# Chapitre 12

# Espaces préhilbertiens
# Espaces euclidiens

Dans ce chapitre, $\mathcal{H}$ désigne un $\mathbb{R}$-espace vectoriel, et $E$ un $\mathbb{R}$-espace vectoriel de dimension finie.

## I. Produit scalaire

**Définition**

• Un **produit scalaire** sur $\mathcal{H}$ est une forme bilinéaire symétrique définie positive sur $\mathcal{H}$, c'est-à-dire, une application $f : \mathcal{H} \times \mathcal{H} \to \mathbb{R}$ vérifiant les propriétés suivantes :

- **Bilinéarité** : pour tout $(a,b) \in \mathcal{H} \times \mathcal{H}$, les applications $x \mapsto f(x,b)$ et $y \mapsto f(a,y)$ sont linéaires.
- **Symétrie** : pour tout $(x,y) \in \mathcal{H} \times \mathcal{H}$, $f(x,y) = f(y,x)$.
- **Définie positivité** pour tout $x \in \mathcal{H}$, $f(x,x) \geqslant 0$, et on a l'équivalence :

$$f(x,x) = 0 \ \Leftrightarrow x = 0.$$

Si $f$ est un produit scalaire sur $\mathcal{H}$, on note le plus souvent, pour $(x,y) \in \mathcal{H}^2$,

$$f(x,y) = (x \,|\, y)\,, \quad \text{ou} \quad \langle x, y\rangle, \quad \text{ou} \quad x \cdot y.$$

• Si $\mathcal{H}$ est muni d'un produit scalaire $(\cdot \,|\, \cdot)$, on dit que $(\mathcal{H}, (\cdot \,|\, \cdot))$ (ou simplement $\mathcal{H}$ s'il n'y a pas d'ambiguité sur le produit scalaire) est un **espace préhilbertien** (réel).

• Un **espace euclidien** est un espace préhilbertien de dimension finie.

Remarques

• Du fait de la symétrie, il suffit en fait d'imposer la linéarité par rapport à une seule des deux variables.

• Si $(\cdot \,|\, \cdot)$ est un produit scalaire sur $\mathcal{H}$, alors pour tout $(a,b) \in \mathcal{H}^2$, $(a \,|\, 0) = (0 \,|\, b) = 0$.

• Si $E$ est un sous-espace vectoriel de dimension finie de $\mathcal{H}$, et si $(\cdot \,|\, \cdot)$ est un produit scalaire sur $\mathcal{H}$, alors $(\cdot \,|\, \cdot)$ induit par restriction un produit scalaire sur $E$ qui est donc un espace euclidien.

Exemples

• L'application $f_1$ définie sur $\mathbb{R}^2 \times \mathbb{R}^2$ par $f_1((x_1,x_2),(y_1,y_2)) = x_1y_1 + 2x_2y_2$ est un produit scalaire sur $\mathbb{R}^2$, mais pas l'application $f_2$ définie sur $\mathbb{R}^2 \times \mathbb{R}^2$ par $f((x_1,x_2),(y_1,y_2)) = x_1y_1 - 2x_2y_2$, cette dernière ne vérifiant pas la propriété de définie positivité : en effet, $f_2((0,1),(0,1)) = -2 < 0$.

• L'application $(\cdot\,|\,\cdot)$ définie sur $\mathbb{R}^n \times \mathbb{R}^n$ par

$$(x\,|\,y) = \sum_{i=1}^{n} x_i y_i$$

(où $x = (x_1, \dots, x_n)$ et $y = (y_1, \dots, y_n)$) est un produit scalaire sur $\mathbb{R}^n$. Il est appelé produit scalaire **canonique** sur $\mathbb{R}^n$.

• En fait, si $E$ est un $\mathbb{R}$-espace vectoriel de dimension finie, on peut toujours munir $E$ d'une structure d'espace euclidien. En effet, soit $\mathcal{B} = (e_1, \dots, e_n)$ une base de $E$ ; on définit alors, pour $x = x_1 e_1 + \cdots + x_n e_n$ et $y = y_1 e_1 + \cdots + y_n e_n$ éléments de $E$,

$$(x\,|\,y) = \sum_{i=1}^{n} x_i y_i.$$

Ceci définit un produit scalaire sur $E$.

• L'application $g$ définie sur $\mathscr{M}_n(\mathbb{R})^2$ par $g(A,B) = \mathrm{Tr}\,({}^t\!AB)$ est un produit scalaire sur $\mathscr{M}_n(\mathbb{R})$. Si $A = (a_{i,j})$ et $B = (b_{i,j})$, alors, pour tout $(i,j) \in [\![1,n]\!]^2$, le coefficient en position $(i,j)$ de la matrice ${}^t\!AB$ est

$$\sum_{k=1}^{n} a_{k,i}\, b_{k,j},$$

et donc

$$g(A,B) = \sum_{i=1}^{n} \sum_{k=1}^{n} a_{k,i}\, b_{k,i} = \sum_{i,j=1}^{n} a_{i,j}\, b_{i,j}$$

après changement d'indices muets. On est donc dans la situation du point précédent, pour le choix de la base canonique de $\mathscr{M}_n(\mathbb{R})$.

• Soit $\omega : [a,b] \to \mathbb{R}_+^*$ une application continue. L'application $(\cdot\,|\,\cdot)$ définie sur $\mathcal{C}^0([a,b],\mathbb{R})^2$ par

$$(f\,|\,g) = \int_a^b f(x) g(x)\, \omega(x)\, dx,$$

est un produit scalaire sur $\mathcal{C}^0([a,b],\mathbb{R})$ (qui, munit de ce produit scalaire, est un espace préhilbertien réel, mais pas un espace euclidien).

• Soit $I$ un intervalle de $\mathbb{R}$ et $\mathcal{H} = L^2(I,\mathbb{R}) \cap \mathcal{C}^0(I,\mathbb{R})$. L'application

$$(\cdot\,|\,\cdot) : \left\{ \begin{array}{ccc} \mathcal{H} \times \mathcal{H} & \to & \mathbb{R} \\ (f,g) & \mapsto & \displaystyle\int_I f\, g \end{array} \right.$$

est un produit scalaire sur $\mathcal{H}$.

• L'application $(\cdot\,|\,\cdot)$ définie sur $\mathbb{R}_n[X]^2$ par

$$(P\,|\,Q) = \sum_{i=0}^{n} P(i) Q(i),$$

est un produit scalaire sur $\mathbb{R}_n[X]$. Pour la définie positivité, on remarque qu'un polynôme $P$ de $\mathbb{R}_n[X]$ vérifie $(P\,|\,P) = 0$ si et seulement si $P(i) = 0$ pour tout $i \in [\![0,n]\!]$, ce qui équivaut à $P = 0$ (si $P(i) = 0$ pour tout $i \in [\![0,n]\!]$, $P$ possède au moins $n+1$ racines, or $P$ est de degré au plus $n$).

**Théorème – Inégalité de Cauchy-Schwarz**

Soit $(\cdot\,|\,\cdot)$ un produit scalaire sur $\mathcal{H}$. Alors, pour tout $(x,y) \in \mathcal{H}^2$,

$$|\,(x\,|\,y)\,| \leqslant \sqrt{(x\,|\,x)}\,\sqrt{(y\,|\,y)},$$

avec égalité si et seulement si $x$ et $y$ sont colinéaires.

Démonstration – Fixons $x$ et $y$ dans $\mathcal{H}$ et définissons sur $\mathbb{R}$ l'application

$$P : \lambda \mapsto (\lambda x + y \,|\, \lambda x + y)\,.$$

Pour tout $\lambda \in \mathbb{R}$, par bilinéarité et symétrie,

$$P(\lambda) = \lambda^2\,(x\,|\,x) + \lambda\,(x\,|\,y) + \lambda\,(y\,|\,x) + (y\,|\,y) = \lambda^2\,(x\,|\,x) + 2\lambda\,(x\,|\,y) + (y\,|\,y)\,.$$

La fonction $P$ ne prend que des valeurs positives d'après la propriété de définie positivité.

Si $x \neq 0$, $(x\,|\,x) \neq 0$ pour la même raison, et $P$ est une fonction polynomiale de degré 2 ; on en déduit que le discriminant du polynôme $P$ est négatif ou nul, c'est-à-dire

$$(2\,(x\,|\,y))^2 - 4\,(x\,|\,x)\,(y\,|\,y) \leqslant 0, \quad \text{d'où} \quad (x\,|\,y)^2 \leqslant (x\,|\,x)\,(y\,|\,y)\,.$$

Le résultat suit en composant cette inégalité par la fonction croissante racine carrée.

Si $x = 0$, $P$ est une fonction affine partout positive, donc le coefficient directeur associé est nul, c'est-à-dire $(x\,|\,y) = 0$. L'inégalité est également vérifiée dans ce cas.

En ce qui concerne le cas d'égalité : si $x$ et $y$ sont colinéaires, il est immédiat que l'égalité est vérifiée ; par exemple s'il existe $\alpha \in \mathbb{R}$ tel que $y = \alpha x$, on a

$$(x\,|\,y) = (x\,|\,\alpha x) = \alpha\,(x\,|\,x)\,,$$

et

$$\sqrt{(x\,|\,x)}\sqrt{(y\,|\,y)} = \sqrt{(x\,|\,x)}\sqrt{(\alpha x\,|\,\alpha x)} = |\alpha|\sqrt{(x\,|\,x)}\sqrt{(x\,|\,x)} = |\alpha|\,(x\,|\,x)\,,$$

donc on a égalité dans l'inégalité de Cauchy-Schwarz (on procède de même s'il existe $\alpha \in \mathbb{R}$ tel que $x = \alpha\,y$). Réciproquement, supposons que $x \neq 0$ et que $|\,(x\,|\,y)\,| = \sqrt{(x\,|\,x)}\sqrt{(y\,|\,y)}$. En reprenant la démonstration précédente, on voit que le discriminant de $P$ est nul, donc $P$ possède une racine réelle (double) $\lambda$, et on a donc $P(\lambda) = (\lambda x + y\,|\,\lambda x + y) = 0$. Par définie positivité, il s'ensuit que $\lambda x + y = 0$ et donc $x$ et $y$ sont colinéaires. Si $x = 0$, $x$ et $y$ sont également colinéaires.
$\square$

**Propriété/Définition**

- Si $(\cdot\,|\,\cdot)$ est un produit scalaire sur $\mathcal{H}$, l'application $\|\cdot\| : x \mapsto \sqrt{(x\,|\,x)}$ est une norme sur $\mathcal{H}$, dite **norme associée** à $(\cdot\,|\,\cdot)$. Une norme associée à un produit scalaire sur $\mathcal{H}$ est appelée norme **euclidienne**.
- L'application $d$ définie sur $\mathcal{H}^2$ par $d(x,y) = \|x - y\|$ est appelée **distance associée** à $(\cdot\,|\,\cdot)$.

Démonstration du fait que $\|\cdot\|$ est une norme.
L'application $\|\cdot\|$ est bien définie car $(x\,|\,x) \geqslant 0$ pour tout $x \in \mathcal{H}$.
**Homogénéité :** pour tout $x \in \mathcal{H}$ et $\lambda \in \mathbb{R}$,

$$\|\lambda x\| = \sqrt{(\lambda x\,|\,\lambda x)} = \sqrt{\lambda^2\,(x\,|\,x)} = |\lambda|\sqrt{(x\,|\,x)} = |\lambda|\,\|x\|.$$

**Séparation :** pour tout $x \in \mathcal{H}$,

$$\|x\| = 0 \quad \Leftrightarrow \quad (x\,|\,x) = 0 \quad \Leftrightarrow \quad x = 0,$$

car $(\cdot\,|\,\cdot)$ est définie positive.
**Inégalité triangulaire :** comme on l'a remarqué dans le chapitre **Espaces vectoriels normés**, elle résulte de l'inégalité de Cauchy-Schwarz, qui se réécrit

$$\forall\,(x,y) \in \mathcal{H}^2, \quad |\,(x\,|\,y)\,| \leqslant \|x\|\,\|y\|.$$

Pour tout $(x,y) \in \mathcal{H}^2$, on a en effet

$$\begin{aligned}\|x+y\|^2 = (x+y \,|\, x+y) &= \|x\|^2 + 2\,(x \,|\, y) + \|y\|^2 \\ &\leqslant \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 \\ &= (\|x\| + \|y\|)^2.\end{aligned}$$

Le résultat suit en prenant la racine carrée car les deux membres sont positifs. □

On peut également caractériser le cas d'égalité dans l'inégalité triangulaire :

**Propriété – Cas d'égalité dans l'inégalité triangulaire**

Soit $(\cdot\,|\,\cdot)$ un produit scalaire sur $\mathcal{H}$ et $\|\cdot\|$ la norme associée. Pour tout $(x,y) \in \mathcal{H}^2$, on a l'équivalence :

$$\|x+y\| = \|x\| + \|y\| \quad \Leftrightarrow \quad \text{il existe } \alpha \in \mathbb{R}_+ \text{ tel que } x = \alpha\, y \text{ ou } y = \alpha\, x.$$

Démonstration – Si $y = \alpha\, x$ avec $\alpha \in \mathbb{R}_+$,

$$\|x+y\| = \|(1+\alpha)x\| = (1+\alpha)\|x\|,$$

et

$$\|x\| + \|y\| = \|x\| + \|\alpha x\| = \|x\| + \alpha\,\|x\| = (1+\alpha)\|x\|.$$

On procède de même si $x = \alpha\, y$ avec $\alpha \in \mathbb{R}_+$.

Réciproquement, si $\|x+y\| = \|x\| + \|y\|$, alors en reprenant l'inégalité de la démonstration précédente, on a

$$(x \,|\, y) = \|x\|\|y\|.$$

En particulier, il y a égalité dans l'inégalité de Cauchy-Schwarz, donc $x$ et $y$ sont colinéaires. Si $x$ est non nul, on peut écrire $y = \alpha\, x$ avec $\alpha \in \mathbb{R}$, et on a

$$(x \,|\, y) = \alpha\,(x \,|\, x) = \alpha\|x\|^2,$$

et

$$\|x\|\|y\| = |\alpha|\,\|x\|^2.$$

Sachant que $x \neq 0$, $\|x\| \neq 0$ donc $\alpha = |\alpha|$, c'est-à-dire que $\alpha \in \mathbb{R}_+$. Si $x = 0$, la relation $x = \alpha\, y$ est vérifiée avec $\alpha = 0$. □

Exemples

- La norme associée au produit scalaire canonique sur $\mathbb{R}^n$ est définie par

$$\forall\, x \in \mathbb{R}^n, \quad \|x\| = \left(\sum_{i=1}^{n} x_i^2\right)^{1/2}$$

Elle est appelée norme euclidienne **canonique** sur $\mathbb{R}^n$.

- La norme associée au produit scalaire défini sur $\mathscr{M}_n(\mathbb{R})^2$ par $(A \,|\, B) = \mathrm{Tr}\,({}^t\!AB)$ est donnée par :

$$\forall\, A = (a_{i,j}) \in \mathscr{M}_n(\mathbb{R}), \quad \|A\| = \big(\mathrm{Tr}({}^t\!AA\big)^{1/2} = \left(\sum_{i,j=1}^{n} (a_{i,j})^2\right)^{1/2}$$

- La norme associée au produit scalaire défini sur $\mathcal{C}^0([a,b],\mathbb{R})$ par $(f \,|\, g) = \int_a^b f(t)\, g(t)\, dt$ est donnée par :

$$\forall\, f \in \mathcal{C}^0([a,b],\mathbb{R}), \quad \|f\| = \left(\int_a^b f(x)^2\, dx\right)^{1/2}.$$

Le résultat suivant montre qu'une norme euclidienne provient d'un unique produit scalaire, que l'on peut retrouver à partir d'elle.

**Propriété – Identité de polarisation**

Soit $(\cdot\,|\,\cdot)$ un produit scalaire sur $\mathcal{H}$ et $\|\cdot\|$ la norme associée. Alors, pour tout $(x,y) \in \mathcal{H}^2$,

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

Démonstration – Pour $(x,y) \in \mathcal{H}^2$, on a, par bilinéarité et symétrie,

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

et de même

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

On en déduit facilement le premier résultat en retranchant la seconde égalité à la première, et le second résultat en utilisant la première égalité. □

Remarque – Pour tout $(x,y) \in \mathcal{H}^2$, on a en additionnant les deux égalités de la démonstration précédente,

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

Cette égalité est appelée identité du parallélogramme. Géométriquement, cette identité signifie que la somme des carrés des longueurs des diagonales d'un parallélogramme est égale à la somme des carrés de ses côtés.

# II. Orthogonalité

Dans cette partie, $(\mathcal{H}, (\cdot\,|\,\cdot))$ désigne un espace préhilbertien réel.

## 1. Familles orthogonales de vecteurs

**Définition**

- Si $x \in \mathcal{H}$, on dit que $x$ est **unitaire** (ou normé) si $\|x\| = 1$.
- Si $x$ et $y$ appartiennent à $\mathcal{H}$, on dit que $x$ et $y$ sont **orthogonaux** si $(x\,|\,y) = 0$.
- Si $(x_i)_{i \in I}$ est une famille de vecteurs de $\mathcal{H}$ ($I$ étant un ensemble d'indices), on dit que cette famille est :

– **normée** si pour tout $i \in I$, $\|x_i\| = 1$.
– **orthogonale** si pour tout $(i,j) \in I$ tel que $i \neq j$, $(x_i\,|\,x_j) = 0$.
– **orthonormale** (ou orthonormée) si elle est orthogonale et normée.
Ceci équivaut au fait que $(x_i\,|\,x_j) = \delta_{i,j}$ pour tout $(i,j) \in I^2$.

**Propriété**

Une famille orthogonale finie de vecteurs **tous non nuls** de $\mathcal{H}$ est libre.

Démonstration – Soit $(x_1, \ldots, x_p)$ une famille orthogonale de vecteurs tous non nuls de $\mathcal{H}$ et $(\lambda_1, \ldots, \lambda_p)$ une famille de scalaires telle que

$$\lambda_1 x_1 + \cdots + \lambda_p x_p = 0_{\mathcal{H}}.$$

Alors pour tout $i \in [\![1,p]\!]$,

$$\begin{aligned} & (x_i\,|\,\lambda_1 x_1 + \cdots + \lambda_p x_p) = 0 \\ \textit{i.e.}\quad & \lambda_1\,(x_i\,|\,x_1) + \cdots + \lambda_p\,(x_i\,|\,x_p) = 0, \end{aligned}$$

et donc $\lambda_i = 0$ car la famille est orthogonale et $x_i \neq 0$, d'où $(x_i\,|\,x_i) \neq 0$. □

Exemple – Définissons sur $\mathbb{R}$, pour tout $k \in \mathbb{N}$, la fonction $c_k : x \mapsto \cos(kx)$. Alors, pour tout $n \in \mathbb{N}$, la famille $(c_0, \dots, c_n)$ est libre dans $\mathcal{C}^0([0,2\pi],\mathbb{R})$, car est elle composée de vecteurs tous non nuls, et orthogonale pour le produit scalaire usuel sur $\mathcal{C}^0([0,2\pi],\mathbb{R})$. En effet, pour tous $p$ et $q$ distincts dans $\mathbb{N}$, on a $p - q \neq 0$ et $p + q \neq 0$, donc

$$\begin{aligned}\int_0^{2\pi} \cos(px)\cos(qx)\,dx &= \int_0^{2\pi} \frac{1}{2}(\cos((p+q)x) + \cos((p-q)x))\,dx \\ &= \frac{1}{2}\left[\frac{\sin((p+q)x)}{p+q} + \frac{\sin((p-q)x)}{p-q}\right]_0^{2\pi} = 0.\end{aligned}$$

**Théorème de Pythagore**

Soit $(x_1, \dots, x_p)$ une famille orthogonale de vecteurs de $\mathcal{H}$.
Alors
$$\|x_1 + \cdots + x_p\|^2 = \|x_1\|^2 + \cdots + \|x_p\|^2.$$

Démonstration – C'est immédiat puisque les termes $2\,(x_i \,|\, x_j)$ dans le développement de $\|x_1 + \cdots + x_p\|^2$ sont nuls par orthogonalité de la famille $(x_1, \dots, x_p)$. □

**Définition – Base orthonormée**

Soit $E$ un espace euclidien et $\mathcal{B} = (e_1, \dots, e_n)$ une famille de vecteurs de $E$.
On dit que $\mathcal{B}$ est une **base orthonormée** de $E$ si $\mathcal{B}$ est une base de $E$ et une famille orthonormale.

**Propriété – Calculs dans une base orthonormée**

Soit $(E, (\cdot\,|\,\cdot))$ un espace euclidien, et $\mathcal{B} = (e_1, \dots, e_n)$ une base orthonormée de $E$. Soient $x = x_1e_1 + \cdots + x_ne_n$ et $y = y_1e_1 + \cdots + y_ne_n$ deux vecteurs de $E$.
Alors :
$$(x\,|\,y) = \sum_{i=1}^{n} x_i\,y_i \quad \text{et} \quad \|x\| = \left(\sum_{i=1}^{n} |x_i|^2\right)^{1/2}.$$
Si $X = {}^t\begin{pmatrix} x_1 & \cdots & x_n\end{pmatrix}$ et $Y = {}^t\begin{pmatrix} y_1 & \cdots & y_n\end{pmatrix}$ sont les vecteurs-colonnes des coordonnées de $x$ et $y$ dans la base $\mathcal{B}$, on a (en identifiant une matrice de $\mathscr{M}_1(\mathbb{R})$ à son unique coefficient)
$$(x\,|\,y) = {}^tX\,Y \quad \text{et} \quad \|x\| = ({}^tX X)^{1/2}.$$

Démonstration – Il suffit de montrer le premier point. Or, par bilinéarité de $(\cdot\,|\,\cdot)$,

$$\begin{aligned}(x\,|\,y) &= (x_1e_1 + \cdots + x_ne_n \,|\, y_1e_1 + \cdots + y_ne_n) \\ &= \sum_{i,j=1}^{n} x_i\,y_j\,(e_i\,|\,e_j) \\ &= \sum_{i=1}^{n} x_i\,y_i,\end{aligned}$$

car la base $\mathcal{B}$ est orthonormée. □

Remarque – Dans $\mathscr{M}_{n,1}(\mathbb{R})$, l'expression du produit scalaire canonique entre deux vecteurs $X$ et $Y$ s'écrit simplement $(X\,|\,Y) = {}^tX\,Y$.

**Propriété – Matrice d'une application linéaire dans une base orthonormée**

Soit $(E, (\cdot\,|\,\cdot))$ un espace euclidien, $u \in \mathscr{L}(E)$ et $\mathcal{B} = (e_1, \ldots, e_n)$ une base orthonormée de $E$. Alors

$$\mathrm{Mat}_{\mathcal{B}}(u) = \big( (e_i \,|\, u(e_j)) \big)_{1\leqslant i,j\leqslant n}$$

Démonstration – Notons $a_{i,j}$ les coefficients de la matrice $\mathrm{Mat}_{\mathcal{B}}(u)$. Pour tout $j \in [\![1,n]\!]$, on a donc

$$u(e_j) = \sum_{k=1}^{n} a_{k,j} e_k.$$

Le produit scalaire $(e_i \,|\, u(e_j))$ est donc égal à

$$\left( e_i \,\middle|\, \sum_{k=1}^{n} a_{k,j} e_k \right) = \sum_{k=1}^{n} a_{k,j}\,(e_i \,|\, e_k) = a_{i,j}$$

car $\mathcal{B}$ est une famille orthonormée. D'où le résultat. □

Les résultats précédents montrent l'intérêt, pour la simplicité des calculs, de travailler dans des bases orthonormées. On va donc chercher à construire de telles bases orthonormées.

## 2. Orthonormalisation

**Théorème – Procédé d'orthonormalisation de Gram-Schmidt**

Soit $(e_1, \ldots, e_p)$ une famille libre de vecteurs de $\mathcal{H}$ et $F = \mathrm{Vect}(e_1, \ldots, e_p)$.
Alors il existe une base orthonormée $(\varepsilon_1, \ldots, \varepsilon_p)$ de $F$ telle que pour tout $i \in [\![1,p]\!]$,

$$\mathrm{Vect}(\varepsilon_1, \ldots, \varepsilon_i) = \mathrm{Vect}(e_1, \ldots, e_i).$$

Démonstration – On procède par récurrence sur $p$.

**Initialisation :** pour $p = 1$, on remarque que $e_1 \neq 0$ car la famille $(e_1)$ est libre. Il suffit alors de poser

$$\varepsilon_1 = \frac{e_1}{\|e_1\|}.$$

On a évidemment $\|\varepsilon_1\| = 1$ et $\mathrm{Vect}(\varepsilon_1) = \mathrm{Vect}(e_1)$.

**Hérédité :** supposons la propriété vraie pour un entier $p$ et considérons une famille libre $(e_1, \ldots, e_{p+1})$. Par hypothèse de récurrence, on peut supposer $\varepsilon_1, \ldots, \varepsilon_p$ construits.

*Analyse :* le vecteur $\varepsilon_{p+1}$ doit vérifier $\varepsilon_{p+1} \in \mathrm{Vect}(e_1, \ldots, e_{p+1}) = \mathrm{Vect}(\varepsilon_1, \ldots, \varepsilon_n, e_{p+1})$, donc il doit exister $(\lambda_1, \ldots, \lambda_{p+1}) \in \mathbb{R}^{p+1}$ tel que

$$\varepsilon_{p+1} = \lambda_1 \varepsilon_1 + \cdots + \lambda_p \varepsilon_p + \lambda_{p+1} e_{p+1}.$$

Alors pour tout $i \in [\![1,p]\!]$,

$$0 = (\varepsilon_i \,|\, \varepsilon_{p+1}) = \sum_{j=1}^{p} \lambda_j\ (\varepsilon_i \,|\, \varepsilon_j) + \lambda_{p+1}\,(\varepsilon_i \,|\, e_{p+1}) = \lambda_i + \lambda_{p+1}\,(\varepsilon_i \,|\, e_{p+1})\,,$$

car la famille $(\varepsilon_1, \ldots, \varepsilon_{p+1})$ doit être orthonormée. On en déduit que

$$\varepsilon_{p+1} = \lambda_{p+1} \left( e_{p+1} - \sum_{i=1}^{p} (\varepsilon_i \,|\, e_{p+1})\, \varepsilon_i \right).$$

*Synthèse :* on sait que $\mathrm{Vect}(\varepsilon_1, \ldots, \varepsilon_p) = \mathrm{Vect}(e_1, \ldots, e_p)$ ; de plus, la famille $e_1, \ldots, e_{p+1}$ étant libre, le vecteur $f_{p+1} = e_{p+1} - \sum_{i=1}^{p} (\varepsilon_i \,|\, e_{p+1})\, \varepsilon_i$ est non nul. On peut donc poser

$$\varepsilon_{p+1} = \frac{f_{p+1}}{\|f_{p+1}\|}.$$

Tout d'abord, la famille $(\varepsilon_1,\dots,\varepsilon_{p+1})$ est normée. Elle est également orthogonale : en effet, soit $(j,k)\in[\![1,p+1]\!]^2$ avec $j\neq k$. Si $j\leqslant p$ et $k\leqslant p$, alors $(\varepsilon_j\,|\,\varepsilon_k)=0$ par hypothèse de récurrence. Si $j\leqslant p$ et $k=p+1$, alors

$$\begin{aligned}(\varepsilon_j\,|\,\varepsilon_{p+1}) &= \frac{1}{\|f_{p+1}\|}\left(\varepsilon_j\,\middle|\,e_{p+1}-\sum_{i=1}^{p}(\varepsilon_i\,|\,e_{p+1})\,\varepsilon_i\right)\\ &= \frac{1}{\|f_{p+1}\|}\left((\varepsilon_j\,|\,e_{p+1})-\sum_{i=1}^{p}(\varepsilon_i\,|\,e_{p+1})\;(\varepsilon_j\,|\,\varepsilon_i)\right)\\ &= \frac{1}{\|f_{p+1}\|}\left((\varepsilon_j\,|\,e_{p+1})-(\varepsilon_j\,|\,e_{p+1})\right),\end{aligned}$$

car seul le terme correspondant à $i=j$ est éventuellement non nul, et $(\varepsilon_i\,|\,\varepsilon_i)=1$. Donc $(\varepsilon_j\,|\,\varepsilon_{p+1})=0$.

Ensuite, montrons que pour tout $i\in[\![1,p+1]\!]$,

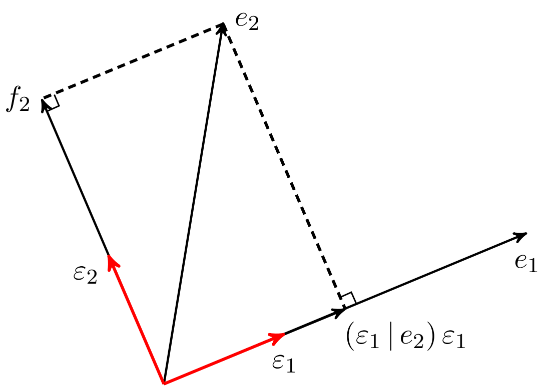
$$\mathrm{Vect}(\varepsilon_1,\dots,\varepsilon_i)=\mathrm{Vect}(e_1,\dots,e_i).$$

C'est vrai si $i\leqslant p$ par hypothèse de récurrence. Il suffit donc de le montrer pour $i=p+1$. Or

$$\varepsilon_{p+1}=\frac{1}{\|f_{p+1}\|}e_{p+1}+y$$

avec $y\in\mathrm{Vect}(\varepsilon_1,\dots,\varepsilon_p)=\mathrm{Vect}(e_1,\dots,e_p)$. On en déduit le résultat par double inclusion immédiate. □

Illustrons les différentes étapes de ce procédé dans le plan :




Remarques

• On peut aussi montrer que l'on peut imposer que $(\varepsilon_i\,|\,e_i)\in\mathbb{R}_+^*$ pour tout $i$. La famille $(\varepsilon_1,\dots,\varepsilon_p)$ est alors unique.

• Cette démonstration est constructive : elle donne un algorithme qui permet de construire explicitement une famille $(\varepsilon_1,\dots,\varepsilon_p)$. En particulier, elle est programmable sur ordinateur. En pratique, on pourra procéder ainsi : on remarque qu'à chaque étape, si $f_1,\dots,f_i$ sont construits, $f_{i+1}$ est de la forme

$$f_{i+1}=e_{i+1}+\lambda_i f_i+\cdots+\lambda_1 f_1$$

où $\lambda_1,\dots,\lambda_i$ sont des scalaires. Il suffit alors d'imposer les conditions

$$(f_{i+1}\,|\,f_1)=\cdots=(f_{i+1}\,|\,f_i)=0$$

pour déterminer ces scalaires. À la fin de la procédure, on pose alors $\varepsilon_i=f_i/\|f_i\|$ et l'on obtient une famille qui répond au problème. Avec cette façon de faire, on peut ainsi ne normer les vecteurs qu'à la fin de la procédure, ce qui évite des erreurs de calculs.

On peut procéder de même en cherchant $f_{i+1}$ sous la forme $e_{i+1}+\mu_i e_i+\cdots+\mu_1 e_1$, car $\mathrm{Vect}(e_1,\dots,e_i)=\mathrm{Vect}(f_1,\dots,f_i)$.

Exemple – Soit $\mathcal{B} = (e_0,e_1,e_2)$ la base canonique de $\mathbb{R}_2[X]$, muni du produit scalaire défini par

$$(P \mid Q) = \int_0^1 P(x)\, Q(x)\, dx.$$

Orthonormalisons la base $\mathcal{B}$.

- On pose $f_0 = e_0 = 1$.
- On choisit $f_1$ de la forme $f_1 = e_1 + \alpha f_0$ ($\alpha$ réel) de sorte que $(f_1 \mid f_0) = 0$, ce qui équivaut à

$$\int_0^1 (t+\alpha)\, dt = 0 \quad \Leftrightarrow \quad \alpha = -\frac{1}{2}.$$

On pose donc $f_1 = X - \dfrac{1}{2}$.

- On choisit $f_2$ de la forme $f_2 = e_2 + \beta f_1 + \gamma f_0$ ($\beta$ et $\gamma$ réels) de sorte que $(f_2 \mid f_0) = 0$ et $(f_2 \mid f_1) = 0$, ce qui équivaut à

$$\left\{\begin{array}{r} \displaystyle\int_0^1 \left(t^2 + \beta\left(t - \frac{1}{2}\right) + \gamma\right) dt = 0 \\ \displaystyle\int_0^1 \left(t^2 + \beta\left(t - \frac{1}{2}\right) + \gamma\right)\left(t - \frac{1}{2}\right) dt = 0 \end{array}\right.$$

$$i.e.\text{ à } \quad \left\{\begin{array}{c} \dfrac{1}{3} + \gamma = 0 \\ \dfrac{1}{4} - \dfrac{1}{6} + \beta\dfrac{1}{12} = 0 \end{array}\right. \quad i.e.\text{ à } \quad \left\{\begin{array}{l} \beta = -1 \\ \gamma = -\dfrac{1}{3} \end{array}\right.$$

On pose donc $f_2 = X^2 - f_1 - \dfrac{1}{3} f_0 = X^2 - X + \dfrac{1}{6}$.

- On norme enfin les vecteurs $f_0, f_1$ et $f_2$ :

$$\begin{aligned} \|f_0\| &= 1 \\ \|f_1\| &= \left(\int_0^1 \left(t - \frac{1}{2}\right)^2 dt\right)^{1/2} = \sqrt{\frac{1}{12}} = \frac{1}{2\sqrt{3}} \\ \|f_2\| &= \left(\int_0^1 \left(t^2 - t + \frac{1}{6}\right)^2 dt\right)^{1/2} = \frac{1}{\sqrt{180}}. \end{aligned}$$

On obtient une famille $(\varepsilon_0,\varepsilon_1,\varepsilon_2)$ qui convient.

**Corollaire**

Soit $(E, (\cdot \mid \cdot))$ un espace euclidien.
Il existe des bases orthonormées de $E$.
Toute famille orthonormale de $E$ peut être complétée en une base orthonormée de $E$.

Démonstration – Pour le premier point, il suffit d'appliquer le procédé d'othonormalisation de Gram-Schmidt à une base quelconque de $E$. On obtient alors une famille génératrice de $E$ et libre (car orthonormale), c'est-à-dire une base de $E$. Pour le second, on sait que toute famille orthonormale est libre, on peut la compléter en une base de $E$ puis orthonormaliser cette base par le procédé de Gram-Schmidt, ce qui ne modifie pas la famille initiale. $\square$

## 3. Sommes orthogonales

**Définition – Sous-espaces orthogonaux**

Soient $F$ et $G$ deux sous-espaces vectoriels de $\mathcal{H}$.
On dit que $F$ et $G$ sont **orthogonaux** si :

$$\forall\,(x,y) \in F \times G, \quad (x\,|\,y) = 0.$$

Ceci se note également : $F \perp G$.

**Propriété**

Soient $F_1,\dots,F_p$ des sous-espaces vectoriels de $\mathcal{H}$, deux à deux orthogonaux.
Alors la somme $F_1 + \cdots + F_p$ est directe.

Démonstration – Soit $(x_1,\dots,x_p) \in F_1 \times \cdots \times F_p$ tel que $x_1 + \cdots + x_p = 0$. En faisant le produit scalaire de cette expression avec $x_i$ pour $i \in [\![1,p]\!]$, on obtient

$$0 = (x_i\,|\,x_1 + \cdots + x_p) = (x_i\,|\,x_1) + \cdots + (x_i\,|\,x_p)\,.$$

Les $F_j$ étant deux à deux orthogonaux, on en déduit que $(x_i\,|\,x_i) = 0$ et donc $x_i = 0$, et ce pour tout $i$. D'où le résultat. □

**Définition**

• Soient $F_1,\dots,F_p$ des sous-espaces vectoriels de $\mathcal{H}$, deux à deux orthogonaux.
La somme $\bigoplus_{i=1}^{p} F_i$ est appelée **somme directe orthogonale** des $F_i$ (on dit aussi que les $F_i$ sont en somme directe orthogonale).
• Soient $F$ et $G$ deux sous-espaces vectoriels de $\mathcal{H}$.
On dit que $F$ et $G$ sont **supplémentaires orthogonaux** si $F \perp G$ et $F \oplus G = \mathcal{H}$.
Ceci se note parfois $\mathcal{H} = F \oplus^{\perp} G$.

Remarques

• Soient $F_1,\dots,F_p$ des sous-espaces vectoriels deux à deux orthogonaux d'un espace euclidien $E$. Alors leur somme est directe, donc d'après un résultat du chapitre **Espaces vectoriels et applications linéaires**, on a

$$\dim(F_1 \oplus \cdots \oplus F_p) = \dim(F_1) + \cdots + \dim(F_p)$$

et pour que $E = F_1 \oplus \cdots \oplus F_p$, il faut et il suffit que

$$\dim(E) = \dim(F_1) + \cdots + \dim(F_p).$$

• Si $F$ et $G$ sont deux sous-espaces vectoriels de $\mathcal{H}$, pour montrer que $F$ et $G$ sont supplémentaires orthogonaux, il suffit de montrer que $F \perp G$ et $\mathcal{H} \subset F + G$. En effet, d'après la propriété précédente, si $F \perp G$, l'aspect direct de la somme $F + G$ est acquis (notamment, $F \cap G = \{0_E\}$).

## 4. Orthogonal d'un sous-espace vectoriel

**Propriété/Définition**

Soit $F$ un sous-espace vectoriel de $\mathcal{H}$. On appelle **orthogonal** de $F$ l'ensemble

$$F^{\perp} = \{y \in \mathcal{H};\ \forall\,x \in F,\ (x\,|\,y) = 0\}.$$

C'est un sous-espace vectoriel de $\mathcal{H}$, orthogonal à $F$.

Démonstration – On a $F^\perp \subset \mathcal{H}$, et le vecteur nul de $\mathcal{H}$ est orthogonal à tout vecteur donc appartient à $F^\perp$. Si $y$ et $z$ appartiennent à $F^\perp$ et si $\lambda \in \mathbb{R}$, alors pour tout $x \in F$,

$$(x \,|\, \lambda y + z) = \lambda \,(x \,|\, y) + (x \,|\, z) = 0$$

donc $\lambda y + z \in F^\perp$. Ainsi $F^\perp$ est un sous-espace vectoriel de $\mathcal{H}$. Il est orthogonal à $F$, car par définition, si $x \in F$ et $y \in F^\perp$, $(x \,|\, y) = 0$. □

Exemple – Dans $\mathbb{R}^n$ ($n \geqslant 1$) muni du produit scalaire canonique, soit $a = (a_1, \ldots, a_n)$ un vecteur non nul. Alors $\mathrm{Vect}(a)^\perp$ est l'ensemble des vecteurs $x = (x_1, \ldots, x_n)$ tels que

$$\sum_{i=1}^{n} a_i \, x_i = 0.$$

Il s'agit du noyau de la forme linéaire $\varphi$ définie sur $\mathbb{R}^n$ par

$$\varphi(x_1, \ldots, x_n) = \sum_{i=1}^{n} a_i \, x_i,$$

qui est non nulle car $a$ est non nul. En particulier, $\mathrm{Vect}(a)^\perp$ est un hyperplan de $\mathbb{R}^n$.

Remarque – Si $F$ et $G$ sont deux sous-espaces vectoriels de $\mathcal{H}$, alors on a les équivalences

$$F \perp G \quad \Leftrightarrow \quad F \subset G^\perp \quad \Leftrightarrow \quad G \subset F^\perp.$$

Par contre, lorsque $F \perp G$ on n'a pas toujours les égalités $F = G^\perp$ et $G = F^\perp$.

**Propriété**

On a $\mathcal{H}^\perp = \{0_\mathcal{H}\}$ et $\{0_\mathcal{H}\}^\perp = \mathcal{H}$.

Démonstration – En effet, si $y \in \mathcal{H}$ vérifie $(x \,|\, y) = 0$ pour tout $x \in \mathcal{H}$, alors pour le choix de $x = y$ on obtient $(y \,|\, y) = 0$ et donc $y = 0_\mathcal{H}$. L'autre inclusion (et la seconde égalité) vient simplement du fait que $(x \,|\, 0_\mathcal{H}) = 0$ pour tout $x \in \mathcal{H}$. □

Remarque – Soient $x$ et $y$ deux éléments de $\mathcal{H}$ tels que pour tout $z \in \mathcal{H}$, $(x \,|\, z) = (y \,|\, z)$. Alors $x = y$.

En effet, l'hypothèse entraîne que $(x - y \,|\, z) = 0$ pour tout $z \in \mathcal{H}$, et donc $x - y \in \mathcal{H}^\perp = \{0_\mathcal{H}\}$. D'où le résultat.

**Propriété**

Soit $F$ un sous-espace vectoriel de dimension finie de $\mathcal{H}$ et $(e_1, \ldots e_p)$ une famille génératrice de $F$.

Pour tout $x \in \mathcal{H}$, on a l'équivalence :

$$x \in F^\perp \quad \Leftrightarrow \quad \forall\, i \in [\![1,p]\!],\ (e_i \,|\, x) = 0.$$

Démonstration – Si $x \in F^\perp$, alors pour tout $i \in [\![1,p]\!]$, $(e_i \,|\, x) = 0$, car $e_i \in F$. Réciproquement, si $(e_i \,|\, x) = 0$ pour tout $i \in [\![1,p]\!]$, alors pour tout $(\lambda_1, \ldots, \lambda_p) \in \mathbb{R}^p$,

$$\sum_{i=1}^{p} \lambda_i \,(e_i \,|\, x) = 0, \quad i.e. \quad \left( \sum_{i=1}^{n} \lambda_i e_i \,\middle|\, x \right) = 0,$$

par linéarité à gauche de $(\cdot \,|\, \cdot)$. Comme $F = \mathrm{Vect}(e_1, \ldots, e_p)$, on a bien $x \in F^\perp$. □

**Propriété**

Soit $F$ un sous-espace vectoriel de $\mathcal{H}$. Alors :
- $F \subset (F^\perp)^\perp$.
- $F$ et $F^\perp$ sont en somme directe orthogonale. En particulier, $F \cap F^\perp = \{0_\mathcal{H}\}$.
- Si $G$ est un supplémentaire orthogonal de $F$, alors $G = F^\perp$.

Démonstration

- Soit $x \in F$. Alors, pour tout $y \in F^\perp$, $(x \,|\, y) = 0$, donc $x \in (F^\perp)^\perp$.
- C'est une conséquence de la propriété du paragraphe précédent, car $F \perp F^\perp$.
- Soit $G$ un supplémentaire orthogonal de $F$. Montrons que $G = F^\perp$. Tout d'abord, $F \perp G$ donc $G \subset F^\perp$. Réciproquement, soit $x \in F^\perp$. On peut décomposer $x$ sous la forme $y + z$ avec $y \in F$ et $z \in G$. Alors $y = x - z \in F^\perp$ car $x \in F^\perp$ et $z \in G \subset F^\perp$. Donc $y \in F \cap F^\perp = \{0_\mathcal{H}\}$ et $x = z \in G$. D'où l'égalité $G = F^\perp$. Ainsi, $F$ a au plus un supplémentaire orthogonal, qui ne peut être que $F^\perp$. $\square$

Remarque – Il est important de remarquer que l'inclusion réciproque du premier point est fausse en général. Par exemple, soit $\mathcal{H} = \mathcal{C}^0([0,1],\mathbb{R})$ muni du produit scalaire usuel défini par

$$(f \,|\, g) = \int_0^1 f(t)g(t)\,dt.$$

Considérons le sous-espace vectoriel $F = \{f \in E;\ f(0) = 0\}$ de $\mathcal{H}$. Soit $f \in F^\perp$ ; alors, la fonction $g : t \mapsto t\,f(t)$ étant un élément de $F$, on a $(f \,|\, g) = 0$, *i.e.*

$$\int_0^1 t\,f(t)^2\,dt = 0.$$

La fonction $t \mapsto t\,f(t)^2$ étant de plus continue et positive, elle est nulle, donc $f(t) = 0$ pour tout $t \in ]0,1]$. Par continuité de $f$, on a également $f(0) = 0$, et finalement, $f = 0_\mathcal{H}$. On en déduit que $F^\perp = \{0_\mathcal{H}\}$. Ainsi, dans ce cas, on a $(F^\perp)^\perp = \{0_\mathcal{H}\}^\perp = \mathcal{H} \neq F$.

On remarque également que la somme $F \oplus F^\perp$ n'est pas toujours égale à $\mathcal{H}$ : dans l'exemple précédent, on a $F \oplus F^\perp = F \neq \mathcal{H}$. En général, $F$ et $F^\perp$ ne sont donc pas toujours supplémentaires orthogonaux.

En revanche, les résultats sont vrais lorsque $F$ est de dimension finie :

**Théorème – Supplémentaire orthogonal d'un sous-espace de dimension finie**

Soit $F$ un sous-espace vectoriel **de dimension finie** de $\mathcal{H}$. Alors :
- $\mathcal{H} = F \oplus F^\perp$.
- $(F^\perp)^\perp = F$.

Remarque – D'après le premier point, si $F$ est un sous-espace vectoriel de dimension finie de $\mathcal{H}$, $F^\perp$ est un supplémentaire orthogonal de $F$, et on sait d'après la propriété précédente que c'est alors l'unique supplémentaire orthogonal de $F$.

Démonstration

- On sait déjà que la somme est directe, il suffit de montrer que $\mathcal{H} \subset F + F^\perp$. Soit $\mathcal{B} = (\varepsilon_1, \ldots, \varepsilon_n)$ une base orthonormée de $F$ (qui existe d'après le procédé de Gram-Schmidt). Pour tout $x \in \mathcal{H}$, on cherche à écrire $x = y + z$ avec $y \in F$ et $z \in F^\perp$.

*Analyse :* supposons qu'une telle décomposition existe, et soit $y = \sum_{i=1}^n \lambda_i \varepsilon_i$ la décomposition de $y$ dans la base $\mathcal{B}$. Alors $x - y = z \in F^\perp$, donc pour tout $j \in [\![1,n]\!]$, $(\varepsilon_j \,|\, x - y) = 0$, c'est-à-dire

$$(\varepsilon_j \,|\, x) = (\varepsilon_j \,|\, y) = \sum_{i=1}^n \lambda_i\,(\varepsilon_j \,|\, \varepsilon_i) = \lambda_j$$

car $\mathcal{B}$ est orthonormée. Ainsi $y$ est nécessairement donné par : $y = \sum_{i=1}^{n} (\varepsilon_i \,|\, x)\, \varepsilon_i$.

*Synthèse :* définissons donc $y$ par cette formule. Alors $y \in F$ et $x - y \in F^\perp$ car pour tout $j \in [\![1,n]\!]$, $(\varepsilon_j \,|\, x - y) = 0$ en reprenant le calcul précédent. On a donc bien la décomposition souhaitée avec $z = x - y$.

• Le premier point montre que $F^\perp$ a un supplémentaire orthogonal, à savoir $F$. Le dernier point de la propriété précédente (appliqué avec $F^\perp$ à la place de $F$ et $F$ à la place de $G$) montre alors que $F = (F^\perp)^\perp$. $\square$

**Théorème/Définition – Projection orthogonale**

Soit $F$ un sous-espace vectoriel de dimension finie de $\mathcal{H}$.

La projection $p_F$ sur $F$ parallèlement à $F^\perp$ est bien définie car $\mathcal{H} = F \oplus F^\perp$, elle est appelée **projection orthogonale** sur $F$.

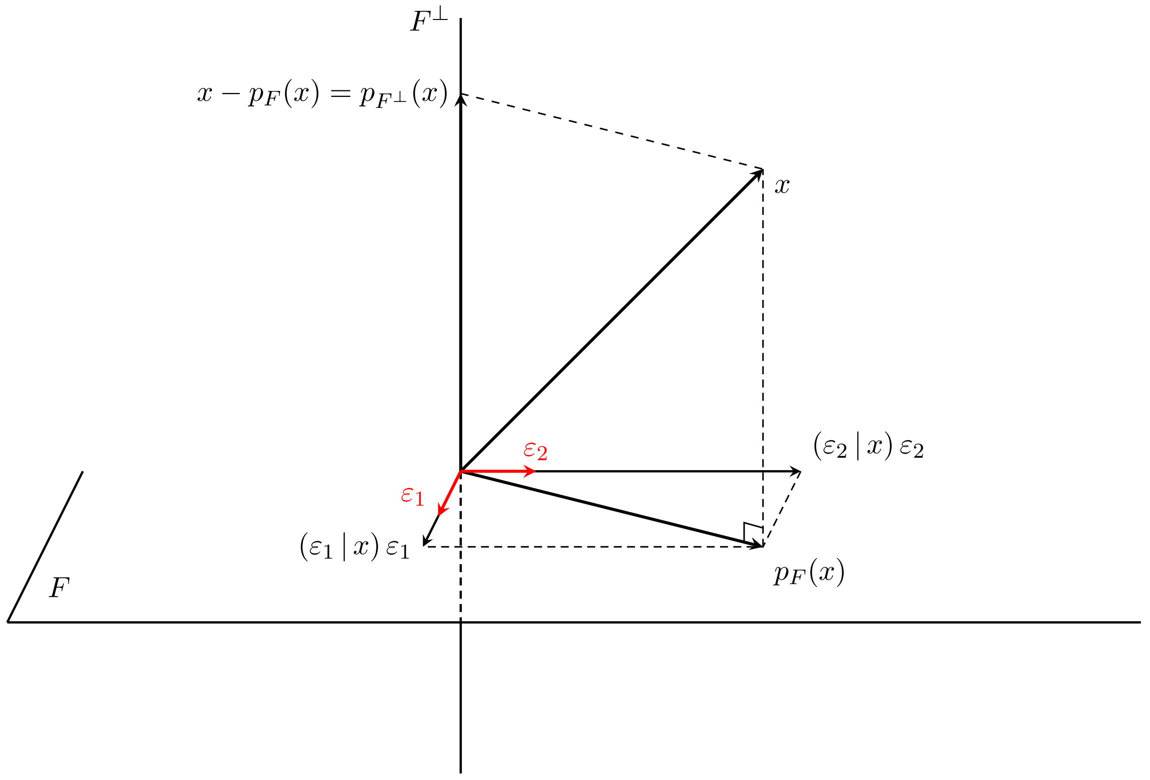
Si $(\varepsilon_1, \dots, \varepsilon_n)$ est une base orthonormée de $F$, alors pour tout $x \in \mathcal{H}$,

$$p_F(x) = \sum_{i=1}^{n} (\varepsilon_i \,|\, x)\ \varepsilon_i.$$

Le vecteur $p_F(x)$ est appelé le **projeté orthogonal** de $x$ sur $F$.

Démonstration – La formule donnant $p_F(x)$ a été démontrée dans le théorème précédent. $\square$

Voici une illustration de la situation :




Remarques

• Pour déterminer le projeté orthogonal de $x$ sur $F$, il n'est pas nécessaire de disposer d'une base *orthonormée* de $F$. En effet, il suffit de remarquer que $p_F(x)$ est entièrement caractérisé par : $p_F(x) \in F$ et $x - p_F(x) \in F^\perp$. Si l'on dispose d'une famille génératrice quelconque $(e_1, \dots, e_p)$ de $F$, alors d'après une propriété précédente, $x - p_F(x) \in F^\perp$ si et seulement si

$$\forall\, i \in [\![1,p]\!], \quad (e_i \,|\, x - p_F(x)) = 0,$$

ce qui peut s'écrire comme un système linéaire dont les inconnues sont les scalaires d'une décomposition de $p_F(x)$ sur la famille $(e_1,\ldots,e_p)$.

En revanche, pour que la formule explicite de $p_F(x)$ de la propriété précédente soit vraie, il est essentiel que $(\varepsilon_1,\ldots,\varepsilon_n)$ soit une base orthonormée de $F$.

• Si $F$ est un sous-espace vectoriel de dimension finie de $\mathcal{H}$, on appelle symétrie orthogonale par rapport à $F$ la symétrie $s_F$ par rapport à $F$, parallèlement à $F^\perp$. On a la relation $\mathrm{Id}_{\mathcal{H}} + s_F = 2p_F$. Si $E$ est euclidien et $F$ est un hyperplan de $E$, on dit que $s_F$ est la réflexion par rapport à $F$.

**Propriété**

Soit $\mathcal{B} = (\varepsilon_1,\ldots,\varepsilon_n)$ une base orthonormée d'un espace euclidien $(E,(\cdot\,|\,\cdot))$.
Alors, la décomposition d'un vecteur $x \in E$ dans la base $\mathcal{B}$ est

$$x = \sum_{i=1}^{n} (\varepsilon_i \,|\, x)\ \varepsilon_i.$$

Démonstration – C'est une conséquence immédiate de la formule du théorème précédent, avec le choix particulier de $F = E$; dans ce cas, bien sûr, le projeté orthogonal de $x$ sur $E$ est $x$ lui-même. □

Remarque – En particulier, pour tout $(x,y) \in E^2$,

$$(x \,|\, y) = \sum_{i=1}^{n} (\varepsilon_i \,|\, x)\,(\varepsilon_i \,|\, y) \quad \text{et} \quad \|x\| = \left(\sum_{i=1}^{n} (\varepsilon_i \,|\, x)^2\right)^{1/2}.$$

**Théorème**

• Soit $F$ un sous-espace vectoriel d'un espace euclidien $E$.
Alors $E = F \oplus F^\perp$. En particulier,

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

• Si $F$ et $G$ sont deux sous-espaces vectoriels de $E$, pour que $F$ et $G$ soient supplémentaires orthogonaux, il faut et il suffit que

$$F \perp G \quad \text{et} \quad \dim(F) + \dim(G) = \dim(E).$$

Démonstration

• On a montré que le résultat $E = F \oplus F^\perp$ est toujours vrai si $F$ est de dimension finie, ce qui est le cas dans la situation présente, $E$ étant de dimension finie. La formule des dimensions vient de la première remarque du paragraphe précédent.

• Cela vient aussi de la première remarque du paragraphe précédent. □

Remarque – Si $F$ est un sous-espace vectoriel de $E$, $F$ et $F^\perp$ sont de dimension finie, les projections orthogonales sur $F$ et $F^\perp$ sont bien définies et on a la relation $p_F + p_{F^\perp} = \mathrm{Id}$, c'est-à-dire que pour tout $x \in E$, $x - p_F(x) = p_{F^\perp}(x)$.

## III. Distance

Étant donné un vecteur $x$ de $\mathcal{H}$ et $F$ un sous-espace vectoriel de $\mathcal{H}$ de dimension finie, on cherche un vecteur de $F$ qui soit le plus proche de $x$ au sens de la distance associée au produit scalaire $(\cdot\,|\,\cdot)$ sur $\mathcal{H}$.

**Théorème/Définition**

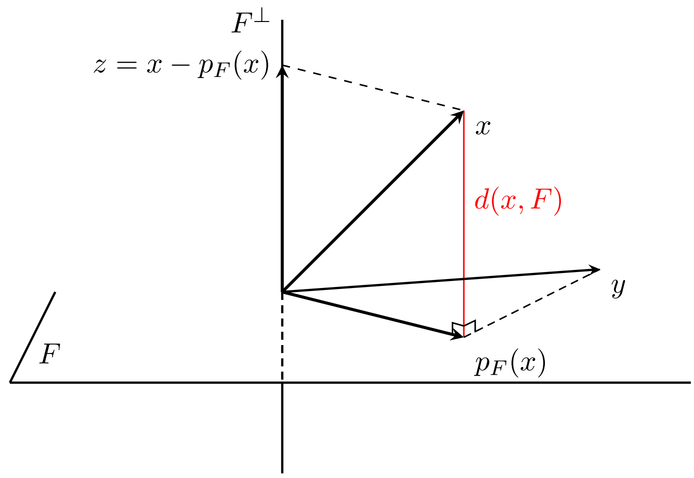
Soit $x \in \mathcal{H}$ et $F$ un sous-espace vectoriel de dimension finie de $\mathcal{H}$.
Alors la fonction
$$\left\{ \begin{array}{ccc} F & \to & \mathbb{R}_+ \\ y & \mapsto & \|x-y\| \end{array} \right.$$
a un minimum sur $F$, qui est atteint en $p_F(x)$ et uniquement en ce point.
Autrement dit, il existe un unique vecteur $y_0$ de $F$ tel que

$$\|x - y_0\| = \min_{y \in F} \|x-y\|,$$

et ce vecteur est $p_F(x)$.
Le réel positif $\|x - p_F(x)\|$ est appelé **distance** de $x$ à $F$, noté $d(x,F)$ :

$$d(x,F) = \|x - p_F(x)\| = \min_{y \in F} \|x - y\|.$$

Démonstration – Voici une illustration de la situation et de la démonstration :




Comme $F$ est de dimension finie, on sait que $\mathcal{H} = F \oplus F^\perp$. On peut donc écrire $x = p_F(x) + z$ avec $p_F(x) \in F$ et $z \in F^\perp$. Alors pour tout $y \in F$, $p_F(x) - y \in F$ et donc $x - p_F(x) = z$ est orthogonal à $p_F(x) - y$. D'après le théorème de Pythagore, on a donc

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

avec égalité si et seulement si $\|p_F(x) - y\|^2 = 0$ c'est-à-dire $y = p_F(x)$. $\square$

**Propriété**

Soit $F$ un sous-espace vectoriel de dimension finie de $\mathcal{H}$ et $(\varepsilon_1, \ldots, \varepsilon_n)$ une base orthonormée de $F$.
La distance de $x$ à $F$ est donnée par les formules

$$d(x,F)^2 = \|x\|^2 - \|p_F(x)\|^2 = \|x\|^2 - \sum_{i=1}^{n} (\varepsilon_i \,|\, x)^2 .$$

Démonstration – La famille $(\varepsilon_1, \ldots, \varepsilon_n)$ est une base orthonormée de $F$, donc pour tout $x \in \mathcal{H}$, on connaît l'expression explicite de $p_F(x)$ :

$$p_F(x) = \sum_{i=1}^{n} (\varepsilon_i \,|\, x)\, \varepsilon_i,$$

et on a également

$$\|p_F(x)\|^2 = \left\| \sum_{i=1}^{n} (\varepsilon_i \,|\, x)\ \varepsilon_i \right\|^2 = \sum_{i=1}^{n} (\varepsilon_i \,|\, x)^2 .$$

De plus, les vecteurs $p_F(x)$ et $x - p_F(x)$ sont orthogonaux, donc d'après le théorème de Pythagore,

$$\|x\|^2 = \|p_F(x) + (x - p_F(x))\|^2 = \|p_F(x)\|^2 + \|x - p_F(x)\|^2 = \|p_F(x)\|^2 + d(x,F)^2.$$

On en déduit les deux formules. □

**Corollaire – Inégalité de Bessel**

Soit $F$ un sous-espace vectoriel de dimension finie de $\mathcal{H}$ et $(\varepsilon_1, \ldots, \varepsilon_n)$ une base orthonormée de $F$.

Pour tout $x \in \mathcal{H}$, on a

$$\|p_F(x)\| \leqslant \|x\|.$$

Démonstration – En effet, la différence $\|x\|^2 - \|p_F(x)\|^2$ est égale à $d(x,F)^2 \geqslant 0$. □

Exemple – Déterminons le polynôme de degré au plus 2 qui soit le plus proche de $X^3$ au sens de la norme associée au produit scalaire défini sur $\mathbb{R}[X]$ par

$$(P \,|\, Q) = \int_0^1 P(x)\, Q(x)\, dx.$$

Nous avons déterminé ci-dessus une base orthonormée $(\varepsilon_0, \varepsilon_1, \varepsilon_2)$ de $\mathbb{R}_2[X]$ pour ce produit scalaire. L'unique polynôme qui répond au problème est le projeté orthogonal de $X^3$ sur $\mathbb{R}_2[X]$, c'est-à-dire le polynôme

$$\begin{aligned}
P(X) &= \left(\varepsilon_0 \,|\, X^3\right) \varepsilon_0 + \left(\varepsilon_1 \,|\, X^3\right) \varepsilon_1 + \left(\varepsilon_2 \,|\, X^3\right) \varepsilon_2 \\
&= \left( \int_0^1 t^3\, dt \right) + \left(2\sqrt{3}\right)^2 \left( \int_0^1 \left(t - \frac{1}{2}\right) t^3\, dt \right) \left(X - \frac{1}{2}\right) \\
&\quad + 180 \left( \int_0^1 \left(t^2 - t + \frac{1}{6}\right) t^3\, dt \right) \left(X^2 - X + \frac{1}{6}\right) \\
&= \frac{1}{4} + 12 \left(\frac{1}{5} - \frac{1}{2} \times \frac{1}{4}\right) \left(X - \frac{1}{2}\right) + 180 \left(\frac{1}{6} - \frac{1}{5} + \frac{1}{6} \times \frac{1}{4}\right) \left(X^2 - X + \frac{1}{6}\right).
\end{aligned}$$

Après simplifications, on obtient $P(X) = \dfrac{3}{2}X^2 - \dfrac{3}{5}X + \dfrac{1}{20}$.

Comme indiqué dans une remarque de la partie précédente, on peut aussi déterminer $P(X)$ en résolvant le système

$$\begin{cases}
\left(X^3 - aX^2 - bX - c \,|\, 1\right) = 0 \\
\left(X^3 - aX^2 - bX - c \,|\, X\right) = 0 \\
\left(X^3 - aX^2 - bX - c \,|\, X^2\right) = 0
\end{cases}$$

ce qui aboutit bien sûr à la même valeur de $P(X)$, et ne nécessite pas de disposer de la famille $(\varepsilon_0, \varepsilon_1, \varepsilon_2)$.

On peut alors déterminer la distance de $X^3$ à $\mathbb{R}_2[X]$, c'est-à-dire la racine carrée de la quantité

$$\inf_{(a,b,c)\in\mathbb{R}^3} \int_0^1 \left(t^3 - at^2 - bt - c\right)^2 \, dt;$$

en effet cette borne inférieure est un minimum, qui est atteint pour $a = \dfrac{3}{2}$, $b = -\dfrac{3}{5}$ et $c = \dfrac{1}{20}$ et uniquement pour ces valeurs.

D'après la propriété ci-dessus, on peut également calculer cette valeur en utilisant la formule

$$\inf_{(a,b,c)\in\mathbb{R}^3} \int_0^1 \left(t^3 - at^2 - bt - c\right)^2 \, dt = \|X^3\|^2 - \sum_{i=0}^{2} \left(\varepsilon_i \,|\, X^3\right)^2.$$

## IV. Formes linéaires sur un espace euclidien

Dans cette partie, $(E, (\cdot \,|\, \cdot))$ désigne un espace euclidien.

**Théorème – Représentation des formes linéaires sur un espace euclidien**

Soit $f$ une forme linéaire sur $E$. Alors il existe un unique vecteur $a \in E$ tel que :

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

On dit parfois que le vecteur $a$ représente $f$ *via* le produit scalaire $(\cdot \,|\, \cdot)$.

Démonstration – Soit $\mathcal{B} = (e_1, \ldots, e_n)$ une base orthonormée de $E$, et soit $x = x_1e_1 + \cdots + x_ne_n$ un vecteur de $E$. Alors

$$f(x) = \sum_{i=1}^{n} x_i f(e_i),$$

qui est le produit scalaire entre $x$ et le vecteur $a = f(e_1)e_1 + \cdots + f(e_n)e_n$ car $\mathcal{B}$ est orthonormée. Ceci prouve l'existence de $a$.

Quant à l'unicité, supposons que deux vecteurs $a$ et $b$ vérifient, pour tout $x \in E$,

$$f(x) = (a \,|\, x) = (b \,|\, x)\,.$$

Alors, pour tout $x \in E$, $(a - b \,|\, x) = 0$ et donc $a - b \in E^\perp = \{0_E\}$. On en déduit que $a = b$. □

Remarque – Réciproquement, si $a \in E$, l'application $x \mapsto (a \,|\, x)$ est linéaire, par linéarité à droite du produit scalaire. Le résultat précédent signifie donc que dans un espace euclidien, on sait décrire entièrement les formes linéaires : il s'agit exactement des applications de la forme $x \mapsto (a \,|\, x)$ où $a$ est un vecteur de $E$, chaque forme linéaire $f$ sur $E$ étant associée à un unique vecteur $a$.

Exemples

- Dans le cas de la forme linéaire définie sur $\mathbb{R}^3$ (muni du produit scalaire canonique) par $f(x,y,z) = x + 2y + 3z$, $a$ est le vecteur $(1,2,3)$.
- Les formes linéaires sur $\mathscr{M}_n(\mathbb{R})$ sont exactement les applications de la forme $M \mapsto \mathrm{Tr}(AM)$ où $A \in \mathscr{M}_n(\mathbb{R})$.

**Propriété/Définition – Vecteur normal à un hyperplan**

Soit $H$ un hyperplan de $E$ et $f$ une forme linéaire non nulle sur $E$ telle que $H = \mathrm{Ker}(f)$. Il existe $a \in E$ non nul tel que $f : x \mapsto (a \,|\, x)$. Ainsi, pour $x \in E$, on a l'équivalence

$$x \in H \quad \Leftrightarrow \quad (a \,|\, x) = 0.$$

On dit que $a$ est **un vecteur normal** à $H$.
Avec les notations précédentes, en notant $a = a_1 e_1 + \cdots + a_n e_n$ la décomposition de $a$ dans la base orthonormée $\mathcal{B}$, on a

$$(a \,|\, x) = a_1 x_1 + \cdots + a_n x_n.$$

Ainsi, $H$ a pour équation

$$a_1 x_1 + \cdots + a_n x_n = 0$$

dans la base $\mathcal{B}$.

Remarque – Avec les notations précédentes, l'ensemble des formes linéaires caractérisant $H$ est $\mathrm{Vect}(f) \setminus \{0\}$. De la même façon, l'ensemble des vecteurs normaux à $H$ est $\mathrm{Vect}(a) \setminus \{0\}$. Il est en effet évident que pour tout $\lambda \in \mathbb{R}^*$, $\lambda f$ est représentée par le vecteur $\lambda a$. Les équations de $H$ sont donc exactement les équations $(\lambda a \,|\, x) = 0$ où $\lambda \in \mathbb{R}^*$.

Si l'on travaille dans une base orthonormée $\mathcal{B} = (e_1, \ldots, e_n)$, et si $H$ a pour équation

$$a_1 x_1 + \cdots + a_n x_n = 0$$

dans la base $\mathcal{B}$, les vecteurs normaux à $H$ sont exactement les vecteurs $\lambda(a_1 e_1 + \cdots + a_n e_n)$ où $\lambda \in \mathbb{R}^*$.

**Propriété – Distance d'un vecteur à un hyperplan ou une droite**

• Soit $H$ un hyperplan de $E$ et $a$ un vecteur normal à $H$. Alors, pour tout $x \in E$, la distance de $x$ à $H$ est donnée par

$$d(x,H) = \frac{|\,(x \,|\, a)\,|}{\|a\|}.$$

• Soit $D$ une droite vectorielle de $E$ et $a$ un vecteur non nul de $D$. Alors, pour tout $x \in E$, la distance de $x$ à $D$ est donnée par

$$d(x,D) = \sqrt{\|x\|^2 - \frac{(x \,|\, a)^2}{\|a\|^2}}.$$

Démonstration
• La distance de $x$ à $H$ est donnée par $d(x,H) = \|x - p_H(x)\|$, le vecteur $p_H(x)$ étant entièrement caractérisé par : $p_H(x) \in H$ et $x - p_H(x) \in H^\perp = \mathrm{Vect}(a)$. Ainsi, $p_H(x)$ est l'unique vecteur de la forme $x - \lambda a$, où $\lambda \in \mathbb{R}$, qui appartienne à $H$, *i.e.* tel que $(x - \lambda a \,|\, a) = 0$, ce qui équivaut à : $(x \,|\, a) - \lambda \|a\|^2 = 0$. On a alors

$$d(x,H) = \|x - p_H(x)\| = \|\lambda a\| = \frac{|\,(x \,|\, a)\,|}{\|a\|}.$$

• D'après le théorème de Pythagore, on a $d(x,D)^2 = \|x\|^2 - d(x,D^\perp)^2$, la distance $d(x,D^\perp)$ étant donnée par le premier point, car $a$ est un vecteur normal à l'hyperplan $D^\perp$. On en déduit la formule. □