

33 | Espaces préhilbertiens réels

Dans ce chapitre, on travaille seulement avec le corps de base \mathbb{R} .

I – Produit scalaire et norme

1 – Définitions et généralités

Définition 33.1 – Soit E un \mathbb{R} -espace vectoriel. On appelle **produit scalaire sur E** une application

$$\begin{aligned} E \times E &\longrightarrow \mathbb{R} \\ (x, y) &\longmapsto \langle x, y \rangle \end{aligned}$$

- **bilinéaire**, c'est à dire linéaire à gauche :

$$\forall x_1, x_2, y \in E, \quad \forall \lambda \in \mathbb{R}, \quad \langle \lambda x_1 + x_2, y \rangle = \lambda \langle x_1, y \rangle + \langle x_2, y \rangle,$$

et linéaire à droite :

$$\forall x, y_1, y_2 \in E, \quad \forall \lambda \in \mathbb{R}, \quad \langle x, \lambda y_1 + y_2 \rangle = \lambda \langle x, y_1 \rangle + \langle x, y_2 \rangle.$$

- **symétrique** : $\forall (x, y) \in E^2, \quad \langle x, y \rangle = \langle y, x \rangle$.
- **positive** : $\forall x \in E, \quad \langle x, x \rangle \geq 0$.
- **définie** : $\forall x \in E, \quad \langle x, x \rangle = 0 \implies x = 0_E$. (propriété de **séparation**),

Un espace vectoriel réel muni d'un produit scalaire est appelé un **espace préhilbertien réel**. Un espace préhilbertien réel **DE DIMENSION FINIE** est appelé un **espace euclidien**.

Notation 33.2 – Le produit scalaire $\langle x, y \rangle$ est aussi parfois noté $\langle x | y \rangle$, $(x | y)$ ou bien sûr $x \cdot y$.

Remarque 33.3 –

1. Si l'on commence par montrer la symétrie, il suffit de montrer que l'application est linéaire à gauche. La bilinéarité découle de la symétrie. Prenez le temps de l'écrire pour vous en convaincre!
2. Un cas particulier de la bilinéarité est $\langle x, \lambda y \rangle = \lambda \langle x, y \rangle$ avec $\lambda = 0$. On obtient alors que pour tout x on a (en utilisant également la symétrie) :

$$\forall x \in E, \quad \langle x, 0_E \rangle = \langle 0_E, x \rangle = 0.$$

Définition-Théorème 33.4 –

1. L'application $(X, Y) \longmapsto X^T Y = \sum_{k=1}^n x_k y_k$ est un produit scalaire sur \mathbb{R}^n appelé son **produit scalaire canonique**.
2. L'application $(A, B) \longmapsto \text{tr}(A^T B) = \sum_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}} a_{ij} b_{ij}$ est un produit scalaire sur $\mathcal{M}_{n,p}(\mathbb{R})$ appelé son **produit scalaire canonique**.

Remarque 33.5 –

- Pourquoi *canonique*? Parce qu'il s'agit du produit scalaire le plus *naturel* à poser sur ces espaces! Pensez au cas de \mathbb{R}^2 pour vous en convaincre. La première définition donne dans ce cas :

$$\forall (x_1, x_2), (y_1, y_2) \in \mathbb{R}^2, \quad (x_1, x_2) \cdot (y_1, y_2) = x_1 y_1 + x_2 y_2,$$

qui correspond bien à la définition de produit scalaire dans le plan à laquelle vous avez été habitués.

- Pour tous $X, Y \in \mathbb{R}^n$, le produit matriciel $X^T Y$ est une matrice carrée de taille 1, i.e. un réel. Plus généralement, $A^T B$ est une matrice carrée de taille p pour tous $A, B \in \mathcal{M}_{n,p}(\mathbb{R})$, que l'on convertit en un réel grâce à la trace.

Démonstration.

1. Comme indiqué dans la remarque ci-dessus, pour $X, Y \in \mathbb{R}^n$, le produit scalaire $X^T Y$ est une matrice carrée de taille 1, donc automatiquement symétrique.

- **Symétrie :** Pour tous $X, Y \in \mathbb{R}^n$: $\langle X, Y \rangle = X^\top Y = (X^\top Y)^\top = Y^\top X = \langle Y, X \rangle$.
- **Bilinéarité :** Par symétrie, la linéarité par rapport à la première variable suffit. Pour tous $X_1, X_2, Y \in \mathbb{R}^n$, et $\lambda \in \mathbb{R}$, par linéarité de la transposition et du produit matriciel,

$$\langle \lambda X_1 + X_2, Y \rangle = (\lambda X_1 + X_2)^\top Y = (\lambda X_1^\top + X_2^\top) Y = \lambda X_1^\top Y + X_2^\top Y = \lambda \langle X_1, Y \rangle + \langle X_2, Y \rangle$$

- **Positivité et séparation :** Pour $X = (x_1, \dots, x_n) \in \mathbb{R}^n$, on a $\langle X, X \rangle = \sum_{k=1}^n x_k^2 \geq 0$. De plus, si $\langle X, X \rangle = 0$, comme on somme des réels positifs : $x_k = 0$ pour tout $k \in \llbracket 1, n \rrbracket$, i.e $X = 0_{\mathbb{R}^n}$.

2. Pour commencer, pour tous $A, B \in \mathcal{M}_{n,p}(\mathbb{R})$:

$$A^\top B \in \mathcal{M}_p(\mathbb{R}), \quad \text{donc } \text{tr}(A^\top B) = \sum_{j=1}^p (A^\top B)_{jj} = \sum_{j=1}^p \sum_{i=1}^n (A^\top)_{ji} b_{ij} = \sum_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}} a_{ij} b_{ij}$$

- **Symétrie :** Pour tous $A, B \in \mathcal{M}_{n,p}(\mathbb{R})$: $\text{tr}(A^\top B) = \text{tr}((A^\top B)^\top) = \text{tr}(B^\top A)$.
- **Bilinéarité :** Par symétrie, la linéarité par rapport à la première variable suffit. Pour tous $A, B, C \in \mathcal{M}_n(\mathbb{R})$ et $\lambda \in \mathbb{R}$, par bilinéarité du produit matriciel et linéarité de la trace :

$$\text{tr}(A^\top (\lambda B + C)) = \text{tr}(\lambda (A^\top B) + (A^\top C)) = \lambda \text{tr}(A^\top B) + \text{tr}(A^\top C)$$

- **Positivité et séparation :** Pour tout $A \in \mathcal{M}_{n,p}(\mathbb{R})$: $\text{tr}(A^\top A) = \sum_{i,j} a_{ij}^2 \geq 0$, et si $\text{tr}(A^\top A) = 0$, alors comme on somme des réels positifs : $a_{ij} = 0$ pour tous $i, j \in \llbracket 1, n \rrbracket$, i.e. $A = 0$. □

Exemple 33.6 – De nombreux produits scalaires peuvent exister sur un même espace vectoriel. L'application $(X, Y) \mapsto X^\top \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} Y$ est par exemple un produit scalaire sur \mathbb{R}^2 distinct du produit scalaire canonique.

Démonstration. Remarquons encore une fois que le produit matriciel donne ici une matrice de taille 1, i.e un réel.

- **Symétrie :** pour tous $X, Y \in \mathbb{R}^2$,

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

- **Bilinéarité :** Il suffit de montrer (par exemple) la linéarité à droite, qui vient directement de la linéarité du produit matriciel.
- **Positivité et séparation :** pour tout $X = (x, y) \in \mathbb{R}^2$:

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

et si $X^\top \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} X = 0$, alors $x = y = x + y = 0$ car x^2, y^2 et $(x+y)^2$ sont positifs, donc $X = (0, 0)$. □

Exemple 33.7 – Soient $a, b \in \mathbb{R}$ avec $a < b$. L'application $(f, g) \mapsto \int_a^b f(t)g(t) dt$ est un produit scalaire sur $\mathcal{C}([a, b], \mathbb{R})$.

Démonstration. La symétrie est évidente et la bilinéarité provient de la linéarité de l'intégrale. Ensuite, pour tout $f \in \mathcal{C}([a, b], \mathbb{R})$: $\int_a^b f^2(t) dt \geq 0$ et si $\int_a^b f^2(t) dt = 0$, f étant CONTINUE ET POSITIVE, alors $f^2 = 0$ sur $[a, b]$ donc $f = 0$. □



ATTENTION ! Muni du produit scalaire défini ci-dessus, $\mathcal{C}([a, b], \mathbb{R})$ n'est pas un espace EUCLIDIEN car ce n'est pas un \mathbb{R} -espace vectoriel de dimension FINIE. C'est seulement un ESPACE PRÉHILBERTIEN RÉEL.

Exemple 33.8 – Soient $x_0, \dots, x_n \in \mathbb{R}$ distincts. L'application $(P, Q) \mapsto \sum_{k=0}^n P(x_k) Q(x_k)$ est un produit scalaire sur $\mathbb{R}_n[X]$.

Démonstration.

- **Symétrie et bilinéarité** : Symétrie évidente, donc la linéarité par rapport à la première variable suffit. Pour tous $P, Q, R \in \mathbb{R}_n[X]$ et $\lambda, \mu \in \mathbb{R}$:
$$\sum_{k=0}^n (\lambda P + \mu Q)(x_k) R(x_k) = \lambda \sum_{k=0}^n P(x_k) R(x_k) + \mu \sum_{k=0}^n Q(x_k) R(x_k).$$
- **Positivité et séparation** : Pour tout $P \in \mathbb{R}_n[X]$: $\sum_{k=0}^n P(x_k)^2 \geq 0$, et si $\sum_{k=0}^n P(x_k)^2 = 0$, alors comme on somme des réels positifs : $P(x_k) = 0$ pour tout $k \in \llbracket 0, n \rrbracket$, i.e. x_0, \dots, x_n sont des racines de P . Le polynôme P de degré inférieur ou égal à n possède ainsi $n + 1$ racines distinctes, donc est nul. □

Définition 33.9 – Soit E un espace préhilbertien réel.

- **Norme** : On appelle **norme** sur E (associée au produit scalaire $\langle \cdot, \cdot \rangle$) l'application $\| \cdot \| : E \rightarrow \mathbb{R}_+$ définie pour tout $x \in E$ par : $\|x\| = \sqrt{\langle x, x \rangle}$. On dit qu'un vecteur x de E est **unitaire** si $\|x\| = 1$.
- **Distance** : Pour tous $x, y \in E$, le réel $\|x - y\|$, souvent noté $d(x, y)$, est appelé la **distance** entre x et y (associée au produit scalaire $\langle \cdot, \cdot \rangle$).



ATTENTION ! La notion de distance n'est pas forcément celle qu'on croit ! La distance dépend d'un choix de produit scalaire. Par exemple, pour le produit scalaire $(X, Y) \mapsto X^T \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} Y$ sur \mathbb{R}^2 : $\|(1, 0)\| = \sqrt{2} \neq 1$.

Définition 33.10 – Un vecteur $x \in E$ tel que $\|x\| = 1$ est appelé vecteur unitaire.

Remarque 33.11 –

- Soit $\| \cdot \|$ la norme associée au produit scalaire $\langle \cdot, \cdot \rangle$ sur E . Alors, pour tout $x \in E$ et tout $\lambda \in \mathbb{R}$,

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

Cette propriété est aussi appelée propriété d'homogénéité.

- On note aussitôt que si $x \neq 0_E$ alors $y = \frac{x}{\|x\|}$ est un vecteur unitaire colinéaire à x .

Proposition 33.12 – Identités remarquables

Soit $(E, \langle \cdot, \cdot \rangle)$ un espace préhilbertien et $\| \cdot \|$ la norme associée au produit scalaire. Alors, pour tout $(x, y) \in E^2$,

$$\begin{aligned} \|x + y\|^2 &= \|x\|^2 + \|y\|^2 + 2\langle x, y \rangle \\ \|x - y\|^2 &= \|x\|^2 + \|y\|^2 - 2\langle x, y \rangle \\ \|x\|^2 - \|y\|^2 &= \langle x - y, x + y \rangle \end{aligned}$$

Démonstration.

- Montrons la première identité. Soit $(x, y) \in E^2$. Utilisons la bilinéarité du produit scalaire :

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

puis la définition de la norme associée et la symétrie du produit scalaire donne $\|x + y\|^2 = \|x\|^2 + \|y\|^2 + 2\langle x, y \rangle$.

- La deuxième identité se déduit de la première en remplaçant y par $-y$.
- La dernière identité s'obtient comme la première en développant par bilinéarité la quantité $\langle x - y, x + y \rangle$. □

Ces identités sont des généralisations des identités remarquables sur \mathbb{R} étudiées au lycée avec $\langle x, y \rangle = xy$ et la norme associée $\|x\|^2 = x^2$.

Remarque 33.13 – On peut aussi « inverser » ces relations et récupérer le produit scalaire en fonction de la norme. On obtient alors ce qu'on appelle des **identités de polarisation**. Pour tout $(x, y) \in E^2$,

$$\begin{aligned} \langle x, y \rangle &= \frac{1}{2} (\|x + y\|^2 - \|x\|^2 - \|y\|^2) \\ \langle x, y \rangle &= \frac{1}{2} (\|x\|^2 + \|y\|^2 - \|x - y\|^2) \\ \langle x, y \rangle &= \frac{1}{4} (\|x + y\|^2 - \|x - y\|^2) \end{aligned}$$

Morale : connaissant la norme, on peut retrouver le produit scalaire, et inversement.

2 – Inégalité de Cauchy-Schwarz et inégalité triangulaire

Proposition 33.14 – Inégalité de Cauchy-Schwarz

Soit $(E, \langle \cdot, \cdot \rangle)$ un espace préhilbertien et $\|\cdot\|$ la norme associée au produit scalaire. Alors,

$$\forall (x, y) \in E^2, \quad |\langle x, y \rangle| \leq \|x\| \|y\|.$$

avec égalité, si et seulement si, les vecteurs x et y sont colinéaires.

Démonstration.

- Pour tous $(x, y) \in E^2$ et $t \in \mathbb{R}$, $\|x + ty\|^2 \geq 0$. En développant cette expression, on trouve que la fonction polynomiale en la variable t , $t \mapsto \|y\|^2 t^2 + 2\langle x, y \rangle t + \|x\|^2$, est toujours positif.
 - Si $\|y\|^2 = 0$, alors ce polynôme est de degré au plus 1, et pour être positif, il doit être égal à une constante positive, d'où $\langle x, y \rangle = 0$ et l'inégalité est vérifiée.
 - Sinon le polynôme est de degré 2 et donc son discriminant Δ doit être négatif. Or $\Delta = 4\langle x, y \rangle^2 - 4\|x\|^2\|y\|^2$, d'où l'inégalité.
- Cas d'égalité : Le cas d'égalité est trivial si $y = 0_E$. Dans le cas contraire, en reprenant la preuve qui précède, l'inégalité est une égalité si et seulement si le discriminant calculé est nul, i.e. si et seulement si la fonction $t \mapsto \|x + ty\|^2$ s'annule. Or cette annulation est équivalente à l'existence d'un réel $t_0 \in \mathbb{R}$ pour lequel $x + t_0 y = 0_E$, i.e. à la colinéarité de x et y . □

Exemple 33.15 – Traduisons l'inégalité de Cauchy-Schwarz pour le produit canonique sur \mathbb{R}^n ; pour tous (x_1, \dots, x_n) et $(y_1, \dots, y_n) \in \mathbb{R}^n$,

$$\left| \sum_{k=1}^n x_k y_k \right| \leq \sqrt{\sum_{k=1}^n x_k^2} \sqrt{\sum_{k=1}^n y_k^2}.$$

En particulier, en prenant $(y_1, \dots, y_n) = (1, \dots, 1)$, on a pour tous $x_1, \dots, x_n \in \mathbb{R}$: $\sum_{k=1}^n x_k \leq \sqrt{n} \sqrt{\sum_{k=1}^n x_k^2}$, avec égalité si et seulement si $x_1 = \dots = x_n$.

Exemple 33.16 – Traduisons l'inégalité de Cauchy-Schwarz pour le produit canonique sur $\mathcal{C}^0([a, b], \mathbb{R})$:

$$\forall f, g \in \mathcal{C}^0([a, b], \mathbb{R}), \quad \left| \int_a^b f(t)g(t) dt \right| \leq \sqrt{\int_a^b f(t)^2 dt} \sqrt{\int_a^b g(t)^2 dt}.$$

Exemple 33.17 – Pour toute fonction $f \in \mathcal{C}^1([0, 1], \mathbb{R})$ pour laquelle $f(0) = 0$: $f(1)^2 \leq 2 \sqrt{\int_0^1 f(t)^2 dt} \sqrt{\int_0^1 f'(t)^2 dt}$.

Démonstration. Appliquons l'inégalité de Cauchy-Schwarz à f et f' dans l'espace préhilbertien réel $\mathcal{C}([0, 1], \mathbb{R})$ muni du produit scalaire $(u, v) \mapsto \int_0^1 u(t)v(t) dt$: $\left| \int_0^1 f(t)f'(t) dt \right| \leq \sqrt{\int_0^1 f(t)^2 dt} \sqrt{\int_0^1 f'(t)^2 dt}$. On conclut en calculant l'intégrale de gauche, qui vaut $\frac{1}{2}(f(1)^2 - f(0)^2) = \frac{1}{2}f(1)^2$. □

Remarque 33.18 – Si nous avons défini le produit scalaire à partir de normes et d'angles, l'inégalité de Cauchy-Schwarz serait une pure trivialité : $|\vec{u} \cdot \vec{v}| = \|\vec{u}\| \times \|\vec{v}\| \times |\cos(\vec{u}, \vec{v})| \leq \|\vec{u}\| \cdot \|\vec{v}\|$, mais dans le contexte de ce chapitre, cette inégalité est remarquable justement parce que nous n'avons pas encore de définition propre des angles.

Proposition 33.19 – Inégalité triangulaire

Soit $(E, \langle \cdot, \cdot \rangle)$ un espace préhilbertien et $\|\cdot\|$ la norme associée au produit scalaire. Alors,

$$\forall x, y \in E, \quad \|x + y\| \leq \|x\| + \|y\|.$$

avec égalité si et seulement si x et y sont colinéaires **DE MÊME SENS** (on dit aussi *positivement liés*).

Démonstration. Soit $(x, y) \in E^2$. D'après l'inégalité de Cauchy-Schwarz,

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

Le résultat suit en appliquant la fonction croissante $x \mapsto \sqrt{x}$. □

Corollaire 33.20 – Inégalité triangulaire généralisée

Soit $(E, \langle \cdot, \cdot \rangle)$ un espace préhilbertien et $\|\cdot\|$ la norme associée au produit scalaire. Alors,

$$\forall (x, y) \in E^2, \quad \|x - y\| \geq \left| \|x\| - \|y\| \right|.$$

Démonstration. Il suffit d'écrire : $\|x\| = \|(x + y) + (-y)\| \leq \|x + y\| + \|-y\| = \|x + y\| + \|y\|$, donc $\|x\| - \|y\| \leq \|x + y\|$, et de même $\|y\| - \|x\| \leq \|x + y\|$. □

Exemple 33.21 – Dans le cadre de l'espace préhilbertien $\mathcal{C}([a, b], \mathbb{R})$ vu précédemment, l'inégalité triangulaire s'écrit

$$\forall f, g \in \mathcal{C}^0([a, b], \mathbb{R}), \quad \sqrt{\int_0^1 (f + g)^2} \leq \sqrt{\int_0^1 f^2} + \sqrt{\int_0^1 g^2}$$

II – Orthogonalité

1 – Définitions

Définition 33.22 – Soit $(E, \langle \cdot, \cdot \rangle)$ un espace préhilbertien réel, $x, y \in E$, X et Y deux parties de E et $(x_i)_{i \in I}$ une famille de vecteurs de E . On dit que :

- x et y sont **orthogonaux**, ce que l'on note $x \perp y$, ssi $\langle x, y \rangle = 0$;
- les parties X et Y sont **orthogonales**, ce que l'on note $X \perp Y$, si pour tout $x \in X$ et $y \in Y$: $\langle x, y \rangle = 0$;
- la famille $(x_i)_{i \in I}$ est **orthogonale** si pour tous $i, j \in I$ distincts : $\langle x_i, x_j \rangle = 0$;
-
- la famille $(x_i)_{i \in I}$ est **orthonormale** si elle est orthogonale et constituée de vecteurs unitaires, i.e si pour tous $i, j \in I$: $\langle x_i, x_j \rangle = \delta_{ij}$.

Remarque 33.23 –

- Jusqu'ici, la notion d'orthogonalité reposait sur la notion d'angle, qui elle-même permettait de définir un produit scalaire via la formule $\vec{u} \cdot \vec{v} = \|\vec{u}\| \times \|\vec{v}\| \times \cos(\vec{u}, \vec{v})$. Ici, on raisonne dans l'autre sens : la notion d'orthogonalité repose sur la définition préalable d'un produit scalaire. En particulier, à chaque produit scalaire est associée une notion d'orthogonalité, ce qui fait que les angles droits ne sont pas droits absolument, mais relativement.
- La propriété de séparation énonce que **LE VECTEUR NUL EST LE SEUL VECTEUR ORTHOGONAL À LUI MÊME**. En particulier, seul le vecteur nul est orthogonal à tout vecteur.

Exemple 33.24 –

- Pour le produit scalaire canonique de \mathbb{R}^n , la base canonique $(E_i)_{1 \leq i \leq n}$ de \mathbb{R}^n est orthonormale :

$$\forall 1 \leq i, j \leq n, \quad \langle E_i, E_j \rangle = \sum_{k=1}^n \delta_{ik} \delta_{jk} = \delta_{ij}$$

- La base canonique de \mathbb{R}^2 n'est pas orthonormale pour le produit scalaire $(X, Y) \mapsto X^T \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} Y$ sur \mathbb{R}^2 , mais la famille $\left(\frac{(1,0)}{\sqrt{2}}, \frac{(1,-2)}{\sqrt{6}} \right)$ l'est.

Rappelons que pour $X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ et $Y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$, on a, pour ce produit scalaire, $\langle X, Y \rangle = 2x_1 y_1 + x_1 y_2 + x_2 y_1 + 2x_2 y_2$.

Dès lors, $\left\langle \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\rangle = 0 + 1 + 0 + 0 = 1$. Donc la base canonique n'est pas orthogonale pour ce produit scalaire. Par ailleurs,

$$\begin{aligned} \left\| \frac{(1,0)}{\sqrt{2}} \right\|^2 &= 2 \times \left(\frac{1}{\sqrt{2}} \right)^2 = 1 \\ \left\| \frac{(1,-2)}{\sqrt{6}} \right\|^2 &= 2 \times \left(\frac{1}{\sqrt{6}} \right)^2 + 2 \times \frac{-2}{6} + 2 \times \left(\frac{-2}{\sqrt{6}} \right)^2 = \frac{1}{3} - \frac{2}{3} + \frac{4}{3} = 1 \\ \left\langle \frac{(1,0)}{\sqrt{2}}, \frac{(1,-2)}{\sqrt{6}} \right\rangle &= 2 \times \frac{1}{\sqrt{12}} + \frac{-2}{\sqrt{12}} + 0 + 0 = 0 \end{aligned}$$

Donc la famille $\left(\frac{(1,0)}{\sqrt{2}}, \frac{(1,-2)}{\sqrt{6}} \right)$ est bien orthonormale pour ce produit scalaire.

- Pour le produit scalaire canonique de $\mathcal{M}_{n,p}(\mathbb{R})$, la base canonique de $\mathcal{M}_{n,p}(\mathbb{R})$ est orthonormale. Pour tous $i \in \llbracket 1, n \rrbracket$ et $j \in \llbracket 1, p \rrbracket$, notons E_{ij} la matrice de $\mathcal{M}_{n,p}(\mathbb{R})$ dont tous les coefficients sont nuls à l'exception du coefficient de position (i, j) , égal à 1. Fixons $i, k \in \llbracket 1, n \rrbracket$ et $j, l \in \llbracket 1, p \rrbracket$. Seules la l -ième colonne et la j -ième ligne de la matrice $E_{ij}^T E_{kl}$ sont éventuellement non nulles. Cela revient à dire que tous les coefficients de cette matrice sont nuls, sauf peut-être son coefficient de position (j, l) . Et que vaut-il? Réponse : δ_{ik} , donc $E_{ij}^T E_{kl} = \delta_{ik} E_{jl}$. **CE CALCUL EST IMPORTANT ET MÉRITE QUE VOUS LE REFASIEZ SEULS .** Finalement : $\langle E_{ij}, E_{kl} \rangle = \text{tr} \left(E_{ij}^T E_{kl} \right) = \text{tr} \left(\delta_{ik} E_{jl} \right) = \delta_{ik} \delta_{jl} = \delta_{(i,j),(k,l)}$.

Exemple 33.25 – La famille de fonctions $c_k : x \mapsto \cos(kx)$, avec $k \in \mathbb{N}^*$, est orthonormale dans $\mathcal{C}([0, 2\pi], \mathbb{R})$ pour le produit scalaire $\langle f, g \rangle = \frac{1}{\pi} \int_0^{2\pi} f(t)g(t) dt$.

En effet,

$$\forall p, q \in \mathbb{N}, \quad \langle c_p, c_q \rangle = \frac{1}{\pi} \int_0^{2\pi} \cos(pt) \cos(qt) dt = \frac{1}{2\pi} \int_0^{2\pi} (\cos(p-q)t + \cos(p+q)t) dt$$

Or, pour tout $k \in \mathbb{N}$, $\int_0^{2\pi} \cos(kt) dt = 0$ si $k \neq 0$ et $\int_0^{2\pi} \cos(0t) dt = 2\pi$.

Ainsi $\langle c_p, c_q \rangle = \delta_{p,q}$ donc la famille est bien orthonormale.

On montre de même que les fonctions $d_k : x \mapsto \sin(kx)$ ($k \in \mathbb{N}^*$) sont orthonormales pour le même produit scalaire.

Exemple 33.26 – Soit a_0, \dots, a_n des réels deux à deux distincts. Considérons, pour tout $i \in \llbracket 0, n \rrbracket$, le polynôme interpolateur de Lagrange

$$L_i(X) = \prod_{j=0, j \neq i}^n \frac{X - a_j}{a_j - a_i}$$

Alors la famille $(L_i)_{i \in \llbracket 0, n \rrbracket}$ est une famille orthonormale pour le produit scalaire $\langle P, Q \rangle = \sum_{i=0}^n P(a_i)Q(a_i)$ sur $\mathbb{R}_n[X]$.

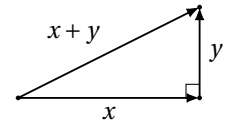
En effet, il suffit de remarquer que pour tous $i, j \in \llbracket 0, n \rrbracket$, $L_i(a_j) = \delta_{i,j}$ et les propriétés suivent directement.

Exemple 33.27 – Soit x et y deux vecteurs de même norme. Alors $x - y$ et $x + y$ sont orthogonaux car $\langle x - y, x + y \rangle = \|x\|^2 - \|y\|^2 = 0$. On retrouve l'une des propriétés des losanges.

Théorème 33.28 – Propriétés des familles orthogonales

Soit E un espace préhilbertien réel.

- Théorème de Pythagore :** Pour tous $x, y \in E$, x et y sont orthogonaux si et seulement si $\|x + y\|^2 = \|x\|^2 + \|y\|^2$.



En outre, pour toute famille orthogonale $(x_1, \dots, x_n) \in E$: $\left\| \sum_{i=1}^n x_i \right\|^2 = \sum_{i=1}^n \|x_i\|^2$.

- Liberté :** Toute famille orthogonale de vecteurs **NON NULS** de E est libre. En particulier, toute famille orthonormale de E est libre.

Démonstration.

- Le théorème de Pythagore vient directement de l'identité remarquable : $\|x + y\|^2 = \|x\|^2 + \|y\|^2 + 2\langle x, y \rangle$.

Pour le deuxième point, de manière similaire : $\left\| \sum_{i=1}^n x_i \right\|^2 = \sum_{i=1}^n \|x_i\|^2 + 2 \sum_{1 \leq i < j \leq n} \underbrace{\langle x_i, x_j \rangle}_{=0} = \sum_{i=1}^n \|x_i\|^2$.

- Soient (x_1, \dots, x_n) une famille orthogonale de vecteurs non nuls de E et $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ pour lesquels $\sum_{k=1}^n \lambda_k x_k = 0_E$. Pour

tout $i \in \llbracket 1, n \rrbracket$: $0 = \langle 0_E, x_i \rangle = \left\langle \sum_{k=1}^n \lambda_k x_k, x_i \right\rangle = \sum_{k=1}^n \lambda_k \langle x_k, x_i \rangle = \lambda_i \|x_i\|^2$, or $\|x_i\| \neq 0$ par séparation, donc $\lambda_i = 0$. La famille (x_1, \dots, x_n) est ainsi libre. □

Remarque 33.29 – Il est important de noter qu'avec plus de deux vecteurs, la réciproque du théorème de Pythagore est fautive. considérons en effet les vecteurs $e_1 = (2, 0)$, $e_2 = (0, 2)$ et $e_3 = (-1, 2)$ de \mathbb{R}^2 muni du produit scalaire canonique. Alors

$$\|e_1 + e_2 + e_3\|^2 = \|e_1\|^2 + \|e_2\|^2 + \|e_3\|^2,$$

et pourtant ces vecteurs ne sont pas orthogonaux entre eux.

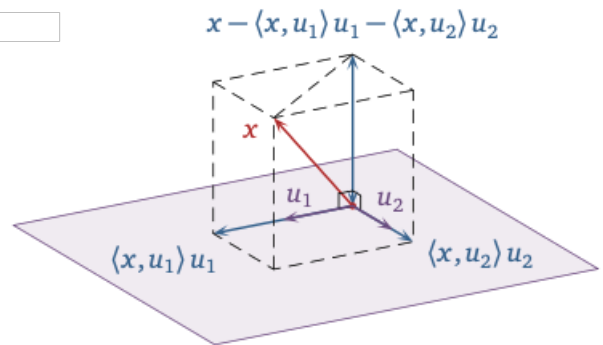
Terminons par un petit résultat qui nous sera bien utile dans la suite du chapitre.

Théorème 33.30

Soient E un espace préhilbertien réel, (u_1, \dots, u_n) une famille **ORTHONORMALE** de E et $x \in E$. Le vecteur $x - \sum_{k=1}^n \langle x, u_k \rangle u_k$ est orthogonal à u_1, \dots, u_n .

Démonstration. Pour tout $i \in \llbracket 1, n \rrbracket$:

$$\begin{aligned} \left\langle x - \sum_{k=1}^n \langle x, u_k \rangle u_k, u_i \right\rangle &= \langle x, u_i \rangle - \sum_{k=1}^n \langle x, u_k \rangle \langle u_k, u_i \rangle \\ &= \langle x, u_i \rangle - \sum_{k=1}^n \langle x, u_k \rangle \delta_{ki} = \langle x, u_i \rangle - \langle x, u_i \rangle = 0 \end{aligned}$$



Quelconque, x n'a aucune raison d'être orthogonal à u_1, \dots, u_n . Mais si on le « redresse » correctement, il le devient.

2 – Coordonnées dans une base orthonormale

Nous démontrerons bientôt que tout espace euclidien possède une base orthonormale, mais en attendant, quelques remarques simples sur les coordonnées dans une telle base.

Théorème 33.31 – Coordonnées dans une base orthonormale

Soient E un espace euclidien, (e_1, \dots, e_n) une base **ORTHONORMALE** de E et $x \in E$.

Alors, $x = \sum_{k=1}^n \langle x, e_k \rangle e_k$. En d'autres termes, les coordonnées de x dans (e_1, \dots, e_n) sont $(\langle x, e_1 \rangle, \dots, \langle x, e_n \rangle)$.

Démonstration. D'après le théorème qui précède, $x - \sum_{k=1}^n \langle x, e_k \rangle e_k$ est orthogonal à e_1, \dots, e_n , donc à tout vecteur de E par linéarité, donc ce vecteur est nul.

On peut aussi voir les choses autrement. Notons (x_1, \dots, x_n) les coordonnées de x dans (e_1, \dots, e_n) . Pour tout $k \in \llbracket 1, n \rrbracket$:

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

Théorème 33.32 – Expression du produit scalaire et de la norme dans une base orthonormale

Soient E un espace euclidien et $x, y \in E$ de coordonnées respectives $X = (x_1, \dots, x_n)$ et $Y = (y_1, \dots, y_n)$ dans une certaine base **ORTHONORMALE** de E .

$$\langle x, y \rangle = \sum_{k=1}^n x_k y_k = X^T Y \quad \text{et} \quad \|x\| = \sqrt{\sum_{k=1}^n x_k^2} = \sqrt{X^T X}.$$

Démonstration. Notons (e_1, \dots, e_n) la base orthonormale considérée.

Tout simplement : $\langle x, y \rangle = \left\langle \sum_{i=1}^n x_i e_i, \sum_{j=1}^n y_j e_j \right\rangle = \sum_{1 \leq i, j \leq n} x_i y_j \langle e_i, e_j \rangle = \sum_{1 \leq i, j \leq n} x_i y_j \delta_{ij} = \sum_{k=1}^n x_k y_k = X^T Y. \quad \square$

En d'autres termes, le produit scalaire canonique sur \mathbb{R}^n est un modèle pour tous les produits scalaires des espaces euclidiens. Calculer le produit scalaire $\langle x, y \rangle$ dans un espace euclidien abstrait revient à calculer le produit scalaire canonique des coordonnées des vecteurs x et y dans une base orthonormale quelconque.



ATTENTION ! Ces formules sont fausses en général pour des coordonnées dans une base non orthonormale !

3 – Algorithme d'orthonormalisation de Gram-Schmidt

Théorème 33.33 – Gram-Schmidt

Soit (e_1, \dots, e_n) une famille libre d'un espace préhilbertien E . Il existe une unique famille orthonormale $(\varepsilon_1, \dots, \varepsilon_n)$ telle que pour tout $p \in \llbracket 1, n \rrbracket$,

- (e_1, \dots, e_p) et $(\varepsilon_1, \dots, \varepsilon_p)$ engendrent le même sous-espace ;
- $\langle e_p, \varepsilon_p \rangle > 0$.

La première condition est essentielle pour la construction qui suit alors que la seconde condition ne sert intuitivement qu'à garantir l'unicité (en conservant « l'orientation » de la base).

Démonstration. On construit la famille $(\varepsilon_1, \dots, \varepsilon_n)$ par une récurrence finie sur $p \in \llbracket 1, n \rrbracket$. Le cas $n = 1$ est traité dans l'initialisation et on suppose par la suite que $n \geq 2$.

- Le vecteur ε_1 doit être unitaire, colinéaire à e_1 et vérifier la condition $\langle e_1, \varepsilon_1 \rangle > 0$: ε_1 est nécessairement égal à $\frac{1}{\|e_1\|} e_1$. La propriété est donc vraie pour $p = 1$ (et a fortiori pour $n = 1$).
- Soit $p \in \llbracket 1, n - 1 \rrbracket$. Supposons établie l'existence et l'unicité de la famille $(\varepsilon_1, \dots, \varepsilon_p)$ satisfaisant les conditions de l'énoncé. Si un vecteur ε_{p+1} satisfait aux conditions du théorème alors

$$\varepsilon_{p+1} \in \text{Vect}(e_1, \dots, e_p) \oplus \text{Vect}(e_{p+1}) = \text{Vect}(\varepsilon_1, \dots, \varepsilon_p) \oplus \text{Vect}(e_{p+1}).$$

Par conséquent, ε_{p+1} est de la forme

$$\sum_{i=1}^p \alpha_i \varepsilon_i + \alpha e_{p+1},$$

avec $\alpha \neq 0$ (car sinon $\varepsilon_{p+1} \in \text{Vect}(\varepsilon_1, \dots, \varepsilon_p)$ et la famille n'est pas libre et ne peut donc être orthonormale).

Des relations d'orthogonalité, pour $i \leq p$, on obtient $\alpha_i = -\alpha \langle e_{p+1}, \varepsilon_i \rangle$. De plus, comme ε_{p+1} est unitaire, on obtient

$$|\alpha| = \frac{1}{\left\| e_{p+1} - \sum_{i=1}^p \langle e_{p+1}, \varepsilon_i \rangle \varepsilon_i \right\|}.$$

Enfin la condition $\langle e_{p+1}, \varepsilon_{p+1} \rangle > 0$ se réécrit

$$\alpha \left(\|e_{p+1}\|^2 - \sum_{i=1}^p \langle e_{p+1}, \varepsilon_i \rangle^2 \right) > 0;$$

Or les vecteurs $e_{p+1} - \sum_{i=1}^p \langle e_{p+1}, \varepsilon_i \rangle \varepsilon_i$ et $\sum_{i=1}^p \langle e_{p+1}, \varepsilon_i \rangle \varepsilon_i$ sont orthogonaux d'après le Théorème 33.30. Le théorème de Pythagore appliqué à ces deux vecteurs donne alors :

$$\|e_{p+1}\|^2 - \sum_{i=1}^p \langle e_{p+1}, \varepsilon_i \rangle^2 = \|e_{p+1} - \sum_{i=1}^p \langle e_{p+1}, \varepsilon_i \rangle \varepsilon_i\|^2 > 0$$

d'où $\alpha > 0$, i.e $\alpha = \frac{1}{\left\| e_{p+1} - \sum_{i=1}^p \langle e_{p+1}, \varepsilon_i \rangle \varepsilon_i \right\|}$.

Réciproquement, le vecteur $\frac{e_{p+1} - \sum_{i=1}^p \langle e_{p+1}, \varepsilon_i \rangle \varepsilon_i}{\|e_{p+1} - \sum_{i=1}^p \langle e_{p+1}, \varepsilon_i \rangle \varepsilon_i\|}$ satisfait bien les conditions. Ce qui achève de démontrer l'existence et l'unicité du vecteur ε_{p+1} . □

L'algorithme de Gram-Schmidt construit ε_p en « redressant » e_{p+1} , i.e en le rendant orthogonal à $\varepsilon_1, \dots, \varepsilon_{p-1}$ (voir le Théorème 33.30) puis en le rendant unitaire.

Exemple 33.34 – Soit $e_1 = (3, 4, 0)$, $e_2 = (5, 0, 0)$, $e_3 = (0, 0, 1)$ des vecteurs qui forment une base de \mathbb{R}^3 (muni du produit scalaire canonique). Déterminons la base orthonormée $(\varepsilon_1, \varepsilon_2, \varepsilon_3)$ associée par l'algorithme de Gram-Schmidt.

Commençons par normaliser e_1 pour obtenir ε_1 . $\|e_1\| = 5$, donc $\varepsilon_1 = \left(\frac{3}{5}, \frac{4}{5}, 0\right)$.

On sait ensuite que ε_2 est de la forme $\lambda(e_2 + \mu\varepsilon_1)$ avec $\mu = -\langle e_2, \varepsilon_1 \rangle = -3$. Pour obtenir ε_2 , il suffit alors de normaliser $e_2 - 3\varepsilon_1$. Comme $\|e_2 - 3\varepsilon_1\| = 4$, on obtient $\varepsilon_2 = \left(\frac{4}{5}, -\frac{3}{5}, 0\right)$.

On recommence la même procédure avec e_3 en l'écrivant comme combinaison linéaire de $e_3, \varepsilon_1, \varepsilon_2$. Cependant on peut observer que e_3 est déjà orthogonal aux deux premiers vecteurs ε_1 et ε_2 et est de norme 1, ainsi $\varepsilon_3 = e_3$.

La base recherchée est donc :

$$\left(\frac{3}{5}, \frac{4}{5}, 0\right), \left(\frac{4}{5}, -\frac{3}{5}, 0\right), (0, 0, 1).$$

Exemple 33.35 – La famille $(1, \sqrt{3}(2X - 1), \sqrt{5}(6X^2 - 6X + 1))$ est orthonormale pour le produit scalaire

$(P, Q) \mapsto \int_0^1 P(t)Q(t) dt$ sur $\mathbb{R}[X]$. Exercice en passant : justifier que c'est un produit scalaire sur $\mathbb{R}[X]$.

Exploitions l'algorithme de Gram-Schmidt pour construire une famille orthonormale (U_0, U_1, U_2) à partir de la famille libre $(1, X, X^2)$.

• **Construction de U_0** : $\|1\|^2 = \int_0^1 dt = 1$, donc on pose $U_0 = \frac{1}{\|1\|} = 1$.

• **Construction de U_1** : $\langle X, U_0 \rangle = \int_0^1 t dt = \frac{1}{2}$, donc $X - \langle X, U_0 \rangle U_0 = X - \frac{1}{2}$.

Ensuite : $\left\| X - \frac{1}{2} \right\|^2 = \int_0^1 \left(t - \frac{1}{2}\right)^2 dt = \frac{1}{12}$, donc on pose $U_1 = \frac{X - \langle X, U_0 \rangle U_0}{\|X - \langle X, U_0 \rangle U_0\|} = \sqrt{3}(2X - 1)$.

• **Construction de U_2** : $\langle X^2, U_0 \rangle = \int_0^1 t^2 dt = \frac{1}{3}$ et $\langle X^2, U_1 \rangle = \int_0^1 t^2 \times \sqrt{3}(2t - 1) dt = \frac{1}{2\sqrt{3}}$, donc

$$X^2 - \langle X^2, U_0 \rangle U_0 - \langle X^2, U_1 \rangle U_1 = X^2 - X + \frac{1}{6}.$$

Ensuite : $\left\| X^2 - X + \frac{1}{6} \right\|^2 = \int_0^1 \left(t^2 - t + \frac{1}{6}\right)^2 dt = \frac{1}{180}$, donc on pose

$$U_2 = \frac{X^2 - \langle X^2, U_0 \rangle U_0 - \langle X^2, U_1 \rangle U_1}{\|X^2 - \langle X^2, U_0 \rangle U_0 - \langle X^2, U_1 \rangle U_1\|} = \sqrt{5}(6X^2 - 6X + 1). \text{ Ouf!}$$

Essentiel en pratique, l'algorithme d'orthonormalisation de Gram-Schmidt a aussi des conséquences théoriques.

Théorème 33.36 – Existence de bases orthonormales en dimension finie

1. Tout espace euclidien possède une base orthonormale.
2. **Théorème de la base orthonormale incomplète :** Soit E un espace euclidien. Toute famille orthonormale de E peut être complétée en une base orthonormale de E .

Démonstration.

1. Tout \mathbb{R} -espace vectoriel de dimension finie possède une base - éventuellement vide - donc une base orthonormale grâce à l'algorithme de Gram-Schmidt.
2. Libre, toute famille orthonormale \mathcal{F} de E peut être complétée en une base de E . Or appliqué à cette base, l'algorithme de Gram-Schmidt n'affecte pas les vecteurs de \mathcal{F} (ils sont déjà orthonormés!) et on peut donc considérer que \mathcal{F} a été complétée en une base orthonormale de E .

□

III – Sous-espace orthogonal et projection orthogonale

1 – Supplémentaire orthogonal d'un sous-espace vectoriel

Définition-Théorème 33.37 – Soit E un espace préhilbertien réel et X une partie de E . On appelle **orthogonal de X dans E** l'ensemble $X^\perp = \{y \in E \mid \forall x \in X, \langle x, y \rangle = 0\}$.

1. X^\perp est un sous-espace vectoriel de E , orthogonal à X .
2. Si X est un sous-espace vectoriel de E , X et X^\perp sont en somme directe.
3. $X^\perp = \text{Vect}(X)^\perp$ et $X \subset (X^\perp)^\perp$.

Démonstration.

1. D'abord $0_E \in X^\perp$ car $\langle 0_E, x \rangle = 0$ pour tout $x \in X$. Pour la stabilité par combinaison linéaire, soient $y_1, y_2 \in X^\perp$ et $\lambda \in \mathbb{R}$. Pour tout $x \in X$: $\langle x, \lambda y_1 + y_2 \rangle = \lambda \langle x, y_1 \rangle + \langle y_2, x \rangle = \lambda \cdot 0 + 0 = 0$, donc $\lambda y_1 + y_2 \in X^\perp$.
2. Si X est un sous-espace vectoriel de E , $X \cap X^\perp$ contient 0_E . Inversement, pour tout $x \in X \cap X^\perp$: $x \perp x$ donc $\langle x, x \rangle = 0$, donc $x = 0_E$ par séparation.
3. Par linéarité à droite du produit scalaire, il est équivalent pour un vecteur d'être orthogonal à X ou à ses combinaisons linéaires, autrement dit à $\text{Vect}(X)$, donc $X^\perp = \text{Vect}(X)^\perp$. Enfin, l'inclusion $X \subset (X^\perp)^\perp$ signifie juste que tout vecteur de X est orthogonal à tout vecteur de X^\perp - ce qui est évident par définition de X^\perp .

□

Remarque 33.38 –

- On fera bien attention à noter que X^\perp est un espace vectoriel pour toute partie X , même si X n'est pas un sous-espace vectoriel.
- L'égalité $X^\perp = \text{Vect}(X)^\perp$ signifie que pour déterminer l'orthogonal d'un sous-espace vectoriel F , il n'est pas nécessaire d'exiger l'orthogonalité à **TOUS** les vecteurs de F , l'orthogonalité à tout vecteur d'une partie génératrice suffit.



ATTENTION ! Pour un sous-espace vectoriel F de E , F et F^\perp sont **TOUJOURS** en somme directe. En revanche, il n'est pas vrai en général que F et F^\perp sont supplémentaires dans E , ils sont seulement en somme directe. Contrairement aux apparences, il n'est pas non plus vrai en général que $F^{\perp\perp} = F$. Nous y reviendrons.

Exemple 33.39 – Pour tout espace préhilbertien réel E : $\{0_E\}^\perp = E$ et $E^\perp = \{0_E\}$.

En effet, pour tout $x \in E$, on a $\langle 0_E, x \rangle = 0$ donc $E \subset \{0_E\}^\perp$. L'inclusion réciproque étant toujours vraie, on a bien $\{0_E\}^\perp = E$. De plus, si $x \in E^\perp$ alors pour tout $y \in E$, on a $\langle x, y \rangle = 0$. En particulier, $\langle x, x \rangle = 0$ d'où $x = 0_E$ par séparation. Donc, $E^\perp \subset \{0_E\}$, et l'inclusion réciproque étant toujours vraie (E^\perp est un sev!), on a bien $E^\perp = \{0_E\}$.

Exemple 33.40 – Dans l'espace euclidien canonique \mathbb{R}^3 , le plan vectoriel d'équation $3x - y + 2z = 0$ n'est jamais que l'orthogonal $\{(3, -1, 2)\}^\perp$ car pour tout $(x, y, z) \in \mathbb{R}^3$: $\langle (3, -1, 2), (x, y, z) \rangle = 3x - y + 2z$. En termes géométriques, le vecteur $(3, -1, 2)$ est normal au plan d'équation $3x - y + 2z = 0$.

Exemple 33.41 – On considère $\mathbb{R}_2[X]$ muni du produit scalaire $(P, Q) \mapsto P(-1)Q(-1) + P(0)Q(0) + P(1)Q(1)$ (c'est bien un produit scalaire puisqu'il s'agit d'un cas particulier de l'exemple 33.8). Montrer que $\mathbb{R}_1[X]^\perp = \text{Vect}(3X^2 - 2)$. Pour tout $P = aX^2 + bX + c \in \mathbb{R}_2[X]$:

$$P \in \mathbb{R}_1[X]^\perp \stackrel{\mathbb{R}_1[X] = \text{Vect}(1, X)}{\iff} \langle P, 1 \rangle = 0 \text{ et } \langle P, X \rangle = 0$$

$$\iff \begin{cases} (a - b + c) \times 1 + c \times 1 + (a + b + c) \times 1 = 0 \\ (a - b + c) \times (-1) + c \times 0 + (a + b + c) \times 1 = 0 \end{cases} \iff 2a + 3c = 0 \text{ et } 2b = 0.$$

Le point important de cet exemple, c'est que l'orthogonalité à 1 et X a suffi car $\mathbb{R}_1[X] = \text{Vect}(1, X)$.

Définition-Théorème 33.42 – Soient E un espace préhilbertien réel et F un sous-espace vectoriel **DE DIMENSION FINIE** de E .

- Supplémentaire orthogonal** : F^\perp est un supplémentaire de F dans E , orthogonal à F , et c'est même le seul. On l'appelle par conséquent **LE supplémentaire orthogonal** de F dans E . Pour résumer, on écrit souvent les choses ainsi : $E = F \oplus F^\perp$.
En particulier, si E lui-même est de dimension finie : $\dim F^\perp = \dim E - \dim F$.
- Bi-orthogonal** : $F^{\perp\perp} = F$.

Démonstration.

- Nous savons déjà que $F \cap F^\perp = \{0_E\}$, et bien sûr $F \perp F^\perp$. Pour montrer que $E = F \oplus F^\perp$, il nous reste à montrer que $E \subset F + F^\perp$. Comme F est de dimension finie, nous pouvons nous en donner une base orthonormale (f_1, \dots, f_n) — éventuellement vide. Pour tout $x \in E$, le vecteur $x - \sum_{k=1}^n \langle x, f_k \rangle f_k$ est orthogonal à f_1, \dots, f_n , donc appartient à $\text{Vect}(f_1, \dots, f_n)^\perp = F^\perp$, donc $x \in F + F^\perp$.
Montrons que F^\perp est le seul supplémentaire de F dans E orthogonal à F . Soit F' un tel supplémentaire. Aussitôt $F' \subset F^\perp$ car $F \perp F'$. Inversement, soit $x \in F^\perp$. Comme $E = F + F'$: $x = f + f'$ pour certains $f \in F$ et $f' \in F'$, donc $\langle f, f' \rangle = \langle f + f', f \rangle = \langle x, f \rangle = 0$, puis $f = 0_E$ par séparation, et enfin $x = f' \in F'$.
- Nous savons déjà que $F \subset F^{\perp\perp}$. Inversement, pour tout $x \in F^{\perp\perp}$, disons $x = f + f'$ avec $f \in F$ et $f' \in F^\perp$: $\langle f', f' \rangle = \langle f + f', f' \rangle = \langle x, f' \rangle = 0$, donc $f' = 0_E$ par séparation, i.e. $x = f \in F$.

□



ATTENTION! Il est ici essentiel que F soit **DE DIMENSION FINIE**. Par ailleurs, il existe un unique supplémentaire orthogonal mais tout plein de supplémentaires en toute généralité.

Exemple 33.43 – Dans l'espace euclidien canonique \mathbb{R}^4 , le plan vectoriel P d'équations $x - y - z - t = 0$ et $2x + y + z - t = 0$ admet $\text{Vect}((1, -1, -1, -1), (2, 1, 1, -1))$ pour orthogonal.

Démonstration.

$$P = \{(x, y, z, t) \in \mathbb{R}^4 \mid \langle (1, -1, -1, -1), (x, y, z, t) \rangle = 0 \text{ et } \langle (2, 1, 1, -1), (x, y, z, t) \rangle = 0\}$$

$$= \{(1, -1, -1, -1), (2, 1, 1, -1)\}^\perp = \text{Vect}((1, -1, -1, -1), (2, 1, 1, -1))^\perp$$

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

Exemple 33.44 – On munit $\mathcal{M}_n(\mathbb{R})$ de sa structure euclidienne canonique - cela veut dire qu'on travaille avec le produit scalaire canonique de $\mathcal{M}_n(\mathbb{R})$. Dans ce contexte : $\mathcal{S}_n(\mathbb{R})^\perp = \mathcal{A}_n(\mathbb{R})$.

Démonstration. Pour tous $S \in \mathcal{S}_n(\mathbb{R})$ et $A \in \mathcal{A}_n(\mathbb{R})$:

$$\langle A, S \rangle = \text{tr}(A^\top S) = -\text{tr}(AS) = -\text{tr}(SA) = -\text{tr}(S^\top A) = -\langle S, A \rangle = -\langle A, S \rangle, \text{ donc } \langle A, S \rangle = 0,$$

donc $\mathcal{S}_n(\mathbb{R})$ et $\mathcal{A}_n(\mathbb{R})$ sont orthogonaux - donc en somme directe. Bref : $\mathcal{A}_n(\mathbb{R}) \subset \mathcal{S}_n(\mathbb{R})^\perp$, et notez bien que ce n'est qu'une inclusion à ce stade. Cela dit, pour tout $M \in \mathcal{M}_n(\mathbb{R})$: $M = \frac{M + M^\top}{2} + \frac{M - M^\top}{2}$ avec $\frac{M + M^\top}{2} \in \mathcal{S}_n(\mathbb{R})$ et $\frac{M - M^\top}{2} \in \mathcal{A}_n(\mathbb{R})$, donc $\mathcal{M}_n(\mathbb{R}) = \mathcal{S}_n(\mathbb{R}) + \mathcal{A}_n(\mathbb{R})$. Finalement $\mathcal{M}_n(\mathbb{R}) = \mathcal{S}_n(\mathbb{R}) \oplus \mathcal{A}_n(\mathbb{R})$, donc par unicité du supplémentaire orthogonal : $\mathcal{A}_n(\mathbb{R}) = \mathcal{S}_n(\mathbb{R})^\perp$. □

2 – Projection orthogonale sur un sous-espace vectoriel de dimension finie

Définition 33.45 – Soient E un espace préhilbertien réel et F un sous-espace vectoriel **DE DIMENSION FINIE** de E . On appelle **projection orthogonale sur F** ou **projecteur orthogonal sur F** la projection sur F parallèlement à F^\perp .

Comme F est de dimension finie, d'après le Théorème 33.42, on a $E = F \oplus F^\perp$, et il est donc légitime de parler de la projection sur F parallèlement à F^\perp . On pourrait également s'intéresser à la symétrie par rapport à F parallèlement à F^\perp .



Théorème 33.46 – Expression d'un projeté orthogonal dans une base orthonormale

Soient E un espace préhilbertien réel, F un sous-espace vectoriel **DE DIMENSION FINIE** de E et (f_1, \dots, f_n) une base orthonormale de F .

Pour tout $x \in E$, le projeté orthogonal de x sur F vaut $\sum_{k=1}^n \langle x, f_k \rangle f_k$.

Démonstration. D'après nos résultats précédents : $x = \underbrace{\sum_{k=1}^n \langle x, f_k \rangle f_k}_{\in F} + \underbrace{(x - \sum_{k=1}^n \langle x, f_k \rangle f_k)}_{\in F^\perp}$. □

Méthode 33.47 –

On peut calculer essentiellement de deux manières un projeté orthogonal. Donnons-nous E un espace préhilbertien réel, F un sous-espace vectoriel de dimension finie de E , (f_1, \dots, f_n) une base pas forcément orthonormale de F et $x \in E$. Comment calculer le projeté orthogonal $p(x)$ de x sur F ? Si la base (f_1, \dots, f_n) est orthonormale, on peut bien sûr utiliser le théorème précédent, mais sinon?

- **Première stratégie** : On s'y ramène de force en orthonormalisant (f_1, \dots, f_n) grâce à l'algorithme de Gram-Schmidt.
- **Deuxième stratégie** : Le projeté $p(x)$ est caractérisé par deux assertions :

$$p(x) \in F \quad \text{et} \quad x - p(x) \in F^\perp.$$

Concrètement, on introduit les coordonnées $(\lambda_1, \dots, \lambda_n)$ de $p(x)$ dans (f_1, \dots, f_n) : $p(x) = \lambda_1 f_1 + \dots + \lambda_n f_n$, puis on les calcule grâce aux relations $\langle x - p(x), f_j \rangle = 0$, j décrivant $\llbracket 1, n \rrbracket$. Ces relations expriment l'appartenance de $x - p(x)$ à F^\perp et fournissent un système de n équations à n inconnues qu'on n'a plus qu'à résoudre.

Illustrons ces deux stratégies sur deux exemples.



Exemple 33.48 – Dans \mathbb{R}^4 muni de son produit scalaire canonique, on considère F le sous-espace vectoriel défini par

$$F = \{(x, y, z, t) \in \mathbb{R}^4 : x + y + t = 0 \text{ et } x + y + 2z - t = 0\}.$$

Déterminer le projeté orthogonal de $u = (1, 8, 1, 1)$ sur F .

On commence par rechercher une base de F . Pour cela on écrit que

$$(x, y, z, t) \in F \iff \begin{cases} x + y + t = 0 \\ x + y + 2z - t = 0 \end{cases} \xrightarrow{L_2 - L_1} \begin{cases} x + y + t = 0 \\ 2z - 2t = 0 \end{cases} \iff \begin{cases} x = -y \\ y = -t \\ z = t \\ t = t \end{cases}$$

Ainsi, si on pose $u_1 = (-1, 1, 0, 0)$ et $u_2 = (-1, 0, 1, 1)$, on trouve que (u_1, u_2) est une base de F .

Notons ensuite $p_F(u)$ le projeté orthogonal de u sur F et illustrons les deux méthodes pour le calculer.

- **Première stratégie** : on va orthonormaliser la base (u_1, u_2) . Puisque $\|u_1\| = \sqrt{2}$, on pose $v_1 = \frac{1}{\sqrt{2}}(-1, 1, 0, 0)$.

On cherche ensuite $u'_2 = u_2 + \alpha v_1$ de sorte que $\langle u'_2, v_1 \rangle = 0$. Ceci donne $\frac{1}{\sqrt{2}} + \alpha = 0 \iff \alpha = -\frac{1}{\sqrt{2}}$.

On obtient $u'_2 = (-1, 0, 1, 1) + \left(\frac{1}{2}, -\frac{1}{2}, 0, 0\right) = \left(-\frac{1}{2}, -\frac{1}{2}, 1, 1\right)$. D'autre part, $\|u'_2\|^2 = \frac{1}{4} + \frac{1}{4} + 1 + 1 = \frac{10}{4}$

et donc on pose $v_2 = \frac{u'_2}{\|u'_2\|} = \frac{1}{\sqrt{10}}(-1, -1, 2, 2)$ On sait ensuite que $p_F(u) = \langle u, v_1 \rangle v_1 + \langle u, v_2 \rangle v_2$.

Or, $\langle u, v_1 \rangle = \frac{7}{\sqrt{2}}$ et $\langle u, v_2 \rangle = \frac{-5}{\sqrt{10}}$ de sorte que $\langle u, v_1 \rangle v_1 = \left(-\frac{7}{2}, \frac{7}{2}, 0, 0\right)$ et $\langle u, v_2 \rangle v_2 = \left(\frac{1}{2}, \frac{1}{2}, -1, -1\right)$.

Après un dernier petit calcul, on trouve $p_F(u) = (-3, 4, -1, -1)$.

- **Deuxième stratégie** : on écrit que $p_F(u) = au_1 + bu_2 = (-a - b, a, b, b)$ de sorte que $u - p_F(u) = (1 + a + b, 8 - a, 1 - b, 1 - b)$. On sait que $u - p_F(u) \perp u_1$. Calculant le produit scalaire, on trouve

$$-1 - a - b + 8 - a = 0 \iff 2a + b = 7.$$

On sait aussi que $u - p_F(u) \perp u_2$ et toujours avec l'aide du produit scalaire :

$$-1 - a - b + 1 - b + 1 - b = 0 \iff a + 3b = 1.$$

Ainsi, (a, b) est solution du système suivant, que l'on va résoudre :

$$\begin{cases} a + 3b = 1 \\ 2a + b = 7 \end{cases} \iff \begin{cases} a + 3b = 1 \\ -5b = 5 \end{cases} \iff \begin{cases} a = 4 \\ b = 7 \end{cases}$$

On retrouve $p_F(u) = (-3, 4, -1, -1)$.

Exemple 33.49 – On note F le sous-espace vectoriel $\text{Vect}(\cos, \sin)$ de $\mathcal{C}([0, 2\pi], \mathbb{R})$. Le projeté orthogonal de l'identité Id sur F pour le produit scalaire $(f, g) \mapsto \int_0^{2\pi} f(t)g(t) dt$ est la fonction $t \mapsto -2 \sin t$.

Démonstration. Nous aurons besoin d'un certain nombre de produits scalaires, calculons-les de prime abord pour que l'essentiel des stratégies adoptées soit bien lisible ensuite.

$$\|\cos\|^2 = \int_0^{2\pi} \cos^2 t dt = \int_0^{2\pi} \frac{1 + \cos(2t)}{2} dt = \left[\frac{t}{2} + \frac{\sin(2t)}{4} \right]_{t=0}^{t=2\pi} = \pi, \quad \text{et de même } \|\sin\|^2 = \pi$$

En outre : $\langle \cos, \sin \rangle = \int_0^{2\pi} \cos t \sin t dt = \left[\frac{\sin^2 t}{2} \right]_{t=0}^{t=2\pi} = 0$, donc la famille (\cos, \sin) est orthogonale. Enfin :

$$\int_0^{2\pi} t e^{it} dt = \left[t \times (-i) e^{it} \right]_{t=0}^{t=2\pi} - \int_0^{2\pi} (-i) e^{it} dt = -2i\pi + i \left[(-i) e^{it} \right]_{t=0}^{t=2\pi} = -2i\pi, \quad \text{donc :}$$

$$\langle \text{Id}, \sin \rangle = \int_0^{2\pi} t \sin t dt = \text{Im} \left(\int_0^{2\pi} t e^{it} dt \right) = \text{Im}(-2i\pi) = -2\pi \quad \text{et de même } \langle \text{Id}, \cos \rangle = 0$$

- **Première stratégie :** On orthonormalise la base (\cos, \sin) de F grâce à l’algorithme de Gram-Schmidt. Cette famille étant déjà orthogonale, $\left(\frac{\cos}{\|\cos\|}, \frac{\sin}{\|\sin\|}\right) = \left(\frac{\cos}{\sqrt{\pi}}, \frac{\sin}{\sqrt{\pi}}\right)$ est une base orthonormale de F . Le projeté orthogonal de Id sur F est finalement la fonction :

$$\left\langle \text{Id}, \frac{\cos}{\sqrt{\pi}} \right\rangle \frac{\cos}{\sqrt{\pi}} + \left\langle \text{Id}, \frac{\sin}{\sqrt{\pi}} \right\rangle \frac{\sin}{\sqrt{\pi}} = \frac{\langle \text{Id}, \cos \rangle}{\pi} \cos + \frac{\langle \text{Id}, \sin \rangle}{\pi} \sin = -2 \sin.$$

- **Deuxième stratégie :** Notons $p(\text{Id})$ le projeté orthogonal de Id sur F et (λ, μ) ses coordonnées dans la base (\cos, \sin) de F : $p(\text{Id}) = \lambda \cos + \mu \sin$. Appuyons-nous sur le fait que $\text{Id} - p(\text{Id}) \in F^\perp$.

$$0 = \langle \text{Id} - p(\text{Id}), \cos \rangle = \langle \text{Id} - \lambda \cos - \mu \sin, \cos \rangle = \langle \text{Id}, \cos \rangle - \lambda \|\cos\|^2 - \mu \langle \sin, \cos \rangle = -\lambda \pi$$

$$\text{et } 0 = \langle \text{Id} - p(\text{Id}), \sin \rangle = \langle \text{Id} - \lambda \cos - \mu \sin, \sin \rangle = \langle \text{Id}, \sin \rangle - \lambda \langle \cos, \sin \rangle - \mu \|\sin\|^2 = -2\pi - \mu \pi$$

Conclusion : $\lambda = 0$ et $\mu = -2$, donc $p(\text{Id}) = -2 \sin$.

□

3 – Distance à un sous-espace de dimension finie

Définition 33.50 – Soit E un espace préhilbertien réel, F un sev de E de dimension finie et $x \in E$. On appelle **distance de x à F** , le réel noté $d(x, F)$ défini par $d(x, F) = \inf_{y \in F} \|x - y\|$.

Intuitivement, la distance d’un vecteur x à un sous-espace vectoriel F est la plus petite distance séparant x d’un élément de F , mais qui nous dit qu’une telle plus petite distance existe? Elle n’existe justement pas forcément et c’est pour cela que la définition introduit une borne inférieure. Ici, l’ensemble $\{\|x - y\| \mid y \in F\}$ est une partie de \mathbb{R} non vide (car $F \neq \emptyset$) et minorée par 0, donc possède bien une borne inférieure d’après la propriété de la borne inférieure.

Proposition 33.51

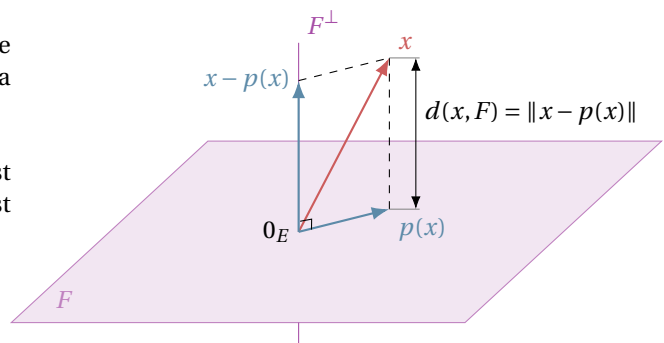
Soit E un espace préhilbertien, $x \in E$ et F un sous-espace de E de dimension finie. Alors $d(x, F) = \|x - p_F(x)\|$, où p_F est la projection orthogonale sur le sous-espace F .

Théorème 33.52 – Distance à un sous-espace vectoriel de dimension finie

Soient E un espace préhilbertien réel, F un sous-espace vectoriel **DE DIMENSION FINIE** de E et $x \in E$. On note p la projection orthogonale sur F .

La distance de x à F est mieux qu’une borne inférieure, c’est un minimum : $d(x, F) = \|x - p(x)\|$, et ce minimum n’est atteint qu’en $p(x)$.

Par ailleurs : $d(x, F)^2 = \|x\|^2 - \|p(x)\|^2$.



Démonstration.

- Pour tout $f \in F$: $x - f = \overbrace{(x - p(x))}^{\in \text{Ker } p = F^\perp} + \overbrace{(p(x) - f)}^{\in \text{Im } p = F}$, donc $\|x - f\|^2 = \|x - p(x)\|^2 + \|p(x) - f\|^2$ d’après le théorème de Pythagore.
- Notons à présent \mathcal{D} l’ensemble $\{\|x - f\| \mid f \in F\}$. Montrons que $\|x - p(x)\| = \min \mathcal{D}$, cela prouvera que $\|x - p(x)\| = \inf \mathcal{D} = d(x, F)$. Or d’une part $\|x - p(x)\| \in \mathcal{D}$ car $p(x) \in F$, et d’autre part, pour tout $f \in F$: $\|x - f\| = \sqrt{\|x - p(x)\|^2 + \|p(x) - f\|^2} \geq \|x - p(x)\|$, donc \mathcal{D} est minoré par $\|x - p(x)\|$.
- Pour finir, pour tout $f \in F \setminus \{p(x)\}$: $\|x - f\| = \sqrt{\|x - p(x)\|^2 + \|p(x) - f\|^2} > \|x - p(x)\|$ car $\|p(x) - f\| > 0$ par séparation. Comme voulu, la distance $d(x, F)$ n’est atteinte qu’en $p(x)$.

□

Exemple 33.53 – Voyons un problème de minimisation classique. Objectif : calculer $\inf_{a,b,c \in \mathbb{R}} \int_0^1 (t^3 - (at^2 + bt + c))^2 dt$.

Première étape : Traduire cela en un problème de distance, c'est-à-dire le voir sous la forme $\inf_{y \in F} d(x, y)$ ou éventuellement, $\inf_{y \in F} d(x, y)^2$. Pour cela :

- Sur quel espace E travaille-t-on? Ici, $E = C^0([0, 1])$ ou les fonctions polynomiales.
- Avec quel produit scalaire? avec le produit scalaire intégral $\langle f, g \rangle = \int_0^1 f(t)g(t) dt$.
- Qui joue le rôle de x ? (C'est la partie qui ne dépend pas des paramètres qui varient.) Ici, il s'agit de la fonction $f : t \mapsto t^3$.
- Qui joue le rôle de y ? (C'est la partie qui varie en fonction des paramètres.) Ici, il s'agit de la fonction $g : t \mapsto at^2 + bt + c$.
- Quel est le sous-espace vectoriel F ? (C'est l'ensemble qui est décrit par les vecteurs y quand les paramètres varient.) Ici, $F = \mathbb{R}_2[X]$.

Deuxième étape : On sait que cet inf est atteint lorsque l'on choisit pour g le projeté orthogonal de f sur F . Le nombre cherché est alors $\|f - p_F(f)\|^2$.

Il nous reste donc à déterminer ce projeté orthogonal par l'une des méthodes décrites précédemment.

Ici, cherchons une base orthonormée de $F = \text{Vect}(g_0, g_1, g_2)$ avec $g_k : t \mapsto t^k$. On effectue donc un procédé de Gram-Schmidt. On obtient au bout d'un temps fini $f_0 = 1, f_1 = \sqrt{3}(2t - 1), f_2 = \sqrt{5}(6t^2 - 6t + 1)$.

Finalement $p_F(f) = \frac{3}{2}t^2 - \frac{3}{5}t + \frac{1}{20}$. Il n'y a plus qu'à calculer

$$\|f - p_F(f)\|^2 = \|f\|^2 - \|p_F(f)\|^2 = \frac{1}{7} - \frac{57}{400} = \frac{1}{2800},$$

et donc $\min_{a,b,c \in \mathbb{R}} \int_0^1 (t^3 - (at^2 + bt + c))^2 dt = \frac{1}{2800}$.