasles :* $\overrightarrow{AC} = \overrightarrow{AB} + \overrightarrow{BC}$.

La *dimension* de $\mathcal{E}$ est par définition celle de E.

EXERCICE 3.4. La relation de Chasles a pour conséquence immédiate : $\overrightarrow{AA} = \overrightarrow{0}$ et $\overrightarrow{BA} = -\overrightarrow{AB}$.

EXAMPLE. Tout sous-espace affine un espace vectoriel.

Réciproquement : Fixons $O \in \mathcal{E}$ que l'on appellera « origine, » et posons $v(A) = \overrightarrow{OA}$. Alors $v : \mathcal{E} \to E$ est une bijection qui identifie la structure affine de $\mathcal{E}$ avec la structure affine canonique de E : $\overrightarrow{v(A)v(B)} = v(B) - v(A) = \overrightarrow{OB} - \overrightarrow{OA} = \overrightarrow{AB}$ - conséquence de la relation de Chasles. Le choix d'origine donc "vectorialise" l'espace affine. Réciproquement, on peut dire qu'un espace affine est un espace vectoriel avec l'origine (zéro) « oubliée ».

Ainsi, dans la suite, chaque fois que cela nous convient nous prétendrons que $\mathcal{E}$ est un sous espace affine d'un espace vectoriel. Dans ce dernier cas on a $v = u + (v - u) = u + \overrightarrow{uv}$. Par conséquent :

DÉFINITION 3.5. Pour $A \in \mathcal{E}$ et $w \in E$ on définit $A + u$, qui est la **translation** de A par u, par :

$$B = A + u \qquad \Longleftrightarrow \qquad u = \overrightarrow{AB}$$

(un unique tel B existe, par définition d'un espace affine). Ainsi :

$$A + \overrightarrow{AB} = B.$$

2. Barycentres, combinaisons affines

Commençons avec un exemple :

Soit $v \in E$ et $A \in \mathcal{E}$. La **droite affine** passant par A dirigée par v est l'ensemble de points $M = A + \lambda v$, $\lambda \in \mathbf{R}$. La droite affine passant par deux points A et B est dirigée par le vecteur $v = \overrightarrow{AB}$; ses points sont $M = A + \lambda\overrightarrow{AB}$, $\lambda \in \mathbf{R}$.

Choisissons une "origine" O. Alors $M = A + \lambda\overrightarrow{AB}$ si et seulement si $\overrightarrow{OM} = \overrightarrow{OA} + \lambda\overrightarrow{AB} = (1 - \lambda)\overrightarrow{OA} + \lambda\overrightarrow{OB}$. Autrement dit, la droite affine passant par A et B consiste en tous les points M tels que $\overrightarrow{OM} = \mu\overrightarrow{OA} + \lambda\overrightarrow{OB}$ avec $\lambda + \mu = 1$. On appelle M le **barycentre** des points A et B affectés de poids μ et λ.

LEMME 3.6. *Soit $A_1, ..., A_k$ des points affectés des poids $\lambda_1, ..., \lambda_k$, tels que $\sum_i \lambda_i \neq 0$. Soit aussi $O \in \mathcal{E}$ un point quelconque (une « origine »). Les propriétés suivantes sont équivalentes pour un point $M \in \mathcal{E}$:*

(i) $\sum_i \lambda_i \overrightarrow{MA_i} = 0$.

(ii) $(\sum_i \lambda_i)\overrightarrow{OM} = \sum_i \lambda_i \overrightarrow{OA_i}$.

(iii) $\overrightarrow{OM} = \frac{1}{\sum_i \lambda_i} \sum_i \lambda_i \overrightarrow{OA_i}$.

(iv) $M = O + \frac{1}{\sum_i \lambda_i} \sum_i \lambda_i \overrightarrow{OA_i}$. *En particulier, un unique M ayant ces propriétés existe.*

De surcroît, un unique tel point M existe toujours.

DÉMONSTRATION. (i) $\Longleftrightarrow$ (ii). Puisque

$$\sum_i \lambda_i \overrightarrow{OA_i} - \left(\sum_i \lambda_i\right) \overrightarrow{OM} = \sum_i \lambda_i \left(\overrightarrow{OA_i} - \overrightarrow{OM}\right) = \sum_i \lambda_i \overrightarrow{MA_i} = 0.$$

(ii) $\iff$ (iii) $\iff$ (iv). Clair. $\square$

DÉFINITION 3.7. Le point M défini dans le lemme s'appelle le **barycentre** des points $A_1, ..., A_k$ affectés des coefficients $\lambda_1, ..., \lambda_k$. On notera $M = Bar\{(\lambda_i, A_i)\}$.

Si $M = Bar\{(\lambda_i, A_i)\}$ et $\sum_i \lambda_i = 1$, on écrit aussi $M = \sum_i \lambda_i A_i$.

En d'autres mots, $M = \sum_i \lambda_i A_i$ si et seulement si $\sum \lambda_i = 1$ et $\overrightarrow{OM} = \sum_i \lambda_i \overrightarrow{OA_i}$ (et ceci quel que soit O).

Si $\mathcal{E} = E$, le barycentre v des vecteurs $u_1, ..., u_k$ affectés des poids $\lambda_1, ..., \lambda_k$ est $v = \frac{1}{\sum_i \lambda_i} \sum_i \lambda_i u_i$. En particulier, si $\sum_i \lambda_i = 1$, alors $v = \sum_i \lambda_i u_i$. (Comparer avec la notation pour le barycentre : $M = \sum_i \lambda_i A_i$.)

Vu que v est une combinaison linéaire des vecteurs u_i, on peut appeler le barycentre « combinaison affine des points A_i ».

DÉFINITION 3.8 (Repère affine). On appelle **repère affine** de $\mathcal{E}$ la donnée de $n+1$ points $(A_0, \ldots, A_n)$ tels que chaque $B \in \mathcal{E}$ s'exprime d'une manière unique comme $B = \sum_0^n x_i A_i$ (où $\sum_0^n x_i = 1\,!$). On appelle $(x_0, \ldots, x_n)$ les **coordonnées barycentriques** de B dans ce repère.

Cette définition est symétrique, dans le sens qu'une permutation d'un repère est un repère. Brisons cette symétrie en donnant à A_0 un rôle particulier et posant $e_i = \overrightarrow{A_0A_i}$. On a donc $B = \sum_0^n x_i A_i$ si et seulement si $\overrightarrow{A_0B} = \sum_0^n x_i \overrightarrow{A_0A_i} = \sum_1^n x_i \overrightarrow{A_0A_i}$. Ainsi, $(A_0, \ldots, A_n)$ est un repère si et seulement si chaque vecteur $v \in E$ s'exprime d'une manière unique comme $\sum_1^n x_i e_i$. Autrement dit :

LEMME 3.9. *Une famille de $n+1$ vecteurs $(A_0, \ldots, A_n)$ est un repère affine si et seulement si $(e_1, \ldots, e_n)$ est une base de E. En particulier, $n = \dim E$.*

DÉFINITION 3.10. Les **coordonnées affines** d'un point B dans le repère $(A_0, \ldots, A_n)$ sont par définition les coordonnées du vecteur $\overrightarrow{A_0B}$ dans la base $(e_1, \ldots, e_n)$.

Le lien : les coordonnées barycentriques de B dans un repère $(A_0, \ldots, A_n)$ sont $(x_0, x_1, \ldots, x_n)$ si et seulement si $(x_1, \ldots, x_n)$ sont les coordonnées de $\overrightarrow{A_0B}$ dans la base $(e_1, \ldots, e_n)$ et $x_0 = 1 - \sum_1^n x_i$.

La propriété suivante de "l'associativité du barycentre" est très utile.

PROPOSITION 3.11. *Pour chaque $i = 1, \ldots, k$ soit $B_i = \sum_{j=1}^m \lambda_{i,j} A_{i,j}$, où $\sum_j \lambda_{i,j} = 1$. Soit $C = \sum_{i=1}^k \mu_i B_i$, où $\sum \mu_i = 1$. Alors $\sum_{i,j} \mu_i \lambda_{i,j} = 1$, et $C = \sum_{i,j} \mu_i \lambda_{i,j} A_{i,j}$.*

DÉMONSTRATION. On a
$$\sum_j \lambda_{i,j} \overrightarrow{B_iA_{i,j}} = 0, \qquad \sum_i \mu_i \overrightarrow{CB_i} = 0.$$
Ainsi
$$0 = \sum_i \mu_i \left(\sum_i \lambda_{i,j} \right) \overrightarrow{CB_i} + \sum_i \mu_i \sum_j \lambda_{i,j} \overrightarrow{B_iA_{i,j}} = \sum_{i,j} \mu_i \lambda_{i,j} (\overrightarrow{CB_i} + \overrightarrow{B_iA_{i,j}}) = \sum_{i,j} \mu_i \lambda_{i,j} \overrightarrow{CA_{i,j}}.$$
$\square$

3. Sous-espace affine

DÉFINITION 3.12. Une partie non vide $\mathcal{F}$ de l'espace affine $\mathcal{E}$ est un **sous-espace affine** si pour tout $A, B \in \mathcal{F}$, $\mathcal{F}$ contient la droite passant par ces deux points.

LEMME 3.13. *Les conditions suivantes sont équivalentes pour une partie $\mathcal{F} \subseteq \mathcal{E}$:*

(i) *$\mathcal{F}$ est un sous-espace affine de $\mathcal{E}$.*

(ii) *Pour tout $O \in \mathcal{F}$ l'ensemble $F = \{\overrightarrow{OM}, M \in \mathcal{F}\}$ est un sous-espace vectoriel de E. Dans ce cas, F ne dépend du choix de O, si bien que $\mathcal{F}$ est à son tour un espace affine dirigé par F.*

(iii) *On a $\mathcal{F} = \{O + v, v \in F\}$ où F est un sous-espace vectoriel de E et O un point de $\mathcal{F}$.*

(iv) *Si $A_1, ..., A_k \in \mathcal{F}$ alors tout barycentre des points $A_1, ..., A_k$ est dans $\mathcal{F}$. (Toute combinaison affine des points de $\mathcal{F}$ est dans $\mathcal{F}$.)*

Une *droite* est un sous-espace affine de dimension 1.
Un *hyperplan* est un sous-espace affine de codimension 1 (de dimension $\dim \mathcal{E} - 1$).

Lemme 3.14. *L'intersection d'une famille de sous-espaces affines de $\mathcal{E}$ et un sous-espace affine* **si cette intersection est non vide**.

Soit $\mathcal{F}$ un sous-espace affine dirigé par F ; soit $O \in \mathcal{F}$ et $(v_1, ..., v_k)$ une base de F. On a donc un repère affine de $\mathcal{F}$ et tout point B de $\mathcal{F}$ admet l'expression unique : $B = O + \sum_{i=1}^{k} s_i v_i$, $(s_1, ..., s_k) \in R^k$.

Soit $(x_1, ..., x_n)$ un système de coordonnées affines dans $\mathcal{E}$. Soit $(c_1, ..., c_n)$ les coordonnées du point O et $(a_{1i} ..., a_{ni})$ les coordonnées du vecteur v_i. Alors les coordonnées du point $B = O + \sum_{i=1}^{k} s_i v_i$ sont $x_j = c_j + \sum_{i=1}^{k} a_{ji} s_i$, $j = 1, ..., n$.

Position relative de deux sous-espaces

Lemme 3.15. (i) *Soit $\mathcal{F}$ et $\mathcal{G}$ deux sous-espaces de $\mathcal{E}$ d'intersection non-vide. Alors $\mathcal{F} \cap \mathcal{G}$ est dirigé par $F \cap G$.*

(ii) *Si $F + G = E$ alors $\mathcal{F} \cap \mathcal{G}$ est non vide.*

(iii) *Si les sous-espaces vectoriels F et G sont supplémentaires ($E = F \oplus G$), alors $\mathcal{F} \cap \mathcal{G}$ est un seul point.*

Définition 3.16. On dit que deux sous-espaces $\mathcal{F}$ et $\mathcal{G}$ de $\mathcal{E}$ sont **parallèles** si $F \subset G$ ou $G \subset F$. Dans ce cas que soit $\mathcal{F}$ et $\mathcal{G}$ sont disjoints, soit l'un contient l'autre.

Sous-espace affine engendré par une partie.

Soit $S \subset \mathcal{E}$ une partie non-vide. Le sous-espace affine engendré par S, noté Aff(S), est l'intersection de tous les sous-espaces affines qui contiennent S.

Lemme 3.17. (i) *Aff(S) est l'ensemble de toutes les combinaisons affines (barycentres) des points de S.*

(ii) *Soit $O \in S$ et soit F le sous-espace vectoriel de $\mathcal{E}$ engendré par les vecteurs $\overrightarrow{OA}$, $A \in S$. Alors l'espace directeur de Aff(S) est F : Aff(S) $= \{O + v, v \in F\}$.*

(iii) *$(A_0, ..., A_n)$ est un repère affine si et seulement si $n = \dim \mathcal{E}$ et Aff$(A_0, ..., A_n) = \mathcal{E}$.*

4. Applications affines

Définition 3.18. Une application $f : \mathcal{E} \to \mathcal{F}$ est **affine** s'il existe une origine O et une application linéaire $L_f \colon E \to F$ telles que pour tout $A \in \mathcal{E}$:

$$f(A) = f(O) + L_f(\overrightarrow{OA}),$$

ou, d'une manière équivalente, pour tout $v \in E$:

$$f(O + v) = f(O) + L_f(u).$$

On observe d'abord que le fait que f soit affine ne dépend pas du choix de l'origine O. En effet, si $B = O + u$ est une autre origine possible, alors :

$$f(B+v) = f(O+u+v) = f(O) + L_f(u+v) = f(O) + L_f(u) + L_f(v) = f(B) + L_f(v).$$

Deuxièmement, pour toute origine O, on peux retrouver L_f par :

$$L_f(v) = \overrightarrow{f(O)f(O+v)}.$$

DÉFINITION 3.19. On appelle L_f **la partie linéaire** de f.

Ainsi, on peut dire qu'une application affine est une application linéaire plus une constante.

EXAMPLE. Soit $v \in E$. On définit la **translation par** v :

$$\begin{array}{rccc} T_v\colon & \mathcal{E} & \to & \mathcal{E}, \\ & A & \mapsto & A+v \end{array}$$

Alors T_v est une application affine, de partie linéaire id_E. Réciproquement, **toute** application affine $f\colon \mathcal{E} \to \mathcal{E}$ telle que $L_f = \mathrm{id}_E$ est une translation.

On remarque que $T_0 = \mathrm{id}_{\mathcal{E}}$ et $T_{v+u} = T_v T_u$. Il en suit que $T_v\colon \mathcal{E} \to \mathcal{E}$ est une bijection, avec $T_v^{-1} = T_{-v}$.

EXERCICE 3.20. Les conditions suivantes sont équivalentes pour une application $f\colon \mathcal{E} \to \mathcal{F}$:

(i) f est une application affine.

(ii) f préserve les barycentres et les combinaisons affines :

$$f(Bar\{(\lambda_i, A_i)\}) = Bar\{(\lambda_i, f(A_i))\} \qquad \text{lorsque } \sum \lambda_i \neq 0$$
$$f\left(\sum \lambda_i A_i\right) = \sum \lambda_i f(A_i) \qquad \text{lorsque } \sum \lambda_i = 1$$

(iii) On a $f(\lambda A + \mu B) = \lambda f(A) + \mu f(B)$ pour tout $A, B \in \mathcal{E}$ et tout λ, μ tels que $\lambda + \mu = 1$.

4.1. Écriture en coordonnées. Fixons les repères affines $R = (A_0; e_1, ..., e_n)$ dans $\mathcal{E}$ et $R' = (B_0; e'_1, ..., e'_k)$ dans $\mathcal{F}$.
Soit $(x_1, ..., x_n)$ les coordonnées du point M dans R et $(y_1, ..., y_k)$ les coordonnées de $f(M)$ dans R'.
Alors $y_i = c_i + \sum_{j=1}^n a_{ij} x_j$, où (a_{ij}) est la matrice de L_f dans les bases des repères R, R' et $(c_1, ..., c_k)$ les coordonnées du point $f(A_0)$.
Fonction affine $f : \mathcal{E} \to \mathbf{R}$. Dans un système de coordonnées affines f s'écrit $f(x_1, ..., x_n) = c + \sum_i a_i x_i$.
.

LEMME 3.21. *Soit $f : \mathcal{E} \to \mathcal{F}$ une application affine. Alors*

(i) *Si $\mathcal{E}_1$ un sous-espace affine de $\mathcal{E}$ alors $f(\mathcal{E}_1)$ est un sous-espace affine de $\mathcal{F}$.*

(ii) *Si $\mathcal{F}_1$ un sous-espace affine de $\mathcal{F}$ et $f^{-1}(\mathcal{F}_1)$ est non-vide, alors, $f^{-1}(\mathcal{F}_1)$ est un sous-espace affine de $\mathcal{E}$.*

La *formule du rang* pour les applications linéaires s'adapte aux applications affines : Soit $B \in \mathrm{img}(f)$. Alors

$$\dim \mathcal{E} = \dim \mathrm{img}\, f + \dim f^{-1}(B).$$

Equation d'un hyperplan : tout hyperplan est défini par une équation affine $\sum_i a_i x_i = c$.

Exercice 3.22. Soit $f_i : \mathcal{E} \to \mathcal{F}$, $i = 1, ..., k$ des applications affines et $\lambda_1 + ... + \lambda_k = 1$. On définit $f = \sum_i \lambda_i f_i$ par $f(A) = \sum_i \lambda_i f_i(A)$.

Montrer que f est une application affine.

Exercice 3.23. Définir une structure de l'espace affine sur l'ensemble de toutes les applications affines $f : \mathcal{E} \to \mathcal{F}$, compatible avec les combinaisons affines de l'exercice précédent.

5. Applications affines de $\mathcal{E}$ dans $\mathcal{E}$. Transformations affines.

Les propriérés d'une application affine f sont déterminées par les propriétés de sa partie linéaire L_f.

Proposition 3.24. (i) *Si L_f n'admet pas de vecteur fixe non-nul (i.e., si 1 n'est pas une valeur propre de L_f), alors f admet un unique point fixe.*

(ii) *Si L_f admet des vecteurs fixes non-nuls (i.e., si 1 est une valeur propre de L_f), alors l'ensemble des points fixes de f est soit vide soit un sous-espace affine dirigé par $\{v \in E : L_f(v) = v\}$ (l'espace propre correspondant à 1).*

Si f admet un point fixe O et on prend $O = f(O)$ pour origine, f s'écrit $f(O+v) = O + L_f(v)$: donc, en vectorialisant $\mathcal{E}$ en O, on identifie f avec l'application linéaire L_f.

Définition 3.25. f est une **homothétie** de centre O et de rapport k si $f(O) = O$ et $L_f = kId$.

Proposition 3.26. (i) *$f : \mathcal{E} \to \mathcal{E}$ est une translation si et seulement si $L_f = \mathrm{Id}$.*

(ii) *Si $L_f = kId$ avec $k \neq 1$, alors f est une homothétie.*

Projections et symétries affines. Soit $\mathcal{F}$ un sous-espace affine de $\mathcal{E}$, F l'espace vectoriel associé, et soit G un sous-espace vectoriel de E supplémentaire à F : $E = F \oplus G$.

.

Soit $\Pi : E \to E$ le projecteur vectoriel sur F parallèlement à G : $\Pi(v) = v$ si $v \in F$ et $\Pi(v) = 0$ si $v \in G$.

Définition 3.27. Soit $O \in F$. La **projection affine** p sur $\mathcal{F}$ parallèlement à G est définie par $p(A) = O + \Pi(\overrightarrow{OA})$.

Lemme 3.28. (i) *La projection affine sur $\mathcal{F}$ parallèlement à G ne dépend par du choix du point O. Ainsi, elle vérifie $p(A) = A$ pour tout $A \in \mathcal{F}$ et $dp = \Pi$, et son image est $\mathcal{F}$.*

(ii) *Soit f une transformation affine de $\mathcal{E}$. Alors f est un projecteur affine si et seulement si $f^2 = f$. Dans ce case c'est le projecteur affine sur $\mathcal{F} = f(\mathcal{E})$, parallèlement à $\ker L_f$.*

Soit $\sigma : E \to E$ la symétrie vectorielle par rapport à F parallèlement à G : $\sigma(v) = v$ si $v \in F$ et $\sigma(v) = -v$ si $v \in G$.

2.7. Définition. La **symétrie affine** s par rapport à $\mathcal{F}$ parallèlement à G est définie par les conditions $s(O) = O$ si $O \in \mathcal{F}$ et $ds = \sigma$; donc $p(M) = O + \sigma(\overrightarrow{OM})$.

Lemme. s est une symétrie affine si et seulement si $f \cdot f = Id$.

6. Espaces affines euclidiens

Définition 3.29. Un **espace affine euclidien** est un espace affine $\mathcal{E}$ dirigé par une espace vectoriel euclidien E.

On définit la distance dans $\mathcal{E}$ par $d(A, B) = \|\overrightarrow{AB}\|$.

Définition 3.30. Un repère affine $(A_0, A_1, ..., A_n)$ est dit **orthonormé** si la base vectorielle $(\overrightarrow{A_0A_1}, ..., \overrightarrow{A_0A_n})$ est orthonormée.

Définition 3.31. Deux sous-espaces affine $\mathcal{F}$ et $\mathcal{G}$ sont **orthogonaux** si leurs espaces directeurs F et G sont orthogonaux.

6.1. Isométries.

DÉFINITION 3.32. Soit X un espace métrique. On appelle **isométrie** de X une bijection $f : X \to X$ qui préserve la distance : $d(fx, fy) = d(x, y)$ pour tous $x, y \in X$.

REMARQUE 3.33. Les isométries de X forment un groupe par rapport à la composition.

LEMME 3.34. *Une application affine $f : \mathcal{E} \to \mathcal{E}$ est une isométrie si et seulement si sa partie linéaire $L_f : E \to E$ est une isométrie de E.*

Rappel [une variante de la formule de polarisation] : dans un espace euclidien on a toujours :

$$\langle x, y\rangle = \frac{1}{2}\left[\|x\|^2 + \|y\|^2 - \|x - y\|^2\right] = \frac{1}{2}\left[\|x\|^2 + \|y\|^2 - d(x, y)^2\right].$$

THÉORÈME 3.35. *Toute isométrie d'un espace affine euclidien est une application affine.*

Remarque. Une application affine $f : \mathcal{E} \to \mathcal{E}$ est une isométrie si et seulement si f transforme un (ou tout) repère orthonormé en repère orthonormé.

DÉFINITION 3.36. On appelle **déplacement** ou *isométrie directe* (respectivement, **antidéplacement** ou *isométrie indirecte*) de $\mathcal{E}$ toute isométrie f telle que $\det L_f = 1$ (respectivement, $\det L_f = -1$).

Les translations sont évidemment des déplacements. Les déplacements forment un sous-groupe dans $\mathrm{Iso}(\mathcal{E})$ d'indice 2 ; le produit de deux antidéplacements est un déplacement.

DÉFINITION 3.37. Soit $\mathcal{F}$ un sous-espace affine de $\mathcal{E}$. La **projection orthogonale** sur $\mathcal{F}$, notée $P_{\mathcal{F}} \colon \mathcal{E} \to \mathcal{F}$, est la projection affine sur $\mathcal{F}$ parallèlement à $F^{\perp}$.
Formule : on fixe une origine $O \in \mathcal{F}$, et

$$P_{\mathcal{F}}(A) = O + P_F(\overrightarrow{OA}).$$

DÉFINITION 3.38. Soit $\mathcal{F}$ un sous-espace affine de $\mathcal{E}$.
La **symétrie orthogonale** par rapport à $\mathcal{F}$, notée $s_{\mathcal{F}} \colon \mathcal{E} \to \mathcal{E}$ est la symétrie affine par rapport à $\mathcal{F}$ parallèlement à $F^{\perp}$.
Formule : on fixe une origine $O \in \mathcal{F}$, et

$$s_{\mathcal{F}}(A) = O + s_F(\overrightarrow{OA}).$$

Une symétrie orthogonale est une isométrie, car s_F est une isométrie de E. On a $\det s_F = (-1)^{\dim E - \dim F} = (-1)^{\dim \mathcal{E} - \dim \mathcal{F}}$. Ainsi, $s_{\mathcal{F}}$ est directe si et seulement si la codimension de $\mathcal{F}$ est paire (c'est à dire, $dim(\mathcal{E}) - dim(\mathcal{F})$ est paire).
On appelle **reflexion** une symétrie orthogonale par rapport à un hyperplan ; une reflexion est une isométrie indirecte (un antidéplacement).
Étant donné deux points distincts A et B de $\mathcal{E}$, l'hyperplan médiateur du segment $[AB]$ est donné par : $\mathcal{F} = \frac{1}{2}A + \frac{1}{2}B + \overrightarrow{AB}^{\perp} = \{\frac{1}{2}A + \frac{1}{2}B + v : v \perp \overrightarrow{AB}\}$ (i.e., l'unique s.e.a. passant par $\frac{1}{2}(A + B)$ et dirigé par $\overrightarrow{AB}^{\perp}$). La réflexion orthogonale $s_{\mathcal{F}}$ échange A et B.

EXERCICE 3.39. La réflexion orthogonale $s_{\mathcal{F}}$ est l'unique réflexion orthogonale échangeant A et B.

Vu que toute isométrie vectorièle dans R^n s'écrit comme un produit d'au plus n réflexions, on a :

PROPOSITION 3.40. *Toute isométrie de $\mathcal{E}$ s'écrit comme un produit d'au plus $1 + \dim \mathcal{E}$ réflexions.)*

LEMME 3.41. *Soit E un espace euclidien, $f \in \mathrm{End}(E)$ un endomorphisme orthogonal. Alors pour tout $w \in E$ il existe $v, u \in E$ tels que :*

$$v = fv = w + fu - u.$$

DÉMONSTRATION. Posons $g = f - \mathrm{id}$ et $G = \mathrm{img}\, g$. Alors

$$G^{\perp} = \ker g^* = \ker(f^{-1} - \mathrm{id}) = \{x \in E : fx = x\} = \ker g,$$

i.e., $\ker g \oplus \mathrm{img}\, g = E$. Ainsi on peut exprimer $w = v + u_0$ où $v \in \ker g$ et $u_0 \in \mathrm{img}\, g$. Autrement dit, $fv = v$ et $u_0 = fu - u$ pour un certain $u \in E$. □

PROPOSITION 3.42. *Soit $\mathcal{E}$ un espace affine euclidien et $f\colon \mathcal{E} \to \mathcal{E}$ une isométrie. Alors il existe un point A tel que $L_f(v) = v$ où $v = \overrightarrow{Af(A)}$.*

Dans ce cas on pose $f_A(A + x) = A + L_f(x)$ et on a

— $L_{f_A} = L_f$ *(en particulier, f_A est une isométrie).*

— $f = T_v f_A = f_A T_v$.

— $f_A(A) = A$.

DÉMONSTRATION. Fixons $O \in \mathcal{E}$ arbitrairement. D'après le Lemme il existe $v, u \in E$ tels que

$$v = L_f(v) = \overrightarrow{Of(O)} + L_f(u) - u.$$

Posons $A = O + u$. Alors $f(A) = f(O) + L_f(u) = O + u + \overrightarrow{Of(O)} + L_f(u) - u = A + v$, donc $v = \overrightarrow{Af(A)}$. On a déjà $L_f(v) = v$.

Par définition de f_A on a $L_{f_A} = L_f$ et $L_{f_A}(A) = A$. Il reste :

$$T_v f_A(A + x) = T_v\big(A + L_f(x)\big) = A + L_f(x) + v = f(A) + L_f(x) = f(A)$$
$$f_A T_v(A + x) = f_A(A + x + v) = A + L_f(x + v) = A + L_f(x) + v = \ldots = f(A).$$ □

3.10. Définition. Une transformation affine $f : \mathcal{E} \to \mathcal{E}$ est une **similitude** de rapport k si f multiplie toutes les distances par k :

$d(f(x), f(y)) = kd(x, y)$ pour tous $x, y \in \mathcal{E}$.

Une homothétie de rapport k est une similitude de rapport $\mid k \mid$.

Lemme. Une application affine $f : \mathcal{E} \to \mathcal{E}$ est une similtude de rapport k si et seulement si sa partie linéaire $L_f : E \to E$ est une similitude vectorielle : $\|L_f(v)\| = \pm k\|v\|$.

Remarque : les similitudes de $\mathcal{E}$ forment un groupe par rapport à la composition. Si f_i est une similitude de rapport k_i, $i = 1, 2$, alors $f_1 f_2$ est une similitude de rapport $k_1 k_2$.

3.11. Lemme. (i) Toute similitude de rapport $k \neq 1$ admet un point fixe (unique).

(ii) Toute similitude de rapport k est la composée d'une homothetie de rapport k et d'une isométrie.

.

Classification des isométries en dimension 2 et 3.

La classification des isométries affines repose sur la classifcation des isométries vectorielles.

Lemme. Soit U une isométrie de l'espace vectoriel E.

1. dim $E = 2$. a) dét $U = 1$: U est une rotation.

b) dét $U = -1$: U est une réflexion.

2. dim $E = 3$. a) dét $U = 1$: U est une rotation autour d'un axe.

b) dét $U = -1$: U est une réflexion composée avec une rotation autour de l'axe orthogonal au plan de réflexion.

.

Isométries affines du plan

4.1. Proposition. (i) Tout déplacement du plan est une translation ou une rotation autour d'un point ; ici la rotation commute avec la translation.

(ii) Toute antidéplacement est le produit d'une réflexion par rapport à une droite avec une translation parallèle à cette droite (une "réflexion-translation") ; ici la reflexion commute avec la translation.

.

Isométries affines de l'espace.

4.2. Proposition. (i) Tout déplacement de l'espace est soit une translation soit le produit (commutatif) d'une rotation autour d'un axe et d'une translation parallèle à l'axe de rotation ("vissage").

(ii) Toute antidéplacement est soit le produit (commutatif) d'une réflexion par rapport à un plan avec une translation parallèle à ce plan ("réflexion-translation"), soit le produit (commutatif) d'une réflexion par rapport à un plan avec une rotation autour d'un axe orthogonal à ce plan ("réflexion-rotation").

.

Similitudes planes et nombres complèxes.

On identifie le plan euclidien R^2 avec le plan complèxe $\mathbf{C}$.

4.3. Proposition. Toute similitude directe de $\mathbf{C}$ s'écrit $s(z) = az + b$, avec $a \neq 0$. Toute similitude indirecte s'écrit $s(z) = a\bar{z} + b$, $a \neq 0$.

Le rapport d'une telle similitude est $\mid a \mid$.

En particulier, c'est une isométrie si et seulement si $\mid a \mid = 1$.