Lycée Jules Ferry, Cannes
Année 2019-2020

Classe préparatoire TSI 2e année
Mathématiques - D. Broizat

Chapitre 12
Espaces préhilbertiens et euclidiens

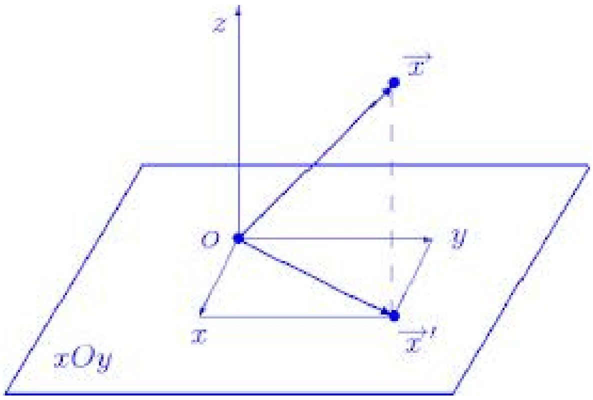

Table des matières

E désigne un espace vectoriel **réel** ($\mathbb{K} = \mathbb{R}$), pas nécessairement de dimension finie.
On rappelle que $E \times E$ désigne l'ensemble des couples (x, y), où x et y sont deux vecteurs de E.

I Produits scalaires

1) Définition

Définition 1 (Produit scalaire).
On appelle **produit scalaire sur** E *toute application* $\varphi : E \times E \to \mathbb{R}$ *qui est :*

(i) *symétrique :* $\forall (x, y) \in E^2,\ \varphi(y, x) = \varphi(x, y)$.

(ii) *bilinéaire :* $\forall (x, y, z) \in E^3,\ \forall \lambda \in \mathbb{R}$ *:*
$\varphi(\lambda x + y, z) = \lambda \varphi(x, z) + \varphi(y, z)$ *(linéarité "à gauche")* ;
$\varphi(x, \lambda y + z) = \lambda \varphi(x, y) + \varphi(x, z)$ *(linéarité "à droite")* ;

(iii) *positive :* $\forall x \in E,\ \varphi(x, x) \geq 0$;

(iv) *définie :* $\forall x \in E,\ (\varphi(x, x) = 0 \Longrightarrow x = 0_E)$.

Vocabulaire
On dit qu'un produit scalaire est une "forme bilinéaire symétrique définie positive".

Remarque (TRES UTILE)
Si une application $\varphi : E \times E \to \mathbb{R}$ est symétrique et linéaire à gauche, alors la linéarité à droite est automatique.

ATTENTION !
La propriété "φ est positive" ne signifie pas que $\varphi(x, y) \geq 0$ pour tout $(x, y) \in E \times E$.
Ce n'est pas possible de toute façon puisque $\varphi(x, -y) = -\varphi(x, y)$ par linéarité à droite.

Notation
On notera souvent $\langle \boldsymbol{x}, \boldsymbol{y} \rangle$ un produit scalaire appliqué à deux vecteurs de E, plutôt que $\varphi(x, y)$. Les notations $(\boldsymbol{x}|\boldsymbol{y})$ et $\boldsymbol{x} \cdot \boldsymbol{y}$ sont aussi employées.

Remarque (Formule fondamentale)
Si $\langle\ ,\ \rangle : E \times E \to \mathbb{R}$ est un produit scalaire sur E, alors on a (par bilinéarité)

$$\left\langle \sum_{i=1}^{n} \alpha_i x_i, \sum_{j=1}^{p} \beta_j y_j \right\rangle = \sum_{i=1}^{n} \sum_{j=1}^{p} \alpha_i \beta_j \langle x_i, y_j \rangle,$$

pour tous vecteurs $x_1, \cdots x_n, y_1, \cdots, y_p$ de E et tous réels $\alpha_1, \cdots, \alpha_n, \beta_1, \cdots, \beta_p$.

Définition 2 (Espace préhilbertien réel, espace euclidien).

(i) *On appelle espace préhilbertien réel tout $\mathbb{R}$-espace vectoriel E muni d'un produit scalaire $\langle\ ,\ \rangle : E \times E \to \mathbb{R}$.*

(ii) *On appelle espace euclidien tout espace préhilbertien réel de dimension finie.*

Proposition 3 (Restriction de la structure préhilbertienne à un sev).
Soit $(E, \langle\ ,\ \rangle)$ un espace préhilbertien réel et soit F un sous-espace vectoriel de E. Alors :

(i) *la restriction $\langle\ ,\ \rangle_{F\times F} : F \times F \to \mathbb{R}$ du produit scalaire de E est un produit scalaire sur F.*

(ii) *muni de ce produit scalaire restreint, F est un espace préhilbertien réel.*

Preuve : La restriction $\langle\ ,\ \rangle_{F\times F}$ reste bilinéaire, symétrique et définie positive. $\square$

Remarque
Ceci signifie que "tout sev d'un espace préhilbertien réel est lui-même un espace préhilbertien réel" (pour le même produit scalaire).

2) Exemples fondamentaux de produits scalaires

- **Produit scalaire "canonique" sur $E = \mathbb{R}^n$** :

  pour $x = (x_1, \cdots, x_n)$ et $y = (y_1, \cdots, y_n)$ dans $\mathbb{R}^n$, on pose $\langle x, y\rangle = \sum_{i=1}^{n} x_i y_i$.

  (on l'appelle aussi **produit scalaire euclidien**).

  Montrons que cette application est bien un produit scalaire :

  - Elle est symétrique car :
  $$\langle y, x\rangle = \sum_{i=1}^{n} y_i x_i = \sum_{i=1}^{n} x_i y_i = \langle x, y\rangle.$$
  - Elle est linéaire à gauche car :
  $$\langle \lambda x + y, z\rangle = \sum_{i=1}^{n} (\lambda x_i + y_i) z_i = \lambda \sum_{i=1}^{n} x_i z_i + \sum_{i=1}^{n} y_i z_i = \lambda\langle x, z\rangle + \langle y, z\rangle.$$

  On a donc une forme bilinéaire symétrique. De plus, pour tout $x \in E$:
  $$\langle x, x\rangle = \sum_{i=1}^{n} x_i^2.$$

  Donc

- $\langle x, x\rangle \geq 0$.
- $\langle x, x\rangle = 0 \implies x_i = 0 \ \forall i \in \{1, \cdots, n\} \implies x = 0_E$.

ce qui montre que cette forme bilinéaire symétrique est définie positive, c'est donc un produit scalaire sur $E = \mathbb{R}^n$.

Remarque (Expression matricielle du produit scalaire canonique)

si $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, $Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in \mathcal{M}_{n,1}(\mathbb{R})$, alors on a

$$\langle x, y\rangle = \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = X^T\, Y.$$

- **Produit scalaire "intégral" sur $E = \mathcal{C}^0([a;b], \mathbb{R})$** :

si $a < b$ et $f, g : [a, b] \to \mathbb{R}$ sont deux fonctions **continues**, alors en posant

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

on a un produit scalaire sur l'espace (de dimension infinie) $\mathcal{C}^0([a;b], \mathbb{R})$.

En effet :

- $\langle g, f\rangle = \int_a^b g(t)f(t)dt = \int_a^b f(t)g(t)dt = \langle f, g\rangle$, donc cette application est symétrique.
- $(\lambda f_1 + f_2, g\rangle = \int_a^b (\lambda f_1(t) + f_2(t))g(t)dt = \lambda \int_a^b f_1(t)g(t)dt + \int_a^b f_2(t)g(t)dt = \lambda\langle f_1, g\rangle + \langle f_2, g\rangle$, donc elle est linéaire à gauche. On a donc une forme bilinéaire symétrique.
- $\langle f, f\rangle = \int_a^b f^2(t)dt \geq 0$ (par positivité de l'intégrale, vu que f^2 est une fonction positive).
- $\langle f, f\rangle = 0 \implies \int_a^b f^2(t)dt = 0$. Puisque la fonction f^2 est continue et positive, cela entraîne $f^2(t) = 0$ pour tout $t \in [a;b]$, et donc f est la fonction nulle sur $[a;b]$.

Ainsi, on a une forme bilinéaire symétrique définie-positive.

ATTENTION !
La condition de **continuité** est ici **indispensable** pour assurer le caractère "défini-positif". L'application $(f,g) \mapsto \int_a^b f(t)g(t)dt$ **n'est pas** un produit scalaire sur $\mathcal{C}_M([a,b],\mathbb{R})$ (espace des fonctions réelles continues par morceaux).

Remarque
L'application définie par $\langle f,g\rangle = \dfrac{1}{b-a}\displaystyle\int_a^b f(t)g(t)dt$, est aussi un produit scalaire sur $\mathcal{C}^0([a;b],\mathbb{R})$ (utilisé pour le calcul des coefficients de Fourier d'une fonction).

- **Produit scalaire "intégral" sur $\boldsymbol{E} = \mathbb{R}[\boldsymbol{X}]$** :

si $a < b$ et P, Q sont deux **polynômes**, alors en posant $\langle P,Q\rangle := \displaystyle\int_a^b P(t)Q(t)dt,$

on a un produit scalaire sur l'espace (de dimension infinie) $\mathbb{R}[X]$.

En effet, vu que tout polynôme définit en particulier une fonction continue $[a,b] \to \mathbb{R}$, les calculs faits dans l'exemple précédent montrent que :

- l'application $\langle .\ ,\ .\rangle$ est une forme bilinéaire symétrique ;
- $\langle P,P\rangle \geq 0$ pour tout $P \in \mathbb{R}[X]$.

Mais pour vérifier le dernier point, l'argument est plus complexe :

$$\langle P,P\rangle = 0 \implies \int_a^b P^2(t)dt = 0 \underset{P^2 \text{ continue et positive}}{\implies} P^2 = 0 \text{ sur } [a;b]$$

D'où $\langle P,P\rangle = 0$ entraîne que le polynôme P s'annule sur tout l'intervalle $[a;b]$. Il possède donc une infinité de racines, ce qui le conduit à être identiquement nul (sinon, il aurait moins de racines que son degré). Finalement, on a la propriété :

$$\langle P,P\rangle = 0 \implies P = 0_{\mathbb{R}[X]},$$

d'où la forme bilinéaire symétrique est définie-positive.

- **Produit scalaire "canonique" sur $\boldsymbol{E = \mathcal{M}_n(\mathbb{R})}$** :
pour $A, B \in \mathcal{M}_n(\mathbb{R})$, en posant

$$\langle A, B\rangle := Tr(A^T\ B) = \sum_{i=1}^{n}\sum_{j=1}^{n} A[i,j]B[i,j],$$

on a un produit scalaire sur l'espace $\mathcal{M}_n(\mathbb{R})$ (de dimension n^2).

- Tout d'abord, montrons la formule $Tr(A^T\ B) = \sum_{i=1}^{n}\sum_{j=1}^{n} A[i,j]B[i,j]$ (elle sera utile pour la suite).

$$\langle A, B\rangle = Tr(A^T\ B) = \sum_{i=1}^{n}(A^T B)[i,i] = \sum_{i=1}^{n}\left(\sum_{j=1}^{n} A^T[i,j]B[j,i]\right),$$

c'est-à-dire

$$\langle A, B\rangle = \sum_{i=1}^{n}\left(\sum_{j=1}^{n} A[j,i]B[j,i]\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} A[j,i]B[j,i] \underset{i\leftrightarrow j}{=} \sum_{i=1}^{n}\sum_{j=1}^{n} A[i,j]B[i,j].$$

- Ensuite, $\langle .\ ,\ .\rangle$ est une forme bilinéaire symétrique car :
  * $\langle B, A\rangle = Tr(B^T\ A) = Tr((B^T\ A)^T) = Tr(A^T\ B) = \langle A, B\rangle$.
  * $\langle \lambda A_1 + A_2, B\rangle = Tr((\lambda A_1 + A_2)^T\ B) = Tr(\lambda A_1^T B + A_2^T B)$.
    Par linéarité de la trace, on en déduit :

$$\langle \lambda A_1 + A_2, B\rangle = \lambda Tr(A_1^T\ B) + Tr(A_2^T\ B) = \lambda\langle A_1, B\rangle + \langle A_2, B\rangle.$$

- Enfin, $\langle . \ , \ . \rangle$ est définie positive car :
  * $\langle A, A\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} A[i,j]^2 \geq 0$.
  * $\langle A, A\rangle = 0 \implies \sum_{i=1}^{n}\sum_{j=1}^{n} A[i,j]^2 = 0$.
    Puisque tous les coefficients $A[i,j]^2$ sont positifs (carrés réels), on en déduit que

$$\langle A, A\rangle = 0 \implies A[i,j] = 0 \ \forall (i,j) \in \{1, \cdots, n\}^2.$$

3) Inégalité de Cauchy-Schwarz

Proposition 4 (Inégalité de Cauchy-Schwarz).
Soit $(E, \langle \ , \ \rangle)$ un espace préhilbertien réel, et soit deux vecteurs x, y de E. Alors :

(i) *On a $|\langle x, y\rangle| \leq \langle x, x\rangle^{1/2}\langle y, y\rangle^{1/2}$.*

(ii) *Dans l'inégalité précédente, on a égalité si et seulement si la famille (x, y) est liée .*

Remarque
Pour tout $x \in E$, la quantité $\langle x, x\rangle^{1/2}$ est bien définie car $\langle x, x\rangle \geq 0$.

Preuve : Posons, pour tout $t \in \mathbb{R}$, $f(t) = \langle tx + y; tx + y\rangle$.
En développant, on obtient, par bilinéarité et symétrie :

$$f(t) = t^2\langle x; x\rangle + 2t\langle x; y\rangle + \langle y; y\rangle,$$

donc, **si $\boldsymbol{x \neq 0_E}$**, alors f est un polynôme du second degré en la variable t.
On traitera le cas $x = 0_E$ à part.

(i)
- Si $x \neq 0_E$, alors la propriété $\langle u; u\rangle \geq 0$ entraîne que $f(t) \geq 0$ pour tout $t \in \mathbb{R}$. Ainsi, la fonction f (qui est un polynôme du second degré) s'annule au plus une fois sur $\mathbb{R}$. Son discriminant est donc négatif, c'est-à-dire :

$$(2\langle x; y\rangle)^2 - 4\langle x; x\rangle\langle y; y\rangle \leq 0.$$

  Ceci équivaut à $(\langle x; y\rangle)^2 \leq \langle x; x\rangle\langle y; y\rangle$, d'où en prenant la racine carrée :

$$|\langle x; y\rangle| \leq \sqrt{\langle x; x\rangle}\sqrt{\langle y; y\rangle}.$$

- Si $x = 0_E$, alors on a $|\langle x; y\rangle| = 0$ (par bilinéarité) et aussi $\sqrt{\langle x; x\rangle}\sqrt{\langle y; y\rangle} = 0 \times \sqrt{\langle y; y\rangle} = 0$, donc l'inégalité reste vraie (on a égalité).

(ii) Montrons que $|\langle x, y\rangle| = \langle x, x\rangle^{1/2}\langle y, y\rangle^{1/2} \iff (x; y)$ est liée.
- Si $x \neq 0_E$, alors la propriété $|\langle x, y\rangle| = \langle x, x\rangle^{1/2}\langle y, y\rangle^{1/2}$ équivaut à dire que le discriminant du polynôme $f(t)$ est nul, c'est-à-dire $\exists t_0 \in \mathbb{R}$ tel que $f(t_0) = 0$, ce qui équivaut successivement à
  $\langle t_0x + y; t_0x + y\rangle = 0$, puis $y = -t_0x$. Ainsi, on a

$$|\langle x, y\rangle| = \langle x, x\rangle^{1/2}\langle y, y\rangle^{1/2} \iff y \in Vect(x) \underset{\text{car } x \neq 0_E}{\iff} (x; y) \text{ liée}$$

- Si $x = 0_E$, alors l'équivalence est triviale, puisque les deux assertions sont toujours vraies (l'égalité est toujours vérifiée pour $x = 0_E$ et la famille $(0; y)$ est toujours liée).

$\square$

Exemple

- Pour le produit scalaire euclidien de $\mathbb{R}^n$, l'inégalité de Cauchy-Schwarz s'écrit :
$$\left|\sum_{i=1}^n x_i y_i\right| \leq \left(\sum_{i=1}^n x_i^2\right)^{1/2} \left(\sum_{i=1}^n y_i^2\right)^{1/2}.$$
- Dans l'espace $\mathcal{C}^0([a,b],\mathbb{R})$, muni du produit scalaire $\langle f, g\rangle := \int_a^b f(t)g(t)dt$,
elle devient : $$\left|\int_a^b f(t)g(t)dt\right| \leq \left(\int_a^b f(t)^2\ dt\right)^{1/2} \left(\int_a^b g(t)^2\ dt\right)^{1/2}.$$

4) Norme associée à un produit scalaire

Proposition 5 (Norme associée à un produit scalaire).
Soit $(E, \langle\ ,\ \rangle)$ un espace préhilbertien réel. Alors, l'application $N : E \to \mathbb{R}_+$ définie par $N(x) = \sqrt{\langle x, x\rangle}$ est **une norme sur *E***, *c'est-à-dire qu'elle vérifie :*

(i) $\forall x \in E,\ N(x) = 0_{\mathbb{R}} \implies x = 0_E$ *(propriété de séparation) ;*

(ii) $\forall x \in E,\ \forall \lambda \in \mathbb{R},\ N(\lambda x) = |\lambda| N(x)$ *(propriété d'homogénéité) ;*

(iii) $\forall (x,y)^2 \in E,\ N(x+y) \leq N(x) + N(y)$ *(inégalité triangulaire).*

Preuve :

(i) Si $N(x) = 0_{\mathbb{R}}$, alors $N(x)^2 = 0_{\mathbb{R}}$, c'est-à-dire $\langle x; x\rangle = 0_{\mathbb{R}}$, ce qui entraîne $x = 0_E$ (puisque le produit scalaire est défini-positif).

(ii) Si $x \in E$ et $\lambda \in \mathbb{R}$, alors

$$N(\lambda x) = \sqrt{\langle \lambda x, \lambda x\rangle} = \sqrt{\lambda^2 \langle x; x\rangle} = \sqrt{\lambda^2}\sqrt{\langle x; x\rangle} = |\lambda| N(x).$$

(iii) Si $x, y \in E$, on a, par bilinéarité

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

Par symétrie, on en déduit :

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

On utilise alors l'inégalité de Cauchy-Schwarz :

$$\langle x, y\rangle \leq |\langle x, y\rangle| \leq \langle x, x\rangle^{1/2}\langle y, y\rangle^{1/2},$$

ce qui amène :

$$N(x+y)^2 \leq \langle x, x\rangle + 2\langle x, x\rangle^{1/2}\langle y, y\rangle^{1/2} + \langle y, y\rangle,$$

c'est-à-dire

$$N(x+y)^2 \leq N(x)^2 + 2N(x)N(y) + N(y)^2 = (N(x) + N(y))^2.$$

On conclut en prenant la racine carrée : puisque $N(x+y)$ et $N(x)+N(y)$ sont positifs, on déduit :

$$N(x+y) \leq N(x) + N(y).$$

$\square$

Notation

Conformément à l'usage, on aura tendance à noter $\|\,.\,\|$ la norme associée à un produit scalaire : $\forall x \in E$, $\|x\| = \sqrt{\langle x, x\rangle}$.

Remarque

- Avec cette notation, les propriétés de la norme se réécrivent donc
  - $\forall x \in E,\ \|x\| \geq 0$ (positivité).
  - $\forall x \in E,\ \|x\| = 0 \implies x = 0_E$ (séparation).
  - $\forall x \in E,\ \forall \lambda \in \mathbb{R},\ \|\lambda x\| = |\lambda| \|x\|$ (homogénéité).
  - $\forall x, y \in E,\ \|x + y\| \leq \|x\| + \|y\|$ (inégalité triangulaire).
- L'implication réciproque du deuxième point est toujours vraie : en effet, si $x = 0_E$, alors d'après la propriété d'homogénéité, on a :

$$\|0_E\| = \|0_{\mathbb{R}} \times 0_E\| = |0_{\mathbb{R}}| \times \|0_E\| = 0_{\mathbb{R}}.$$

- La propriété d'homogénéité entraîne aussi que $\forall x \in E, \|-x\| = \|x\|$.
- L'inégalité de Cauchy-Schwarz s'écrit :

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

avec égalité si et seulement si (x, y) est liée.

Exemple (Norme euclidienne sur $\mathbb{R}^n$)
Sur $E = \mathbb{R}^n$, on retrouve ainsi la "norme euclidienne" classique :

$$\forall x = (x_1, \cdots, x_n) \in \mathbb{R}^n, \qquad \|x\| = \sqrt{\langle x, x\rangle} = \sqrt{\sum_{i=1}^{n} x_i^2},$$

ce qui s'écrit matriciellement :

$$\|x\| = \sqrt{X^T\ X},$$

en notant $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathcal{M}_{n,1}(\mathbb{R})$.

Proposition 6 (Conséquences de l'inégalité triangulaire).
Soit $(E, \langle\ ,\ \rangle)$ un espace préhilbertien réel. On note $\|\ .\ \|$ la norme associée. On a alors

$$\forall (x, y) \in E^2, \qquad \Big| \|x\| - \|y\| \Big| \leq \|x - y\| \leq \|x\| + \|y\|.$$

. . . .

Preuve : Fixons x, y dans E.

- $\|x - y\| = \|x + (-y)\| \leq \|x\| + \|-y\| = \|x\| + |-1| \|y\| = \|x\| + \|y\|$.

- $\|x\| = \|(x-y)+y\| \leq \|x-y\| + \|y\|$, donc $\|x\| - \|y\| \leq \|x-y\|$.
  De même, $\|y\| \leq \|y-x\| + \|x\| = \|x-y\| + \|x\|$, donc $\|x\| - \|y\| \geq -\|x-y\|$.
  On a donc $-\|x-y\| \leq \|x\| - \|y\| \leq \|x-y\|$, c'est-à-dire

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

$\square$

ATTENTION !
L'inégalité $\|x-y\| \leq \|x\| - \|y\|$ est bien entendu **fausse** !

Définition 7 (Vecteur unitaire).
Dans un espace préhilbertien réel, un vecteur x est dit **unitaire** *si $\|x\| = 1$.*

Remarque
Un vecteur unitaire est nécessairement **non nul**, puisque $(x = 0_E \iff \|x\| = 0)$.

Proposition 8 (Vecteurs unitaires sur une droite).
Soit E un espace préhilbertien réel et $\mathcal{D}$ une droite vectorielle de E (i.e. un sev de dimension 1). Alors, la droite $\mathcal{D}$ contient exactement deux vecteurs unitaires et ils sont opposés.

Preuve : Si $\overrightarrow{u}$ est un vecteur non nul de $\mathcal{D}$, alors il forme une base de $\mathcal{D}$: tous les vecteurs de $\mathcal{D}$ s'écrivent de façon unique $\overrightarrow{x} = \alpha \overrightarrow{u}$ avec $\alpha \in \mathbb{R}$. Parmi ces vecteurs, on a

$$\|\overrightarrow{x}\| = 1 \iff |\alpha| \times \|\overrightarrow{u}\| = 1 \iff \alpha = \pm \frac{1}{\|\overrightarrow{u}\|}.$$

Ceci montre que $\mathcal{D}$ possède deux vecteurs unitaires distincts : $\overrightarrow{x_1} = \frac{\overrightarrow{u}}{\|\overrightarrow{u}\|}$ et $\overrightarrow{x_2} = -\frac{\overrightarrow{u}}{\|\overrightarrow{u}\|}$.
$\square$

5) Expressions reliant produit scalaire et norme

En plus des propriétés classiques des normes (inégalités triangulaires), on dispose de **formules de développement** pour les normes issues d'un produit scalaire.

Proposition 9 (Formules de développement).
Soit $(E, \langle\ ,\ \rangle)$ un espace préhilbertien réel et soit x, y deux vecteurs de E.
On note $\|\ .\ \|$ la norme associée au produit scalaire. Alors,

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

(ii) $\|x-y\|^2 = \|x\|^2 - 2\langle x, y\rangle + \|y\|^2$;

(iii) $\|x\|^2 - \|y\|^2 = \langle x+y, x-y\rangle$;

(iv) $\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2$ *(identité du parallélogramme).*

Preuve : Ecrire $\|x+y\|^2 = \langle x+y, x+y\rangle$ et développer…

Illustrer l'identité du parallélogramme par un dessin. □

Proposition 10 (Formules de polarisation).
Soit $(E, \langle\ ,\ \rangle)$ un espace préhilbertien réel et soit x, y deux vecteurs de E.
On note $\|\ .\ \|$ la norme associée au produit scalaire. Alors :

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

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

Preuve :

$\square$

Remarque
Ces "formules de polarisation" montrent donc qu'un produit scalaire peut s'exprimer **uniquement en fonction de sa norme** associée.

6) Distance associée à un produit scalaire

Définition 11 (Distance associée à un produit scalaire).
Soit $(E, \langle\ ,\ \rangle)$ un espace préhilbertien réel et $\|\ .\ \|$ la norme associée au produit scalaire.
Etant donnés x, y de E, on appelle distance entre x et y le réel positif
$d(x, y) = \|x - y\|$.

Remarque
La distance est donc une application $d : E \times E \to \mathbb{R}_+$.

Proposition 12 (Propriétés de la distance).
Soit $(E, \langle\ ,\ \rangle)$ un espace préhilbertien réel. La distance $d : E \times E \to \mathbb{R}_+$ vérifie les propriétés suivantes :

(i) $\forall (x, y) \in E^2,\ d(x, y) = 0 \iff x = y$ *(propriété de séparation) ;*

(ii) $\forall (x, y) \in E^2,\ d(y, x) = d(x, y)$ *(propriété de symétrie) ;*

(iii) $\forall (x, y, z) \in E^3,\ d(x, y) \leq d(x, z) + d(z, y)$ *(inégalité triangulaire).*

Preuve :

(i) D'après la propriété de séparation de la norme, on a

$$d(x,y)=0 \iff \|x-y\|=0 \iff x-y=0_E \iff x=y.$$

(ii) D'après la propriété d'homogénéité de la norme, on a

$$d(y,x)=\|y-x\|=\|-(x-y)\|=|-1|\times\|x-y\|=\|x-y\|=d(x,y).$$

(iii) D'après l'inégalité triangulaire pour la norme, on a

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

□

Dessin de l'inégalité triangulaire :

II Orthogonalité

Dans cette section, E désigne un espace préhilbertien réel (pas nécessairement de dimension finie). On note $\langle\ ,\ \rangle$ son produit scalaire.

1) Vecteurs orthogonaux, sous-espaces orthogonaux

Définition 13 (Vecteurs orthogonaux).
Deux vecteurs x et y de E sont dits **orthogonaux** *si $\langle x,y\rangle=0$* .

Remarque

- Dans cette définition, x et y ont des rôles symétriques puisque $\langle x,y\rangle=\langle y,x\rangle$.
- Le vecteur nul est orthogonal à tout vecteur de E.
  En effet, par linéarité à droite du produit scalaire :
  $$\forall x\in E,\ \langle x,0_E\rangle=\langle x,0_{\mathbb{R}}\times 0_E\rangle=0_{\mathbb{R}}\times\langle x,0_E\rangle=0_{\mathbb{R}}.$$
- Le vecteur nul est le seul vecteur orthogonal à lui-même.
  En effet, $\forall x\in E,\ (\langle x,x\rangle=0 \implies x=0_E)$.

Remarque (Lien avec la géométrie élémentaire du plan ou de l'espace)
On généralise ainsi la notion d'orthogonalité déjà connue pour les vecteurs du plan $\mathbb{R}^2$ ou de l'espace $\mathbb{R}^3$.
Rappelons que dans ces deux espaces vectoriels, on avait donné (en TSI 1) une **définition géométrique du produit scalaire** :

$$\overrightarrow{u} \cdot \overrightarrow{v} = \|\overrightarrow{u}\| \times \|\overrightarrow{v}\| \times \cos(\alpha),$$

où $\alpha \in [0, \pi]$ est l'angle non-orienté entre les deux vecteurs non nuls $\overrightarrow{u}$ et $\overrightarrow{v}$.

On avait alors l'équivalence : $\overrightarrow{u} \cdot \overrightarrow{v} = 0 \iff \alpha = \frac{\pi}{2}$.
Mais dans un espace préhilbertien quelconque, il n'y a plus de notion intuitive "d'angle"...

Généralisation maintenant la notion de droites et plans orthogonaux (déjà connue dans le plan $\mathbb{R}^2$ et l'espace $\mathbb{R}^3$) :

Définition 14 (Sous-espaces orthogonaux).
Deux sous-espaces vectoriels F et G de E sont dits **orthogonaux** *si*

$$\forall x \in F, \quad \forall y \in G, \qquad \langle x, y \rangle = 0.$$

Remarque
Cela signifie que tout vecteur de F est orthogonal à tout vecteur de G (ou l'inverse !).

Lemme 15 (Deux sev orthogonaux sont en somme directe).
Si *deux sous-espaces F et G de E sont orthogonaux,* **alors** $F \cap G = \{0_E\}$.

Preuve : Si $x \in F \cap G$, alors, puisque F et G sont orthogonaux, on a $\langle x, x \rangle = 0$ (le vecteur x étant à la fois dans F et dans G, il est orthogonal à lui-même), donc $x = 0_E$ (puisque le produit scalaire est défini-positif). $\square$

ATTENTION !

- **La réciproque du lemme est fausse !** Deux sev d'intersection nulle ne sont pas nécessairement orthogonaux.
- Deux sev orthogonaux de E ne sont **pas nécessairement supplémentaires**.

Exemple
Dans $E = \mathbb{R}^3$ muni de son produit scalaire canonique, les ensembles $\mathcal{D}_1 = Vect\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$ et $\mathcal{D}_2 = Vect\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ sont des sev orthogonaux mais pas supplémentaires.

2) Orthogonal d'une partie de E

Définition 16 (Orthogonal d'une partie).
Soit $X \subset E$ (pas nécessairement un sev). On appelle **orthogonal de X**
(noté $X^\perp$) l'ensemble des vecteurs de E qui sont orthogonaux à tous les $x \in X$:
$X^\perp = \{y \in E,\ \forall x \in X\ \ \langle x, y\rangle = 0\}$.

Proposition 17 (Propriétés de l'orthogonalité).
(i) *On a $\{0_E\}^\perp = E$ et $E^\perp = \{0_E\}$* .

(ii) *Si $X \subset Y \subset E$, alors $Y^\perp \subset X^\perp \subset E$*

(iii) *Pour toute partie $X \subset E$, l'orthogonal $X^\perp$ est un sous-espace vectoriel de E et on a $(Vect(X))^\perp = X^\perp$* .

ATTENTION !
Bien saisir la nuance : même si X n'est pas un sev de E, l'ensemble $X^\perp$, lui, en est toujours un.

Preuve :

(i) Tout vecteur $x \in E$ est orthogonal au vecteur nul, donc $E \subset \{0_E\}^\perp$, et l'inclusion réciproque est évidente.
Si $x \in E^\perp$, alors x est orthogonal à tous les vecteurs de E, donc en particulier à lui-même. Donc $x = 0_E$, ce qui montre l'inclusion $E^\perp \subset \{0_E\}$.
L'inclusion réciproque est évidente, car le vecteur nul est orthogonal à tout vecteur de E.

(ii) Supposons que $X \subset Y$. Si $z \in Y^\perp$, le vecteur z est orthogonal à tous les éléments de Y, donc *a fortiori* orthogonal à tous ceux de X (puisque Y contient X). D'où $z \in X^\perp$, ce qui montre l'inclusion $Y^\perp \subset X^\perp$.

(iii) Soit X une partie de E. Montrons que l'ensemble $X^\perp$ est un sous-espace vectoriel de E :

- le vecteur 0_E est dans $X^\perp$ car il est orthogonal à tout vecteur de X.
- si y et z sont dans $X^\perp$, alors pour tout $\lambda \in \mathbb{R}$, le vecteur $\lambda y + z$ est dans $X^\perp$, puisque :

$$\forall x \in X, \qquad \langle \lambda y + z, x\rangle = \lambda \underbrace{\langle y, x\rangle}_{=0} + \underbrace{\langle z, x\rangle}_{=0} = 0.$$

Montrons maintenant que $(Vect(X))^\perp = X^\perp$:

- Puisque $X \subset Vect(X)$, on a $(Vect(X))^\perp \subset X^\perp$.

- Si $y \in X^{\perp}$, alors y est orthogonal à tous les éléments de X. Par bilinéarité du produit scalaire, on en déduit que y est orthogonal à toute combinaison linéaire d'éléments de X, puisque pour toute famille finie $(x_j)_{j\in J}$ de vecteurs de X et pour toute famille finie de réels $(\lambda_j)_{j\in J}$:

$$\langle y, \sum_{j\in J} \lambda_j x_j \rangle = \sum_{j\in J} \lambda_j \underbrace{\langle y, x_j \rangle}_{=0} = 0_E.$$

□

Remarque (TRES UTILE)
L'égalité $(Vect(X))^{\perp} = X^{\perp}$ signifie notamment qu'un vecteur est orthogonal à un sev si et seulement si il est **orthogonal à n'importe quelle base** de ce sev.

Exemple
Dans $\mathbb{R}^3$, pour montrer qu'un vecteur $\overrightarrow{x}$ est orthogonal au plan vectoriel engendré par la famille libre $(\overrightarrow{u}, \overrightarrow{v})$, il suffit de montrer que $\langle \overrightarrow{x}, \overrightarrow{u} \rangle = \langle \overrightarrow{x}, \overrightarrow{v} \rangle = 0$.

Proposition 18 (Un sev et son orthogonal sont en somme directe).
Si F est un sous-espace vectoriel de E, alors $F^{\perp}$ aussi et $F + F^{\perp} = F \oplus F^{\perp}$.
. . . .

Preuve : F est un sev par hypothèse, et $F^{\perp}$ en est un d'après la prop. 19. Par définition, les sev F et $F^{\perp}$ sont orthogonaux donc en somme directe, d'après le lemme 17. □

Remarque
Si F est un sev de E, alors $F^{\perp}$ est le plus grand sev de E orthogonal à F.
En fait, pour tout sev G de E, on a l'équivalence :

$$\text{F et G sont orthogonaux} \iff G \subset F^{\perp} \iff F \subset G^{\perp}.$$

ATTENTION !
On n'a pas toujours $\boldsymbol{F \oplus F^{\perp} = E}$
(F et $F^{\perp}$ ne sont pas toujours supplémentaires dans le cas général).
Ils seront supplémentaires lorsque F est de dimension finie (voir plus loin).

3) Familles orthogonales, familles orthonormées

Définition 19 (Famille orthogonale).
Une **famille** $(e_i)_{i\in I}$ *(éventuellement infinie) de vecteurs de* E *est dite* **orthogonale** *si* $\forall (i,j) \in I^2,\ (i \neq j \implies \langle e_i, e_j \rangle = 0)$.

Proposition 20 (Orthogonalité et indépendance linéaire).
Toute famille **orthogonale** *formée de vecteurs* **non nuls** *est libre* .

Preuve : Soit $(e_i)_{i\in I}$ une famille orthogonale formée de vecteurs non nuls, et soit une famille finie de réels $(\lambda_j)_{j\in J}$ (avec $J \subset I$) telle que $\sum_{j\in J} \lambda_j e_j = 0_E$. Pour tout indice $i \in J$, on a :

$$\langle \sum_{j\in J} \lambda_j e_j, e_i \rangle = \langle 0_E, e_i \rangle = 0,$$

c'est-à-dire (par bilinéarité) :

$$\sum_{j\in J} \lambda_j \langle e_j, e_i \rangle = 0_E.$$

Mais $\langle e_j, e_i \rangle = 0$ dès que $j \neq i$, il reste donc $\lambda_i \underbrace{\langle e_i, e_i \rangle}_{\neq 0} = 0_E$, c'est-à-dire $\lambda_i = 0$.

Ceci étant vrai pour tout $i \in J$ et pour toute partie finie $J \subset I$, on en déduit que la famille $(e_i)_{i\in I}$ est libre.

□

Proposition 21 (Théorème de Pythagore).
Si $(e_1, \ldots, e_p)$ *est une famille* **orthogonale**, *alors* $\left\| \sum_{i=1}^{p} e_i \right\|^2 = \sum_{i=1}^{p} \|e_i\|^2$.

Preuve : Par bilinéarité, on a

$$\left\| \sum_{i=1}^{p} e_i \right\|^2 = \langle \sum_{i=1}^{p} e_i, \sum_{j=1}^{p} e_j \rangle = \sum_{i=1}^{p} \sum_{j=1}^{p} \underbrace{\langle e_i, e_j \rangle}_{=0 \text{ si } i\neq j} = \sum_{i=1}^{p} \langle e_i, e_i \rangle = \sum_{i=1}^{p} \|e_i\|^2$$

(dans la double somme, il ne reste que les p termes correspondant aux couples (i,i)).
□

Définition 22 (Famille orthonormée).
Une famille $(e_i)_{i\in I}$ (éventuellement infinie) de vecteurs de E est dite
....
orthonormée *si* $\forall (i,j) \in I^2,\ \langle e_i, e_j \rangle = \begin{cases} 0 & si\ i \neq j \\ 1 & si\ i = j \end{cases}$.
....

Remarque

- Cela revient à dire que les $(e_i)_{i\in I}$ sont deux à deux orthogonaux et de norme 1.
- On peut abréger ceci avec **le symbole de Kronecker** :
  en posant $\delta_{i,j} = \begin{cases} 0 & \text{si } i \neq j \\ 1 & \text{si } i = j \end{cases}$, la famille $(e_i)_{i\in I}$ est orthonormée si et seulement si $\forall (i,j) \in I^2,\ \langle e_i, e_j \rangle = \delta_{i,j}$.
- Toute famille orthonormée est libre, et la réciproque est fausse.

Définition 23 (Base orthonormée).
On appelle **base orthonormée de E** *toute famille orthonormée $(e_i)_{i\in I}$ qui est une base de E.*

Remarque
Si une famille orthonormée $(e_i)_{i\in I}$ est génératrice de E, alors c'est automatiquement une base orthonormée de E (puisque orthonormée $\Longrightarrow$ libre).

Exemple
La base canonique de $\mathbb{R}^n$, notée $(\overrightarrow{e_1}, \cdots, \overrightarrow{e_n})$ est une base orthonormée de l'espace euclidien $\mathbb{R}^n$ (pour le produit scalaire canonique).
En effet, pour tous i, j compris entre 1 et n :

$$\langle \overrightarrow{e_i}, \overrightarrow{e_j} \rangle = \sum_{k=1}^{n} \underbrace{\overrightarrow{e_i}[k]}_{=0 \text{ si } k\neq i} \times \overrightarrow{e_j}[k] = \underbrace{\overrightarrow{e_i}[i]}_{=1} \times \overrightarrow{e_j}[i] = \overrightarrow{e_j}[i] = \begin{cases} 1 \text{ si } j = i \\ 0 \text{ sinon} \end{cases}.$$

ATTENTION !
La base canonique de $\mathbb{R}[X]$ (i.e. la famille infinie $(X^k)_{k\in\mathbb{N}}$) **n'est pas orthonormée** pour le produit scalaire défini par $\langle P, Q \rangle := \int_0^1 P(t)Q(t)dt$.
En effet, $\langle 1, X \rangle = \int_0^1 t\ dt = \frac{1}{2} \neq 0$, donc la base n'est même pas orthogonale.

Moralité : ce n'est pas parce qu'une base est "naturelle" qu'elle est orthonormée : tout dépend du produit scalaire choisi !

4) Orthonormalisation de Gram-Schmidt

On va voir dans cette section **comment construire des bases orthonormées** d'espaces vectoriels à partir de bases quelconques.

Exemple
Dans $\mathbb{R}^4$ muni de son produit scalaire canonique, on considère la famille libre :

$$u_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad u_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad u_3 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

On considère les sous-espaces vectoriels $(E_k)_{1 \leq k \leq 3}$ définis par

$$E_1 = Vect(u_1), \quad E_2 = Vect(u_1, u_2), \quad E_3 = Vect(u_1, u_2, u_3).$$

1. Construire une base orthonormée (e_1) de la droite E_1.
2. Compléter (e_1) en une base orthonormée (e_1, e_2) du plan E_2.
   On pourra d'abord chercher un vecteur e_2' orthogonal à e_1 sous la forme $e_2' = u_2 + \lambda e_1$.
3. Compléter (e_1, e_2) en une base orthonormée (e_1, e_2, e_3) du sous-espace E_3.
   On pourra d'abord chercher un vecteur e_3' orthogonal à e_1 et e_2 sous la forme $e_3' = u_3 + \alpha e_1 + \beta e_2$.

1.

2.

3.

Cette méthode se généralise. On dispose du théorème suivant :

Théorème 24 (Orthonormalisation de Gram-Schmidt).
Soit $(E, \langle\ ,\ \rangle)$ un espace préhilbertien réel, et soit $(u_1, \cdots, u_n)$ une famille libre
de vecteurs de E (avec $n \in \mathbb{N}^*$). Alors, il existe une famille orthonormée
$(e_1, \cdots, e_n)$ de E telle que $\forall p \in [\![1, n]\!]$, $Vect(e_1, \cdots, e_p) = Vect(u_1, \cdots, u_p)$.

Méthode (Mise en oeuvre de l'algorithme de Gram-Schmidt)
Etant donnée une famille libre $(u_1, \cdots, u_n)$ dans E :

- On normalise u_1 en posant $e_1 = \frac{u_1}{\|u_1\|}$.
- On construit un vecteur $e'_2 = u_2 + \lambda e_1$, avec $\lambda \in \mathbb{R}$ tel que $\langle e'_2, e_1\rangle = 0$.
- On normalise e'_2 en posant $e_2 = \frac{e'_2}{\|e'_2\|}$.
- On construit $e'_3 = u_3 + \alpha e_1 + \beta e_2$, avec $\alpha, \beta \in \mathbb{R}$ tel que $\langle e'_3, e_1\rangle = \langle e'_3, e_2\rangle = 0$.
- On normalise e'_3 en posant $e_3 = \frac{e'_3}{\|e'_3\|}$.
- Et ainsi de suite… A chaque étape, on complète la famille orthonormée $(e_1, \cdots, e_p)$ en construisant $e'_{p+1} = u_{p+1} + \displaystyle\sum_{k=1}^{p} \alpha_k e_k$ avec les α_k tels que $\langle e'_{p+1}, e_k\rangle = 0$ pour tout $k \in [\![1, p]\!]$. Puis on normalise e'_{p+1} en posant $e_{p+1} = \frac{e'_{p+1}}{\|e'_{p+1}\|}$.

A la fin, on obtient une famille orthonormée $(e_1, \cdots, e_n)$ qui vérifie bien
$Vect(e_1, \cdots, e_p) = Vect(u_1, \cdots, u_p)$ pour tout $p \in [\![1, n]\!]$.

Remarque
En fait, lorsqu'on construit $e'_{p+1} = u_{p+1} + \displaystyle\sum_{k=1}^{p} \alpha_k e_k$, les conditions d'orthogonalité donnent :

$$\forall k \in [\![1, p]\!], \quad \langle e'_{p+1}, e_k\rangle = 0 \iff \forall k \in [\![1, p]\!], \quad \alpha_k = -\langle u_{p+1}, e_k\rangle,$$

donc on dispose d'une **formule de récurrence explicite** pour construire la base orthonormée $(e_1, \cdots, e_n)$:

$$e_1 = \frac{u_1}{\|u_1\|}, \qquad \forall p \in [\![1, n-1]\!], \quad e_{p+1} = \frac{u_{p+1} - \displaystyle\sum_{k=1}^{p} \langle u_{p+1}, e_k\rangle e_k}{\left\| u_{p+1} - \displaystyle\sum_{k=1}^{p} \langle u_{p+1}, e_k\rangle e_k \right\|}.$$

Mais on n'est pas obligé de connaître cette formule.

Exemple
On munit l'espace des polynômes $\mathbb{R}[X]$ du produit scalaire $\langle P, Q\rangle = \displaystyle\int_0^1 P(t)Q(t)dt$.
Construire une base orthonormée (Q_0, Q_1, Q_2) de $\mathbb{R}_2[X]$ à partir de sa base canonique $(P_0, P_1, P_2) = (1, X, X^2)$.

On obtient la base orthonormée :

$$(Q_0, Q_1, Q_2) = \big(1\ ,\ \sqrt{3} * (2X-1)\ ,\ \sqrt{5} * (6X^2 - 6X + 1)\big)\,.$$

Corollaire 25 (Existence de bases orthonormées en dimension finie).

(i) *Dans un espace préhilbertien réel E, tout sous-espace F de dimension finie possède au moins une base orthonormée.*

(ii) *Tout espace euclidien E possède au moins une base orthonormée.*

(iii) *Dans un espace euclidien, toute famille orthonormée peut-être complétée en une base orthonormée de E.*

Preuve : Cela résulte du théorème précédent. $\square$

5) Expressions du produit scalaire et de la norme en base orthonormée

Proposition 26 (Expressions en base orthonormée en dimension finie).
Soit E un espace euclidien, et $(e_1, \cdots, e_n)$ une base **orthonormée** *de E.*

. . . .

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

. . . .

Preuve : Par bilinéarité du produit scalaire :

$$\langle x, y\rangle = \langle \sum_{i=1}^{n} x_i e_i, \sum_{j=1}^{n} y_j e_j\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} x_i y_j \underbrace{\langle e_i, e_j\rangle}_{=\delta_{i,j}} = \sum_{i=1}^{n} x_i y_i$$

(dans la somme double il ne reste plus que les termes correspondant à $i = j$).

On en déduit $\|x\| = \sqrt{\langle x, x\rangle} = \sqrt{\sum_{i=1}^{n} x_i x_i} = \sqrt{\sum_{i=1}^{n} x_i^2}$. $\square$

Remarque
En dimension finie, quand on travaille avec une base orthonormée $\mathcal{B}$, on a donc une **expression directe** du produit scalaire et de la norme **en fonction des coordonnées** :

$$\langle x, y\rangle = X^T\, Y, \quad \|x\| = \sqrt{X^T\, X}, \quad \text{avec } X = [x]_{\mathcal{B}} \text{ et } Y = [y]_{\mathcal{B}} \quad .$$

ATTENTION !
C'est faux en général pour une base non orthonormée !

Exemple
Soit $E = \mathbb{R}_1[X]$, muni du produit scalaire $\langle P, Q\rangle_E = \int_0^1 P(t)Q(t)dt$.
On pose $P = 1 + X$ et $Q = 2X$. On a d'une part

$$\langle P, Q\rangle_E = \int_0^1 (1+t)2tdt = 2\int_0^1 (t+t^2)dt = 2(1/2+1/3) = 5/3.$$

D'autre part, les coordonnées de P dans la base $(1, X)$ sont $U_P = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, celles de Q sont $U_Q = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$, et $U_P^T\, U_Q = \begin{pmatrix} 1 & 1 \end{pmatrix}\begin{pmatrix} 0 \\ 2 \end{pmatrix} = 1\times 0 + 1 \times 2 = 2.$
Ceci montre que $\langle P, Q\rangle_E \neq U_P^T\, U_Q$, et confirme le fait que la base $(1, X)$ n'est pas orthonormée pour le produit scalaire $\langle\ ,\ \rangle_E$.

Proposition 27 (Expression des coordonnées dans une base orthonormée).
Soit E un espace **euclidien** *et $\mathcal{B} = (e_1, \ldots, e_n)$ une base* **orthonormée** *de E.*

Pour tout $x \in E$, la décomposition de x sur la base $\mathcal{B}$ est $x = \sum_{i=1}^{n} \langle x, e_i\rangle e_i$.

Preuve : $\mathcal{B}$ est une base de E, donc il existe une unique liste de coordonnées $(x_1, \cdots, x_n)$ telles que $x = \sum_{k=1}^n x_k e_k$. Pour tout $i \in \{1, \cdots, n\}$, on a donc

$$\langle x, e_i\rangle = \sum_{k=1}^{n} x_k \underbrace{\langle e_k, e_i\rangle}_{\delta_{k,i}} = x_i,$$

si bien que

$$x = \sum_{i=1}^{n} x_i e_i = \sum_{i=1}^{n} \langle x, e_i\rangle e_i.$$

□

6) Supplémentaire orthogonal d'un sev de dimension finie

Théorème 28 (L'orthogonal d'un sev de dim finie est un supplémentaire). *Soit E un espace préhilbertien, et F un sous-espace vectoriel de dimension finie* . ...
Alors le sev $F^{\perp}$ est un supplémentaire de F dans E (i.e. $E = F \oplus F^{\perp}$) . ...

Preuve : On a $F \cap F^{\perp} = \{0_E\}$, car les sous-espaces F et $F^{\perp}$ sont orthogonaux. Montrons ensuite que $E \subset F + F^{\perp}$. On choisit une base orthonormée $(e_1, \dots, e_p)$ du sev F (elle existe d'après ce qui précède), et on procède par analyse-synthèse pour montrer que tout vecteur $x \in E$ se décompose sur le sev somme $F + F^{\perp}$.

- Si $x = x_1 + x_2$ avec $x_1 \in F = Vect(e_1, \cdots, e_p)$, et $x_2 \in F^{\perp} = \{e_1, \cdots, e_p\}^{\perp}$, alors on a
$$x_1 = \sum_{k=1}^{p} \lambda_k e_k,$$
avec les $\lambda_k \in \mathbb{R}$. On détermine alors les λ_k comme dans la preuve de la proposition 27 : pour tout $i \in [\![1; p]\!]$, on a
$$\langle x, e_i \rangle = \left\langle \sum_{k=1}^{p} \lambda_k e_k + x_2, e_i \right\rangle = \sum_{k=1}^{p} \lambda_k \underbrace{\langle e_k, e_i \rangle}_{=\delta_{k,i}} + \underbrace{\langle x_2, e_i \rangle}_{=0} = \lambda_i.$$

On en déduit donc
$$x_1 = \sum_{i=1}^{p} \langle x, e_i \rangle e_i, \qquad x_2 = x - \sum_{i=1}^{p} \langle x, e_i \rangle e_i.$$

- Pour $x \in E$, posons

$$x_1 = \sum_{i=1}^{p} \langle x, e_i \rangle e_i, \qquad x_2 = x - \sum_{i=1}^{p} \langle x, e_i \rangle e_i.$$

On observe que $x_1 \in F$, et $x_2 \in F^\perp = \{e_1, \ldots, e_p\}^\perp$, car $\forall j \in \{1, \ldots p\}$,

$$\langle x_2, e_j \rangle = \langle x, e_j \rangle - \sum_{i=1}^{p} \langle x, e_i \rangle \langle e_i, e_j \rangle = \langle x, e_j \rangle - \sum_{i=1}^{p} \langle x, e_i \rangle \delta_{i,j} = 0.$$

Donc on a l'existence de la décomposition voulue.

$\square$

Remarque
Pour construire une base orthonormée de E adaptée à la somme $E = F \oplus F^\perp$, il suffit donc de **réunir une base orthonormée de F et une base orthonormée de $F^\perp$**.

ATTENTION !
En général, le sev F possède **d'autres supplémentaires** que $F^\perp$.

Corollaire 29 (Dimension de l'orthogonal en dimension finie).
Soit E un espace **euclidien**. *Pour tout sous-espace vectoriel F de E, on a*

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

Preuve : On a $E = F \oplus F^\perp$ et E de dim. finie, donc $\dim(E) = \dim(F) + \dim(F^\perp)$. $\square$

Exemple
Dans l'espace euclidien $\mathbb{R}^3$, l'orthogonal d'une droite vectorielle est un plan vectoriel, et réciproquement.

Corollaire 30 (Double orthogonal d'un sev de dim finie).
Soit E un espace préhilbertien, et **F un sous-espace de dimension finie**. *Alors*

$$(F^\perp)^\perp = F.$$

Preuve :

- L'inclusion $F \subset (F^\perp)^\perp$ est facile, puisque si $x \in F$, alors x est orthogonal à tous les éléments de $F^\perp$, on a donc $x \in (F^\perp)^\perp$.

- L'inclusion réciproque utilise le théorème précédent. Soit $x \in \left(F^\perp\right)^\perp$. Alors, puisque $E = F \oplus F^\perp$, le vecteur x se décompose de manière unique :

$$x = x_1 + x_2, \qquad (x_1, x_2) \in F \times F^\perp.$$

On a donc $x_2 = x - x_1 \in (F^\perp)^\perp + F \subset (F^\perp)^\perp + (F^\perp)^\perp \subset (F^\perp)^\perp$, puisque $(F^\perp)^\perp$ étant un sous-espace vectoriel de E, il est stable par somme. Ceci montre que $x_2 \in (F^\perp)^\perp$. Mais on a aussi $x_2 \in F^\perp$. Vu que les sous-espaces $F^\perp$ et $(F^\perp)^\perp$ sont orthogonaux, leur intersection est nulle. D'où $x_2 = 0_E$, c'est-à-dire $x = x_1 \in F$.

$\square$

III Projection orthogonale, distance à un sous-espace

Dans cette section, E désigne un espace préhilbertien réel (pas nécessairement de dimension finie). On note $\langle\ ,\ \rangle$ son produit scalaire.

1) Projection orthogonale sur un sev de dimension finie

Définition 31 (Projection orthogonale).
Soit F un sev de dimension finie de E. On appelle projection orthogonale sur F
le projecteur sur F parallèlement à $F^\perp$. On notera $p_F : E \to E$ cette application.

Remarque

- Cette définition a un sens car $E = F \oplus F^\perp$, d'après le résultat de la section précédente. Tout vecteur $x \in E$ se décompose donc de manière unique en $x = x_1 + x_2$, avec $x_1 \in F$ et $x_2 \in F^\perp$.
Par définition d'un projecteur, on a $p_F(x) = x_1$.
- On dispose donc aussi de la proj. orthogonale sur $F^\perp$, et on a $p_F + p_{F^\perp} = Id_E$.

ATTENTION !
Si E est de dimension infinie et F de dimension finie, alors $F^\perp$ est de dimension infinie.

Dessin en dimension 2 :

Dessin en dimension 3 :

Proposition 32 (Propriétés d'une projection orthogonale).
Soit F un sev **de dimension finie** *de E , et p_F la projection orthogonale sur F.*

(i) *On a $p_F \in \mathcal{L}(E,E)$ et $p_F \circ p_F = p_F$.*

(ii) *On a la décomposition $E = Im(p_F) \oplus Ker(p_F)$, avec*

$$Im(p_F) = Ker(p_F - Id_E) = F \quad et \quad Ker(p_F) = F^\perp = Im(p_F)^\perp.$$

(iii) *$\forall x \in E$, $p_F(x)$ est l'unique vecteur de E tel que* $\left\{ \begin{array}{l} p_F(x) \in F \\ x - p_F(x) \in F^\perp \end{array} \right.$.

(iv) *Si $\{0_E\} \subsetneq F \subsetneq E$, alors p_F n'est ni l'endomorphisme nul, ni l'identité. Dans ce cas, p_F possède deux valeurs propres distinctes : 0 et 1, et les sous-espaces propres associés sont $E_1 = F$ et $E_0 = F^\perp$.*

Preuve : Evident en utilisant les propriétés des projecteurs. □

Remarque

- Le noyau de p_F n'est donc pas nécessairement de dimension finie.
- Si E est de dimension finie, la projection orthogonale p_F est **toujours diagonalisable** (comme tout projecteur en dimension finie), et les sous-espaces propres sont supplémentaires et orthogonaux.

Proposition 33 (Expression de p_F en base orthonormée).
Soit F un sev **de dimension finie** *de E, et $(e_1, \cdots, e_p)$ une base* **orthonormée** *de F. Notons p_F la projection orthogonale sur F. Alors on a l'expression :*

$$\forall x \in E, \qquad p_F(x) = \sum_{k=1}^{p} \langle x, e_k \rangle e_k.$$

Preuve : On reprend la preuve du théorème 28, où l'on a montré que $E = F \oplus F^\perp$: tout vecteur $x \in E$ se décompose de manière unique sous la forme

$$x = x_1 + x_2, \qquad x_1 = \sum_{k=1}^{p} \langle x, e_k \rangle e_k \in F, \quad x_2 = x - x_1 \in F^\perp.$$

Par définition de la projection orthogonale sur F, on a $p_F(x) = x_1 = \sum_{k=1}^{p} \langle x, e_k \rangle e_k$. $\square$

ATTENTION !
Dans l'expression du projeté orthogonal de x sur F, l'entier p (nombre de termes de la somme) est **la dimension du sous-espace $\boldsymbol{F}$**, pas celle de E !

Exemple
Dans $E = \mathbb{R}^3$ muni de sa structure euclidienne canonique, calculer la projection orthogonale sur le plan $F = Vect\left(\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}\right)$, notée $p_F : \mathbb{R}^3 \to \mathbb{R}^3$.

<u>Deux méthodes :</u>

1) On construit une base orthonormée de F et on utilise l'expression d'une projection orthogonale en base orthonormée.
   On trouve $p_F(x, y, z) = \frac{1}{3}\begin{pmatrix} 2x + y - z \\ x + 2y + z \\ -x + y + 2z \end{pmatrix}$.

2) Pour $v \in \mathbb{R}^3$, on utilise la caractérisation du projeté $p_F(v)$: c'est l'unique vecteur de $\mathbb{R}^3$ tel que

$$\begin{cases} p_F(v) \in F \\ v - p_F(v) \in F^{\perp}. \end{cases}$$

Remarque
Cette seconde méthode est plus générale, elle sert à calculer aussi des projecteurs non orthogonaux, alors que la formule utilisée dans la première méthode ne fonctionne que parce qu'on projette parallèlement à l'orthogonal.

2) Symétrie orthogonale, réflexion

Définition 34 (Symétrie orthogonale).
Soit F un sev **de dimension finie** *de E. On appelle* **symétrie orthogonale par rapport à $\boldsymbol{F}$** *la symétrie par rapport à F et parallèlement à $F^{\perp}$.*
On notera $\sigma_F : E \to E$ cette application.

Dessin en dimension 2 :

Dessin en dimension 3 :

Proposition 35 (Propriétés d'une symétrie orthogonale).
Soit F un sev **de dim. finie** *de E, et σ_F la symétrie orthogonale par rapport à F.*

(i) On a $\sigma_F \in \mathcal{L}(E,E)$ et $\sigma_F \circ \sigma_F = Id_E$.

(ii) On a la décomposition $Ker(\sigma_F - Id_E) \oplus Ker(\sigma_F + Id_E) = E$, avec

$$F = Ker(\sigma_F - Id_E) \quad \text{et} \quad F^{\perp} = Ker(\sigma_F + Id_E).$$

(iii) Pour tout $x \in E$, on a $\sigma_F(x) = 2p_F(x) - x$. D'où $\sigma_F = 2p_F - Id_E$.

Preuve : Cela résulte des propriétés des symétries. $\square$

Remarque
Comme pour toute symétrie, on a $\sigma_F \in GL(E)$ et ${\sigma_F}^{-1} = \sigma_F$.
De plus, si E est de dimension finie, σ_F est **diagonalisable**, et $sp(\sigma_F) \subset \{-1, 1\}$.

Définition 36 (Réflexion par rapport à un hyperplan).
Si E est de dimension finie (espace euclidien) et si H est un **hyperplan** *de E (i.e. $\dim(H) = \dim(E) - 1$), alors la symétrie orthogonale par rapport à H est appelée* **réflexion par rapport à $\boldsymbol{H}$**.

Exemple
Dans $E = \mathbb{R}^3$ muni de sa structure euclidienne canonique, calculer la réflexion par rapport au plan $F = Vect\left(\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}\right)$.

On a déjà calculé la projection orthogonale sur le plan F (voir p. 58) :

$$v = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \implies p_F(v) = \frac{1}{3}\begin{pmatrix} 2x + y - z \\ x + 2y + z \\ -x + y + 2z \end{pmatrix}.$$

On en déduit que

$$\sigma_F(v) = 2p_F(v) - v = \frac{2}{3}\begin{pmatrix} 2x + y - z \\ x + 2y + z \\ -x + y + 2z \end{pmatrix} - \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \frac{1}{3}\begin{pmatrix} x + 2y - 2z \\ 2x + y + 2z \\ -2x + 2y + z \end{pmatrix}.$$

3) Distance d'un point à un sev de dimension finie

Proposition 37 (La projection orthogonale est un minimiseur).
Soit F un sev **de dimension finie** *de E, et p_F la projection orthogonale sur F. Alors,*

(i) *Pour tout vecteur $x \in E$, on a $\forall y \in F,\ \|x-y\| \geq \|x-p_F(x)\|$.*
Donc $\|x-p_F(x)\| = \inf_{y\in F}\|x-y\| = \min_{y\in F}\|x-y\|$.

(ii) *Etant donné $x \in E$, si un vecteur $z \in F$ vérifie*
$\forall y \in F,\ \|x-y\| \geq \|x-z\|$, alors $z = p_F(x)$.

Remarque
Le projeté orthogonal $p_F(x)$ est donc le seul "point" qui **minimise** la distance de x aux "points" du sev F.

Preuve :

(i) Pour $x \in E$ et $y \in F$, on a

$$\langle x-p_F(x), p_F(x)-y\rangle = 0$$

(car $x-p_F(x) \in F^\perp$ et $p_F(x)-y \in F$), donc, par le théorème de Pythagore,

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

(ii) On a, d'après le théorème de Pythagore,

$$\|z-p_F(x)\|^2 = \|z-x\|^2 - \|x-p_F(x)\|^2$$

(puisque $\langle x-p_F(x), p_F(x)-z\rangle = 0$, vu que $z \in F$).
Mais par hypothèse, on a $\|x-p_F(x)\| \geq \|x-z\|$ (puisque $p_F(x) \in F$). Donc

$$\|z-p_F(x)\|^2 \leq 0,$$

ce qui amène $\|z-p_F(x)\| = 0$ et $z = p_F(x)$.

$\square$

Définition 38 (Distance d'un point à un sous-espace de dimension finie).
Etant donné un point $x \in E$ et un sous-espace F de dimension finie, on appelle **distance de x à F** *le réel positif :*

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

. . . .

Remarque
On a $d(x; F) = 0 \iff x = p_F(x) \iff x \in F$.

Exemple
Dans l'exemple précédent $\boxed{E = \mathbb{R}^3 \text{ et } F = Vect\left((1,1,0),(0,1,1)\right)}$:
pour tout $M = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3$, on a $p_F(M) = \dfrac{1}{3}\begin{pmatrix} 2x+y-z \\ x+2y+z \\ -x+y+2z \end{pmatrix}$, donc

$$d(M, F) = \|M - p_F(M)\| = \|\overrightarrow{p(M)M}\| = \frac{1}{3}\left\|\begin{pmatrix} x-y+z \\ -x+y-z \\ x-y+z \end{pmatrix}\right\| = \frac{|x-y+z|}{\sqrt{3}}.$$

. . . .

Remarque
Dans cet exemple, une équation cartésienne du plan F est $x - y + z = 0$, car les points $M = (x, y, z)$ qui appartiennent à F sont ceux qui annulent la distance $d(M, F)$.