Lycée Henri IV
MPSI

Cours
Durée : 8 h 30

Espaces euclidiens

Les énoncés et parties suivis du symbole [✠] ne seront pas traités en cours.

MPSI – Lycée Henri IV – Sébastien Kerner

Prérequis

Revoir les chapitres sur :

- les espaces vectoriels ;
- les applications linéaires ;
- la dimension finie ;
- les déterminants.

Dans ce cours, le corps de base de tous les espaces vectoriels est $\mathbb{R}$.

La lettre n désigne un entier naturel non nul.

Les trois lettres AQT signifient «Âne Qui Trotte» et sont utilisées pour désigner une démonstration facile laissée au lecteur.

A. Espaces préhilbertiens réels

A.1. Produit scalaire

Définition 1

Soit E un $\mathbb{R}$-espace vectoriel. On appelle produit scalaire sur E toute forme bilinéaire symétrique définie positive, c'est-à-dire toute application $\langle \cdot, \cdot \rangle : E^2 \longrightarrow \mathbb{R}$ telle que

(i) $\forall x, x', y \in E,\ \forall \lambda, \mu \in \mathbb{R},\ \ \langle \lambda x + \mu x', y \rangle = \lambda \langle x, y \rangle + \mu \langle x', y \rangle$ (linéarité à gauche) ;

(i') $\forall x, y, y' \in E,\ \forall \lambda, \mu \in \mathbb{R},\ \ \langle x, \lambda y + \mu y' \rangle = \lambda \langle x, y \rangle + \mu \langle x, y' \rangle$ (linéarité à droite) ;

(ii) $\forall x, y \in E,\ \ \langle x, y \rangle = \langle y, x \rangle$ (symétrie) ;

(iii) $\forall x \in E,\ \ \langle x, x \rangle \geqslant 0$ (positivité) ;

(iv) $\forall x \in E,\ \ (\langle x, x \rangle = 0) \Longrightarrow (x = 0_E)$ (caractère « défini »).

Le produit scalaire $\langle x, y \rangle$ est quelquefois noté $(x \mid y)$ ou $x \cdot y$; cette dernière notation étant généralement dévolue aux produits scalaires géométriques.

Le couple $(E, \langle \cdot, \cdot \rangle)$ constitue alors un espace préhilbertien réel.

Cette définition généralise le produit scalaire géométrique que vous avez rencontré en classe de Première ($\overrightarrow{u} \cdot \overrightarrow{v} = \|\overrightarrow{u}\| \times \|\overrightarrow{v}\| \times \cos(\overrightarrow{u}, \overrightarrow{v})$). Toutefois, ici, le produit scalaire n'est pas défini à l'aide des concepts de norme et d'angle, mais comme une notion intrinsèque attachée à la structure de E. Les normes et les angles vont venir après !

Notons que, pour tout $x \in E$, on a $\langle x, 0_E \rangle = \langle 0_E, x \rangle = 0$. Par conséquent, on peut rassembler (iii) et (iv) en une seule assertion : $\forall x \in E,\ (x \neq 0_E) \Longrightarrow (\langle x, x \rangle > 0)$.

Pour démontrer qu'une application $\langle \cdot, \cdot \rangle : E^2 \longrightarrow \mathbb{R}$ est un produit scalaire, on démontre la symétrie avant la bilinéarité ; cela permet de ne démontrer que la linéarité à gauche ou à droite mais pas les deux. Autrement dit, on démontre (ii) puis (i) (ou (i')) puis (iii) et enfin (iv).

Notons que si $\langle \cdot, \cdot \rangle$ est un produit scalaire sur E, alors pour tout $\lambda \in \mathbb{R}_+^*$, $\lambda \langle \cdot, \cdot \rangle$ en est un autre.

Exemples :

- Soient $\alpha_0, \alpha_1, \ldots, \alpha_n \in \mathbb{R}$ distincts (il y en a $n+1$). Le produit scalaire de Lagrange est l'application $\langle \cdot, \cdot \rangle : \mathbb{R}_n[X]^2 \longrightarrow \mathbb{R}$ définie, pour tous $P, Q \in \mathbb{R}_n[X]$, par

$$\langle P, Q \rangle = P(\alpha_0)Q(\alpha_0) + P(\alpha_1)Q(\alpha_1) + \cdots + P(\alpha_n)Q(\alpha_n).$$

  Vérifions que c'est bien un produit scalaire.
  - La symétrie est évidente.
  - La linéarité à gauche (ou à droite) découle de la linéarité du symbole $\sum$. On en déduit la bilinéarité (grâce à la symétrie).
  - Soit $P \in \mathbb{R}_n[X]$. On a $\langle P, P \rangle = P(\alpha_0)^2 + P(\alpha_1)^2 + \cdots + P(\alpha_n)^2 \geqslant 0$, d'où la positivité.
  - Supposons que $\langle P, P \rangle = 0$, c'est-à-dire $P(\alpha_0)^2 + P(\alpha_1)^2 + \cdots + P(\alpha_n)^2 = 0$. On a alors $P(\alpha_0) = P(\alpha_1) = \cdots = P(\alpha_n) = 0$, ce qui impose que $P = 0$ puisque P possède $n+1$ racines alors qu'il est de degré au plus n. Cela démontre que $\langle \cdot, \cdot \rangle$ est défini.

- Soit $(\Omega, \mathscr{T}, P)$ un espace probabilisé. Sur l'espace vectoriel des variables aléatoires réelles définies sur $(\Omega, \mathscr{T}, P)$, la covariance est une forme bilinéaire symétrique positive. En revanche, elle n'est pas définie puisque une variable de variance nulle est seulement certaine (et pas nulle). Ce n'est donc pas un produit scalaire (mais il s'en faut de peu).

On donne ci-après une liste d'exemples fondamentaux d'espaces préhilbertiens réels. Ces exemples de référence sont à bien connaître. Leur structure préhilbertienne est naturelle et l'on ne vous demandera jamais de redémontrer que ce sont bien des espaces préhilbertiens.

On commence naturellement par le produit scalaire historique : celui de la géométrie.

Exemple fondamental

Sur le plan ou l'espace usuel, le produit scalaire géométrique est défini par

$$\vec{u} \cdot \vec{v} = \|\vec{u}\| \times \|\vec{v}\| \times \cos(\vec{u}, \vec{v}).$$

■ Ce n'est pas si simple que cela... Soient $\vec{u}, \vec{v}, \vec{w}$ trois vecteurs et λ un nombre réel.

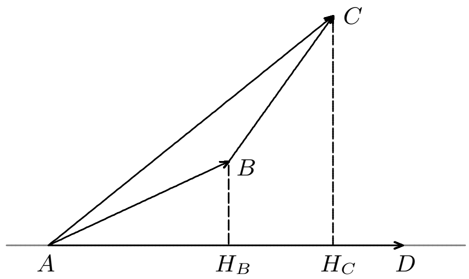
▷ Le caractère symétrique découle du fait que $(\vec{u}, \vec{v}) \equiv -(\vec{v}, \vec{u})\ [2\pi]$ et de la parité du cos.

▷ Démontrons la linéarité à gauche en deux temps.

∗ Démontrons que $(\lambda\vec{u}) \cdot \vec{v} = \lambda(\vec{u} \cdot \vec{v})$. Si $\lambda = 0$, c'est évident. Sinon, comme $\|\lambda\vec{u}\| = |\lambda|.\|\vec{u}\|$, on a $(\lambda\vec{u}) \cdot \vec{v} = |\lambda|.\|\vec{u}\|.\|\vec{v}\|.\cos(\lambda\vec{u}, \vec{v})$. Si $\lambda > 0$, on a $|\lambda| = \lambda$ et $\cos(\lambda\vec{u}, \vec{v}) = \cos(\vec{u}, \vec{v})$. Si $\lambda < 0$, on a $|\lambda| = -\lambda$ et $\cos(\lambda\vec{u}, \vec{v}) = \cos\left(\pi - (\vec{u}, \vec{v})\right) = -\cos(\vec{u}, \vec{v})$. Il s'ensuit que $(\lambda\vec{u}) \cdot \vec{v} = \lambda.\|\vec{u}\|.\|\vec{v}\|.\cos(\vec{u}, \vec{v}) = \lambda(\vec{u} \cdot \vec{v})$.

∗ Démontrons ensuite que $(\vec{u} + \vec{v}) \cdot \vec{w} = \vec{u} \cdot \vec{w} + \vec{v} \cdot \vec{w}$. Considérons des points A, B, C, D tels que $\vec{u} = \overrightarrow{AB}$, $\vec{v} = \overrightarrow{BC}$, $\vec{w} = \overrightarrow{AD}$ et orientons la droite (AD). Notons H_B le projeté orthogonal de B sur (AD) et H_C celui de C sur (AD). On a

$$\begin{aligned}
(\vec{u} + \vec{v}) \cdot \vec{w} &= \left(\overrightarrow{AB} + \overrightarrow{BC}\right) \cdot \overrightarrow{AD} \\
&= \overrightarrow{AC} \cdot \overrightarrow{AD} \\
&= \overline{AH_C} \times \overline{AD} \\
&= (\overline{AH_B} + \overline{H_BH_C}) \times \overline{AD} \\
&= \overline{AH_B} \times \overline{AD} + \overline{H_BH_C} \times \overline{AD} \\
&= \vec{u} \cdot \vec{w} + \vec{v} \cdot \vec{w}
\end{aligned}$$



▷ Le caractère défini positif découle immédiatement du fait que $\vec{u} \cdot \vec{u} = \|\vec{u}\|^2$. ■

On poursuit avec la généralisation, en dimension n, de la formule $(x, y) \cdot (x', y') = xx' + yy'$.

Exemple fondamental

Le produit scalaire canonique de $\mathbb{R}^n$ est défini par

$$\langle (x_1, \ldots, x_n), (y_1, \ldots, y_n) \rangle = \sum_{k=1}^{n} x_k y_k.$$

En identifiant les n-uplets et les vecteurs colonnes $n \times 1$ (ce qui revient à confondre $\mathbb{R}^n$ et $\mathscr{M}_{n,1}(\mathbb{R})$), le produit scalaire canonique de $\mathbb{R}^n$ se reformule ainsi :

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

■ AQT ■

Il existe bien sûr d'autres produits scalaires sur $\mathbb{R}^n$. Par exemple : $\langle x, y \rangle = \sum_{k=1}^{n} k x_k y_k$.

Plus généralement, si E désigne un $\mathbb{R}$-espace vectoriel de dimension finie n muni d'une base $\mathscr{B} = (e_1, \ldots, e_n)$, on définit un produit scalaire en posant

$$\left\langle \sum_{k=1}^{n} x_k e_k, \sum_{k=1}^{n} y_k e_k \right\rangle = \sum_{k=1}^{n} x_k y_k \qquad \text{c'est-à-dire} \qquad \langle x, y \rangle = {}^t \operatorname{Mat}_{\mathscr{B}}(x) \operatorname{Mat}_{\mathscr{B}}(y).$$

Il n'y a pas de raison de se limiter à des vecteurs colonnes : on peut naturellement munir les matrices d'un produit scalaire !

Exemple fondamental

Le produit scalaire canonique de $\mathscr{M}_{n,m}(\mathbb{R})$ est défini par

$$\langle A, B\rangle = \mathrm{Tr}\left({}^t\!AB\right).$$

■ Soient $A, B \in \mathscr{M}_{n,m}(\mathbb{R})$.

On a $\langle A, B\rangle = \mathrm{Tr}\left({}^t\!AB\right) = \mathrm{Tr}\left({}^t({}^t\!BA)\right) = \mathrm{Tr}\left({}^t\!BA\right) = \langle B, A\rangle$. Cela démontre la symétrie.

La linéarité à gauche (ou à droite) découle de la linéarité de la trace.

Si $A = (a_{i,j})_{1\leqslant i\leqslant n, 1\leqslant j\leqslant m}$, on a ${}^t\!AA = \left(\sum_{k=1}^{n} a_{k,i}a_{k,j}\right)_{1\leqslant i\leqslant n, 1\leqslant j\leqslant m}$, d'où $\mathrm{Tr}\left({}^t\!AA\right) = \sum_{j=1}^{m}\sum_{k=1}^{n} {a_{k,j}}^2$. On a donc $\langle A, A\rangle \geqslant 0$, ce qui démontre la positivité.

Si $\langle A, A\rangle = 0$, on a $\sum_{j=1}^{m}\sum_{k=1}^{n} {a_{k,j}}^2 = 0$, c'est-à-dire $\forall k \in [\![1;n]\!]\,,\ \forall j \in [\![1;m]\!]\,,\ a_{k,j} = 0$ ou encore $A = 0_{n,m}$. Cela démontre que $\langle\cdot,\cdot\rangle$ est défini. ■

Il est à noter que sur $\mathscr{M}_{n,1}(\mathbb{R})$, on a $\langle X, Y\rangle = \mathrm{Tr}\left({}^t\!XY\right) = {}^t\!XY$ puisque ${}^t\!XY$ est une matrice 1×1. Le produit scalaire canonique sur $\mathbb{R}^n$ apparaît ainsi comme un cas particulier du produit scalaire matriciel.

Mais on peut inverser le point de vue ! Si $A = (a_{i,j})_{1\leqslant i\leqslant n, 1\leqslant j\leqslant m}$ et $B = (b_{i,j})_{1\leqslant i\leqslant n, 1\leqslant j\leqslant m}$, un calcul similaire à celui de la démonstration ci-dessus dit que $\langle A, B\rangle = \sum_{(i,j)\in[\![1;n]\!]\times[\![1;m]\!]} a_{i,j}b_{i,j}$. Cela démontre que, si l'on regarde les matrices de $\mathscr{M}_{n,m}(\mathbb{R})$ comme des éléments de $\mathbb{R}^{nm}$ (on forme un nm-uplet en lisant, ligne après ligne, les coefficients de la matrice), le produit scalaire matriciel sur $\mathscr{M}_{n,m}(\mathbb{R})$ n'est rien d'autre que le produit scalaire canonique sur $\mathbb{R}^{nm}$.

Voici maintenant le produit scalaire naturel sur les fonctions continues sur un segment.

Exemple fondamental

Soient $a, b \in \mathbb{R}$ tels que $a < b$. Le produit scalaire canonique de $\mathscr{C}^0([a;b],\mathbb{R})$ est défini par

$$\langle f, g\rangle = \int_a^b f(t)g(t)\,\mathrm{d}t.$$

■ La symétrie est évidente.

La linéarité à gauche (ou à droite) découle de la linéarité de l'intégrale.

La positivité découle de la positivité de l'intégrale.

Enfin, si $f \in \mathscr{C}^0([a;b]\,,\mathbb{R})$ est telle que $\langle f, f\rangle = 0$, cela signifie que f^2 est une fonction continue, positive, d'intégrale nulle sur $[a;b]$ et donc que f est la fonction nulle sur $[a;b]$. Cela prouve le caractère défini. ■

Les produits scalaires fonctionnels sont légions. En voici un exemple.

Exemples :

- Sur l'espace vectoriel $\mathscr{C}^0_{2\pi}(\mathbb{R},\mathbb{R})$ des fonctions continues et 2π-périodiques sur $\mathbb{R}$, on dispose du produit scalaire canonique :

$$\langle f, g\rangle = \frac{1}{\pi}\int_0^{2\pi} f(t)g(t)\,\mathrm{d}t.$$

  Les propriétés de symétrie, de bilinéarité et de positivité sont en tout point similaires à ce qui vient d'être fait dans la démonstration ci-dessus. Soit $f \in \mathscr{C}^0_{2\pi}(\mathbb{R},\mathbb{R})$ telle que $\langle f, f\rangle = 0$. Cela signifie que f^2 est une fonction continue, positive, d'intégrale nulle sur $[0;2\pi]$ et donc que f est la fonction nulle sur $[0;2\pi]$. Comme f est 2π-périodique, cela démontre que f est la fonction nulle sur $\mathbb{R}$ tout entier. D'où le caractère défini.

On termine avec le produit scalaire le plus classique sur l'espace des polynômes réels.

Exemple fondamental

Sur $\mathbb{R}[X]$, on dispose du produit scalaire canonique, défini par

$$\left\langle \sum_{j\geqslant 0} a_j X^j \, , \sum_{j\geqslant 0} b_j X^j \right\rangle = \sum_{j\geqslant 0} a_j b_j,$$

où les sommes sont finies puisque les coefficients d'un polynôme stationnent à la valeur 0.

■ AQT ■

Le produit scalaire canonique de $\mathbb{R}[X]$ est loin d'être le seul digne d'intérêt sur cet espace. On peut citer le produit scalaire de Lagrange (vu au début de ce cours). On peut aussi utiliser le fait que les polynômes réels s'identifient aux fonctions polynomiales pour adapter le produit scalaire fonctionnel au cas de $\mathbb{R}[X]$. C'est ce qui est fait dans l'exemple suivant.

Exemples :

- Sur $\mathbb{R}[X]$, on a le produit scalaire suivant :

$$\langle P, Q\rangle = \int_0^{\pi/17} P(t)Q(t)\,\mathrm{d}t.$$

  Les propriétés de symétrie, de bilinéarité et de positivité sont évidentes. Pour le caractère défini, on utilise le fait que seul le polynôme nul possède une infinité de racines (je vous laisse écrire les choses en détail).

A.2. Norme et angles

a) Norme euclidienne

La donnée d'un produit scalaire permet de définir une norme (c'est l'objet de la définition ci-dessous). Nous verrons plus tard que la réciproque n'est pas vraie : certaines normes d'espaces vectoriels ne découlent pas d'un produit scalaire. Autrement dit, la structure préhilbertienne induit une structure d'espace vectoriel normé, mais le contraire n'est pas toujours vrai.

Définition 2

Soit $(E, \langle \cdot, \cdot \rangle)$ un espace préhilbertien réel. On appelle norme euclidienne associée au produit scalaire $\langle \cdot, \cdot \rangle$ l'application $\|\cdot\|$ définie par

$$\|\cdot\| \left\{ \begin{array}{ccc} E & \longrightarrow & \mathbb{R}_+ \\ x & \longmapsto & \|x\| = \sqrt{\langle x, x \rangle} \end{array} \right.$$

On dit qu'un vecteur x est normé ou unitaire lorsque $\|x\| = 1$.

Pour tout vecteur x non nul, il existe exactement deux vecteurs normés colinéaires à x : ce sont les vecteurs opposés y et $-y$, où $y = x / \|x\|$.

Notons immédiatement deux propriétés évidentes de la norme, sur lesquelles nous reviendrons. D'une part, la bilinéarité du produit scalaire implique que $\forall x \in E,\ \forall \lambda \in \mathbb{R},\ \ \|\lambda x\| = |\lambda| . \|x\|$ et, d'autre part, son caractère « défini » dit que $\forall x \in E,\ (\|x\| = 0) \Longrightarrow (x = 0_E)$.

Le lemme suivant donne la version euclidienne des fameuses « identités remarquables ». Ces formules sont très utiles en pratique.

Proposition 1

Soit $(E, \langle \cdot, \cdot \rangle)$ un espace préhilbertien réel. Pour tous $x, y \in E$, on a les identités remarquables euclidiennes suivantes :

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

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

■ Soient $x, y \in E$. Pour la première et la troisième formule, on a

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

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

La deuxième formule découle de la première en remplaçant y par $-y$. ■

Exemples :

- Soit $(E, \langle \cdot, \cdot \rangle)$ un espace préhilbertien réel et $x, y \in E$. En additionnant les formules de développement de $\|x + y\|^2$ et $\|x - y\|^2$, on obtient l'identité du parallélogramme suivante :

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

  Dans le plan, cette identité nous dit que « la somme des carrés des quatre côtés d'un parallélogramme est égale à la somme des carrés des deux diagonales ». Elle permet donc de calculer la longueur de la médiane d'un triangle connaissant les longueurs de ses trois côtés.

La norme euclidienne est définie à partir du produit scalaire. Réciproquement, si l'on connaît la norme euclidienne, on peut retrouver le produit scalaire à l'aide des formules ci-dessous.

Proposition 2

Soit $(E, \langle \cdot, \cdot \rangle)$ un espace préhilbertien réel. Pour tous vecteurs $x, y \in E$, on a les identités de polarisation suivantes :

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

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

■ AQT ■

Retenez cette idée : les formules de polarisation constituent l'outil privilégié pour établir des résultats relatifs au produit scalaire lorsque l'on possède seulement des propriétés sur la norme euclidienne associée.

Exemples :

- Pour tout $(x, y) \in \mathbb{R}^2$, on pose $N(x, y) = \sqrt{x^2 + 2xy + 3y^2}$. Déterminons, s'il existe, le produit scalaire sur $\mathbb{R}^2$ associée à la norme.

  Analyse : S'il existe, ce produit scalaire ne peut être que

  $$\langle (x, y), (x', y') \rangle = \frac{N(x+x', y+y')^2 - N(x-x', y-y')^2}{4} = xx' + xy' + yx' + 3yy',$$

  d'après l'identité de polarisation.

  Synthèse : On vérifie sans mal que l'application $\langle \cdot, \cdot \rangle : (\mathbb{R}^2)^2 \longrightarrow \mathbb{R}$, définie par $\langle (x, y), (x', y') \rangle = xx' + xy' + yx' + 3yy'$, est bilinéaire symétrique définie positive sur $\mathbb{R}^2$ et que N est la norme euclidienne associée.

b) *Inégalité de Cauchy-Schwarz*

Théorème 1

Soit $(E, \langle \cdot, \cdot \rangle)$ un espace préhilbertien réel. Pour tous $x, y \in E$, on a l'inégalité de Cauchy–Schwarz :

$$|\langle x, y\rangle| \leqslant \|x\| . \|y\| .$$

Il y a égalité dans cette inégalité si, et seulement si, x et y sont colinéaires.

■ Soient $x, y \in E$.

▷ Commençons par démontrer l'inégalité. Il est notable de remarquer que, pour cette partie de la preuve, nous allons utiliser le fait que le produit scalaire est une forme bilinéaire symétrique positive mais que nous n'utiliserons pas le caractère « défini » (même si cela simplifierait un peu les choses).

Considérons l'application $P : \mathbb{R} \longrightarrow \mathbb{R}$ définie par $\forall \lambda \in \mathbb{R},\ P(\lambda) = \|x + \lambda y\|^2$.

On note tout d'abord que P est une fonction positive sur $\mathbb{R}$.

Par ailleurs, pour tout $\lambda \in \mathbb{R}$, on a $P(\lambda) = \|x\|^2 + 2\langle x, y\rangle \lambda + \|y\|^2 \lambda^2$. On distingue alors deux cas.

- ∗ Supposons que $\|y\| = 0$. L'application P est alors une fonction affine. Comme elle doit rester positive sur $\mathbb{R}$ tout entier, cette fonction affine ne peut être que constante. On a donc $\langle x, y\rangle = 0$. Par conséquent, l'inégalité de Cauchy–Schwarz est évidente dans ce cas puisqu'elle est équivalente à $0 \leqslant 0$. On peut d'ailleurs noter que l'on est alors dans un cas d'égalité.
- ∗ Supposons que $\|y\| \neq 0$. L'application P est alors une fonction polynomiale du second degré. Comme elle est positive sur $\mathbb{R}$ tout entier, elle ne peut pas admettre deux racines réelles distinctes (sinon elle serait strictement négative entre ces deux racines). Par conséquent, son discriminant $\Delta = 4\langle x, y\rangle^2 - 4\|x\|^2 . \|y\|^2$ est négatif ou nul, c'est-à-dire $\langle x, y\rangle^2 \leqslant \|x\|^2 . \|y\|^2$. D'où le résultat en appliquant la racine carrée (c'est elle qui fait apparaître les valeurs absolues).

▷ Traitons ensuite le cas d'égalité. Cette fois, nous n'éviterons pas l'utilisation du caractère « défini ».

Nous venons de voir que, dans le cas où $\|y\| = 0$, on a bien égalité. La condition $\|y\| = 0$ signifie que $y = 0_E$ et l'on sait bien que 0_E est colinéaire à tout autre vecteur. Tout va bien dans ce cas !

Lorsque $\|y\| \neq 0$, la démonstration de l'inégalité nous dit qu'il y a égalité si, et seulement si, le discrimimant de P est nul, c'est-à-dire si, et seulement si, P admet une racine double $\lambda_0 \in \mathbb{R}$. C'est équivalent à dire que $\|x + \lambda_0 y\|^2 = 0$ ou encore que $x + \lambda_0 y = 0_E$. Cette égalité est équivalente à la colinéarité de x et y (car $y \neq 0_E$). Le cas d'égalité est donc également établi dans ce cas. ■

L'inégalité de Cauchy-Schwarz s'écrit indifféremment $|\langle x, y\rangle| \leqslant \|x\|.\|y\|$ ou $\langle x, y\rangle^2 \leqslant \|x\|^2.\|y\|^2$. Apprenez l'une ou l'autre de ces formulations, mais choisissez une fois pour toutes ! Lorsqu'on essaye de manipuler tantôt l'une, tantôt l'autre, on finit toujours par écrire une formule fausse !

Exemples :

- Pour le produit scalaire géométrique du plan ou de l'espace, l'inégalité de Cauchy–Schwarz $\overrightarrow{u} \cdot \overrightarrow{v} \leqslant \|\overrightarrow{u}\|\,\|\overrightarrow{v}\|$ découle de l'inégalité triviale : $|\cos(\overrightarrow{u}, \overrightarrow{v})| \leqslant 1$.
- Soient $(x_1, \dots, x_n), (y_1, \dots, y_n) \in \mathbb{R}^n$. Pour le produit scalaire canonique de $\mathbb{R}^n$, on a
$$\left|\sum_{k=1}^{n} x_k y_k\right| \leqslant \sqrt{\sum_{k=1}^{n} {x_k}^2}\,\sqrt{\sum_{k=1}^{n} {y_k}^2}.$$
- Soient f, g deux fonctions continues sur $[a; b]$ avec $a < b$. Sur $\mathscr{C}^0([a; b], \mathbb{R})$ muni de son produit scalaire canonique, l'inégalité de Cauchy–Schwarz donne
$$\left|\int_a^b fg\right| \leqslant \sqrt{\int_a^b f^2}\,\sqrt{\int_a^b g^2}.$$
- La covariance n'est certes pas un produit scalaire mais c'est une forme bilinéaire symétrique positive. L'inégalité de Cauchy–Schwarz s'applique donc (mais pas son cas d'égalité). Dans le cours sur les vecteurs aléatoires, il est effectivement démontré une inégalité de Cauchy–Schwarz pour la covariance, à savoir $|\mathrm{Cov}(X, Y)| \leqslant \sigma(X)\sigma(Y)$. C'est ce qui permet d'affirmer que le coefficient de corrélation $\rho(X, Y)$ reste compris entre -1 et 1.

c) *Angles non orientés*

Définition 3

Soient $(E, \langle\cdot,\cdot\rangle)$ un espace préhilbertien réel et x, y deux vecteurs non nuls de E. On appelle angle non orienté de x et y le nombre réel de l'intervalle $[0;\pi]$, noté $(\widehat{x,y})$, défini par

$$(\widehat{x,y}) = \arccos \frac{\langle x,y\rangle}{\|x\| \,.\, \|y\|}.$$

Lorsque $(\widehat{x,y})$ appartient à $]0;\pi/2[$, on dit que l'angle est aigu. Lorsque $(\widehat{x,y})$ appartient à $]\pi/2;\pi[$, on dit que l'angle est obtu. Lorsque $(\widehat{x,y})$ vaut 0 ou π, on dit que l'angle est plat. Enfin, lorsque $(\widehat{x,y})$ vaut $\pi/2$, on dit que l'angle est droit.

C'est bien connu, l'angle droit bout à 90° ;-)

■ Pour que cet angle existe bel et bien, il est nécessaire de savoir que $\langle x,y\rangle / (\|x\| \,.\, \|y\|)$ appartient à $[-1;1]$. C'est précidément ce que dit l'inégalité de Cauchy–Schwarz. ■

En conséquence de cette définition, on retrouve la formule $\langle x,y\rangle = \|x\| \,.\, \|y\| \,.\cos(\widehat{x,y})$. Mais attention, cette formule n'est pas la définition du produit scalaire à l'aide des normes et des angles, puisque ces deux notions ont précisément été introduites à partir du produit scalaire.

La définition de l'angle non orienté $(\widehat{x,y})$ associée à une identité de polarisation nous dit que

$$(\widehat{x,y}) = \arccos \frac{\|x+y\|^2 - \|x-y\|^2}{4\,\|x\| \,.\, \|y\|}.$$

Cette (affreuse) formule est amusante puisqu'elle signifie que l'on peut mesurer un angle avec... un double décimètre ! En effet, il suffit de connaître les normes des vecteurs x, y, $x+y$ et $x-y$ pour être capable de déterminer l'angle $(\widehat{x,y})$ (à condition d'avoir une calculatrice pour l'arc cosinus !).

Il est aisé de constater que, si $\lambda, \mu \in \mathbb{R}^*$, alors on a $(\widehat{\lambda x, \mu y}) = (\widehat{x,y})$ lorsque λ et μ sont de même signe et $(\widehat{\lambda x, \mu y}) = \pi - (\widehat{x,y})$ lorsque λ et μ sont de signes différents.

Les angles permettent d'orienter la colinéarité, comme l'explique la définition ci-dessous.

Définition 4

Soit $(E, \langle\cdot,\cdot\rangle)$ un espace préhilbertien réel. Deux vecteurs non nuls de E sont colinéaires si, et seulement si, l'angle $(\widehat{x,y})$ est plat (c'est-à-dire $(\widehat{x,y}) \in \{0;\pi\}$). Dans le cas où $(\widehat{x,y}) = 0$, on dit que x et y sont colinéaires de même sens et, dans le cas où $(\widehat{x,y}) = \pi$, on dit que x et y sont colinéaires de sens contraire.

■ AQT ■

Il est facile de voir que deux vecteurs non nuls x et y sont colinéaires de même sens lorsqu'il existe $\lambda > 0$ tel que $x = \lambda y$ et de sens contraire lorsqu'il existe $\lambda < 0$ tel que $x = \lambda y$. C'est pourquoi, lorsque deux vecteurs sont colinéaires de même sens (respectivement colinéaires de sens contraire), on dit aussi parfois qu'ils sont positivement liés (respectivement négativement liés).

d) *Propriétés de la norme euclidienne*

Énonçons maintenant une des conséquences principales de l'inégalité de Cauchy–Schwarz.

Théorème 2

Soit $(E, \langle\cdot,\cdot\rangle)$ un espace préhilbertien réel. Pour tous $x, y \in E$, on a l'inégalité de Minkowski :

$$\|x+y\| \leqslant \|x\| + \|y\|\,.$$

Il y égalité dans cette inégalité si, et seulement si, x et y sont colinéaires de même sens.

■ Soient $x, y \in E$.

▷ Commençons par démontrer l'inégalité. On a

$$\|x+y\|^2 = \|x\|^2 + 2\langle x,y\rangle + \|y\|^2 \leqslant \|x\|^2 + 2|\langle x,y\rangle| + \|y\|^2 \leqslant \|x\|^2 + 2\|x\|\,\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2,$$

où la première inégalité est évidente et la seconde provient de l'inégalité de Cauchy–Schwarz. Le résultat en découle puisque des nombres positifs sont rangés dans le même ordre que leurs carrés.

▷ Traitons ensuite le cas d'égalité. Pour avoir égalité dans l'inégalité de Minkowski, il faut et suffit qu'il y ait égalité dans les deux inégalités ci-dessus, c'est-à-dire qu'il faut et suffit que $\langle x,y\rangle \geqslant 0$ et qu'il y ait égalité dans l'inégalité de Cauchy–Schwarz. Autrement dit, il faut et il suffit que x et y soient colinéaires et $\langle x,y\rangle \geqslant 0$. C'est exactement dire que x et y sont colinéaires de même sens. ■

Exemples :

- Soient $(x_1, \ldots, x_n), (y_1, \ldots, y_n) \in \mathbb{R}^n$. Pour le produit scalaire canonique de $\mathbb{R}^n$, l'inégalité de Minkowski donne

$$\sqrt{\sum_{k=1}^{n}(x_k + y_k)^2} \leqslant \sqrt{\sum_{k=1}^{n} {x_k}^2} + \sqrt{\sum_{k=1}^{n} {y_k}^2}.$$

On peut alors dresser la liste des propriétés de la norme.

Proposition 3

Soit $(E, \langle\cdot,\cdot\rangle)$ un espace préhilbertien réel. On a

(i)	$\forall x \in E, \quad \|x\| \geqslant 0$	(positivité) ;
(ii)	$\forall x \in E, \quad \big(\|x\| = 0\big) \Longrightarrow (x = 0_E)$	(séparation) ;
(iii)	$\forall x \in E, \quad \forall \lambda \in \mathbb{R}, \quad \|\lambda x\| = |\lambda|\,\|x\|$	(positive homogénéité) ;
(iv)	$\forall x, y \in E, \quad \|x+y\| \leqslant \|x\| + \|y\|$	(inégalité triangulaire).

■ Les propriétés (i), (ii) et (iii) sont évidentes. La propriété (iv) est l'inégalité de Minkowski. ■

Ces propriétés sont en fait celles qui permettent de définir le concept de norme sur un $\mathbb{R}$-espace vectoriel. Par conséquent, cet énoncé signifie que la norme euclidienne associée à un produit scalaire est une norme ou encore qu'un espace préhilbertien réel est un espace vectoriel normé.

Corollaire 1

Soit $(E, \langle\cdot,\cdot\rangle)$ un espace préhilbertien réel. Pour tout $x, y \in E$, on peut préciser l'inégalité triangulaire en lui adjoignant l'inégalité triangulaire inversée, ce qui donne

$$\big|\,\|x\| - \|y\|\,\big| \leqslant \|x+y\| \leqslant \|x\| + \|y\|\,.$$

■ Soient $x, y \in E$. On a $\|x\| = \|x+y-y\| \leqslant \|x+y\| + \|-y\| = \|x+y\| + \|y\|$, d'où $\|x\| - \|y\| \leqslant \|x+y\|$. En inversant les rôles de x et y, on a $\|y\| - \|x\| \leqslant \|x+y\|$. D'où le résultat. ■

e) Distance euclidienne

Définition 5

Soit $(E, \langle \cdot, \cdot \rangle)$ un espace préhilbertien réel. On appelle distance euclidienne associée au produit scalaire $\langle \cdot, \cdot \rangle$ l'application d définie par

$$\mathrm{d} \left\{ \begin{array}{ccc} E^2 & \longrightarrow & \mathbb{R}_+ \\ (x, y) & \longmapsto & \mathrm{d}(x, y) = \|x - y\| \end{array} \right.$$

Les résultats concernant la norme euclidienne permettent d'énoncer les propriétés de cette distance euclidienne.

Proposition 4

Soit $(E, \langle \cdot, \cdot \rangle)$ un espace préhilbertien réel. On a

(i) $\forall x, y \in E, \quad \mathrm{d}(x, y) \geqslant 0$ (positivité) ;

(ii) $\forall x, y \in E, \quad \mathrm{d}(x, y) = \mathrm{d}(y, x)$ (symétrie) ;

(iii) $\forall x \in E, \quad \big(\mathrm{d}(x, y) = 0\big) \Longrightarrow (x = y)$ (séparation) ;

(iv) $\forall x, y, z \in E, \quad \mathrm{d}(x, z) \leqslant \mathrm{d}(x, y) + \mathrm{d}(y, z)$ (inégalité triangulaire).

■ AQT ■

Là encore, en utilisant ces propriétés, on peut définir une notion plus générale de distance (ou métrique) sur un ensemble. Par conséquent, cet énoncé signifie que la distance euclidienne associée à un produit scalaire est une distance ou encore qu'un espace préhilbertien réel est un espace métrique.

Corollaire 2

Soit $(E, \langle \cdot, \cdot \rangle)$ un espace préhilbertien réel. Pour tout $x, y, z \in E$, on peut préciser l'inégalité triangulaire en lui adjoignant l'inégalité triangulaire inversée, ce qui donne

$$\big|\mathrm{d}(x, y) - \mathrm{d}(y, z)\big| \leqslant \mathrm{d}(x, z) \leqslant \mathrm{d}(x, y) + \mathrm{d}(y, z).$$

2 h 20

■ AQT ■

A.3. Orthogonalité

a) Orthogonalité de vecteurs

Définition 6

Soit $(E, \langle\cdot,\cdot\rangle)$ un espace préhilbertien réel. On dit que deux vecteurs x et y de E sont orthogonaux, et l'on note $x \perp y$, lorsque $\langle x, y\rangle = 0$.

Dire que deux vecteurs non nuls x et y sont orthogonaux revient à dire que $(\widehat{x, y}) = \pi/2$.

Le vecteur nul, quant à lui, est orthogonal à tout autre vecteur, mais cela n'a pas de traduction angulaire.

Dans le cadre général des espaces préhilbertiens réels, l'orthogonalité est donc définie à partir du produit scalaire. Cela implique en particulier que si l'on change de produit scalaire, l'orthogonalité change également !

L'orthogonalité est une piètre relation. Elle est bien symétrique (si $x \perp y$ alors $y \perp x$ d'après la symétrie du produit scalaire). En revanche, l'orthogonalité n'est ni réflexive (le seul vecteur orthogonal à lui-même est le vecteur nul), ni transitive (si $x \perp y$ et $y \perp z$, on ne peut en général rien dire de x et z).

Exemples :

- Dans $\mathbb{R}^n$ muni de son produit scalaire canonique, les vecteurs de la base canonique sont orthogonaux deux à deux.
- Dans l'espace $\mathscr{C}^0_{2\pi}(\mathbb{R},\mathbb{R})$ des fonctions continues 2π-périodiques sur $\mathbb{R}$ muni du produit scalaire canonique $\langle f, g\rangle = \dfrac{1}{\pi}\displaystyle\int_0^{2\pi} f(t)g(t)\,\mathrm{d}t$, les fonctions sin et cos sont orthogonales. En effet, on a

$$\langle \cos, \sin\rangle = \frac{1}{\pi}\int_0^{2\pi} \cos(t)\sin(t)\,\mathrm{d}t = \frac{1}{\pi}\int_0^{2\pi} \frac{\sin(2t)}{2}\,\mathrm{d}t = \frac{1}{\pi}\left[\frac{-\cos(2t)}{4}\right]_0^{2\pi} = 0.$$

On peut caractériser à l'aide des normes l'orthogonalité de deux vecteurs : on retrouve le célèbre théorème de Pythagore.

Théorème 3

Soit $(E, \langle\cdot,\cdot\rangle)$ un espace préhilbertien réel. Pour tous $x, y \in E$, on a

$$(x \perp y) \iff (\|x+y\|^2 = \|x\|^2 + \|y\|^2).$$

« Le carré de l'hypothénuse est égal, si je ne m'abuse, à la somme des carrés des deux autres côtés ».

■ C'est la conséquence directe de l'égalité $\|x+y\|^2 = \|x\|^2 + 2\langle x, y\rangle + \|y\|^2$. ■

Exemples :

- On munit $\mathscr{C}^0_{2\pi}(\mathbb{R},\mathbb{R})$ du produit scalaire donné dans l'exemple précédent. L'orthogonalité de sin et cos donne alors

$$\frac{1}{\pi}\int_0^{2\pi} (\sin t + \cos t)^2\,\mathrm{d}t = \frac{1}{\pi}\int_0^{2\pi} (\sin t)^2\,\mathrm{d}t + \frac{1}{\pi}\int_0^{2\pi} (\cos t)^2\,\mathrm{d}t.$$

On peut généraliser l'orthogonalité à plus de deux vecteurs.

Définition 7

Soit $(E, \langle \cdot, \cdot \rangle)$ un espace préhilbertien réel. On dit qu'une famille $(x_i)_{i \in I}$ de vecteurs de E est orthogonale lorsque ses éléments sont deux à deux orthogonaux, c'est-à-dire

$$\forall i, j \in I, \qquad (i \neq j) \Longrightarrow (\langle x_i, x_j \rangle = 0).$$

On précise que cette famille est orthonormale lorsqu'elle est orthogonale et que ses éléments sont unitaires, c'est-à-dire $\forall i, j \in I,\ \langle x_i, x_j \rangle = \delta_{i,j}$ où $\delta_{i,j}$ désigne le symbole de Kronecker.

Dans l'ancien temps, on disait *orthonormée* pour orthonormale. Mais maintenant, il parait qu'il faut plus...

Le vecteur nul a tout à fait le droit d'apparaître dans une famille orthogonale puisqu'il est orthogonal à tout autre vecteur. En revanche, il est absent dans une famille orthonormale puisqu'il n'est évidemment pas de norme 1.

Exemples :

- Dans $\mathbb{R}[X]$ muni du produit scalaire canonique, la famille des monômes est orthonormale.
- On munit $\mathscr{C}^0_{2\pi}(\mathbb{R}, \mathbb{R})$ de son produit scalaire canonique vu à la page précédente. Pour tout $k \in \mathbb{N}^*$, on considère les éléments c_k et s_k de $\mathscr{C}^0_{2\pi}(\mathbb{R}, \mathbb{R})$ définie par

$$c_k : x \longmapsto \cos(kx) \qquad \text{et} \qquad s_k : x \longmapsto \sin(kx).$$

Vérifions alors que la famille obtenue en juxtaposant les familles $(c_k)_{k \in \mathbb{N}^*}$ et $(s_k)_{k \in \mathbb{N}^*}$ est une famille orthonormale de $\mathscr{C}^0_{2\pi}(\mathbb{R}, \mathbb{R})$.

Pour tous $k, \ell \in \mathbb{N}^*$ tels que $k \neq \ell$, on a

$$\langle c_k, c_\ell \rangle = \frac{1}{\pi} \int_0^{2\pi} \cos(kx)\cos(\ell x)\,\mathrm{d}x = \frac{1}{\pi} \int_0^{2\pi} \frac{\cos((k-\ell)x) + \cos((k+\ell)x)}{2}\,\mathrm{d}x$$

donc

$$\langle c_k, c_\ell \rangle = \frac{1}{2\pi} \left[\frac{\sin((k-\ell)x)}{k-\ell} + \frac{\sin((k+\ell)x)}{k+\ell} \right]_0^{2\pi} = 0$$

et

$$\|c_k\|^2 = \frac{1}{2\pi} \left[x + \frac{\sin(2kx)}{2k} \right]_0^{2\pi} = 1.$$

De même, pour tous $k, \ell \in \mathbb{N}^*$ tels que $k \neq \ell$, on a

$$\langle s_k, s_\ell \rangle = \frac{1}{\pi} \int_0^{2\pi} \sin(kx)\sin(\ell x)\,\mathrm{d}x = \frac{1}{\pi} \int_0^{2\pi} \frac{\cos((k-\ell)x) - \cos((k+\ell)x)}{2}\,\mathrm{d}x$$

donc

$$\langle s_k, s_\ell \rangle = \frac{1}{2\pi} \left[\frac{\sin((k-\ell)x)}{k-\ell} - \frac{\sin((k+\ell)x)}{k+\ell} \right]_0^{2\pi} = 0$$

et

$$\|s_k\|^2 = \frac{1}{2\pi} \left[x - \frac{\sin(2kx)}{2k} \right]_0^{2\pi} = 1.$$

Enfin, pour tous $k, \ell \in \mathbb{N}^*$, on a

$$\langle c_k, s_\ell \rangle = \frac{1}{\pi} \int_0^{2\pi} \cos(kx)\sin(\ell x)\,\mathrm{d}x = \frac{1}{\pi} \int_{-\pi}^{\pi} \cos(kx)\sin(\ell x)\,\mathrm{d}x = 0,$$

où la première égalité provient du fait que l'intégrale d'une fonction 2π-périodique est la même sur toute période et la seconde égalité découle de l'intégration d'une fonction impaire sur un intervalle centré en 0.

Le résultat suivant généralise, dans le cas des familles orthogonales finies, l'une des deux implications du théorème de Pythagore.

Proposition 5

Soit $(E, \langle\cdot,\cdot\rangle)$ un espace préhilbertien réel. Si $(x_1, \ldots, x_p)$ est une famille orthogonale finie de vecteurs de E, alors on a la relation de Pythagore :

$$\left\| \sum_{i=1}^{p} x_i \right\|^2 = \sum_{i=1}^{p} \|x_i\|^2 .$$

■ On a

$$\left\| \sum_{i=1}^{p} x_i \right\|^2 = \left\langle \sum_{i=1}^{p} x_i, \sum_{j=1}^{p} x_j \right\rangle = \sum_{1 \leqslant i,j \leqslant p} \langle x_i, x_j \rangle = \sum_{1 \leqslant i \leqslant p} \langle x_i, x_i \rangle = \sum_{i=1}^{p} \|x_i\|^2 .$$

car $\langle x_i, x_j \rangle = 0$ lorsque $i \neq j$. ■

Lorsque $p \geqslant 3$, la réciproque est fausse. Par exemple, dans $\mathbb{R}^2$ muni de son produit scalaire usuel, les vecteurs $x_1 = (1,2)$, $x_2 = (0,2)$ et $x_3 = (0,-1)$ vérifient la relation de Pythagore sans constituer, pour autant, une famille orthogonale (à vérifier !).

On poursuit avec un résultat fondamental : l'orthogonalité est source de liberté ...

Théorème 4

Dans un espace préhilbertien réel, une famille orthogonale dont aucun des vecteurs n'est nul est libre. En particulier, une famille orthonormale est toujours libre.

■ Soit $(x_i)_{i\in I}$ une famille orthogonale de vecteurs non nuls de E. Considérons la donnée d'un nombre fini d'éléments $x_{i_1}, \ldots, x_{i_p}$ de cette famille tels que $\lambda_{i_1} x_{i_1} + \cdots + \lambda_{i_p} x_{i_p} = 0_E$, où $\lambda_{i_1}, \ldots, \lambda_{i_p} \in \mathbb{R}$. En prenant le produit scalaire des deux membres de cette égalité par x_{i_k}, où k est un entier arbitraire de $[\![1;p]\!]$, on obtient $0 = \left\langle x_{i_k}, \lambda_{i_1} x_{i_1} + \cdots + \lambda_{i_p} x_{i_p} \right\rangle = \lambda_{i_1} \left\langle x_{i_1}, x_{i_k} \right\rangle + \cdots + \lambda_{i_p} \left\langle x_{i_p}, x_{i_k} \right\rangle = \lambda_{i_k} \|x_{i_k}\|^2$. On en déduit que $\lambda_{i_k} = 0$ puisque $x_{i_k} \neq 0_E$. La famille $(x_i)_{i\in I}$ est donc libre.

Le cas particulier d'une famille orthonormale est évident puisqu'un vecteur unitaire n'est pas nul. ■

N'oubliez pas la non nullité des vecteurs de votre famille orthogonale avant de conclure qu'elle est libre. Les correcteurs sont généralement très chatouilleux sur cette hypothèse.

La réciproque est évidemment fausse : une famille libre n'est pas forcément orthogonale. Cependant, le procédé d'orthonormalisation de Gram–Schmidt décrit ci-dessous, nous explique que les vecteurs d'une famille libre peuvent être « redressés » afin d'obtenir une famille orthonormale.

Théorème 5

Soit $(E, \langle\cdot,\cdot\rangle)$ un espace préhilbertien réel. Si $(v_1, \ldots, v_m)$ désigne une famille libre (finie) de vecteurs de E, il existe une unique famille de vecteurs $(e_1, \ldots, e_m)$ telle que :

(i) $(e_1, \ldots, e_m)$ est orthonormale ;

(ii) $\forall p \in [\![1;m]\!]$, $\mathrm{Vect}(e_1, \ldots, e_p) = \mathrm{Vect}(v_1, \ldots, v_p)$;

(iii) $\forall p \in [\![1;m]\!]$, $\langle e_p, v_p \rangle > 0$.

On dit que $(e_1, \ldots, e_m)$ est l'orthonormalisée de Gram–Schmidt de $(v_1, \ldots, v_m)$.

■ Cf Annexe. ■

L'orthonormalisée d'une famille libre est une famille orthonormale, donc libre. Ainsi, la liberté n'est pas perdue. C'est le moins que l'on puisse attendre d'un tel procédé !

La propriété (iii) du théorème précédent est souvent omise. Elle permet d'assurer l'unicité de la famille, mais cette unicité est toute relative : si l'on effectue le procédé de Gram–Schmidt en égrainant les vecteurs dans un ordre différent, on trouve une autre famille orthonormale.

On notera que la matrice de passage de $(v_1, \dots, v_m)$ à $(e_1, \dots, e_m)$ est une matrice triangulaire supérieure à coefficients diagonaux strictement positifs. Cela assure que les deux familles ont la même orientation.

Au delà du résultat théorique énoncé par ce théorème, il faut constater que la démonstration d'icelui (présentée en annexe de ce cours) fournit un procédé concret d'orthonormalisation. Cependant, dans la pratique, on ne procède pas exactement comme on l'a fait dans cette preuve mais en suivant l'algorithme décrit ci-dessous.

Orthonormalisation à la Gram–Schmidt

Soit $(v_1, \dots, v_m)$ une famille libre finie de E. On procède à l'orthonormalisation en deux étapes : on commence par orthogonaliser puis on normalise.

- Orthogonalisation
  - On pose $e_1 = v_1$. Et si e_1 n'est pas à votre goût, vous pouvez prendre n'importe quel vecteur non nul colinéaire à e_1.
  - On cherche e_2 sous la forme $e_2 = v_2 - \lambda e_1$ et on détermine λ pour que l'on ait $e_2 \perp e_1$. Concrètement, on écrit que

    $$(e_2 \perp e_1) \iff (\langle e_2, e_1 \rangle = 0) \iff \big(\langle v_2 - \lambda e_1, e_1 \rangle = 0\big) \iff \left(\lambda = \frac{\langle v_2, e_1 \rangle}{\|e_1\|^2}\right)$$

    où le quotient est licite puisque $e_1 \neq 0$. On obtient

    $$e_2 = v_2 - \frac{\langle v_2, e_1 \rangle}{\|e_1\|^2} e_1.$$

    Pour simplifier, on peut à loisir remplacer e_2 par l'un de ses dilatés non nul.
  - ...
  - À la k-ème étape, on recherche e_k sous la forme $e_k = v_k - (\lambda_1 e_1 + \cdots + \lambda_{k-1} e_{k-1})$, où λ_1 est un nombre réel déterminé par la condition $e_k \perp e_1$, λ_2 est un nombre réel déterminé par la condition $e_k \perp e_2$, ..., λ_{k-1} est un nombre réel déterminé par la condition $e_k \perp e_{k-1}$.
    Si besoin est, on peut choisir, à la place de e_k, un vecteur non nul colinéaire à e_k.
  - ...
  - On s'arrête à l'étape m.
- Normalisation
  Pour tout $k \in [\![1; m]\!]$, on remplace e_k par $\dfrac{e_k}{\|e_k\|}$ pour obtenir une famille orthonormale.

On peut retenir de cet algorithme d'orthogonalisation que, à chaque étape, e_k est construit à partir de v_k en retirant à v_k ses composantes sur $v_1, \dots, v_{k-1}$.

Il faut noter que si les premiers vecteurs de la famille à orthonormaliser forment déjà une famille orthonormale, alors l'algorithme de Gram-Schmidt ne les modifie pas !

On peut étendre la méthode de Gram–Schmidt aux familles libres dénombrables.

Voyons sur un exemple comment mettre en place le procédé d'orthonormalisation.

Exemples :

- Dans $\mathbb{R}^3$ muni de son produit scalaire canonique, orthonormalisons la base

$$u_1 = (1, 1, 0), \qquad u_2 = (1, 0, 1), \qquad u_3 = (0, 1, 1).$$

La famille (u_1, u_2, u_3) est bien une base puisque

$$\det\big(\mathrm{Mat}_{\mathscr{B}_{\mathrm{can}}}(u_1, u_2, u_3)\big) = \begin{vmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{vmatrix} = -2 \neq 0.$$

▷ On pose

$$e_1 = u_1 = (1, 1, 0).$$

▷ On pose ensuite $e_2 = u_2 - \lambda e_1$ et l'on recherche $\lambda \in \mathbb{R}$ telle que $e_2 \perp e_1$. On a

$$\begin{aligned} (e_2 \perp e_1) &\iff (\langle u_2 - \lambda e_1, e_1 \rangle = 0) \\ &\iff (\langle u_2, e_1 \rangle - \lambda \|e_1\|^2 = 0) \\ &\iff \Big(\lambda = \frac{\langle u_2, e_1 \rangle}{\|e_1\|^2} = \frac{1}{2}\Big), \end{aligned}$$

d'où

$$e_2 = u_2 - \frac{1}{2} e_1 = \Big(\frac{1}{2}, -\frac{1}{2}, 1\Big).$$

On préfère prendre

$$e_2 = (1, -1, 2).$$

▷ On pose $e_3 = u_3 - \mu e_1 - \nu e_2$ et l'on recherche $\mu, \nu \in \mathbb{R}$ telles que $e_3 \perp e_1$ et $e_3 \perp e_2$. On a

$$\begin{aligned} (e_3 \perp e_1) &\iff (\langle u_3 - \mu e_1 - \nu e_2, e_1 \rangle = 0) \\ &\iff (\langle u_3, e_1 \rangle - \mu \|e_1\|^2 = 0) \\ &\iff \Big(\mu = \frac{\langle u_3, e_1 \rangle}{\|e_1\|^2} = \frac{1}{2}\Big) \end{aligned}$$

et

$$\begin{aligned} (e_3 \perp e_2) &\iff (\langle u_3 - \mu e_1 - \nu e_2, e_2 \rangle = 0) \\ &\iff (\langle u_3, e_2 \rangle - \nu \|e_2\|^2 = 0) \\ &\iff \Big(\nu = \frac{\langle u_3, e_2 \rangle}{\|e_2\|^2} = \frac{1}{6}\Big), \end{aligned}$$

d'où

$$e_3 = u_3 - \mu e_1 - \nu e_2 = \Big(-\frac{2}{3}, \frac{2}{3}, \frac{2}{3}\Big).$$

On peut prendre

$$e_3 = (-1, 1, 1).$$

▷ En normalisant, on obtient finalement

$$e_1 = \Big(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\Big), \qquad e_2 = \Big(\frac{1}{\sqrt{6}}, -\frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}\Big) \qquad \text{et} \qquad e_3 = \Big(-\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\Big).$$

b) Orthogonalité de parties

Définition 8

Soient $(E, \langle \cdot, \cdot \rangle)$ un espace préhilbertien réel et A, B deux parties de E. On dit que A et B sont orthogonales, et l'on note $A \perp B$, lorsque

$$\forall a \in A, \quad \forall b \in B, \qquad \langle a, b \rangle = 0.$$

En particulier, si x est un élément de E et si A est une partie de E, on dit que x est orthogonal à A, et on note $x \perp A$, lorsque le singleton $\{x\}$ est orthogonal à A, c'est-à-dire lorsque $\forall a \in A, \ \langle x, a \rangle = 0$.

Si la partie A est orthogonale à la partie B, alors toute sous-partie de A est elle-même orthogonale à toute sous-partie de B.

Il faut bien noter que la notion d'orthogonalité pour des ensembles n'est pas limitée aux sous-espaces vectoriels. Les parties A et B ont tout à fait le droit de ne pas avoir de structure vectorielle.

En particulier, cette définition permet de parler d'orthogonalité pour des sous-espaces affines de E. Il n'est alors pas très difficile de constater que deux sous-espaces affines sont orthogonaux si, et seulement si, leurs directions sont orthogonales. De même, on peut constater que si $\mathscr{F}$ est un sous-espace affine (faiblement) parallèle à un sous-espace affine $\mathscr{G}$ et si $\mathscr{G}$ est orthogonal à un troisième sous-espace affine $\mathscr{H}$, alors $\mathscr{F}$ est orthogonal à $\mathscr{H}$.

Pour qu'un vecteur x soit orthogonal à un sous-espace F, il faut et il suffit qu'il soit orthogonal à tous les vecteurs d'une base de ce sous-espace. Cette remarque toute bête est parfois bien utile !

Exemples :

- 0_E est orthogonal à toute partie de E.
- Dans $\mathbb{R}^n$ muni de son produit scalaire canonique, la droite vectorielle $\mathrm{Vect}\{(a_1, \ldots, a_n)\}$ et l'hyperplan d'équation $a_1 x_1 + \cdots + a_n x_n = 0$ sont orthogonaux.
- Dans $\mathscr{C}^0([-1;1], \mathbb{R})$ muni du produit scalaire canonique, le sous-espace $\mathscr{P}$ des fonctions paires est orthogonal au sous-espace $\mathscr{I}$ des fonctions impaires. En effet, si $p \in \mathscr{P}$ et $i \in \mathscr{I}$, on a
$$\langle p, i \rangle = \int_{-1}^{1} p(t) i(t) \, \mathrm{d}t = 0$$
car on intègre la fonction impaire pi sur un intervalle centré en 0. On a donc bien $\mathscr{P} \perp \mathscr{I}$.

Le vecteur nul étant le seul vecteur orthogonal à lui-même, l'intersection de deux parties orthogonales est nécessairement contenue dans $\{0_E\}$. Cette propriété de quasi-évitement des parties orthogonales est l'objet de la proposition suivante.

Proposition 6

Soient $(E, \langle \cdot, \cdot \rangle)$ un espace préhilbertien réel et A, B deux parties de E. On a la propriété, dite d'incompatibilité orthogonale, suivante :

$$(A \perp B) \quad \Longrightarrow \quad \big(A \cap B = \varnothing \text{ ou } A \cap B = \{0_E\}\big).$$

■ Supposons que $A \perp B$. Soit $x \in A \cap B$. Alors $x \in A$ et $x \in B$, ce qui implique que x est orthogonal à lui même. Autrement dit, on a $\langle x, x \rangle = 0$ ou encore $\|x\| = 0$. Cela donne $x = 0_E$ d'après la propriété de séparation de la norme. On a donc $A \cap B \subset \{0_E\}$, ce qui correspond au résultat attendu. ■

L'énoncé suivant explique que l'orthogonalité de sous-espaces entraine que la somme de ces sous-espaces est nécessairement directe. Il faut évidemment faire l'analogie entre ce résultat et celui que nous avions obtenu au théorème 4 : l'orthogonalité de vecteurs non nuls implique leur liberté.

Proposition 7

Soit $(E, \langle \cdot, \cdot \rangle)$ un espace préhilbertien réel.

Soient F, G deux sous-espaces vectoriels de E. Si F et G sont orthogonaux, alors ils sont en somme directe. Dans ce cas, la somme $F+G$ est notée $F \overset{\perp}{\oplus} G$ et l'on dit que F et G sont en somme directe orthogonale. Si, de plus, on a $F \overset{\perp}{\oplus} G = E$, on dit que F et G sont des supplémentaires orthogonaux dans E.

Plus généralement, considérons $(F_i)_{i\in I}$ une famille de sous-espaces vectoriels de E. Si les F_i sont deux à deux orthogonaux, alors ils sont en somme directe. Dans ce cas, la somme $\sum_{i\in I} F_i$ est notée $\overset{\perp}{\bigoplus}_{i\in I} F_i$ et on dit que les F_i sont en somme directe orthogonale.

■ Le cas de deux sous-espaces est une conséquence de la proposition 6.

Soit $(F_i)_{i\in I}$ une famille de sous-espaces vectoriels de E deux à deux orthogonaux. Considérons la donnée d'un nombre fini d'éléments $F_{i_1}, \ldots, F_{i_p}$ de cette famille ainsi que des vecteurs $x_{i_1} \in F_{i_1}, \ldots, x_{i_p} \in F_{i_p}$ tels que $x_{i_1} + \cdots + x_{i_p} = 0_E$. En prenant le produit scalaire des deux membres de cette égalité par x_{i_k}, où k est un entier arbitraire de $[\![1;p]\!]$, il vient $0 = \left\langle x_{i_k}, x_{i_1} + \cdots + x_{i_p} \right\rangle = \langle x_{i_1}, x_{i_k} \rangle + \cdots + \left\langle x_{i_p}, x_{i_k} \right\rangle = \|x_{i_k}\|^2$. On en déduit que $x_{i_k} = 0_E$. Les F_i sont donc en somme directe. ■

On retiendra, en particulier, que pour démontrer que des sous-espaces sont en somme directe orthogonale, il suffit de prouver qu'ils sont orthogonaux. Le caractère direct de la somme vous est alors gracieusement offert ! 4 h 20

c) *Orthogonal d'une partie*

Définition 9

Soient $(E, \langle\cdot,\cdot\rangle)$ un espace préhilbertien réel et A une partie de E. On appelle orthogonal de A la partie de E, noté $A^\perp$, qui contient tous les vecteurs de E qui sont orthogonaux à A, c'est-à-dire

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

En particulier, si a est un vecteur de E, on note $a^\perp$ l'ensemble $\{a\}^\perp$ de tous les vecteurs orthogonaux à a.

Évidemment, A et $A^\perp$ sont deux parties orthogonales (c'est-à-dire $A \perp A^\perp$).
Par conséquent, la proposition 6 nous dit que $A \cap A^\perp = \varnothing$ ou $A \cap A^\perp = \{0_E\}$.

On notera que deux parties A et B sont orthogonales si, et seulement si, $A \subset B^\perp$ ou, ce qui revient au même, $B \subset A^\perp$.

Pour qu'un vecteur x appartienne à l'orthogonal $F^\perp$ d'un sous-espace vectoriel F de E, il faut et il suffit que x soit orthogonal à tous les vecteurs d'une base de F. Encore un petit résultat tout simple mais bien utile !

Exemples :

- Il est clair que $\varnothing^\perp = E$.
- Il est tout aussi clair que ${0_E}^\perp = \{0_E\}^\perp = E$.
- On a aussi $E^\perp = \{0_E\}$ mais là il faut réfléchir un peu.

  L'inclusion $\supset$ est évidente.

  Traitons l'inclusion contraire $\subset$. Soit $x \in E^\perp$. Cela signifie que x est orthogonal à tout autre vecteur de E. Mézalors, x est orthogonal à lui-même, c'est-à-dire $\langle x, x\rangle = 0$ ou encore $\|x\| = 0$. La propriété de séparation de la norme dit alors que $x = 0_E$. Cela démontre l'inclusion $\subset$.
- Si a est un vecteur non nul de E, l'orthogonal $a^\perp$ de a est un hyperplan puisque c'est le noyau de la forme linéaire non nulle $\ell_a : x \longmapsto \langle x, a\rangle$.

Voici le résultat essentiel sur l'orthogonal d'une partie.

Proposition 8

Soient $(E, \langle\cdot,\cdot\rangle)$ un espace préhilbertien réel et A une partie de E. L'orthogonal $A^\perp$ de A est un sous-espace vectoriel de E.

■ On peut le démontrer « à la main ». En effet, il est clair que $A^\perp$ contient le vecteur nul puisque celui-ci est orthogonal à tout vecteur. De plus, si x et y sont dans $A^\perp$, alors, pour tous $\lambda, \mu \in \mathbb{R}$, on a $\forall a \in A,\ \ \langle \lambda x + \mu y, a\rangle = \lambda\,\langle x, a\rangle + \mu\,\langle y, a\rangle = \lambda 0 + \mu 0 = 0$, ce qui démontre que $\lambda x + \mu y \in A^\perp$.

On peut aussi être plus subtil. Pour cela, pour tout $a \in A \setminus \{0_E\}$, on considère à nouveau la forme linéaire non nulle $\ell_a : x \longmapsto \langle x, a\rangle$ et on remarque que $A^\perp = \bigcap_{a \in A\setminus\{0\}} \operatorname{Ker} \ell_a$, ce qui démontre que $A^\perp$ est un sous-espace vectoriel de E en tant qu'intersection d'hyperplans de E. ■

Ce résultat peut paraître surprenant : il énonce que, pour n'importe quelle partie A de E (y compris une partie qui n'a pas de structure vectorielle), l'orthogonal $A^\perp$ est automatiquement un espace vectoriel !

En particulier, lorsque F est un sous-espace vectoriel de E, la proposition ci-dessus nous dit que F et $F^\perp$ sont deux sous-espaces orthogonaux. La proposition 6 permet alors de préciser que F et $F^\perp$ sont en somme directe orthogonale.

L'énoncé suivant énonce la décroissance du passage à l'orthogonal.

Proposition 9

Soient $(E, \langle\cdot,\cdot\rangle)$ un espace préhilbertien réel. Le passage à l'orthogonal est une application décroissante. Autrement dit, si A, B sont deux parties de E telles que $A \subset B$, alors $B^\perp \subset A^\perp$.

■ Soient $A, B \in \mathscr{P}(E)$ telles que $A \subset B$. Soit $b' \in B^\perp$. On a alors $\forall b \in B,\ \langle b', b\rangle = 0$. Comme $A \subset B$, on a, en particulier, $\forall a \in A,\ \langle b', a\rangle = 0$, ce qui signifie que $b' \in A^\perp$. Donc $B^\perp \subset A^\perp$. ■

La décroissance du passage à l'orthogonal peut, par exemple, servir pour établir la propriété suivante.

Corollaire 3

Soient $(E, \langle\cdot,\cdot\rangle)$ un espace préhilbertien réel et A une partie de E. On a $A^\perp = \big(\mathrm{Vect}(A)\big)^\perp$.

■ On procède par double inclusion.

- $\supset$ Comme $A \subset \mathrm{Vect}(A)$, la proposition précédente dit que $A^\perp \supset \big(\mathrm{Vect}(A)\big)^\perp$.
- $\subset$ Tout vecteur qui est orthogonal à A est aussi orthogonal à toute combinaison linéaire de vecteurs de A (par bilinéarité du produit scalaire). Donc $\big(\mathrm{Vect}(A)\big)^\perp \supset A^\perp$. ■

On termine ce paragraphe en parlant du biorthogonal.

Définition 10

Soient $(E, \langle\cdot,\cdot\rangle)$ un espace préhilbertien réel et A une partie de E. On appelle biorthogonal de A la partie, notée $A^{\perp\perp}$, égale à l'orthogonal de l'orthogonal de A, c'est-à-dire $A^{\perp\perp} = (A^\perp)^\perp$.

Le résultat élémentaire mais fondamental sur le biorthogonal est l'inclusion donnée dans l'énoncé suivant.

Proposition 10

Soient $(E, \langle\cdot,\cdot\rangle)$ un espace préhilbertien réel et A une partie de E. On a $A \subset A^{\perp\perp}$.

■ Soit $a \in A$. Pour appartenir à $A^{\perp\perp}$, il faut que a vérifie $\forall a' \in A^\perp,\ \langle a, a'\rangle = 0$. Or, cette propriété est satisfaite du fait de la définition de $A^\perp$. Donc $a \in A^{\perp\perp}$, ce qui démontre que $A \subset A^{\perp\perp}$. ■

L'inclusion $A \subset A^{\perp\perp}$ peut ne pas être une égalité car A n'est pas nécessairement un sous-espace alors que $A^{\perp\perp}$ en est un. Et même dans le cas où F est un sous-espace, il est tout à fait possible que $F^{\perp\perp}$ soit strictement plus gros que F. Nous verrons un exemple en exercice.

B. Espaces euclidiens

B.1. Définition et exemples

Définition 11

On appelle espace vectoriel euclidien un espace préhilbertien réel de dimension finie (c'est-à-dire un espace vectoriel réel de dimension finie muni d'un produit scalaire).

Un sous-espace d'un espace préhilbertien réel hérite évidemment de la structure préhilbertienne (il suffit de restreindre le produit scalaire au sous-espace). Par conséquent, lorsqu'on travaille dans un espace préhilbertien de dimension infinie et qu'on a besoin d'un résultat de dimension finie, il peut être judicieux de se restreindre à sous-espace de dimension finie, qui est alors un espace euclidien.

Exemples :

- Le plan euclidien muni du produit scalaire géométrique $\overrightarrow{u} \cdot \overrightarrow{v} = \|\overrightarrow{u}\| \times \|\overrightarrow{v}\| \times \cos(\overrightarrow{u}, \overrightarrow{v})$ est un espace euclidien de dimension 2.
- $\mathbb{R}^n$ muni de son produit scalaire canonique est un espace euclidien de dimension n.
- $\mathscr{M}_{n,m}(\mathbb{R})$ muni de son produit scalaire canonique est un espace euclidien de dimension nm.
- $\mathbb{R}_n[X]$ muni de n'importe quel produit scalaire (le canonique, celui de Lagrange, celui hérité des fonctions continues, ...) est un espace euclidien de dimension $n+1$. Bien évidemment, la structure euclidienne se modifie dès que l'on change de produit scalaire !

B.2. Bases orthonormales d'un espace euclidien

Définition 12

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

On appelle base orthogonale de E une base de E qui est une famille orthogonale.

On appelle base orthonormale de E (en abrégé b.o.n.) une base de E qui est une famille orthonormale. Autrement dit, $(e_1, \ldots, e_n)$ est une base orthonormale de E lorsque $\forall i, j \in [\![1; n]\!]\,,\ \langle e_i, e_j\rangle = \delta_{i,j}$.

Pour vérifier qu'une famille de vecteurs donnée est une base orthonormale d'un espace euclidien de dimension n, on contrôle que la famille possède n vecteurs et qu'elle est orthonormale. C'est suffisant puisque l'orthonormalité assure la liberté et la condition de cardinalité permet alors d'utiliser le théorème Bonux.

Exemples :

- Dans $\mathbb{R}^n$ muni de son produit scalaire canonique, la base canonique $(\varepsilon_1, \ldots, \varepsilon_n)$ est une base orthonormale.
- Dans $\mathscr{M}_{n,m}(\mathbb{R})$ muni de son produit scalaire canonique, la base canonique des matrices élémentaires $(E_{i,j})_{1\leqslant i\leqslant n, 1\leqslant j\leqslant m}$ est une base orthonormale.
- Dans $\mathbb{R}_{n-1}[X]$ muni du produit scalaire de Lagrange, les polynômes élémentaires de Lagrange forment une base orthonormale.
- Soit E un $\mathbb{R}$-espace vectoriel de dimension finie n muni d'une base $\mathscr{B} = (e_1, \ldots, e_n)$. Il est possible de munir E d'un produit scalaire de sorte que $\mathscr{B}$ soit une base orthonormale pour ce produit scalaire. Il suffit pour cela de considérer le produit scalaire associé à cette base défini par $\langle x, y\rangle = {}^t\,\mathrm{Mat}_{\mathscr{B}}(x)\,\mathrm{Mat}_{\mathscr{B}}(y)$.

Le théorème suivant assure l'existence de bases orthonormales en dimension finie.

Théorème 6

Tout espace vectoriel euclidien possède au moins une base orthonormale.

■ Il suffit de partir d'une base puis de l'orthonormaliser par le procédé de Gram–Schmidt. On obtient alors une famille orthonormale dont le cardinal est la dimension de l'espace. C'est gagné ! ■

Dans un espace vectoriel de dimension infinie, on a dit que le lemme de Zorn permet d'affirmer l'existence de bases. Malheureusement, ce résultat ne se généralise pas pour les bases orthonormales : un espace préhilbertien de dimension infinie peut ne pas avoir de bases orthonormales !

On termine ce paragraphe avec le théorème de la base orthonormale incomplète.

Théorème 7

Dans un espace euclidien, toute famille orthonormale peut être complétée en une base orthonormale.

■ Soient $(E, \langle\cdot,\cdot\rangle)$ un espace euclidien de dimension n et $(e_1, \ldots, e_p)$ une famille orthonormale de E. Cette famille est libre puisqu'elle est orthonormale (c'est d'ailleurs ce qui justifie que cette famille est finie). Le théorème de la base incomplète permet donc de la compléter en une base $(e_1, \ldots, e_p, x_{p+1}, \ldots, x_n)$ de E. Il suffit alors d'appliquer le procédé de Gram–Schmidt à la famille $(e_1, \ldots, e_p, x_{p+1}, \ldots, x_n)$ et de constater que les vecteurs $e_1, \ldots, e_p$ sont conservés. ■

Dans une base orthonormale, l'expression du produit scalaire est très simple.

Proposition 11

Soit $(E, \langle\cdot,\cdot\rangle)$ un espace euclidien de dimension n, muni d'une base orthonormale $\mathscr{B}$.

Soient x, y deux vecteurs de E de coordonnées respectives $(x_1, \dots, x_n)$ et $(y_1, \dots, y_n)$ dans la base $\mathscr{B}$. Alors

$$\langle x, y\rangle = x_1 y_1 + \cdots + x_n y_n \qquad \text{et} \qquad \|x\| = \sqrt{{x_1}^2 + \cdots + {x_n}^2}.$$

Matriciellement, si l'on note X et Y les vecteurs colonnes constitués respectivement des coordonnées de x et y (c'est-à-dire $X = \mathrm{Mat}_{\mathscr{B}}(x)$ et $Y = \mathrm{Mat}_{\mathscr{B}}(y)$), on a

$$\langle x, y\rangle = {}^t XY \qquad \text{et} \qquad \|x\| = \sqrt{{}^t XX}.$$

■ On a

$$\langle x, y\rangle = \left\langle \sum_{k=1}^{n} x_k e_k \,, \sum_{\ell=1}^{n} y_\ell e_\ell \right\rangle = \sum_{1\leqslant k,\ell\leqslant n} x_k y_\ell \left\langle e_k, e_\ell\right\rangle = \sum_{1\leqslant k,\ell\leqslant n} x_k y_\ell\, \delta_{k,\ell} = \sum_{k=1}^{n} x_k y_k.$$

Les autres formules en découlent. ■

On peut retenir que, dans une base orthonormale, les calculs de produits scalaires et de normes s'effectuent, sur les n-uplets de coordonnées de vecteurs, exactement comme agit le produit scalaire canonique de $\mathbb{R}^n$.

En langage plus savant, cela revient à dire que l'application qui associe, à un vecteur de E, ses composantes dans une base orthonormale donnée est un isomorphisme entre E et $\mathbb{R}^n$ qui conserve le produit scalaire et la norme ($\mathbb{R}^n$ est évidemment muni de sa structure euclidienne canonique).

Dans une base orthonormale, le produit scalaire permet d'exprimer facilement les coordonnées d'un vecteur.

Proposition 12

Soient $(E, \langle\cdot,\cdot\rangle)$ un espace euclidien de dimension n et $\mathscr{B} = (e_1, \dots, e_n)$ une base orthonormale de E. Si $(x_1, \dots, x_n)$ désigne la famille des coordonnées d'un vecteur x de E, alors, pour tout $k \in [\![1; n]\!]$, on a

$$x_k = \left\langle x, e_k\right\rangle.$$

Autrement dit, la décomposition de x sur la base $\mathscr{B}$ est donnée par

$$x = \sum_{k=1}^{n} \left\langle x, e_k\right\rangle e_k.$$

■ Soit $k \in [\![1; n]\!]$. En prenant le produit scalaire par e_k des deux membres de l'égalité $x = x_1 e_1 + \cdots + x_n e_n$, on obtient $\langle x, e_k\rangle = 0 + \cdots + 0 + x_k + 0 + \cdots + 0$, ce qui donne le résultat. ■

Les formules des deux dernières proposition sont fausses si l'on ne travaille pas dans une base orthonormale !

<u>Exemples :</u>

- Soient $(E, \langle\cdot,\cdot\rangle)$ un espace euclidien muni d'une base orthonormale $\mathscr{B} = (e_1, \dots, e_n)$ et u un endomorphisme de E. Alors

$$\mathrm{Mat}_{\mathscr{B}}(u) = \big(\left\langle e_i, u(e_j)\right\rangle\big)_{1\leqslant i,j\leqslant n}.$$

6 h 00

C. Projections orthogonales

C.1. Supplémentaire orthogonal

Dans un espace préhilbertien réel E (non nécessairement de dimension finie), on a vu que deux sous-espaces orthogonaux sont en somme directe. En particulier, un sous-espace vectoriel F et son orthogonal $F^\perp$ sont en somme directe. L'énoncé ci-dessous précise ce résultat dans le cas d'un sous-espace de dimension finie : les sous-espaces F et $F^\perp$ sont alors supplémentaires.

Proposition 13

Soient $(E, \langle\cdot,\cdot\rangle)$ un espace préhilbertien réel et F un sous-espace vectoriel de E. Si F est de dimension finie, alors F et $F^\perp$ sont des supplémentaires orthogonaux dans E, c'est-à-dire

$$F \overset{\perp}{\oplus} F^\perp = E.$$

Le sous-espace $F^\perp$ est appelé le supplémentaire orthogonal de F dans E.

■ Nous procédons par analyse/synthèse (pardi !).

Soit $x \in E$. On note $(e_1, \dots, e_p)$ une base orthonormale de F.

▷ Analyse :
Supposons l'existence de $y \in F$ et $z \in F^\perp$ tels que $x = y+z$. Notons $(y_1, \dots, y_p)$ les coordonnées de y dans la base de F de sorte que $x = (y_1e_1+\cdots+y_pe_p)+z$. Soit $\ell \in [\![1;p]\!]$. On prend le produit scalaire par e_ℓ dans cette égalité, ce qui donne $\langle x, e_\ell\rangle = y_1\langle e_1, e_\ell\rangle+\cdots+y_\ell\langle e_\ell, e_\ell\rangle+\cdots+y_p\langle e_p, e_\ell\rangle$. Comme e_ℓ est orthogonal à z et aux autres e_k, il vient $\langle x, e_\ell\rangle = 0+\cdots+0+y_\ell+0+\cdots+0$, c'est-à-dire $y_\ell = \langle x, e_\ell\rangle$. On a donc nécessairement $y = \langle x, e_1\rangle e_1 + \cdots + \langle x, e_p\rangle e_p$ et $z = x - (\langle x, e_1\rangle e_1 + \cdots + \langle x, e_p\rangle e_p)$. Cela justifie l'unicité, sous réserve d'existence, du couple (y, z), ce qui démontre que F et $F^\perp$ sont en somme directe. Mais on le savait déjà !

▷ Synthèse :
On pose $y = \langle x, e_1\rangle e_1 + \cdots + \langle x, e_p\rangle e_p$ et $z = x - (\langle x, e_1\rangle e_1 + \cdots + \langle x, e_p\rangle e_p)$. On a évidemment $x+y = z$ et $x \in F$ (puique x est une combinaison des vecteurs e_i). De plus, pour tout $\ell \in [\![1;p]\!]$, on a $\langle z, e_\ell\rangle = \langle x, e_\ell\rangle - (\langle x, e_1\rangle\langle e_1, e_\ell\rangle + \cdots + \langle x, e_\ell\rangle\langle e_\ell, e_\ell\rangle + \cdots + \langle x, e_p\rangle e_p) = \langle x, e_\ell\rangle - \langle x, e_\ell\rangle = 0$, ce qui démontre que z est orthogonal à tout vecteur de la base $(e_1, \dots, e_p)$ de F et donc que $z \in F^\perp$. Par conséquent, on a $F + F^\perp = E$.

On en conclut que $F \overset{\perp}{\oplus} F^\perp = E$. ■

Tout sous-espace F de dimension finie possède donc un unique supplémentaire orthogonal. Ce supplémentaire se distingue (par son orthogonalité vis-à-vis de F) parmi la ribambelle de supplémentaires « généraux » de F.

Lorsque F est de dimension infinie, F et $F^\perp$ peuvent être supplémentaires... ou pas ! (dans ce cas, $F \overset{\perp}{\oplus} F^\perp$ est un sous-espace strict de E). Autrement dit, sans l'hypothèse « de dimension finie », rien n'assure de l'existence du supplémentaire orthogonal.

En dimension finie, la proposition précédente permet d'obtenir la dimension du supplémentaire orthogonal.

Corollaire 4

Soient $(E, \langle\cdot,\cdot\rangle)$ un espace euclidien et F un sous-espace vectoriel de E. On a

$$\dim F + \dim F^\perp = \dim E.$$

■ AQT ■

Autrement dit, en dimension finie, la dimension de $F^\perp$ est la codimension de F.

La proposition précédente permet d'établir l'égalité entre un sous-espace et son biorthogonal en dimension finie.

Proposition 14

Soient $(E, \langle \cdot, \cdot \rangle)$ un espace préhilbertien réel et F un sous-espace vectoriel de E. Si F et $F^\perp$ sont supplémentaires dans E, en particulier si F est de dimension finie, alors $F^{\perp\perp} = F$.

■ $\supset$ On sait déjà que $F^{\perp\perp} \supset F$.

$\subset$ Soit $x \in F^{\perp\perp}$. Comme $E = F + F^\perp$, il existe $y \in F$ et $z \in F^\perp$ tel que $x = y + z$. Prenons le produit scalaire par z dans cette égalité. Cela donne $\langle x, z \rangle = \langle y, z \rangle + \|z\|^2$, c'est-à-dire $0 = 0 + \|z\|^2$ puisque $x \perp z$ et $y \perp z$. Il s'ensuit que $z = 0_E$ et donc $x = y$, ce qui implique que $x \in F$. Donc $F^{\perp\perp} \subset F$. ■

L'hypothèse « F et $F^\perp$ sont supplémentaires dans E » (ou « F est de dimension finie ») est cruciale !

Dans le cas où E est un espace euclidien, la démonstration de ce résultat se simplifie nettement : on sait que $F^{\perp\perp} \supset F$ et que $\dim F^{\perp\perp} = \dim E - \dim F^\perp = \dim E - \dim E + \dim F = \dim F$, donc $F = F^{\perp\perp}$.

Si E est un espace euclidien et A est une partie de E, alors $A^{\perp\perp} = \mathrm{Vect}(A)$.

C.2. Projections et symétries orthogonales

Définition 13

Soient $(E, \langle \cdot, \cdot \rangle)$ un espace préhilbertien réel et F un sous-espace de E tel que $F \overset{\perp}{\oplus} F^\perp = E$ (c'est le cas, par exemple, lorsque F est de dimension finie).

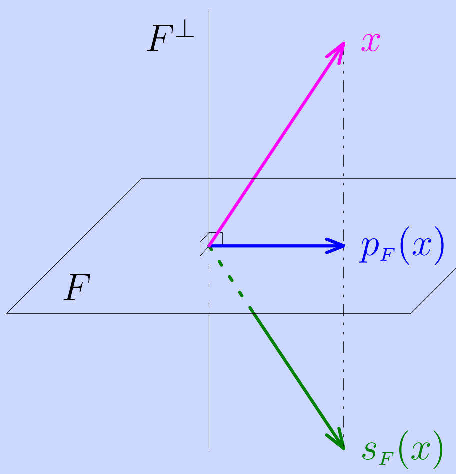
On appelle projection orthogonale sur F la projection p_F sur F dans la direction de $F^\perp$, c'est-à-dire

$$p_F \left\{ \begin{array}{ccc} E = F \overset{\perp}{\oplus} F^\perp & \longrightarrow & E \\ x = y + z & \longmapsto & y \end{array} \right.$$

On appelle symétrie orthogonale par rapport à F la symétrie s_F sur F dans la direction de $F^\perp$, c'est-à-dire

$$s_F \left\{ \begin{array}{ccc} E = F \overset{\perp}{\oplus} F^\perp & \longrightarrow & E \\ x = y + z & \longmapsto & y - z \end{array} \right.$$

Graphiquement, cela donne



On a donc

(i)	$p_F \circ p_F = p_F$	(j)	$s_F \circ s_F = \mathrm{Id}_F$
(ii)	$\operatorname{Im} p_F = F$	(jj)	$\operatorname{Ker}(s_F - \mathrm{Id}_E) = F$
(iii)	$\operatorname{Ker} p_F = F^\perp$	(jjj)	$\operatorname{Ker}(s_F + \mathrm{Id}_E) = F^\perp$
(iv)	$\forall x \in E,\ p_F(x) \perp x - p_F(x)$	(jw)	$\forall x \in E,\ x - s_F(x) \perp x + s_F(x)$.

De plus, la projection orthogonale sur F et la symétrie orthogonale par rapport à F sont reliées par la relation :

$$s_F = 2p_F - \mathrm{Id}_E\,.$$

■ L'existence de p_F et s_F est justifiée par le fait que F et $F^\perp$ sont supplémentaires dans E puisque F est de dimension finie. ■

La condition (iv), couplée au théorème de Pythagore, implique que, pour tout $x \in E$, on a $\|p_F(x)\|^2 + \|x - p_F(x)\|^2 = \|x\|^2$. On en déduit que, pour tout $x \in E$, on a $\|p_F(x)\| \leqslant \|x\|$ avec égalité, si et seulement si, x appartient à F.

Le projecteur associé à p_F est $p_{F^\perp}$ (puisque $F^{\perp\perp} = F$), c'est-à-dire $p_F + p_{F^\perp} = \mathrm{Id}_E$. Ainsi, lorsqu'on connaît p_F, on connaît $p_{F^\perp}$.

Une symétrie par rapport à un hyperplan H est appelée une réflexion (par rapport à H).

Lorsqu'on dispose d'une base orthonormale de F, on peut donner l'expression de p_F et s_F.

Proposition 15

Soit $(E, \langle\cdot,\cdot\rangle)$ un espace préhilbertien réel. On considère F un sous-espace vectoriel de E de dimension finie, muni d'une base orthonormale $(e_1, \dots, e_p)$. On note p_F la projection orthogonale sur F et s_F la symétrie orthogonale par rapport à F. Pour tout $x \in E$, on a

$$p_F(x) = \sum_{k=1}^{p} \langle e_k, x\rangle\, e_k \qquad \text{et} \qquad s_F(x) = 2\sum_{k=1}^{p} \langle e_k, x\rangle\, e_k - x.$$

■ Soit $x \in E$. Comme $E = F \overset{\perp}{\oplus} F^\perp$, il existe $y \in F$ et $z \in F^\perp$ tel que $x = y + z$. Or $(e_1, \dots, e_p)$ est une base de F donc il existe $y_1, \dots, y_p \in \mathbb{R}$ tels que $y = y_1 e_1 + \cdots + y_p e_p$, ce qui donne $x = y_1 e_1 + \cdots + y_p e_p + z$. En prenant le produit scalaire par e_k (où $k \in [\![1;p]\!]$) dans cette égalité, on obtient $\langle x, e_k\rangle = y_k$ car e_k est orthogonal à z et à tous les autres e_j. On a donc bien $y = \langle x, e_1\rangle\, e_1 + \cdots + \langle x, e_p\rangle\, e_p$, c'est-à-dire $p_F(x) = \langle x, e_1\rangle\, e_1 + \cdots + \langle x, e_p\rangle\, e_p$. L'expression de $s_F(x)$ en découle puisque $s_F = 2p_F - \mathrm{Id}_E$. ■

Le résultat de cette proposition permet de calculer le projeté orthogonal d'un vecteur sur un sous-espace (de dimension finie) F quand celui-ci est pouvu d'une base orthonormale.

Mais que faire lorsqu'on connaît une base $(v_1, \dots, v_p)$ de F qui n'est pas orthonormale ?

La première idée qui vient à l'esprit est d'orthonormaliser la base $(v_1, \dots, v_p)$ grâce à la méthode de Gram–Schmidt. Mais cette méthode peut être maladroite car les calculs sont lourds.

Il est souvent préférable de conserver telle quelle la base $(v_1, \dots, v_p)$ de F et d'introduire les coordonnées $\alpha_1, \dots, \alpha_p \in \mathbb{R}$ de $p(x)$ dans cette base puis de les calculer à l'aide des relations $\forall k \in [\![1;p]\!]\,,\ \langle x - p(x), v_k\rangle = 0$. On obtient un système de p équations à p inconnues à résoudre.

Exemples :

- On se place dans $\mathbb{R}^3$ muni de sa structure euclidienne canonique. On pose $F = \mathrm{Vect}(u_1, u_2)$ où $u_1 = (1, 1, 0)$ et $u_2 = (1, 0, 1)$. Déterminons le projeté orthogonal sur F de $u_3 = (0, 1, 1)$.

  On pose $p = p_F$ et on propose les deux méthodes évoquées ci-dessus.

  ∗ Comme u_1 et u_2 ne sont pas colinéaires, (u_1, u_2) est une base de F. Il existe donc $a, b \in \mathbb{R}$ tels que $p(u_3) = au_1 + bu_2 = (a+b, a, b)$ et donc $u_3 - p(u_3) = (-a-b, 1-a, 1-b)$. On a $\langle u_3 - p(u_3), u_1\rangle = 0$ et $\langle u_3 - p(u_3), u_2\rangle = 0$ d'où $(-a-b).1 + (1-a).1 + (1-b).0 = 0$ et $(-a-b).1 + (1-a).0 + (1-b).1 = 0$, c'est-à-dire $2a + b = 1$ et $a + 2b = 1$. On en déduit que $a = 1/3$ et $b = 1/3$. On a donc $p(u_3) = (2/3, 1/3, 1/3)$.

  ∗ Nous avons déjà fait l'orthonormalisation de (u_1, u_2) dans un exemple précédent. On a trouvé que la famille (e_1, e_2), où $e_1 = (1/\sqrt{2}, 1/\sqrt{2}, 0)$ et $e_2 = (1/\sqrt{6}, -1/\sqrt{6}, 2/\sqrt{6})$, est une base orthonormale de F. Alors, par le résultat de la proposition 15, on a $p(u_3) = \langle u_3, e_1\rangle\, e_1 + \langle u_3, e_2\rangle\, e_2$. Or $\langle u_3, e_1\rangle = \big\langle (0,1,1), (1,1,0)/\sqrt{2}\big\rangle = 1/\sqrt{2}$ et $\langle u_3, e_2\rangle = \big\langle (0,1,1), (1,-1,2)/\sqrt{6}\big\rangle = 1/\sqrt{6}$, donc $p(u_3) = (2/3, 1/3, 1/3)$.

- On munit $\mathscr{C}^0_{2\pi}(\mathbb{R}, \mathbb{R})$ de son produit scalaire canonique. On considère $c_k : x \longmapsto \cos(kx)$ et $s_k : x \longmapsto \sin(kx)$ pour tout $k \in \mathbb{N}^*$. On a vu que la famille de ces fonctions est orthonormale. Il n'est pas difficile de voir que si l'on ajoute à cette famille la fonction $\widehat{c_0} : x \longmapsto 1/\sqrt{2}$, la famille reste orthonormale. Posons $F = \mathrm{Vect}(\widehat{c_0}, c_1, \dots, c_n, s_1, \dots, s_n)$. Si f est un élément de $\mathscr{C}^0_{2\pi}(\mathbb{R}, \mathbb{R})$, la proposition 15 dit que le projeté orthogonal de f sur F est la fonction, appelée développement en série de Fourier à l'ordre n de f, définie par

  $$\varphi_n : x \longmapsto \frac{a_0}{2} + \sum_{k=1}^{n} a_k \cos(kx) + b_k \sin(kx)$$

  où, pour tout $k \in [\![0;n]\!]$,

  $$a_k = \langle f, c_k\rangle = \frac{1}{\pi}\int_0^{2\pi} f(t)\cos(kt)\,\mathrm{d}t \qquad \text{et} \qquad b_k = \langle f, s_k\rangle = \frac{1}{\pi}\int_0^{2\pi} f(t)\sin(kt)\,\mathrm{d}t.$$

C.3. Distance à un sous-espace de dimension finie

Définition 14

Soient $(E, \langle\cdot,\cdot\rangle)$ un espace préhilbertien réel, A une partie non vide de E et x un élément de E. On appelle distance de x à A le nombre réel positif ou nul, noté $\mathrm{d}(x, A)$, défini par

$$\mathrm{d}(x, A) = \inf\{\|x - a\| : a \in A\}.$$

■ L'ensemble $\{\|x - a\| : a \in A\}$ est une partie de $\mathbb{R}$ qui est non vide (car $A \neq \varnothing$) et minorée (par 0). Cela justifie l'existence de la borne inférieure. ■

En général, la distance d'un élément à une partie n'est pas atteinte (vous pouvez penser à la distance de 2 à $]0;1[$ dans $\mathbb{R}$ qui vaut 1 mais qui n'est pas atteinte...). Cela rend l'énoncé ci-dessous d'autant plus remarquable : la projection orthogonale réalise le minimum de la distance d'un vecteur à un sous-espace de dimension finie.

Proposition 16

Soient $(E, \langle\cdot,\cdot\rangle)$ un espace préhilbertien réel et F un sous-espace de E tel que $F \overset{\perp}{\oplus} F^\perp = E$ (c'est le cas, par exemple, lorsque F est de dimension finie). On note p_F la projection orthogonale sur F. Pour tout $x \in E$, on a

$$\mathrm{d}(x, F) = \|x - p_F(x)\|$$

De plus, $p_F(x)$ est le seul élément y de F tel que $\mathrm{d}(x, F) = \|x - y\|$.

■ Soit $x \in E$. Pour tout $y \in F$, les vecteurs $p_F(x) - y$ et $x - p_F(x)$ sont orthogonaux (le premiers est dans F et le second est dans $F^\perp$), donc, d'après le théorème de Pythagore, on a

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

Il découle immédiatement de cette relation que

$$\forall y \in F \setminus \{p_F(x)\}, \qquad \|x - y\| > \|x - p_F(x)\| .$$

Cela signifie précisément que $p_F(x)$ est l'unique vecteur de F le plus proche de x. ■

Si F et $F^\perp$ ne sont pas supplémentaires (ce qui peut se produire si F est de dimension infinie), il est tout à fait possible que la distance $\mathrm{d}(x, F)$ ne soit pas atteinte !

Le calcul de la distance à un sous-espace de dimension finie se ramène donc à la détermination d'un projeté orthogonal.

Exemples :

- On se place dans $\mathbb{R}^3$ muni de sa structure euclidienne canonique. On pose $F = \mathrm{Vect}(u_1, u_2)$ où $u_1 = (1, 1, 0)$ et $u_2 = (1, 0, 1)$. D'après les calculs effectués dans l'exemple de la page précédente, la distance de $u_3 = (0, 1, 1)$ à F est donnée par

$$\mathrm{d}(u_3, F) = \|u_3 - p(u_3)\| = \|(0, 1, 1) - (2/3, 1/3, 1/3)\| = \|(-2/3, 2/3, 2/3)\| = \frac{2}{\sqrt{3}}.$$

D. Hyperplans d'un espace euclidien

D.1. Hyperplans vectoriels d'un espace euclidien

Définition 15

Soient $(E, \langle \cdot, \cdot \rangle)$ un espace euclidien et H un hyperplan de E. Le supplémentaire orthogonal $H^\perp$ de H dans E est une droite vectorielle. Tout vecteur non nul de cette droite est appelé un vecteur normal de H.

Si a est un vecteur normal d'un hyperplan H, alors $H = a^\perp$ (car $H = H^{\perp\perp}$), c'est-à-dire que $(x \in H) \iff \big(\langle x, a\rangle = 0\big)$. La relation $\langle x, a\rangle = 0$ est donc une équation cartésienne de H.

Dans un espace euclidien orienté E de dimension n, la donnée d'un vecteur normal à un hyperplan H permet d'orienter l'hyperplan : il suffit de considérer qu'une base $(e_1, \dots, e_{n-1})$ de H est directe lorsque $(e_1, \dots, e_{n-1}, a)$ est une base directe de E.

Exemples :

- Dans $\mathbb{R}^3$ muni de sa structure euclidienne canonique et de sa base orthonormale canonique $\mathscr{B} = (\varepsilon_1, \varepsilon_2, \varepsilon_3)$, l'hyperplan d'équation $ax + by + cz = 0$ (où $a, b, c \in \mathbb{R}$ sont non tous nuls) admet le vecteur (a, b, c) comme vecteur normal. Tout autre vecteur non nul qui lui est colinéaire est aussi un vecteur normal.

On peut alors énoncer le théorème de représentation de Riesz.

Théorème 8

Soit $(E, \langle \cdot, \cdot \rangle)$ un espace euclidien. Pour toute forme linéaire $\ell \in E^*$, il existe un unique vecteur a tel que

$$\forall x \in E, \qquad \ell(x) = \langle x, a\rangle\,.$$

■ Soit $\ell \in E^*$. On propose trois démonstrations (la première est la plus jolie !).

▷ À tout a de E, on peut associer la forme linéaire $\ell_a : x \longmapsto \langle x, a\rangle$.

L'application $L : E \longrightarrow E^*$ définie par $\forall a \in E,\ L(a) = \ell_a$ est alors linéaire puisque, pour tout $a, b \in E$, on a $\forall x \in E,\ \ell_{a+b}(x) = \langle x, a+b\rangle = \langle x, a\rangle + \langle x, b\rangle = \ell_a(x) + \ell_b(x)$.

Son noyau est $\operatorname{Ker} L = \{a \in E : \forall x \in E,\ \ \langle x, a\rangle = 0\} = \{0_E\}$ puisque le seul vecteur orthogonal à tous les autres est le vecteur nul. Ainsi, L est injectif.

Comme, par ailleurs, E et E^* sont deux espaces de même dimension finie (c'est un résultat connu !), le théorème bonux permet d'affirmer que L est un isomorphisme.

Par conséquent, toute forme linéaire ℓ admet un unique antécédent par cet isomorphisme, c'est-à-dire qu'il existe un unique vecteur a tel que, pour tout $x \in E$, on ait $\ell(x) = \langle x, a\rangle$.

▷ Si ℓ est la forme linéaire nulle, on prend $a = 0_E$ et c'est la seule possibilité (sinon on aurait $\ell(a) \neq 0$).

Supposons dorénavant que $\ell \neq \widetilde{0}$. Introduisons l'hyperplan $H = \operatorname{Ker} \ell$ et notons a l'un de ses vecteurs normaux (c'est un vecteur non nul). On possède alors deux équations de H, à savoir $\ell(x) = 0$ et $\langle x, a\rangle = 0$. On a vu que cela impliquait que les formes linéaires ℓ et $x \longmapsto \langle x, a\rangle$ étaient proportionnelles, autrement dit qu'il existe $\lambda \in \mathbb{R}^*$ tel que $\ell = \lambda \langle \bullet, a\rangle$. On a donc $\ell = \langle \bullet, \lambda a\rangle$, ce qui démontre la partie « existence » de notre résultat.

Supposons qu'il existe $a, b \in E$ tels que $\ell = \langle \bullet, a\rangle = \langle \bullet, b\rangle$. On a alors $\forall x \in E,\ \ \langle x, a-b\rangle = 0$. En particulier, on a $\langle a-b, a-b\rangle = 0$, ce qui donne $a - b = 0_E$ et donc $a = b$. Cela démontre l'unicité.

▷ Considérons une base orthonormale $\mathscr{B} = (e_1, \dots, e_n)$ de E. Nous raisonnons par analyse/synthèse.

Analyse : Supposons qu'il existe $a \in E$ tel que $\forall x \in E,\ \ell(x) = \langle x, a\rangle$ et notons $a = a_1 e_1 + \cdots + a_n e_n$ sa décomposition dans la base $\mathscr{B}$. Pour tout $k \in [\![1; n]\!]$, on a $\ell(e_k) = \langle e_k, a_1 e_1 + \cdots + a_n e_n\rangle = a_k$. On a donc $a = \ell(e_1)e_1 + \cdots + \ell(e_n)e_n$. Cela démontre l'unicité de a, sous réserve d'existence.

Synthèse : Posons $a = \ell(e_1)e_1 + \cdots + \ell(e_n)e_n$. Pour tout $x \in E$ tel que $x = x_1 e_1 + \cdots + x_n e_n$, on a $\ell(x) = x_1 \ell(e_1) + \cdots + x_n \ell(e_n) = \langle x, a\rangle$. Cela démontre l'existence. ■

On termine ce paragraphe par un énoncé qui simplifie la proposition 15 dans le cas où l'on projète sur un hyperplan.

Proposition 17

Soient $(E, \langle\cdot,\cdot\rangle)$ un espace euclidien et H un hyperplan de E. On note p_H la projection orthogonale sur H et s_H la réflexion associée. Si a désigne un vecteur normal de H, alors, pour tout $x \in E$, on a

$$p_H(x) = x - \frac{\langle x, a\rangle}{\|a\|^2} a \qquad \text{et} \qquad s_H(x) = x - 2\frac{\langle x, a\rangle}{\|a\|^2} a.$$

■ Soit $x \in E$. Comme $(a/\|a\|)$ est une base de $F^\perp$, la proposition 15 nous dit que

$$p_{H^\perp}(x) = \left\langle x, \frac{a}{\|a\|}\right\rangle \frac{a}{\|a\|} = \frac{\langle x, a\rangle}{\|a\|^2} a$$

Comme $p_H = \mathrm{Id}_E - p_{H^\perp}$ et $s_H = 2p_H - \mathrm{Id}_E = \mathrm{Id}_E - 2p_{H^\perp}$, on en déduit les deux résultats. ■

En substance, cette proposition explique que pour projeter un vecteur sur un hyperplan, le plus simple est de retirer au vecteur sa composante sur la droite qui est le supplémentaire orthogonale de l'hyperplan !

$\mathcal{D}$.2. Hyperplans affines d'un espace euclidien

Définition 16

Soient $(E, \langle\cdot,\cdot\rangle)$ un espace euclidien et $\mathscr{H}$ un hyperplan affine de E. Un vecteur normal de $\mathscr{H}$ est, par définition, un vecteur normal de la direction H de $\mathscr{H}$.

Si a est un vecteur normal d'un hyperplan affine $\mathscr{H}$, alors $a^\perp$ est la direction de $\mathscr{H}$, c'est-à-dire que, si A est l'un des points de $\mathscr{H}$, alors $(M \in \mathscr{H}) \iff \big(\langle \overrightarrow{AM}, a\rangle = 0\big) \iff \big(\langle M, a\rangle = \langle A, a\rangle\big)$. La relation $\langle M, a\rangle = \langle A, a\rangle$ est une équation cartésienne de l'hyperplan affine $\mathscr{H}$.

Comme pour les hyperplans vectoriels, la donnée d'un vecteur normal permet d'orienter un hyperplan affine d'un espace euclidien orienté.

Exemples :

- Dans $\mathbb{R}^3$ muni de sa structure affine euclidienne canonique et de son repère orthonormal canonique $\mathscr{R} = (O; \varepsilon_1, \varepsilon_2, \varepsilon_3)$, l'hyperplan affine d'équation $ax + by + cz = d$ (où $a, b, c \in \mathbb{R}$ sont non tous nuls) admet le vecteur (a, b, c) comme vecteur normal. Tout autre vecteur non nul qui lui est colinéaire est aussi un vecteur normal.

La proposition suivante explique comment calculer la distance d'un point à un hyperplan affine.

Proposition 18

Soient $(E, \langle\cdot,\cdot\rangle)$ un espace euclidien et $\mathscr{H}$ un hyperplan affine de E passant par le point A et de vecteur normal $\overrightarrow{a}$. Pour tout point M de E, on a

$$\mathrm{d}(M, \mathscr{H}) = \frac{\left|\langle \overrightarrow{AM}, \overrightarrow{a}\rangle\right|}{\|\overrightarrow{a}\|}.$$

■ Notons H la direction de $\mathscr{H}$ et p_H la projection orthogonale sur H. On a

$$\mathrm{d}(M, \mathscr{H}) = \mathrm{d}(M, A + H) = \inf\{\|\overrightarrow{OM} - (\overrightarrow{OA} + \overrightarrow{u})\| : \overrightarrow{u} \in H\} = \inf\{\|\overrightarrow{AM} - \overrightarrow{u}\| : \overrightarrow{u} \in H\}.$$

Or, d'après la proposition 16, la projection orthogonale réalise la distance minimale, ce qui donne

$$\mathrm{d}(M, \mathscr{H}) = \|\overrightarrow{AM} - p_H(\overrightarrow{AM})\|.$$

Comme la proposition 17 nous dit que $p_H(\overrightarrow{AM}) = \overrightarrow{AM} - \dfrac{\langle \overrightarrow{AM}, a\rangle}{\|a\|^2} a$, on en conclut que

$$\mathrm{d}(M, \mathscr{H}) = \left\|\frac{\langle \overrightarrow{AM}, a\rangle}{\|a\|^2} a\right\| = \frac{\left|\langle \overrightarrow{AM}, \overrightarrow{a}\rangle\right|}{\|\overrightarrow{a}\|}.$$ ■

Exemples :

- Dans $\mathbb{R}^2$ muni de sa structure affine euclidienne canonique et de son repère orthonormal canonique $\mathscr{R} = (O; \varepsilon_1, \varepsilon_2)$, on considère la droite affine $\mathscr{D}$ d'équation $ax+by = c$ (où $a, b \in \mathbb{R}$ sont non tous nuls) et le point M de coordonnées (x_M, y_M). Alors

$$\mathrm{d}(M, \mathscr{D}) = \frac{|ax_M + by_M - c|}{\sqrt{a^2 + b^2}}.$$

- Dans $\mathbb{R}^3$ muni de sa structure affine euclidienne canonique et de son repère orthonormal canonique $\mathscr{R} = (O; \varepsilon_1, \varepsilon_2, \varepsilon_3)$, on considère le plan affine $\mathscr{P}$ d'équation $ax+by+cz = d$ (où $a, b, c \in \mathbb{R}$ sont non tous nuls) et le point M de coordonnées (x_M, y_M, z_M). Alors

$$\mathrm{d}(M, \mathscr{P}) = \frac{|ax_M + by_M + cz_M - d|}{\sqrt{a^2 + b^2 + c^2}}.$$

8 h 30