Produits scalaires

Cours de É. Bouchet – PCSI

12 juin 2024

Table des matières

Dans ce chapitre, on se limite au cas $\mathbb{K} = \mathbb{R}$.

1 Produits scalaires et normes

1.1 Produits scalaires

Définition 1.1 (Produit scalaire)

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 $\langle , \rangle$ de $E \times E$ dans $\mathbb{R}$ qui est :
- bilinéaire : $\forall (x, y, z) \in E^3$, $\forall \lambda \in \mathbb{R}$, $\langle \lambda x + y, z\rangle = \lambda \langle x, z\rangle + \langle y, z\rangle$ et $\langle x, \lambda y + z\rangle = \lambda \langle x, y\rangle + \langle x, z\rangle$;
- symétrique : $\forall (x, y) \in E^2$, $\langle y, x\rangle = \langle x, y\rangle$;
- définie positive : $\forall x \in E$, $\langle x, x\rangle \geqslant 0$ et $(\langle x, x\rangle = 0 \Rightarrow x = 0_E)$.

Remarque. Suivant les contextes, le produit scalaire $\langle x, y\rangle$ peut aussi se noter $(x|y)$ ou $x \cdot y$.

Remarque. On montre en général la symétrie avant la bilinéarité, parce qu'il suffit alors de montrer la linéarité d'un seul côté (à droite ou à gauche, au choix).

Remarque. Soit $x \in E$, la linéarité à droite donne $\langle x, 0_E\rangle = \langle x, 0 \cdot 0_E\rangle = 0 \langle x, 0_E\rangle = 0$. La linéarité à gauche donne de même $\langle 0_E, x\rangle = 0$.

Définition 1.2 (Espace préhilbertien réel, espace euclidien)

On appelle **espace préhilbertien réel** tout $\mathbb{R}$-espace vectoriel muni d'un produit scalaire. Si l'espace vectoriel est en plus de dimension finie, on parle d'**espace euclidien**.

Proposition 1.3 (Produit scalaire canonique sur $\mathbb{R}^n$)

L'application $\langle , \rangle \ : \big((x_1, \ldots, x_n), (y_1, \ldots, y_n)\big) \mapsto \displaystyle\sum_{k=1}^{n} x_k y_k$ est un produit scalaire sur $\mathbb{R}^n$.

Démonstration. On vérifie (dans l'ordre qui nous arrange) les points de la définition :
- Symétrie : soit $x = (x_1, \ldots, x_n)$ et $y = (y_1, \ldots, y_n)$ dans $\mathbb{R}^n$, $\langle y, x\rangle = \displaystyle\sum_{i=1}^{n} y_i x_i = \sum_{i=1}^{n} x_i y_i = \langle x, y\rangle$.
- Bilinéarité : par symétrie, il suffit de montrer la linéarité à gauche. Soit $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n)$, $z = (z_1, \ldots, z_n)$ dans $\mathbb{R}^n$ et $\lambda \in \mathbb{R}$, on a bien :
$$\langle \lambda x + y, z\rangle = \sum_{i=1}^{n} (\lambda x_i + y_i) z_i = \sum_{i=1}^{n} (\lambda x_i z_i + y_i z_i) = \lambda \sum_{i=1}^{n} x_i z_i + \sum_{i=1}^{n} y_i z_i = \lambda \langle x, z\rangle + \langle y, z\rangle .$$
- Positivité et cas d'annulation : soit $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, $\langle x, x\rangle = \sum_{i=1}^{n} x_i^2 \geqslant 0$, d'où la positivité. On suppose maintenant que $\langle x, x\rangle = 0$. Comme c'est une somme de termes positifs, on en déduit $x_1^2 = \ldots = x_n^2 = 0$, puis $x_1 = \ldots = x_n = 0$ et $x = (0, \ldots, 0)$. Donc $\langle , \rangle$ est définie positive.

Il s'agit donc bien d'un produit scalaire sur $\mathbb{R}^n$. □

Remarque. En particulier, $\mathbb{R}^n$ muni de son produit scalaire canonique est un espace euclidien.

Remarque. On retrouve les formules du lycée : pour tous vecteurs $\vec{u} = (x, y)$ et $\vec{v} = (x', y')$, $\vec{u} \cdot \vec{v} = xx' + yy'$.

Remarque. En utilisant les notations matricielle, le produit matriciel canonique sur $\mathbb{R}^n$ s'écrit $\langle X, Y\rangle = X^\top Y$.

Remarque. Le produit scalaire canonique n'est pas le seul produit scalaire sur $\mathbb{R}^n$.

Exercice 1. Montrer que l'application $\varphi \ : \ (X, Y) \mapsto X^\top \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} Y$ est un produit scalaire sur $\mathbb{R}^2$.

<u>Solution :</u> On raisonne cette fois-ci sous forme matricielle pour alléger les calculs :

— Symétrie : soit X et Y dans $\mathbb{R}^2$. Comme $Y^\top \left(\begin{smallmatrix} 2 & 1 \\ 1 & 2 \end{smallmatrix}\right) X$ est un réel, il est égal à sa transposée, donc :

$$\varphi(Y,X) = Y^\top \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} X = \left(Y^\top \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} X\right)^\top = X^\top \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}^\top Y = X^\top \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} Y = \varphi(X,Y).$$

— Bilinéarité : comme l'application est symétrique, il suffit de montrer la linéarité à gauche. Soit X, Y, Z dans $\mathbb{R}^2$ et $\lambda \in \mathbb{R}$, les propriétés de la transposée et du produit matriciel donnent :

$$\begin{aligned} \varphi(\lambda X + Y, Z) &= (\lambda X + Y)^\top \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} Z \\ &= (\lambda X^\top + Y^\top) \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} Z \\ &= \lambda X^\top \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} Z + Y^\top \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} Z \\ \varphi(\lambda X + Y, Z) &= \lambda \varphi(X,Z) + \varphi(Y,Z). \end{aligned}$$

— Positivité et cas d'annulation : soit $X = \left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) \in \mathbb{R}^2$,

$$\varphi(X,X) = X^\top \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} X = \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} 2x+y \\ x+2y \end{pmatrix} = 2x^2 + 2xy + 2y^2 = x^2 + y^2 + (x+y)^2 \geqslant 0.$$

On suppose maintenant que $\varphi(X,X) = 0$, alors : $x = y = x + y = 0$ car x^2, y^2 et $(x+y)^2$ sont positifs. Donc : $X = \left(\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right)$.

L'application φ est donc bien un produit scalaire sur $\mathbb{R}^2$.

Proposition 1.4 (Produit scalaire canonique sur $C([a,b],\mathbb{R})$)

Soient $(a,b) \in \mathbb{R}^2$ tels que $a < b$. L'application $\langle , \rangle \ : (f,g) \mapsto \int_a^b f(t)g(t)dt$ est un produit scalaire sur $C([a,b],\mathbb{R})$.

Démonstration. On revient à la définition d'un produit scalaire :

— Symétrie : soit f et g deux fonctions continues sur $[a,b]$, $\langle g, f\rangle = \int_a^b g(t)f(t)dt = \int_a^b f(t)g(t)dt = \langle f, g\rangle$.

— Bilinéarité : par symétrie, on se contente d'étudier la linéarité à gauche. Soit f, g, h des fonctions continues sur $[a,b]$ et $\lambda \in \mathbb{R}$. Par linéarité de l'intégrale,

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

— Positivité et cas d'annulation : soit f une fonction continue sur $[a,b]$, $\langle f, f\rangle = \int_a^b f(t)^2 dt \geqslant 0$ par positivité de l'intégrale. On suppose maintenant que $\int_a^b f(t)^2 dt = 0$. Comme f est continue et positive, on obtient que $f^2 = 0$ sur $[a,b]$, donc $f = 0$.

On a donc bien défini un produit scalaire sur $C([a,b],\mathbb{R})$. □

Remarque. Muni du produit scalaire défini ci-dessus, $C([a,b],\mathbb{R})$ est un espace préhilbertien réel. Ce n'est par contre pas un espace euclidien car il n'est pas de dimension finie.

Remarque. Cette définition de produit scalaire s'adapte facilement à $\mathbb{R}[X]$. En effet, un polynôme qui s'annule sur $[a,b]$ s'annule en une infinité de points, donc possède une infinité de racines, donc est le polynôme nul.

1.2 Norme associée à un produit scalaire

Définition 1.5 (Norme associée à un produit scalaire)

Soit $(E, \langle , \rangle)$ un espace préhilbertien réel. On appelle **norme associée à ce produit scalaire** l'application définie de E dans $\mathbb{R}_+$ par : $\forall x \in E$, $\|x\| = \sqrt{\langle x, x\rangle}$.
On dit qu'un vecteur x de E est **unitaire** lorsque $\|x\| = 1$.

Remarque. Comme un produit scalaire est défini positif, on en déduit que pour tout $x \in E$, $\|x\| = 0 \Rightarrow x = 0_E$.

Définition 1.6 (Distance associée à une norme)

Soit $(E, \langle , \rangle)$ un espace préhilbertien réel et $(x, y) \in E^2$. On appelle **distance** entre x et y le réel

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

Remarque. Attention : ces notions de norme et de distance dépendent du produit scalaire associé.
Par exemple, pour le produit scalaire $(X, Y) \longrightarrow X^\top \left(\begin{smallmatrix} 2 & 1 \\ 1 & 2 \end{smallmatrix}\right) Y$ sur $\mathbb{R}^2$,

$$\left\| \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\| = \sqrt{\begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix}} = \sqrt{\begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix}} = \sqrt{2} \neq 1.$$

Proposition 1.7 (Identité remarquable)

Soit $(E, \langle , \rangle)$ un espace préhilbertien réel et $(x, y) \in E^2$, alors $\|x + y\|^2 = \|x\|^2 + 2\langle x, y \rangle + \|y\|^2$.

Démonstration. C'est une conséquence directe de la bilinéarité et de la symétrie :

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

□

Remarque. On montre de même que $\forall (x, y) \in E^2$, $\|x - y\|^2 = \|x\|^2 - 2\langle x, y \rangle + \|y\|^2$.

Proposition 1.8 (Formule de polarisation)

Soit $(E, \langle , \rangle)$ un espace préhilbertien réel et $(x, y) \in E^2$, alors $\langle x, y \rangle = \dfrac{1}{2}\left(\|x + y\|^2 - \|x\|^2 - \|y\|^2\right)$.

Démonstration. D'après l'identité remarquable, $\|x + y\|^2 = \|x\|^2 + 2\langle x, y \rangle + \|y\|^2$, isoler $\langle x, y \rangle$ donne alors directement le résultat. □

Remarque. Connaître les valeurs de la norme permet donc de retrouver le produit scalaire associé, si nécessaire.

Remarque. En combinant les formules de $\|x + y\|^2$ et $\|x - y\|^2$, on montre aussi $\langle x, y \rangle = \frac{1}{4}\left(\|x + y\|^2 - \|x - y\|^2\right)$.

Proposition 1.9 (Inégalité de Cauchy-Schwarz)

Soit $(E, \langle , \rangle)$ un espace préhilbertien réel et $(x, y) \in E^2$, alors $|\langle x, y \rangle| \leqslant \|x\| \|y\|$, avec égalité si et seulement si x et y sont colinéaires.

Démonstration. Pour tout $t \in \mathbb{R}$, l'identité remarquable précédente et les propriétés du produit scalaire donnent :

$$0 \leqslant \|x + ty\|^2 = \|x\|^2 + 2t\langle x, y \rangle + t^2 \|y\|^2.$$

— Si $\|y\| = 0$, alors $y = 0_E$. Donc $|\langle x, y \rangle| = 0 = \|x\| \|y\|$.

— Si $\|y\| \neq 0$, la fonction $t \mapsto \langle x, x \rangle + 2t\langle x, y \rangle + t^2\langle y, y \rangle$ est polynomiale de degré 2 et positive sur $\mathbb{R}$ tout entier, donc son discriminant est négatif. Cela donne $4\langle x, y \rangle^2 - 4\langle x, x \rangle\langle y, y \rangle \leqslant 0$, donc $\langle x, y \rangle^2 \leqslant \|x\|^2 \|y\|^2$ et par passage à la racine (croissante sur $\mathbb{R}_+$), $|\langle x, y \rangle| \leqslant \|x\| \|y\|$.
Il y a alors égalité si et seulement si le discriminant est nul, c'est-à-dire si et seulement si la fonction $t \mapsto \|x + ty\|^2$ s'annule. Cette annulation est équivalente a l'existence d'un réel $t_0 \in \mathbb{R}$ pour lequel $x + t_0 y = 0_E$, c'est-à-dire à la colinéarité de x et y.

□

Remarque. Si nous avions défini le produit scalaire comme au lycée, à partir de normes et d'angles, l'inégalité de Cauchy-Schwarz serait immédiate : $|\vec{u}\cdot\vec{v}| = \|\vec{u}\| \times \|\vec{v}\| \times |\cos(\vec{u},\vec{v})| \leqslant \|\vec{u}\|\,\|\vec{v}\|$.

Exercice 2. Soit $(x_1,\dots,x_n)\in\mathbb{R}^n$. Montrer que $\left(\sum_{k=1}^n x_k\right)^2 \leqslant n\sum_{k=1}^n x_k^2$ et étudier le cas d'égalité.
Solution : On applique l'inégalité de Cauchy-Schwarz aux vecteurs $(x_1,\dots,x_n)$ et $(1,\dots,1)$ pour $\mathbb{R}^n$ muni du produit scalaire canonique :

$$\left|\sum_{i=1}^n x_i\right| \leqslant \sqrt{\sum_{i=1}^n x_i^2}\sqrt{\sum_{i=1}^n 1^2}.$$

Un passage au carré donne alors $\left(\sum_{i=1}^n 1\times x_i\right)^2 \leqslant n\sum_{i=1}^n x_i^2$. Et il y a égalité si et seulement si $(x_1,\dots,x_n)$ et $(1,\dots,1)$ sont colinéaires, c'est-à-dire quand les x_i sont tous égaux.

Exercice 3. Montrer que pour toute fonction $f\in C^1([0,1],\mathbb{R})$, $f(1)^2-f(0)^2 \leqslant 2\sqrt{\int_0^1 f(t)^2dt}\sqrt{\int_0^1 f'(t)^2dt}$.
Solution : On applique l'inégalité de Cauchy-Schwarz à f et f' dans l'espace préhilbertien réel $C([0,1],\mathbb{R})$ muni du produit scalaire usuel :

$$\left|\int_0^1 f(t)f'(t)dt\right| \leqslant \sqrt{\int_0^1 f(t)^2dt}\sqrt{\int_0^1 f'(t)^2dt}.$$

Or $\int_0^1 f(t)f'(t)dt = \left[\frac{f(t)^2}{2}\right]_0^1 = \frac{f(1)^2-f(0)^2}{2}$, d'où le résultat annoncé.

Proposition 1.10 (Inégalité triangulaire)

Soit $(E,\langle,\rangle)$ un espace préhilbertien réel et $(x,y)\in E^2$, alors $\|x+y\|\leqslant\|x\|+\|y\|$, avec égalité si et seulement si x et y sont colinéaires et de même sens.

Démonstration. Les propriétés de la norme et l'inégalité de Cauchy-Schwarz donnent :

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

La croissance de la racine carrée sur $\mathbb{R}_+$ donne alors $\|x+y\|\leqslant\|x\|+\|y\|$.
Étudions maintenant le cas d'égalité. Il y a égalité si et seulement si $\langle x,y\rangle = \|x\|\,\|y\|$, reste à trouver les x et y auxquels ça correspond.

— Supposons $\langle x,y\rangle = \|x\|\,\|y\|$: d'après le cas d'égalité de Cauchy-Schwarz, les vecteurs x et y sont colinéaires. Quitte a les permuter, on peut supposer qu'il existe $\lambda\in\mathbb{R}$ tel que $y=\lambda x$. On obtient alors d'une part :

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

et d'autre part :

$$\|x\|\,\|y\| = \|x\|\,\|\lambda x\| = \|x\|\sqrt{\langle\lambda x,\lambda x\rangle} = \|x\|\sqrt{\lambda^2\langle x,x\rangle} = |\lambda|\,\|x\|^2.$$

Donc soit x est nul, soit $\lambda=|\lambda|$, c'est-à-dire $\lambda\geqslant 0$. Dans les deux cas, x et y sont colinéaires de même sens.

— Réciproquement, si x et y sont colinéaires de même sens, il est immédiat que l'égalité est vérifiée.

□

2 Orthogonalité

2.1 Premières définitions et propriétés

Définition 2.1 (Vecteurs orthogonaux)

Soit $(E,\langle,\rangle)$ un espace préhilbertien réel et $(x,y)\in E^2$. On dit que x et y sont **orthogonaux** et on note $x\perp y$ lorsque $\langle x,y\rangle = 0$.

Remarque. Soit $x \in E$, alors $x \perp x \iff \langle x, x\rangle = 0 \iff x = 0_E$, donc le vecteur nul est le seul vecteur orthogonal à lui-même.

Remarque. Dans ce chapitre, la notion d'orthogonalité dépend du produit scalaire pour lequel elle est définie, et ne correspond donc pas toujours à la définition géométrique vue au lycée.

Définition 2.2 (Orthogonal d'une partie)

Soit $(E, \langle,\rangle)$ un espace préhilbertien réel et X une partie de E. On appelle **orthogonal** de X, noté $X^\perp$, l'ensemble des vecteurs orthogonaux à tous les éléments de X :

$$X^\perp = \{y \in E | \forall x \in X, \langle y, x\rangle = 0\}.$$

Proposition 2.3 (L'orthogonal est un sous-espace vectoriel)

Soit $(E, \langle,\rangle)$ un espace préhilbertien réel et X une partie de E, alors $X^\perp$ est un sous-espace vectoriel de E.

Démonstration.

— $\forall x \in X$, $\langle x, 0_E\rangle = 0$, donc $0_E \in X^\perp$.

— Soit $(y, z) \in (X^\perp)^2$ et $\lambda \in \mathbb{R}$. Alors $\forall x \in X$, $\langle \lambda y + z, x\rangle = \lambda \langle y, x\rangle + \langle z, x\rangle = \lambda \times 0 + 0 = 0$. Donc $\lambda y + z \in X^\perp$.

Donc $X^\perp$ est un sous-espace vectoriel de E. $\square$

Proposition 2.4 (Autres propriétés de l'orthogonal)

Soit $(E, \langle,\rangle)$ un espace préhilbertien réel et X une partie de E. Alors $X^\perp = \mathrm{Vect}(X)^\perp$ et $X \subset X^{\perp\perp}$.

Démonstration. Si un vecteur est orthogonal à tous les éléments de $\mathrm{Vect}(X)$, il est en particulier orthogonal aux éléments de X, donc $\mathrm{Vect}(X)^\perp \subset X^\perp$. Réciproquement, si un vecteur est orthogonal aux éléments de X, par linéarité à droite (ou à gauche) du produit scalaire, il est orthogonal à toutes les combinaisons linéaires d'éléments de X, donc à $\mathrm{Vect}(X)$. Donc $X^\perp \subset \mathrm{Vect}(X)^\perp$. Donc par double inclusion, $X^\perp = \mathrm{Vect}(X)^\perp$.
Pour le deuxième point : Soit $x \in X$, alors $\forall y \in X^\perp$, $\langle x, y\rangle = 0$. Donc $x \in (X^\perp)^\perp$. Donc $X \subset X^{\perp\perp}$. $\square$

Remarque. L'égalité $X^\perp = \mathrm{Vect}(X)^\perp$ signifie que pour déterminer l'orthogonal d'un sous-espace vectoriel, il est suffisant d'exiger l'orthogonalité aux éléments d'une famille génératrice.

Remarque. Contrairement aux apparences, dans le cas général, $F^{\perp\perp} \neq F$. D'ailleurs, F n'est pas nécessairement un espace vectoriel alors que $F^{\perp\perp}$ en est un.

Exemple. Soit E un espace préhilbertien réel, alors $\{0_E\}^\perp = E$ et $E^\perp = \{0_E\}$. En effet, 0_E est le seul vecteur de E orthogonal à tout vecteur.

Exercice 4. Soit $E = \mathbb{R}^2$ muni de son produit scalaire canonique et $X = \mathrm{Vect}((1, 1))$. Déterminer $X^\perp$.
Solution : Soit $u = (u_1, u_2) \in \mathbb{R}^2$, alors :

$$u \in X^\perp \iff u \in \mathrm{Vect}((1, 1))^\perp \iff u \in \{(1, 1)\}^\perp \iff \langle (u_1, u_2), (1, 1)\rangle = 0 \iff u_1 + u_2 = 0.$$

On en déduit :

$$u \in X^\perp \iff u = (u_1, -u_1) \iff u = u_1(1, -1) \iff u \in \mathrm{Vect}((1, -1)).$$

Donc $X^\perp = \mathrm{Vect}((1, -1))$.

Définition 2.5 (Familles orthogonales ou orthonormées)

Soit $(E, \langle,\rangle)$ un espace préhilbertien réel. On dit qu'une famille $(x_i)_{i\in I}$ de vecteurs est

— **orthogonale** lorsque pour tous $(i, j) \in I^2$ tels que $i \neq j$, $\langle x_i, x_j\rangle = 0$.

— **orthonormée (ou orthonormale)** lorsqu'elle est orthogonale et constituée de vecteurs unitaires, c'est-à-dire lorsque pour tous $(i, j) \in I^2$, $\langle x_i, x_j\rangle = \delta_{ij}$.

Exemple. Pour le produit scalaire canonique sur $\mathbb{R}^n$, la base canonique $(e_i)_{1\leqslant i\leqslant n}$ de $\mathbb{R}^n$ est orthonormale. En effet, $\forall(i,j)\in[\![1,n]\!]^2$, $\langle e_i, e_j\rangle = \delta_{ij}$.

Exercice 5. Pour le produit scalaire $(X,Y)\mapsto X^\top\left(\begin{smallmatrix}2&1\\1&2\end{smallmatrix}\right)Y$ sur $\mathbb{R}^2$, montrer que la base canonique n'est pas orthonormale, mais que la famille $\left(\frac{1}{\sqrt{2}}\left(\begin{smallmatrix}1\\0\end{smallmatrix}\right), \frac{1}{\sqrt{6}}\left(\begin{smallmatrix}1\\-2\end{smallmatrix}\right)\right)$ l'est.
Solution : On calcule $\langle\left(\begin{smallmatrix}1\\0\end{smallmatrix}\right),\left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)\rangle = \left(\begin{smallmatrix}1&0\end{smallmatrix}\right)\left(\begin{smallmatrix}2&1\\1&2\end{smallmatrix}\right)\left(\begin{smallmatrix}0\\1\end{smallmatrix}\right) = \left(\begin{smallmatrix}1&0\end{smallmatrix}\right)\left(\begin{smallmatrix}1\\2\end{smallmatrix}\right) = 1\neq 0$, donc la base canonique n'est pas orthonormale (elle n'est même pas orthogonale). Par contre :

— $\left\langle\frac{1}{\sqrt{2}}\left(\begin{smallmatrix}1\\0\end{smallmatrix}\right), \frac{1}{\sqrt{6}}\left(\begin{smallmatrix}1\\-2\end{smallmatrix}\right)\right\rangle = \frac{1}{\sqrt{12}}\left(\begin{smallmatrix}1&0\end{smallmatrix}\right)\left(\begin{smallmatrix}2&1\\1&2\end{smallmatrix}\right)\left(\begin{smallmatrix}1\\-2\end{smallmatrix}\right) = \frac{1}{\sqrt{12}}\left(\begin{smallmatrix}1&0\end{smallmatrix}\right)\left(\begin{smallmatrix}0\\-3\end{smallmatrix}\right) = 0$,

— $\left\|\frac{1}{\sqrt{2}}\left(\begin{smallmatrix}1\\0\end{smallmatrix}\right)\right\|^2 = \frac{1}{2}\left\|\left(\begin{smallmatrix}1\\0\end{smallmatrix}\right)\right\|^2 = \frac{1}{2}\left(\begin{smallmatrix}1&0\end{smallmatrix}\right)\left(\begin{smallmatrix}2&1\\1&2\end{smallmatrix}\right)\left(\begin{smallmatrix}1\\0\end{smallmatrix}\right) = \frac{1}{2}\left(\begin{smallmatrix}1&0\end{smallmatrix}\right)\left(\begin{smallmatrix}2\\1\end{smallmatrix}\right) = \frac{2}{2} = 1$,

— $\left\|\frac{1}{\sqrt{6}}\left(\begin{smallmatrix}1\\-2\end{smallmatrix}\right)\right\|^2 = \frac{1}{6}\left\|\left(\begin{smallmatrix}1\\-2\end{smallmatrix}\right)\right\|^2 = \frac{1}{6}\left(\begin{smallmatrix}1&-2\end{smallmatrix}\right)\left(\begin{smallmatrix}2&1\\1&2\end{smallmatrix}\right)\left(\begin{smallmatrix}1\\-2\end{smallmatrix}\right) = \frac{1}{6}\left(\begin{smallmatrix}1&-2\end{smallmatrix}\right)\left(\begin{smallmatrix}0\\-3\end{smallmatrix}\right) = \frac{6}{6} = 1$.

Donc la famille $\left(\frac{1}{\sqrt{2}}\left(\begin{smallmatrix}1\\0\end{smallmatrix}\right), \frac{1}{\sqrt{6}}\left(\begin{smallmatrix}1\\-2\end{smallmatrix}\right)\right)$ est orthonormale.

Exercice 6. Pour tout $n\in\mathbb{N}^*$, on définit $s_n : t\mapsto \sin(nt)$. Montrer que $(s_n)_{n\in\mathbb{N}^*}$ est orthonormale dans $C([0,2\pi],\mathbb{R})$ pour le produit scalaire $(f,g)\mapsto\frac{1}{\pi}\int_0^{2\pi}f(t)g(t)dt$.
Solution : Soit $n\in\mathbb{N}^*$, les formules trigonométriques donnent :

$$\|s_n\|^2 = \frac{1}{\pi}\int_0^{2\pi}\sin^2(nt)dt = \frac{1}{\pi}\int_0^{2\pi}\frac{1-\cos(2nt)}{2}dt = 1-\frac{1}{\pi}\left[\frac{\sin(2nt)}{4n}\right]_0^{2\pi} = 1.$$

Soit $(m,n)\in(\mathbb{N}^*)^2$ avec $n\neq m$, on trouve de même :

$$\langle s_m, s_n\rangle = \int_0^{2\pi}\frac{\sin(nt)\sin(mt)}{\pi}dt = \int_0^{2\pi}\frac{\cos((m-n)t)-\cos((m+n)t)}{2\pi}dt = \left[\frac{\sin((m-n)t)}{2\pi(m-n)}-\frac{\sin((m+n)t)}{2\pi(m+n)}\right]_0^{2\pi} = 0$$

La famille est donc bien orthonormée.

Proposition 2.6 (Familles orthogonales et liberté)

Soit $(E,\langle,\rangle)$ un espace préhilbertien réel. Toute famille orthogonale de vecteurs de E non nuls est libre.

Démonstration. Soient $(x_1,\dots,x_n)$ une famille orthogonale de vecteurs de E non nuls et soit $(\lambda_1,\dots,\lambda_n)\in\mathbb{R}^n$. On suppose $\sum_{k=1}^n\lambda_kx_k = 0_E$. Alors pour tout $i\in[\![1,n]\!]$,

$$0 = \langle 0_E, x_i\rangle = \left\langle\sum_{k=1}^n\lambda_kx_k, x_i\right\rangle = \sum_{k=1}^n\lambda_k\langle x_k,x_i\rangle = \lambda_i\|x_i\|^2.$$

Or $\|x_i\|\neq 0$ puisque $x_i\neq 0_E$, donc : $\lambda_i = 0$. Donc $(x_1,\dots,x_n)$ est une famille libre. □

Remarque. En particulier, toute famille orthonormale de vecteurs de E est libre (les vecteurs étant unitaires, ils ne peuvent pas être nuls).

Proposition 2.7 (Théorème de Pythagore)

Soit $(E,\langle,\rangle)$ un espace préhilbertien réel. Soit $(x,y)\in E^2$, alors : $x\perp y\iff\|x+y\|^2 = \|x\|^2+\|y\|^2$.

Démonstration. On sait que $\|x+y\|^2 = \|x\|^2+2\langle x,y\rangle+\|y\|^2$ (identité remarquable). Donc :

$$x\perp y\iff\langle x,y\rangle = 0\iff\|x+y\|^2 = \|x\|^2+\|y\|^2.$$

□

Remarque. Si $(x_1,\dots,x_n)$ est une famille orthogonale de E, alors on a de même $\left\|\sum_{i=1}^n x_i\right\|^2 = \sum_{i=1}^n\|x_i\|^2$.

2.2 Algorithme d'orthonormalisation de Gram-Schmidt

Proposition 2.8 (Orthonormalisation de Gram-Schmidt)

Soit $(E, \langle , \rangle)$ un espace préhilbertien réel et $(e_1, \ldots, e_n)$ une famille libre de E. On peut transformer $(e_1, \ldots, e_n)$ en une famille orthonormale de E $(u_1, \ldots, u_n)$ telle que :

$$\forall k \in [\![1, n]\!], \quad \mathrm{Vect}(u_1, \ldots, u_k) = \mathrm{Vect}(e_1, \ldots, e_k).$$

Exemple. Soit (e_1, e_2) une famille libre de E, qu'on souhaite transformer en famille orthonormée.

— On vérifie que $\|e_1\| \neq 0$, puis on pose $u_1 = \frac{e_1}{\|e_1\|}$. Ainsi, u_1 est une famille orthonormale (on vient de normer le vecteur) et $\mathrm{Vect}(u_1) = \mathrm{Vect}(e_1)$.

— Il faut maintenant construire un vecteur orthogonal à u_1. Pour cela, on pose $v = e_2 - \langle e_2, u_1 \rangle u_1$ (faire le dessin justifiant cette définition). On a bien :

$$\langle v, u_1 \rangle = \langle e_2 - \langle e_2, u_1 \rangle u_1, u_1 \rangle = \langle e_2, u_1 \rangle - \langle e_2, u_1 \rangle \langle u_1, u_1 \rangle = \langle e_2, u_1 \rangle - \langle e_2, u_1 \rangle 1 = 0.$$

De plus $\mathrm{Vect}(u_1, v) = \mathrm{Vect}(u_1, e_2 - \langle e_2, u_1 \rangle u_1) = \mathrm{Vect}(u_1, e_2) = \mathrm{Vect}(e_1, e_2)$.

— On vérifie que $\|v\| \neq 0$, puis on pose $u_2 = \frac{v}{\|v\|}$, pour créer un vecteur unitaire conservant les propriétés précédentes.

On a donc bien construit une famille orthonormale (u_1, u_2).

Démonstration. Soit $k \in [\![1, n]\!]$, on pose $P(k)$: « il existe une famille orthonormale de E $(u_1, \ldots, u_k)$ telle que $\mathrm{Vect}(u_1, \ldots, u_k) = \mathrm{Vect}(e_1, \ldots, e_k)$ ».

— (e_1) est libre, donc $e_1 \neq 0_E$, ce qui permet de poser $u_1 = \frac{e_1}{\|e_1\|}$. Par construction, u_1 est unitaire et $\mathrm{Vect}(e_1) = \mathrm{Vect}(u_1)$, donc $P(1)$ est vraie.

— Soit $k \in [\![2, n]\!]$. Supposons que $P(k-1)$ est vraie. Alors il existe une famille orthonormale $(u_1, \ldots, u_{k-1})$ pour laquelle $\mathrm{Vect}(u_1, \ldots, u_{k-1}) = \mathrm{Vect}(e_1, \ldots, e_{k-1})$.
On pose $v_k = e_k - \sum_{i=1}^{k-1} \langle e_k, u_i \rangle u_i$. Comme $(e_1, \ldots, e_k)$ est libre, $v_k \neq 0_E$ (sinon e_k serait combinaison linéaire de $(u_1, \ldots, u_{k-1})$ et donc de $(e_1, \ldots, e_{k-1})$). On peut ainsi poser $u_k = \frac{v_k}{\|v_k\|}$ (ou son opposé).
Montrons que la famille $(u_1, \ldots, u_k)$ est orthonormale. On sait que $(u_1, \ldots, u_{k-1})$ l'est. De plus, u_k est unitaire. Il suffit donc de montrer que u_k est orthogonal à u_1,...,u_{k-1}. Soit $j \in [\![1, k-1]\!]$, on a bien :

$$\langle u_k, u_j \rangle = \frac{1}{\|v_k\|} \langle v_k, u_j \rangle = \frac{1}{\|v_k\|} \left\langle e_k - \sum_{i=1}^{k-1} \langle e_k, u_i \rangle u_i, u_j \right\rangle = \frac{1}{\|v_k\|} \left(\langle e_k, u_j \rangle - \langle e_k, u_j \rangle \langle u_j, u_j \rangle \right) = 0.$$

Enfin, par hypothèse de récurrence, $\mathrm{Vect}(u_1, \ldots, u_{k-1}, u_k) = \mathrm{Vect}(e_1, \ldots, e_{k-1}, u_k)$. Or u_k est combinaison linéaire de e_1,...,e_{k-1} et e_k avec un coefficient non nul devant e_k, donc : $\mathrm{Vect}(u_1, \ldots, u_k) = \mathrm{Vect}(e_1, \ldots, e_k)$.
Donc $P(k)$ est vraie.

Donc $\forall k \in [\![0, n]\!]$, $P(k)$ est vrai, ce qui montre le résultat annoncé. □

Remarque. Cette démonstration fournit au passage une méthode de construction récursive :
poser $u_1 = \dfrac{e_1}{\|e_1\|}$ puis pour tout $k \in [\![2, n]\!]$, $u_k = \dfrac{e_k - \sum_{i=1}^{k-1} \langle e_k, u_i \rangle u_i}{\left\| e_k - \sum_{i=1}^{k-1} \langle e_k, u_i \rangle u_i \right\|}$.

Exercice 7. On se place sur $\mathbb{R}_2[X]$ muni du produit scalaire $\langle P, Q \rangle = \int_0^1 P(t)Q(t)dt$. Déterminer une base orthonormale (P_0, P_1, P_2).

Solution : On applique l'algorithme de Gram-Schmidt à la famille libre (car base canonique) $(1, X, X^2)$:

— $\|1\|^2 = \int_0^1 1^2 dt = 1$, on pose donc $P_0(X) = \frac{1}{\|1\|} = 1$.

— On pose $Q(X) = X - \langle X, P_0(X) \rangle P_0(X)$. Comme $\langle X, P_0(X) \rangle = \int_0^1 t dt = \frac{1}{2}$, on trouve $Q(X) = X - \frac{1}{2}$. De plus. $\|Q(X)\|^2 = \int_0^1 \left(t - \frac{1}{2}\right)^2 dt = \frac{1}{12}$, on pose donc $P_1(X) = \frac{Q(X)}{\|Q(X)\|} = \sqrt{12}(X - \frac{1}{2}) = \sqrt{3}(2X - 1)$.

— On pose $T(X) = X^2 - \left\langle X^2, P_0(X) \right\rangle P_0(X) - \left\langle X^2, P_1(X) \right\rangle P_1(X)$. Comme $\left\langle X^2, P_0(X) \right\rangle = \int_0^1 t^2 dt = \frac{1}{3}$ et $\left\langle X^2, P_1(X) \right\rangle = \int_0^1 t^2 \times \sqrt{3}(2t-1)dt = \frac{1}{2\sqrt{3}}$, on trouve $T(X) = X^2 - X + \frac{1}{6}$. De plus,

$$\|T(X)\|^2 = \int_0^1 \left(t^2 - t + \frac{1}{6} \right)^2 dt = \int_0^1 t^4 - 2t^3 + \frac{4t^2}{3} - \frac{t}{3} + \frac{1}{36} dt = \frac{1}{5} - \frac{2}{4} + \frac{4}{9} - \frac{1}{6} + \frac{1}{36} = \frac{1}{5 \times 6^2}.$$

On pose donc $P_2(X) = \frac{T(X)}{\|T(X)\|} = \sqrt{5}(6X^2 - 6X + 1)$.
L'algorithme de Gram-Schmidt garantit que la famille (P_0, P_1, P_2) est une base orthonormale de $\mathbb{R}_2[X]$.

2.3 Bases orthonormées

Proposition 2.9 (Existence de bases orthonormées dans un espace euclidien)

Soit $(E, \langle , \rangle)$ un espace euclidien de dimension $n \in \mathbb{N}^*$. Alors E possède une base orthonormale.

Démonstration. E est un espace vectoriel de dimension finie non nulle, donc possède une base. Il suffit de l'orthonormaliser par l'algorithme de Gram-Schmidt pour obtenir une base orthonormale. □

Proposition 2.10 (Théorème de la base orthonormée incomplète)

Soit $(E, \langle , \rangle)$ un espace euclidien de dimension $n \in \mathbb{N}^*$. Toute famille orthonormale de E peut être complétée en une base orthonormale de E.

Démonstration. Soit $(e_1, \ldots, e_p)$ une famille orthonormale, donc libre, de E. Comme E est de dimension finie non nulle, le théorème de la base incomplète permet de la compléter en une base de E. On orthonormalise ensuite cette base par l'algorithme de Gram-Schmidt. Par construction,les premiers vecteurs e_1, …, e_p ne seront pas affectés (puisqu'ils forment déjà une famille orthonormale). On a donc complété $(e_1, \ldots, e_p)$ en une base orthonormale de E. □

Proposition 2.11 (Coordonnées dans une base orthonormée)

Soit $(E, \langle , \rangle)$ un espace euclidien de dimension $n \in \mathbb{N}^*$, $(e_1, \ldots, e_n)$ une base orthonormale de E et $x \in E$. Alors les coordonnées de x dans $(e_1, \ldots, e_n)$ sont $(\langle x, e_1 \rangle, \ldots, \langle x, e_n \rangle)$. Autrement dit, $x = \sum_{k=1}^{n} \langle x, e_k \rangle e_k$.

Démonstration. On pose $(x_1, \ldots, x_n)$ les coordonnées de x dans la base $(e_1, \ldots, e_n)$. Alors pour tout $k \in [\![1, n]\!]$,

$$\langle x, e_k \rangle = \left\langle \sum_{i=1}^{n} x_i e_i, e_k \right\rangle = \sum_{i=1}^{n} x_i \langle e_i, e_k \rangle = \sum_{i=1}^{n} x_i \delta_{ik} = x_k.$$

□

Proposition 2.12 (Expression du produit scalaire et de la norme dans une base orthonormée)

Soit $(E, \langle , \rangle)$ un espace euclidien de dimension $n \in \mathbb{N}^*$ et $(x, y) \in E^2$ de coordonnées respectives $(x_1, \ldots, x_n)$ et $(y_1, \ldots, y_n)$ dans une base orthonormale B de E. Alors $\langle x, y \rangle = \sum_{k=1}^{n} x_k y_k$ et $\|x\| = \sqrt{\sum_{k=1}^{n} x_k^2}$.

Démonstration. On note $B = (e_1, \ldots, e_n)$. On a alors :

$$\langle x, y \rangle = \left\langle \sum_{i=1}^{n} x_i e_i, \sum_{j=1}^{n} y_j e_j \right\rangle = \sum_{1 \leqslant i,j \leqslant n} x_i y_j \langle e_i, e_j \rangle = \sum_{1 \leqslant i,j \leqslant n} x_i y_j \delta_{ij} = \sum_{k=1}^{n} x_k y_k.$$

Le résultat sur la norme découle alors de la relation $\|x\| = \sqrt{\langle x, x \rangle}$. □

Remarque. Attention, ces formules sont fausses dans le cas général (pour des coordonnées dans une base non orthonormée).

Remarque. Sous forme matricielle, avec des vecteurs colonnes de coordonnées dans une base orthonormale X et Y, cela donne $\langle X, Y \rangle = X^\top Y$ et $\|X\| = \sqrt{X^\top X}$.

3 Projection orthogonale sur un sous-espace de dimension finie

3.1 Supplémentaire orthogonal d'un sous-espace

Proposition 3.1 (Somme directe avec l'orthogonal)

Soit $(E, \langle , \rangle)$ un espace préhilbertien réel et F un sous-espace vectoriel de E. Alors F et $F^\perp$ sont en somme directe.

Démonstration. Comme F et $F^\perp$ sont des sous-espaces vectoriels de E, on a $\{0_E\} \subset F \cap F^\perp$. Réciproquement, soit $x \in F \cap F^\perp$. Alors $x \perp x$, donc $\langle x, x\rangle = 0$. Par propriétés du produit scalaire, on en déduit $x = 0_E$. Donc $F \cap F^\perp \subset \{0_E\}$. Donc par double inclusion $F \cap F^\perp = \{0_E\}$. □

Remarque. Attention : on a montré que F et $F^\perp$ sont en somme directe, mais ils ne sont pas nécessairement supplémentaires dans E.

Définition 3.2 (Supplémentaire orthogonal d'un sous-espace vectoriel de dimension finie)

Soient $(E, \langle , \rangle)$ un espace préhilbertien réel et F un sous-espace vectoriel de E de dimension finie. Il existe un unique supplémentaire de F dans E orthogonal à F, et c'est $F^\perp$. On l'appelle donc **le supplémentaire orthogonal** de F dans E et on note :

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

Démonstration. Si $F = \{0_E\}$, $F^\perp = E$ et le résultat est immédiat.
Dans le cas contraire, comme F est de dimension finie, il possède une base orthonormale $(f_1, \ldots, f_n)$. On sait déjà que $F \cap F^\perp = \{0_E\}$, il reste à montrer que $E = F + F^\perp$. L'inclusion $F + F^\perp \subset E$ est immédiate, montrons $E \subset F + F^\perp$. Soit $x \in E$, alors $x = (\sum_{k=1}^n \langle x, f_k\rangle f_k) + (x - \sum_{k=1}^n \langle x, f_k\rangle f_k)$, avec $\sum_{k=1}^n \langle x, f_k\rangle f_k \in F$. De plus, pour tout $i \in [\![1, n]\!]$,

$$\left\langle x - \sum_{k=1}^n \langle x, f_k\rangle f_k, f_i \right\rangle = \langle x, f_i\rangle - \sum_{k=1}^n \langle x, f_k\rangle \langle f_k, f_i\rangle = \langle x, f_i\rangle - \sum_{k=1}^n \langle x, f_k\rangle \delta_{ki} = \langle x, f_i\rangle - \langle x, f_i\rangle = 0.$$

Donc $x - \sum_{k=1}^n \langle x, f_k\rangle f_k$ est orthogonal à $f_1, \ldots, f_n$, donc élément de $\{f_1, \ldots, f_n\}^\perp = F^\perp$. La décomposition précédente donne donc $x \in F + F^\perp$, ce qu'on voulait. Donc $E = F \oplus F^\perp$.
Montrons maintenant l'unicité. Soit G un supplémentaire de F dans E orthogonal à F.

- Soit $x \in G$, alors $\forall f \in F$, $\langle x, f\rangle = 0$, donc $x \in F^\perp$. Donc $G \subset F^\perp$
- Réciproquement, soit $x \in F^\perp$. Comme $E = F + G$, il existe $f \in F$ et $g \in G$ tels que $x = f + g$. On a alors $\langle f, f\rangle = \langle f, f\rangle + \langle g, f\rangle = \langle f + g, f\rangle = \langle x, f\rangle = 0$. Par propriété du produit scalaire, on en déduit $f = 0_E$, donc $x = g \in G$. Donc $F^\perp \subset G$.

Donc par double inclusion, $G = F^\perp$. D'où l'unicité. □

Remarque. Attention : F doit être de dimension finie pour que ce résultat s'applique.

Remarque. En particulier, si E lui-même est de dimension finie : $\dim(F^\perp) = \dim(E) - \dim(F)$.

Remarque. Il existe un unique supplémentaire orthogonal, mais de nombreux supplémentaires « tout court ».

Proposition 3.3 (Cas particulier d'un sous-espace vectoriel de dimension finie)

Soient $(E, \langle , \rangle)$ un espace préhilbertien réel et F un sous-espace vectoriel de E de dimension finie. Alors $F^{\perp\perp} = F$.

Démonstration. On sait déjà que $F \subset F^{\perp\perp}$. Soit $x \in F^{\perp\perp}$, comme $E = F \oplus F^\perp$, il existe $f \in F$ et $g \in F^\perp$ tels que $x = f + g$. En particulier $f \perp g$ donc $\langle g, g\rangle = \langle g, g\rangle + \langle f, g\rangle = \langle f + g, g\rangle = \langle x, g\rangle = 0$, où la dernière égalité est vraie car $x \in F^{\perp\perp}$. Le produit scalaire étant défini positif, on en déduit que $g = 0_E$ et donc $x = f \in F$. Donc $F^{\perp\perp} \subset F$ et on conclut par double inclusion. □

Exercice 8. Dans l'espace euclidien canonique $\mathbb{R}^4$, on s'intéresse au plan vectoriel P d'équations : $x-y-z-t=0$ et $2x+y+z-t=0$. Déterminer son supplémentaire orthogonal.
Solution : On commence par constater que :

$$\begin{aligned} P &= \left\{(x,y,z,t) \in \mathbb{R}^4 \,|\, \langle (1,-1,-1,-1),(x,y,z,t)\rangle = 0 \text{ et } \langle (2,1,1,-1),(x,y,z,t)\rangle = 0\right\} \\ &= \{(1,-1,-1,-1),(2,1,1,-1)\}^{\perp} \\ P &= \mathrm{Vect}((1,-1,-1,-1),(2,1,1,-1))^{\perp} \end{aligned}$$

Or $\mathbb{R}^4$ est euclidien, donc $P^{\perp} = \mathrm{Vect}((1,-1,-1,-1),(2,1,1,-1))^{\perp\perp} = \mathrm{Vect}((1,-1,-1,-1),(2,1,1,-1))$.

Définition 3.4 (Vecteur normal à un hyperplan)

Soient $(E,\langle,\rangle)$ un espace euclidien de dimension $n \in \mathbb{N}^*$ et H un hyperplan de E. Le sous-espace $H^{\perp}$ est une droite dont tout vecteur non nul est appelé **vecteur normal** à H.

Démonstration. Comme E est de dimension finie, $E = H \oplus H^{\perp}$ donc $\dim(E) = \dim(H) + \dim(H^{\perp})$. Comme H est un hyperplan, on en déduit $\dim(H^{\perp}) = 1$, $H^{\perp}$ est donc bien une droite vectorielle. □

Remarque. Soit a un vecteur normal à l'hyperplan H, alors $H = \{a\}^{\perp} = \{x \in E \,|\, \langle x,a\rangle = 0\}$. Donc H est le noyau de la forme linéaire non nulle $x \mapsto \langle x,a\rangle$.

3.2 Projections orthogonales et propriétés

Définition 3.5 (Projection orthogonale sur un sous-espace vectoriel de dimension finie)

Soit $(E,\langle,\rangle)$ un espace préhilbertien réel et F un sous-espace vectoriel de E de dimension finie. On appelle **projection orthogonale** sur F la projection sur F parallèlement à $F^{\perp}$.

Remarque. Comme F est de dimension finie, on a bien $E = F \oplus F^{\perp}$, ce qui permet de définir le projecteur associé.

Proposition 3.6 (Expression du projeté orthogonal dans une base orthonormée)

Soit $(E,\langle,\rangle)$ un espace préhilbertien réel, F un sous-espace vectoriel de E de dimension $n \in \mathbb{N}^*$ et $(f_1,\dots,f_n)$ une base orthonormée de F. On note p la projection orthogonale sur F. Alors $\forall x \in E$, $p(x) = \displaystyle\sum_{k=1}^{n} \langle x,f_k\rangle f_k$.

Démonstration. Soit $x \in E$. On a montré dans la démonstration d'existence du supplémentaire orthogonal que $x = \left(\displaystyle\sum_{k=1}^{n} \langle x,f_k\rangle f_k\right) + \left(x - \displaystyle\sum_{k=1}^{n} \langle x,f_k\rangle f_k\right)$, avec $\displaystyle\sum_{k=1}^{n} \langle x,f_k\rangle f_k \in F$ et $x - \displaystyle\sum_{k=1}^{n} \langle x,f_k\rangle f_k \in F^{\perp}$.
Donc $p(x) = \displaystyle\sum_{k=1}^{n} \langle x,f_k\rangle f_k$. □

Remarque. Quand on dispose des coordonnées dans une base orthonormée, on obtient donc facilement les projetés orthogonaux. Dans le cas contraire, on a deux stratégies de calcul :
- Première possibilité : construire (avec Gram-Schmidt) une base orthonormée et se ramener à la formule.
- Deuxième possibilité : utiliser les relations $p(x) \in F$ et $x - p(x) \in F^{\perp}$ (faire le dessin associé).

Exercice 9. Soit $F = \mathrm{Vect}(\sin,\cos)$, qui est un sous-espace vectoriel de $C([0,2\pi],\mathbb{R})$. Déterminer le projeté orthogonal de id sur F pour le produit scalaire $(f,g) \mapsto \int_0^{2\pi} f(t)g(t)dt$.
Solution : On s'intéresse tout d'abord à la famille $(\sin,\cos)$, pour déterminer si c'est une base orthonormée :
- $\langle \sin,\cos\rangle = \int_0^{2\pi} \sin(t)\cos(t)dt = \left[\frac{\sin^2(t)}{2}\right]_0^{2\pi} = 0$. Donc la famille $(\sin,\cos)$ est orthogonale.
- $\|\sin\|^2 = \int_0^{2\pi} \sin^2(t)dt = \int_0^{2\pi} \frac{1-\cos(2t)}{2}dt = \left[\frac{t}{2} - \frac{\sin(2t)}{4}\right]_0^{2\pi} = \pi$.

— $\|\cos\|^2 = \int_0^{2\pi} \cos^2(t)dt = \int_0^{2\pi} \frac{1+\cos(2t)}{2}dt = \left[\frac{t}{2} + \frac{\sin(2t)}{4}\right]_0^{2\pi} = \pi$.

Pour la suite, on aura aussi besoin des valeurs de $\langle \mathrm{id}, \cos\rangle$ et $\langle \mathrm{id}, \sin\rangle$, que l'on calcule donc en prévision. Une intégration par parties (les fonctions étant de classe C^1 sur $[0, 2\pi]$) donne :

$$\int_0^{2\pi} te^{it}dt = \left[t\frac{e^{it}}{i}\right]_0^{2\pi} - \int_0^{2\pi} \frac{e^{it}}{i}dt = -2i\pi + i\left[\frac{e^{it}}{i}\right]_0^{2\pi} = -2i\pi + 0 = -2i\pi,$$

donc $\langle \mathrm{id}, \sin\rangle = \int_0^{2\pi} t\sin(t)dt = \mathrm{Im}\left(\int_0^{2\pi} te^{it}dt\right) = \mathrm{Im}(-2i\pi) = -2\pi$ et $\langle \mathrm{id}, \cos\rangle = \mathrm{Re}\left(\int_0^{2\pi} te^{it}dt\right) = \mathrm{Re}(-2i\pi) = 0$.
On compare maintenant les deux stratégies de calcul disponibles :

— Première stratégie : on orthonormalise cette base de F par Gram-Schmidt. Cette famille étant déjà orthogonale, $\left(\frac{\sin}{\|\sin\|}, \frac{\cos}{\|\cos\|}\right) = \left(\frac{\sin}{\sqrt{\pi}}, \frac{\cos}{\sqrt{\pi}}\right)$ est une base orthonormale de F. Le projeté orthogonal de id sur F est donc :
$$\left\langle \mathrm{id}, \frac{\sin}{\sqrt{\pi}}\right\rangle \frac{\sin}{\sqrt{\pi}} + \left\langle \mathrm{id}, \frac{\cos}{\sqrt{\pi}}\right\rangle \frac{\cos}{\sqrt{\pi}} = \frac{\langle \mathrm{id}, \sin\rangle}{\pi}\sin + \frac{\langle \mathrm{id}, \cos\rangle}{\pi}\cos = -2\sin.$$

— Deuxième stratégie : on note $p(\mathrm{id})$ le projeté orthogonal de id sur F et (λ, μ) ses coordonnées dans la base $(\sin, \cos)$ de F. Alors $p(\mathrm{id}) = \lambda\sin + \mu\cos$. Or on sait que $\mathrm{id} - p(\mathrm{id}) \in F^\perp$, ce qui donne deux relations :
$$0 = \langle \mathrm{id} - p(\mathrm{id}), \sin\rangle = \langle \mathrm{id} - \lambda\sin - \mu\cos, \sin\rangle = \langle \mathrm{id}, \sin\rangle - \lambda\|\sin\|^2 - \mu\langle\cos, \sin\rangle = -2\pi - \lambda\pi,$$
$$0 = \langle \mathrm{id} - p(\mathrm{id}), \cos\rangle = \langle \mathrm{id} - \lambda\sin - \mu\cos, \cos\rangle = \langle \mathrm{id}, \cos\rangle - \lambda\langle\sin, \cos\rangle - \mu\|\cos\|^2 = -\mu\pi.$$
Résoudre ce système donne $\lambda = -2$ et $\mu = 0$, et donc $p(\mathrm{id}) = -2\sin$.

Remarque. De manière générale, quand on se trouve dans un espace euclidien et qu'on veut projeter orthogonalement sur un sous-espace F, on préfère :
— projeter directement sur F si $\dim(F) \leqslant \dim(F^\perp)$,
— projeter d'abord sur $F^\perp$ sinon.

Exercice 10. Soit E un espace euclidien de dimension $n \in \mathbb{N}^*$ et H un hyperplan de E de vecteur normal a. On note p la projection orthogonale sur H. Soit $x \in E$, déterminer $p(x)$.
<u>Solution :</u> Notons p_a la projection orthogonale sur $\mathrm{Vect}(a) = H^\perp$. On sait que $\left(\frac{a}{\|a\|}\right)$ est une base orthonormale de $\mathrm{Vect}(a)$, donc $p_a(x) = \frac{\langle x,a\rangle a}{\|a\|^2}$. Or p et p_a sont des projecteurs associés, donc $p + p_a = \mathrm{id}_E$. On en déduit que $p(x) = x - \frac{\langle x,a\rangle a}{\|a\|^2}$.

3.3 Distance à un sous-espace vectoriel de dimension finie

Définition 3.7 (Distance à une partie)

Soient $(E, \langle,\rangle)$ un espace préhilbertien réel, A une partie non vide de E et $x \in E$. On appelle **distance** de x à A, notée $d(x, A)$, le réel $d(x, A) = \inf_{a\in A} \|x - a\|$.

Démonstration. L'ensemble $\{\|x - a\| \,|a \in A\}$ est une partie de $\mathbb{R}$ non vide (car $A \neq \emptyset$) et minorée par 0. Par propriété de la borne inférieure, le réel $d(x, A)$ est donc bien défini. □

Remarque. Intuitivement, la distance d'un vecteur x à une partie A est la plus petite distance séparant x d'un élément de A. On utilise une borne inférieure et pas un minimum pour éviter de se poser la question de l'existence de cette plus petite distance.

Proposition 3.8 (Distance à un sous-espace vectoriel de dimension finie)

Soient $(E, \langle,\rangle)$ un espace préhilbertien réel, F un sous-espace vectoriel de E de dimension finie et $x \in E$. Alors la projection orthogonale de x sur F, notée $p(x)$, est l'unique élément de F qui réalise la distance de x à F. En particulier, $d(x, F) = \|x - p(x)\|$.

Démonstration. Soit $f \in F$, alors $x-f = (x-p(x))+(p(x)-f)$. Or $x-p(x) \in \mathrm{Ker}(p) = F^{\perp}$ et $p(x)-f \in \mathrm{Im}(p) = F$, donc d'après le théorème de Pythagore, $\|x-f\|^2 = \|x-p(x)\|^2 + \|p(x)-f\|^2$.
On pose $D = \{\|x-f\| \,|\, f \in F\}$, alors :

— $\|x-p(x)\| \in D$ puisque $p(x) \in F$.

— par le calcul précédent, $\forall f \in F$, $\|x-f\| = \sqrt{\|x-p(x)\|^2 + \|p(x)-f\|^2} \geqslant \|x-p(x)\|$. Donc D est minoré par $\|x-p(x)\|$.

Donc $\|x-p(x)\| = \min(D)$, ce qui garantit que $\|x-p(x)\| = \inf(D) = d(x,F)$.
Enfin, pour tout $f \in F \setminus \{p(x)\}$, $\|x-f\| = \sqrt{\|x-p(x)\|^2 + \|p(x)-f\|^2} > \|x-p(x)\|$ car $p(x)-f \neq 0$, donc par propriété de la norme $\|p(x)-f\| > 0$. La distance $d(x,F)$ n'est donc atteinte qu'en $p(x)$. □