

Espaces préhilbertiens réels

I Produit scalaire

1. Définitions et produits scalaires usuels

Définition – Produit scalaire – Espace préhilbertien – Espace euclidien

Si E est un $\mathbb{R}$ -espace vectoriel, on appelle *produit scalaire* sur E une forme bilinéaire symétrique définie positive sur E , c'est-à-dire une application de $E \times E$ dans $\mathbb{R}$, notée $\langle \cdot, \cdot \rangle$, qui est :

- *bilinéaire* : $\forall x, y, z \in E, \forall \lambda, \mu \in \mathbb{R}, \langle \lambda x + \mu y, z \rangle = \lambda \langle x, z \rangle + \mu \langle y, z \rangle$ et $\langle x, \lambda y + \mu z \rangle = \lambda \langle x, y \rangle + \mu \langle x, z \rangle$,
- *symétrique* : $\forall x, y \in E, \langle x, y \rangle = \langle y, x \rangle$,
- *positive* : $\forall x \in E, \langle x, x \rangle \geq 0$,
- *définie* : $\forall x \in E, \langle x, x \rangle = 0 \Rightarrow x = 0$.

Si E est muni d'un produit scalaire, on dit que E est un espace *préhilbertien*. Si de plus il est de dimension finie, on dit qu'il est *euclidien*.

Remarque. Lorsqu'on considère un espace préhilbertien E , il sera parfois sous-entendu que $\langle \cdot, \cdot \rangle$ désigne le produit scalaire dont E est muni. On note parfois aussi $(\cdot | \cdot)$ le produit scalaire.

Produits scalaires canoniques

On peut munir les espaces vectoriels $\mathbb{R}^n$, $\mathcal{M}_{n,p}(\mathbb{R})$ et $\mathcal{C}^0([a, b], \mathbb{R})$ des produits scalaires usuels décrits ci-dessous. Dans la pratique, lorsque le produit scalaire n'est pas précisé, il est sous-entendu qu'il s'agit du produit scalaire canonique.

1. Sur $\mathbb{R}^n$, on définit un produit scalaire en posant

$$\forall x, y \in \mathbb{R}^n, \langle x, y \rangle = \sum_{i=1}^n x_i y_i.$$

En identifiant $\mathbb{R}^n$ à $\mathcal{M}_{n,1}(\mathbb{R})$, le produit scalaire ci-dessus se réécrit $\langle X, Y \rangle = X^\top Y$.

2. Sur $\mathcal{M}_{n,p}(\mathbb{R})$, on définit un produit scalaire en posant

$$\forall A, B \in \mathcal{M}_{n,p}(\mathbb{R}), \langle A, B \rangle = \text{tr}(A^\top B).$$

Remarque. Pour toutes matrices $A, B \in \mathcal{M}_{n,p}(\mathbb{R})$, $\langle A, B \rangle$ n'est autre que la somme des produits coefficient par coefficient de A et B :

$$\langle A, B \rangle = \sum_{j=1}^p (A^\top B)_{j,j} = \sum_{j=1}^p \sum_{i=1}^n (A^\top)_{j,i} b_{i,j} = \sum_{j=1}^p \sum_{i=1}^n a_{i,j} b_{i,j}.$$

Démonstration. La bilinéarité de l'application $\langle \cdot, \cdot \rangle$ provient directement de la linéarité de la trace, ainsi que de la linéarité des applications $M \mapsto A^\top M$ et $M \mapsto MB$.

Par ailleurs, si $A, B \in \mathcal{M}_{n,p}(\mathbb{R})$, $\langle A, B \rangle = \text{tr}(A^\top B) = \text{tr}((A^\top B)^\top) = \text{tr}(B^\top A) = \langle B, A \rangle$, d'où le caractère symétrique. Pour finir, si $A \in \mathcal{M}_{n,p}(\mathbb{R})$, alors

$$\langle A, A \rangle = \sum_{1 \leq i, j \leq n} a_{i,j}^2 \geq 0, \quad \text{donc} \quad \langle A, A \rangle = 0 \Rightarrow \forall (i, j) \in \llbracket 1, n \rrbracket^2, a_{i,j} = 0.$$

Ceci donne le caractère défini positif de $\langle \cdot, \cdot \rangle$, qui est bien un produit scalaire sur $\mathcal{M}_n(\mathbb{R})$. $\square$

On remarque que ceci inclut le cas du produit scalaire dans $\mathcal{M}_{n,1}(\mathbb{R})$ mentionné ci-dessus, et peut donc être vu comme comprenant le cas du produit scalaire dans $\mathbb{R}^n$.

Sur $\mathcal{C}^0([a, b], \mathbb{R})$, on définit un produit scalaire en posant

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

Démonstration. Le caractère bilinéaire, symétrique et positif de $\langle \cdot, \cdot \rangle$ découle directement des propriétés de l'intégrale, et le caractère défini peut être vu en remarquant que :

$$\text{si } \int_a^b f^2(t) dt = 0, \text{ alors comme } f^2 \text{ est positive et continue, } f^2 \text{ est nulle. } \square$$

Remarque. Si $A \in \mathcal{M}_{n,p}(\mathbb{R})$, $X \in \mathbb{R}^p$ et $Y \in \mathbb{R}^n$, alors pour le produit scalaire usuel sur $\mathcal{M}_{n,p}(\mathbb{R})$ introduit ci-dessus,

$$\langle AX, Y \rangle = \langle X, A^T Y \rangle.$$

Démonstration. On a $\langle AX, Y \rangle = (AX)^T Y = X^T A^T Y = X^T (A^T Y) = \langle X, A^T Y \rangle.$ $\square$

Définition - Norme, distance

Soit E est un espace préhilbertien dont on note le produit scalaire $\langle \cdot, \cdot \rangle$.

- Pour tout $x \in E$, on appelle *norme* de x le réel $\|x\| = \sqrt{\langle x, x \rangle}$.
- Pour tous $x, y \in E$, on appelle *distance* entre x et y le réel $d(x, y) = \|x - y\|$.

On dit par ailleurs qu'un vecteur $x \in E$ est *unitaire* si $\|x\| = 1$.

 La norme et la distance sur E dépendent du produit scalaire choisi sur E .

Remarques.

- Si $\lambda \in \mathbb{R}$ et $x \in E$, alors $\|\lambda x\| = |\lambda| \|x\|$.

Démonstration. On a $\|\lambda x\| = \sqrt{\langle \lambda x, \lambda x \rangle} = \sqrt{\lambda^2 \langle x, x \rangle} = |\lambda| \sqrt{\langle x, x \rangle} = |\lambda| \|x\|.$ $\square$

- Si $x \in E$ est non nul, le vecteur $u = \frac{x}{\|x\|}$ est unitaire. On dit alors qu'on a *renormalisé* le vecteur x .

Exemple. Si on munit $E = \mathcal{C}^0([a, b], \mathbb{R})$ du produit scalaire canonique sur E et $f \in E$, alors $\|f\| = \sqrt{\int_a^b f^2(t) dt}.$

Théorème - Identités classiques

Soient E est un espace préhilbertien muni du produit scalaire $\langle \cdot, \cdot \rangle$ et $x, y \in E$.

- i. $\|x + y\|^2 = \|x\|^2 + 2\langle x, y \rangle + \|y\|^2$, et $\|x - y\|^2 = \|x\|^2 - 2\langle x, y \rangle + \|y\|^2$,
- ii. $\|x\|^2 - \|y\|^2 = \langle x + y, x - y \rangle$,
- iii. $\langle x, y \rangle = \frac{1}{4}(\|x + y\|^2 - \|x - y\|^2)$ (*identité de polarisation*),
- iv. $\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2)$ (*identité du parallélogramme*).

Remarque. L'identité de polarisation permet de retrouver le produit scalaire à partir de la norme lorsqu'on sait qu'elle provient d'un produit scalaire.

Démonstration.

- i. Par bilinéarité, on a $\|x + y\|^2 = \langle x + y, x + y \rangle = \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle = \|x\|^2 + 2\langle x, y \rangle + \|y\|^2$.
Par ailleurs, $\|x - y\|^2 = \langle x - y, x - y \rangle = \langle x, x \rangle - \langle x, y \rangle - \langle y, x \rangle + \langle y, y \rangle = \|x\|^2 - 2\langle x, y \rangle + \|y\|^2$.
- ii. Toujours par bilinéarité, $\langle x + y, x - y \rangle = \langle x, x \rangle - \langle x, y \rangle + \langle y, x \rangle - \langle y, y \rangle = \|x\|^2 - \|y\|^2$.
- iii. En soustrayant les égalités de i., on obtient $\|x + y\|^2 - \|x - y\|^2 = 4\langle x, y \rangle$, ce qui conclut.
- iv. En ajoutant cette fois les égalités de i., on obtient le résultat. $\square$

Remarque. On peut généraliser l'identité remarquable au cas de n vecteurs :

$$\left\| \sum_{i=1}^n x_i \right\|^2 = \sum_{i=1}^n \|x_i\|^2 + 2 \sum_{1 \leq i < j \leq n} \langle x_i, x_j \rangle.$$

En effet, $\left\langle \sum_{i=1}^n x_i, \sum_{j=1}^n x_j \right\rangle = \sum_{1 \leq i, j \leq n} \langle x_i, x_j \rangle = \sum_{i=1}^n \|x_i\|^2 + \sum_{i \neq j} \langle x_i, x_j \rangle = \sum_{i=1}^n \|x_i\|^2 + 2 \sum_{1 \leq i < j \leq n} \langle x_i, x_j \rangle.$

2. Inégalité de Cauchy-Schwarz

Dans toute la suite, E désigne un espace préhilbertien muni du produit scalaire $\langle \cdot, \cdot \rangle$, et on note $\| \cdot \|$ la norme associée.

Théorème - Inégalité de Cauchy-Schwarz

Pour tous $x, y \in E$,

$$|\langle x, y \rangle| \leq \|x\| \|y\|,$$

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

Démonstration. Soient $x, y \in E$. On remarque déjà que si $x = 0_E$, alors l'inégalité est claire, et il y a même égalité. On suppose donc $x \neq 0_E$. Pour tout $t \in \mathbb{R}$, on a

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

Ainsi, la fonction $f : t \mapsto \langle tx + y, tx + y \rangle$ est polynomiale de degré 2. Par ailleurs, par positivité du produit scalaire, f est de signe constant positif sur $\mathbb{R}$, donc le discriminant associé Δ est négatif. Ceci s'écrit

$$\Delta = 4 \langle x, y \rangle^2 - 4 \|x\|^2 \|y\|^2 \leq 0,$$

et donne directement l'inégalité de Cauchy-Schwarz.

Par ailleurs, s'il y a égalité le discriminant ci-dessus est nul, ce qui se traduit par l'existence d'un réel t_0 tel que $f(t_0) = 0$, c'est-à-dire $\langle t_0 x + y, t_0 x + y \rangle = 0$. Ainsi, $t_0 x + y = 0$, et les vecteurs x et y sont colinéaires. Réciproquement, si x et y sont colinéaires, il est clair qu'il y a égalité. $\square$

Exemples.

- Si $E = \mathbb{R}^n$ et $(x_1, \dots, x_n), (y_1, \dots, y_n) \in E$, l'inégalité de Cauchy-Schwarz s'écrit :

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

- Si $E = \mathcal{C}^0([a, b], \mathbb{R})$ et $f, g \in E$, l'inégalité de Cauchy-Schwarz s'écrit :

$$\left| \int_a^b f(t)g(t) dt \right| \leq \sqrt{\int_a^b f^2(t) dt} \sqrt{\int_a^b g^2(t) dt}.$$

Théorème - Propriétés de la norme

Si E un espace préhilbertien, alors sa norme $\| \cdot \|$ vérifie les propriétés suivantes :

- *séparation* : $\forall x \in E, \|x\| = 0 \Rightarrow x = 0_E$,
- *homogénéité* : $\forall x \in E, \forall \lambda \in \mathbb{R}, \|\lambda x\| = |\lambda| \|x\|$,
- *inégalité triangulaire* : $\forall x, y \in E, \|x + y\| \leq \|x\| + \|y\|$.

De plus, il y a égalité dans l'inégalité triangulaire si et seulement si x et y sont positivement liés : il existe $\lambda \in \mathbb{R}_+$ tel que $y = \lambda x$ ou $x = \lambda y$.

Démonstration. La séparation provient directement du caractère défini du produit scalaire. L'homogénéité a été vue plus haut. Pour voir l'inégalité triangulaire, il suffit d'écrire

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

On obtient le résultat en composant par la fonction racine. Par ailleurs, s'il y a égalité dans l'inégalité ci-dessus, alors $\langle x, y \rangle = |\langle x, y \rangle| = \|x\| \|y\|$, donc x et y sont colinéaires. Ainsi, soit $x = 0_E$, soit il existe $\lambda \in \mathbb{R}$ tel que $y = \lambda x$. Dans ce dernier cas, comme $\langle x, \lambda x \rangle = |\langle x, \lambda x \rangle|$, on a $\lambda \|x\|^2 = |\lambda| \|x\|^2$, donc $\lambda = |\lambda|$, puis $\lambda \geq 0$, donc x et y sont positivement liés¹. Réciproquement, si x et y sont positivement liés, il y a bien égalité. $\square$

1. On remarque en effet que l'existence de $\lambda \in \mathbb{R}_+$ tel que $y = \lambda x$ ou $x = \lambda y$ se réécrit encore : $x = 0_E$ ou il existe $\lambda \in \mathbb{R}_+$ tel que $y = \lambda x$

Corollaire - Inégalité triangulaire généralisée

Si E est un espace préhilbertien dont la norme est notée $\|\cdot\|$, alors pour tous $x, y \in E$,

$$\|\|x\| - \|y\|\| \leq \|x + y\| \leq \|x\| + \|y\|.$$

Démonstration. La démonstration est la même que dans les cas réels et complexes : si $x, y \in E$, alors on peut écrire $\|x\| = \|(x + y) - y\| \leq \|x + y\| + \|y\|$, donc $\|x\| - \|y\| \leq \|x + y\|$. De même pour l'autre inégalité. $\square$

II Orthogonalité

1. Définitions

Définition - Orthogonalité

Soit E un espace préhilbertien.

- On dit que deux vecteurs x et y de E sont orthogonaux si $\langle x, y \rangle = 0$, on note alors $x \perp y$.
- On dit que deux parties A et B de E sont orthogonales si pour tout $(x, y) \in A \times B$, $\langle x, y \rangle = 0$.

Exemple. On se place sur $\mathbb{R}^n$ est muni de son produit scalaire canonique et on note $(e_1, \dots, e_n)$ la base canonique de $\mathbb{R}^n$. Si $i \neq j$, alors $\langle e_i, e_j \rangle = 0$.

Définition-théorème - Orthogonal d'une partie

Soient E un espace préhilbertien et A une partie de E . On appelle *orthogonal* de A l'ensemble

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

- i. $A^\perp$ est un sous-espace vectoriel de E , et $A^\perp = \text{Vect}(A)^\perp$.
- ii. $A \subset (A^\perp)^\perp$.

Démonstration.

- i. On a déjà $\langle a, 0_E \rangle = 0$ pour tout $a \in A$ donc $0_E \in A^\perp$. Par ailleurs, si $x, y \in A^\perp$, alors pour tout $a \in A$, $\langle a, \lambda x + \mu y \rangle = \lambda \langle a, x \rangle + \mu \langle a, y \rangle = 0$, donc $\lambda x + \mu y \in A^\perp$. Ainsi, $A^\perp$ est un sous-espace vectoriel de E .
Si $x \in \text{Vect}(A)^\perp$, il est clair que $x \in A^\perp$. Réciproquement, si $x \in A^\perp$, et $y \in \text{Vect}(A)$, alors on peut écrire $y = \lambda_1 a_1 + \dots + \lambda_k a_k$, où $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ et $a_1, \dots, a_k \in A$. Ainsi, $\langle x, y \rangle = \lambda_1 \langle x, a_1 \rangle + \dots + \lambda_k \langle x, a_k \rangle = 0$.
- ii. Si $x \in A$ et $y \in A^\perp$, alors on a $\langle x, y \rangle = 0$ par définition, ce qui entraîne que $A \subset (A^\perp)^\perp$. $\square$

Remarques.

- ◊  Si $A \subset E$, on a toujours l'inclusion $A \subset (A^\perp)^\perp$, mais l'égalité n'est pas toujours vraie.
- ◊ Si $A \subset E$, alors $A \cap A^\perp = \{0_E\}$: en effet, si $x \in A \cap A^\perp$ alors en particulier $\langle x, x \rangle = 0$, donc $x = 0_E$.

Exemples. – $\{0_E\}^\perp = E$: pour tout $x \in E$, on a $\langle x, 0_E \rangle = 0$, donc $x \in \{0_E\}^\perp$.
– $E^\perp = \{0_E\}$: si $x \in E^\perp$, alors $\forall y \in E$, $\langle x, y \rangle = 0$. En particulier, $\langle x, x \rangle = 0$, donc $x = 0_E$.

Théorème - Théorème de Pythagore

Soient E un espace préhilbertien et $x, y \in E$. On a $\langle x, y \rangle = 0$ si et seulement si $\|x + y\|^2 = \|x\|^2 + \|y\|^2$.

Démonstration. L'équivalence repose simplement sur l'identité remarquable $\|x + y\|^2 = \|x\|^2 + 2\langle x, y \rangle + \|y\|^2$. $\square$

2. Familles orthogonales

Définition - Famille orthogonale, famille orthonormale

Si E est un espace préhilbertien, on dit qu'une famille $(x_i)_{i \in I}$ de vecteurs de E est :

- *orthogonale* si les x_i sont orthogonaux deux à deux : $\forall i, j \in I, i \neq j \Rightarrow \langle x_i, x_j \rangle = 0$.
- *orthonormale* ou *orthonormée* si les x_i sont de plus unitaires : $\forall i, j \in I, \langle x_i, x_j \rangle = \delta_{i,j}$.

Exemples.

1. Dans $\mathbb{R}^n$, la base canonique $(e_1, \dots, e_n)$ est orthogonale, comme nous l'avons vu précédemment. Elle est par ailleurs orthonormale car les vecteurs e_i sont unitaires : $\langle e_i, e_i \rangle = 1$ pour tout $i \in \llbracket 1, n \rrbracket$.
2. Dans $\mathcal{M}_{n,p}(\mathbb{R})$, la base canonique $(E_{i,j})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}}$ est orthonormale.

Démonstration. Si $i, i' \in \llbracket 1, n \rrbracket$ et $j, j' \in \llbracket 1, p \rrbracket$, alors

$$\langle E_{i,j}, E_{i',j'} \rangle = \sum_{k=1}^p (E_{i,j}^\top E_{i',j'})_{k,k} = \sum_{k=1}^p \sum_{l=1}^n (E_{i,j}^\top)_{k,l} (E_{i',j'})_{l,k} = \sum_{k=1}^p \sum_{l=1}^n (E_{i,j})_{l,k} (E_{i',j'})_{l,k}.$$

On en déduit que $\langle E_{i,j}, E_{i',j'} \rangle = \sum_{k=1}^p \sum_{l=1}^n \delta_{i,l} \delta_{j,k} \delta_{i',l} \delta_{j',k} = \delta_{i,i'} \delta_{j,j'}$, ce qui conclut. $\square$

Exercice 1. On munit $\mathcal{C}^0([0, 2\pi], \mathbb{R})$ de son produit scalaire canonique, et on introduit $f_n : x \mapsto \frac{1}{\sqrt{\pi}} \cos(nx)$ pour tout $n \in \mathbb{N}^*$. Montrer que la famille $(f_n)_{n \in \mathbb{N}^*}$ est orthonormale.

Théorème - Théorème de Pythagore

Si E est un espace préhilbertien et $(x_1, \dots, x_n)$ est une famille orthogonale de E , alors

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

Démonstration. À nouveau, l'égalité repose sur l'identité remarquable $\left\| \sum_{i=1}^n x_i \right\|^2 = \sum_{i=1}^n \|x_i\|^2 + 2 \sum_{1 \leq i < j \leq n} \langle x_i, x_j \rangle$. $\square$

Théorème - Liberté d'une famille orthogonale

Toute famille orthogonale de vecteurs non nuls d'un espace préhilbertien E est libre.

Démonstration. Soit $(x_i)_{i \in I}$ une famille orthogonale de E . On suppose que $\lambda_1 x_{i_1} + \dots + \lambda_n x_{i_n} = 0_E$, où $\lambda_1, \dots, \lambda_n \in \mathbb{R}$. Ainsi, en prenant le produit scalaire avec x_{i_k} , on obtient $\lambda_k \langle x_{i_k}, x_{i_k} \rangle = 0$, ce qui donne $\lambda_k = 0$ car $\langle x_{i_k}, x_{i_k} \rangle \neq 0$. Finalement $\lambda_1 = \dots = \lambda_n = 0$, et la famille est libre. $\square$

3. Orthonormalisation de Gram-Schmidt

Le résultat ci-dessous est à la base de l'algorithme qui suivra, permettant de transformer une famille libre de E en une famille orthonormale.

L'idée sera de retirer successivement à chaque vecteur de la famille ses composantes selon les vecteurs précédents, afin d'assurer que le vecteur obtenu est orthogonal à tous les précédents. Il ne restera alors plus qu'à renormaliser.

Théorème et définition - Composante selon un vecteur

Soient E un espace préhilbertien et $x \in E$.

- Si $u \in E$ est unitaire, alors $x - \langle x, u \rangle u$ est orthogonal à u .
- *Cas général* : si $y \in E \setminus \{0_E\}$, alors $x - \langle x, y \rangle \frac{y}{\|y\|^2}$ est orthogonal à y .

Si $y \in E \setminus \{0_E\}$, on appelle composante de x selon le vecteur y le vecteur $\langle x, y \rangle \frac{y}{\|y\|^2}$.

Démonstration. Si $\|u\| = 1$, on a $\langle x - \langle x, u \rangle u, u \rangle = \langle x, u \rangle - \langle x, u \rangle \langle u, u \rangle = 0$ car $\langle u, u \rangle = \|u\|^2 = 1$.

Si $y \in E \setminus \{0_E\}$, on sait que $\frac{y}{\|y\|}$ est unitaire, et $x - \left\langle x, \frac{y}{\|y\|} \right\rangle \frac{y}{\|y\|} = x - \langle x, y \rangle \frac{y}{\|y\|^2}$. On conclut par le point précédent. $\square$

Théorème - Orthonormalisation de Gram-Schmidt

Soient E un espace préhilbertien et $(x_1, \dots, x_n)$ une famille libre de E . Il existe une famille orthonormale $(u_1, \dots, u_n)$ de E telle que pour tout $k \in \llbracket 1, n \rrbracket$, on a $\text{Vect}(x_1, \dots, x_k) = \text{Vect}(u_1, \dots, u_k)$.

On peut construire une telle famille en posant $u_1 = \frac{x_1}{\|x_1\|}$, puis en posant successivement pour chaque $k \in \llbracket 2, n \rrbracket$:

$$y_k = x_k - \sum_{i=1}^{k-1} \langle x_k, u_i \rangle u_i, \quad \text{puis} \quad u_k = \frac{y_k}{\|y_k\|}.$$

Démonstration. Montrons par récurrence que la famille de vecteur construites ci-dessus convient :

- Le vecteur u_1 est bien unitaire, et $\text{Vect}(x_1) = \text{Vect}(u_1)$.
- Soit $k \in \llbracket 2, n \rrbracket$. On suppose que la famille $(u_1, \dots, u_{k-1})$ construite comme ci-dessus est orthonormale et pour tout $l \in \llbracket 1, k-1 \rrbracket$, $\text{Vect}(x_1, \dots, x_l) = \text{Vect}(u_1, \dots, u_l)$. Pour tout $j \in \llbracket 1, k-1 \rrbracket$, on a alors

$$\langle y_k, u_j \rangle = \left\langle x_k - \sum_{i=1}^{k-1} \langle x_k, u_i \rangle u_i, u_j \right\rangle = \langle x_k, u_j \rangle - \sum_{i=1}^{k-1} \langle x_k, u_j \rangle \underbrace{\langle u_i, u_j \rangle}_{=\delta_{i,j}}.$$

Ainsi, $\langle y_k, u_j \rangle = \langle x_k, u_j \rangle - \langle x_k, u_j \rangle = 0$.

Par ailleurs, y_k n'est pas nul, sinon on a $x_k \in \text{Vect}(u_1, \dots, u_{k-1}) = \text{Vect}(x_1, \dots, x_{k-1})$, ce qui est impossible car $(x_1, \dots, x_k)$ est libre. Par conséquent, u_k est bien défini et unitaire. D'après ce qui précède, $(u_1, \dots, u_k)$ est orthonormale, et $\text{Vect}(u_1, \dots, u_k) = \text{Vect}(x_1, \dots, x_k)$, ce qui conclut. $\square$

Remarque. Si on ajoute la contrainte $\langle x_k, u_k \rangle > 0$ pour tout $k \in \llbracket 1, n \rrbracket$, alors on a même l'unicité de la famille $(u_1, \dots, u_n)$. Dans ce cas, il convient de poser $u_k = \pm \frac{y_k}{\|y_k\|}$ selon le signe de $\langle x_k, y_k \rangle$.

Exemple. Utilisons l'algorithme de Gram-Schmidt pour orthonormaliser la famille $(1, X)$ de $\mathbb{R}[X]$, muni du produit scalaire donné par $\langle P, Q \rangle = \int_0^1 P(t)Q(t) dt$ pour tous $P, Q \in \mathbb{R}[X]$.

- Comme $\|1\| = \int_0^1 1 dt = 1$, on pose $U_1 = 1$.
- On a $\langle X, U_1 \rangle = \langle X, 1 \rangle = \int_0^1 t dt = \frac{1}{2}$. On pose alors $P_2 = X - \langle X, U_1 \rangle U_1 = X - \frac{1}{2}$, puis comme

$$\text{comme } \|P_2\|^2 = \int_0^1 \left(t - \frac{1}{2}\right)^2 = \frac{1}{12}, \quad \text{on pose } U_2 = \frac{P_2}{\|P_2\|} = 2\sqrt{3} \left(X - \frac{1}{2}\right) = \sqrt{3}(2X - 1).$$

On obtient donc la famille $(1, \sqrt{3}(2X - 1))$, qui est une base orthonormale de $\mathbb{R}_1[X]$ pour le produit scalaire ci-dessus.

Corollaire - Base orthonormale d'un espace euclidien

- Tout espace euclidien admet une base orthonormale.
- Toute famille orthonormale d'un espace euclidien E peut être complétée en une base orthonormale de E .

Démonstration.

- On sait déjà que tout espace euclidien E admet une base. Il suffit alors d'utiliser le procédé d'orthonormalisation ci-dessus pour obtenir une base orthonormée de E .
- Si $(u_1, \dots, u_k)$ est une famille orthonormale de E , on sait qu'elle est libre. On peut donc la compléter en une base $(u_1, \dots, u_k, x_{k+1}, \dots, x_n)$. On applique ensuite le procédé d'orthonormalisation de Gram-Schmidt qui laisse les vecteurs $u_1, \dots, u_k$ invariants, et fournit finalement une base orthonormée $(u_1, \dots, u_k, u_{k+1}, \dots, u_n)$. $\square$

Théorème - Coordonnées dans une base orthonormale

Soient E un espace euclidien ayant pour base orthonormale $\mathcal{B} = (u_1, \dots, u_n)$ et $x, y \in E$ ayant pour coordonnées respectives $(x_1, \dots, x_n)$ et $(y_1, \dots, y_n)$ dans la base $\mathcal{B}$.

- On a $x = \sum_{i=1}^n \langle x, u_i \rangle u_i$. Autrement dit, les coordonnées de x dans la base $\mathcal{B}$ sont $(\langle x, u_1 \rangle, \dots, \langle x, u_n \rangle)$.
- On a

$$\langle x, y \rangle = \sum_{i=1}^n x_i y_i, \quad \text{et} \quad \|x\| = \sqrt{\sum_{i=1}^n x_i^2}.$$

Remarque. Ceci se réécrit matriciellement : si $X = \text{Mat}_{\mathcal{B}}(x)$ et $Y = \text{Mat}_{\mathcal{B}}(y)$, alors $\langle x, y \rangle = X^T Y$, et $\|x\| = \sqrt{X^T X}$.

Démonstration.

- Soit $i \in \llbracket 1, n \rrbracket$, on a $\langle x, u_i \rangle = \left\langle \sum_{j=1}^n x_j u_j, u_i \right\rangle = \sum_{j=1}^n x_j \langle u_j, u_i \rangle = x_i \langle u_i, u_i \rangle = x_i$, d'où le résultat.
- On a $\langle x, y \rangle = \left\langle \sum_{i=1}^n x_i u_i, \sum_{j=1}^n y_j u_j \right\rangle = \sum_{1 \leq i, j \leq n} x_i y_j \langle u_i, u_j \rangle = \sum_{i=1}^n x_i y_i$. □

4. Projection orthogonale

Théorème - Supplémentaire orthogonal

Soient E un espace préhilbertien et F un sous-espace vectoriel de dimension finie de E .

- $F^\perp$ est un supplémentaire de F dans E : on a $E = F \oplus F^\perp$. De plus, $F^\perp$ est l'unique supplémentaire de F qui lui est orthogonal, on l'appelle le *supplémentaire orthogonal* de F .
- $F = (F^\perp)^\perp$.

Démonstration.

- On a déjà $F \cap F^\perp = \{0_E\}$, donc la somme $F + F^\perp$ est directe. Par ailleurs, on peut considérer une base orthonormale $(u_1, \dots, u_k)$ de F . Si $x \in E$, on considère le vecteur $x_F = \langle x, u_1 \rangle u_1 - \dots - \langle x, u_k \rangle u_k$. On a bien sûr $x_F \in F$, et on sait que $x - x_F$ est orthogonal à $\text{Vect}(u_1, \dots, u_k)$, ce qui entraîne $x - x_F \in F^\perp$. On a donc $x = x_F + (x - x_F) \in F + F^\perp$, ce qui permet d'affirmer que $E = F + F^\perp$.
– Montrons l'unicité du supplémentaire orthogonal : supposons que G et H sont deux supplémentaires orthogonaux de F et fixons $x \in G$. On peut alors écrire $x = x_F + x_H$, où $x_F \in F$ et $x_H \in H$.
Ainsi, on a $x - x_H = x_F \in F \cap F^\perp$, donc $x - x_H = 0_E$, et $x = x_H \in H$. On a donc montré $G \subset H$, et on montre de même que $H \subset G$, donc $G = H$.
- On a déjà vu $F \subset (F^\perp)^\perp$. Par ailleurs, si $x \in (F^\perp)^\perp$, on note $x = x_F + x_{F^\perp}$ la décomposition de x dans $F \oplus F^\perp$, et on remarque que $0 = \langle x, x_{F^\perp} \rangle = \langle x_F, x_{F^\perp} \rangle + \langle x_{F^\perp}, x_{F^\perp} \rangle = \|x_{F^\perp}\|^2$, donc $x_{F^\perp} = 0_E$, et $x = x_F \in F$. □

Remarques.

- On note parfois $E = F \oplus^\perp G$ la décomposition de E en la somme de F et de son supplémentaire orthogonal. On remarque
- Si de plus E est de dimension finie et F est un sous-espace vectoriel de E , alors $\dim F^\perp = \dim E - \dim F$.

Exemple. Si $n \in \mathbb{N}^*$, le supplémentaire orthogonal de $\mathcal{S}_n(\mathbb{R})$ dans $\mathcal{M}_n(\mathbb{R})$ muni de son produit scalaire canonique est $\mathcal{A}_n(\mathbb{R})$.

Démonstration. Montrons le résultat par double inclusion².

- Si $A \in \mathcal{S}_n(\mathbb{R})^\perp$, alors pour tous $i, j \in \llbracket 1, n \rrbracket$, on a $\langle E_{i,j} + E_{j,i}, A \rangle = 0$, i.e. $\text{tr}((E_{i,j} + E_{j,i})^T A) = 0$.
Ainsi,

$$0 = \sum_{k=1}^n \sum_{l=1}^n (E_{i,j} + E_{j,i})_{k,l} a_{k,l} = a_{i,j} + a_{j,i}.$$

On a donc montré que pour tous $i, j \in \llbracket 1, n \rrbracket$, $a_{i,j} = -a_{j,i}$, ce qui entraîne que $A \in \mathcal{A}_n(\mathbb{R})$, donc $\mathcal{S}_n(\mathbb{R})^\perp \subset \mathcal{A}_n(\mathbb{R})$.

- Par ailleurs, si $A \in \mathcal{S}_n(\mathbb{R})$, alors $\langle A, S \rangle = \text{tr}(A^T S) = -\text{tr}(AS) = -\text{tr}(SA) = -\text{tr}(S^T A) = -\langle S, A \rangle$ donc $\langle A, S \rangle = 0$ pour toute matrice $S \in \mathcal{S}_n(\mathbb{R})$. Par conséquent, $\mathcal{A}_n(\mathbb{R}) \subset \mathcal{S}_n(\mathbb{R})^\perp$. □

² On peut aussi se souvenir que $\mathcal{M}_n(\mathbb{R}) = \mathcal{S}_n(\mathbb{R}) \oplus \mathcal{A}_n(\mathbb{R})$, il ne reste plus qu'à vérifier que $\mathcal{S}_n(\mathbb{R})$ et $\mathcal{A}_n(\mathbb{R})$ sont orthogonaux.

Définition - Projection orthogonale

Soient E un espace préhilbertien et F un sous-espace vectoriel de dimension finie de E . On appelle *projection orthogonale* sur F la projection sur F parallèlement à $F^\perp$.

Théorème - Coordonnées du projeté orthogonal

Soient E un espace préhilbertien et F un sous-espace vectoriel de dimension finie de E et $(u_1, \dots, u_k)$ une base orthonormale de F . On note p_F la projection orthogonale sur F . Pour tout $x \in E$,

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

Démonstration. Nous avons vu plus haut que si $x \in E$ et $x_F = \sum_{i=1}^k \langle x, u_i \rangle u_i$, alors $x_F \in F$ et $x - x_F \in F^\perp$. Par définition de la projection, on a alors $p_F(x) = x_F$. $\square$

Remarque. Si $(e_1, \dots, e_k)$ est une base orthogonale de F (pas nécessairement orthonormale) et $x \in E$, alors

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

**Déterminer le projeté orthogonal d'un vecteur**

Il est possible de déterminer le projeté d'un vecteur dans la projection orthogonale sur un sous-espace vectoriel F de dimension finie de E de deux manières :

- en utilisant le théorème ci-dessus, à partir d'une base orthonormale de F (au besoin, on orthonormalise une base à l'aide de l'algorithme de Gram-Schmidt),
- en remarquant que $p_F(x)$ est l'unique vecteur de F tel que $x - p_F(x) \in F^\perp$.

Exemple. On se place dans $\mathbb{R}[X]$ muni du produit scalaire donné par $\langle P, Q \rangle = \int_0^1 P(t)Q(t) dt$ pour tous $P, Q \in \mathbb{R}[X]$. Déterminons le projeté de X^2 sur $F = \mathbb{R}_1[X]$ avec chacune des deux méthodes décrites ci-dessus.

- À l'aide de Gram-Schmidt : nous avons vu qu'une base orthonormale de $\mathbb{R}_1[X]$ pour le produit scalaire ci-dessus est $(1, \sqrt{3}(2X - 1))$. Pour obtenir le projeté $p_F(X^2)$ de X^2 sur F , il suffit d'écrire

$$p_F(X^2) = \langle X^2, 1 \rangle 1 + \langle X^2, \sqrt{3}(2X - 1) \rangle \sqrt{3}(2X - 1)$$

donc $p_F(X^2) = X - \frac{1}{6}$ après calculs.

- Autre méthode : on sait que $p_F(X^2)$ s'écrit $aX + b$ avec $a, b \in \mathbb{R}$. Par ailleurs, $X^2 - p_F(X^2) = X^2 - aX - b$ est orthogonal à F , donc $\langle X^2 - aX - b, 1 \rangle = 0$ et $\langle X^2 - aX - b, X \rangle = 0$. On en déduit alors

$$\begin{cases} 3a + 6b = 2 \\ 4a + 6b = 3 \end{cases} \quad \text{donc} \quad a = 1, \quad b = -\frac{1}{6}.$$

On retrouve alors $p_F(X^2) = X - \frac{1}{6}$.

Théorème - Inégalité de Bessel

Soit E un espace préhilbertien. Si F est un sous-espace vectoriel de dimension finie de E et p_F la projection orthogonale sur F , alors pour tout $x \in E$, on a $\|p_F(x)\| \leq \|x\|$.

Démonstration. Comme $x - p_F(x)$ est orthogonal à $p_F(x)$, on a $\|x\|^2 = \|p_F(x)\|^2 + \|x - p_F(x)\|^2 \geq \|p_F(x)\|^2$. $\square$

Théorème et définition - Projection orthogonale sur un hyperplan

Soient E un espace euclidien et H un hyperplan de E .

- i. $H^\perp$ est de dimension 1, et tout vecteur u non nul de $H^\perp$ est appelé *vecteur normal* à H . On a alors $H \oplus \text{Vect}(u) = E$.

ii. Si u est un vecteur unitaire normal à H , alors le projeté orthogonal sur H de tout vecteur $x \in E$ est donné par $p_H(x) = x - \langle x, u \rangle u$.

Démonstration.

- i. Comme $E = H \oplus H^\perp$, on a $\dim H^\perp = \dim E - \dim H = 1$ car H est un hyperplan de E .
 ii. Si $x \in E$ et $H^\perp = \text{Vect}(u)$ où u est unitaire, on sait que $p_{H^\perp}(x) = \langle x, u \rangle u$. Comme $p_H(x) = x - p_{H^\perp}(x)$, on obtient bien le résultat. $\square$

Remarque. $\triangle!$ En dimension infinie, on n'a pas toujours $\dim H^\perp = 1$ (voir contre-exemple en TD).

5. Distance à un sous-espace vectoriel

Définition - Distance à un sous-espace vectoriel

Soient E un espace préhilbertien et F un sous-espace vectoriel de E et $x \in E$. On appelle *distance* de x à F le réel

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

Remarque. L'existence du réel $d(x, F)$ est assurée par le fait que $\{\|x - y\|, y \in F\}$ est une partie non vide, car elle contient $\|x - 0_E\| = \|x\|$, et minorée (par 0) de $\mathbb{R}$.

Théorème - Distance à un sous-espace vectoriel de dimension finie

Soient F un sous-espace vectoriel de dimension finie de E et p_F la projection orthogonale sur F . Pour tout $x \in E$,

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

Par ailleurs, $p_F(x)$ est l'unique vecteur $y \in F$ tel que $d(x, F) = \|x - y\|$. On a de plus :

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

Remarque. Si F est de dimension finie, l'infimum de la définition de $d(x, F)$ est atteint, c'est en fait un minimum.

Démonstration. On note $x = x_F + x_{F^\perp}$ la décomposition de x selon $F \oplus F^\perp$. Pour tout $y \in F$, on a

$$\|x - y\|^2 = \|x_F + x_{F^\perp} - y\|^2 = \|(x_F - y) + x_{F^\perp}\|^2 = \|x_F - y\|^2 + \|x_{F^\perp}\|^2$$

par Pythagore. Ainsi, $\|x - y\|^2 \geq \|x_{F^\perp}\|^2$, avec égalité si et seulement si $\|x_F - y\|^2 = 0$, c'est-à-dire $y = x_F$. Comme $x_F \in F$, ceci assure que l'infimum dans la définition de $d(x, F)$ est atteint (c'est donc un minimum) pour le seul vecteur $y = x_F = p_F(x)$.

Par ailleurs, on a $\|x\|^2 = \|p_F(x)\|^2 + \|x - p_F(x)\|^2 = \|p_F(x)\|^2 + d(x, F)^2$, donc $d(x, F)^2 = \|x\|^2 - \|p_F(x)\|^2$. $\square$

Exemple. On se place à nouveau dans $\mathbb{R}[X]$ muni de son produit scalaire usuel, et on s'intéresse à la distance entre X^2 et le sous-espace vectoriel $F = \mathbb{R}_1[X]$. On a alors

$$d(X^2, F)^2 = \|X^2\|^2 - \|p_F(X^2)\|^2 = \int_0^1 t^4 dt - \int_0^1 \left(t - \frac{1}{6}\right)^2 dt = \frac{1}{180}.$$

Ainsi, on a $d(X^2, F) = \frac{1}{\sqrt{180}} = \frac{1}{6\sqrt{5}}$.

Corollaire - Distance à un hyperplan en dimension finie

Si E est un espace euclidien et $H = \text{Vect}(u)^\perp$ est un hyperplan de E avec u unitaire, alors

$$\forall x \in E, d(x, H) = |\langle x, u \rangle|.$$

Démonstration. D'après le théorème, on sait que $d(x, H) = \|x - p_H(x)\| = \|p_{H^\perp}(x)\|$. Comme $H^\perp = \text{Vect}(u)$, on a vu que $p_{H^\perp}(x) = \langle x, u \rangle u$, car u est unitaire. Ainsi, $d(x, H) = |\langle x, u \rangle| \|u\| = |\langle x, u \rangle|$. $\square$