

# ESPACES PRÉHILBERTIENS RÉELS

B. Landelle

## Table des matières

<b>I</b>	<b>Produit scalaire</b>	<b>2</b>
1	Définition . . . . .	2
2	Exemples importants . . . . .	2
3	Norme euclidienne . . . . .	3
<b>II</b>	<b>Orthogonalité</b>	<b>5</b>
1	Vecteurs orthogonaux . . . . .	5
2	Orthogonal d'une partie . . . . .	7
3	Famille orthonormale . . . . .	8
4	Bases orthonormales . . . . .	10
<b>III</b>	<b>Projection orthogonale</b>	<b>11</b>
1	Supplémentaire orthogonal . . . . .	11
2	Projection orthogonale . . . . .	12
3	Distance à un sev de dimension finie . . . . .	16
<b>IV</b>	<b>Exemples importants</b>	<b>17</b>
1	Distance à un hyperplan . . . . .	17
2	Droite des moindres carrés . . . . .	18

Dans tout ce chapitre, l'ensemble  $E$  désigne un  $\mathbb{R}$ -ev non nul.

# I Produit scalaire

## 1 Définition

**Définition 1.** Une application  $\varphi : E^2 \rightarrow \mathbb{R}$  est appelée produit scalaire sur  $E$  si elle est :

1. bilinéaire, i.e. linéaire par rapport à chaque variable :

$$\forall y \in E \quad x \mapsto \varphi(x, y) \in \mathcal{L}(E, \mathbb{R}) \quad \text{et} \quad \forall x \in E \quad y \mapsto \varphi(x, y) \in \mathcal{L}(E, \mathbb{R})$$

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

3. positive :  $\forall x \in E \quad \varphi(x, x) \geq 0$

4. définie :  $\forall x \in E \quad \varphi(x, x) = 0 \iff x = 0$

On dit que  $\varphi$  est une forme bilinéaire symétrique définie positive.

**Notations :** On note  $\varphi(x, y) = \langle x, y \rangle$  ou  $(x|y)$  ou  $\langle x|y \rangle$  ou  $x \cdot y$ .

**Définition 2.** Un  $\mathbb{R}$ -ev  $E$  muni d'un produit scalaire est dit espace préhilbertien réel. Si l'espace  $E$  est de dimension finie, on dit que c'est est un espace euclidien.

**Proposition 1.** Soit  $\varphi : E^2 \rightarrow \mathbb{R}$  une application vérifiant les propriétés suivantes :

1. symétrique ;
2. linéaire par rapport à une des deux variables ;
3. positive ;
4. définie.

Alors l'application  $\varphi$  est un produit scalaire.

*Démonstration.* La symétrie étend la linéarité par rapport aux deux variables.  $\square$

**Proposition 2.** Soit  $E$  espace préhilbertien réel. Soit  $(x_i)_{1 \leq i \leq n}$  et  $(y_j)_{1 \leq j \leq p}$  des familles finies de vecteurs de  $E$ ,  $(\alpha_i)_{1 \leq i \leq n}$  et  $(\beta_j)_{1 \leq j \leq p}$  des familles de scalaires. On a

$$\left\langle \sum_{i=1}^n \alpha_i x_i, \sum_{j=1}^p \beta_j y_j \right\rangle = \sum_{(i,j) \in [\![1;n]\!] \times [\![1;p]\!]} \alpha_i \beta_j \langle x_i, y_j \rangle$$

*Démonstration.* Conséquence immédiate de la bilinéarité.  $\square$

## 2 Exemples importants

**Proposition 3.**

1. L'espace  $E = \mathbb{R}^n$  muni de  $\langle x, y \rangle = \sum_{i=1}^n x_i y_i$  est euclidien.

2. L'espace  $E = \mathcal{M}_{n,1}(\mathbb{R})$  muni de  $\langle X, Y \rangle = X^\top Y$  est euclidien.

3. L'espace  $E = \mathcal{M}_n(\mathbb{R})$  muni de  $\langle A, B \rangle = \text{Tr}(A^\top B)$  ou  $\langle A, B \rangle = \text{Tr}(AB^\top)$  est euclidien.

*Démonstration.* 1. et 2. immédiats.

3. On a  $\langle A, B \rangle = \sum_{j=1}^n \sum_{i=1}^n a'_{j,i} b_{i,j}$  avec  $A^\top = (a'_{i,j})$ . D'où

$$\langle A, B \rangle = \sum_{1 \leq i, j \leq n} a_{i,j} b_{i,j}$$

expression symétrique en les coefficients d'où la symétrie de l'application. Les autres propriétés s'obtiennent alors facilement.  $\square$

**Remarques :** (1) On confond souvent  $\mathbb{R}^n$  et  $\mathcal{M}_{n,1}(\mathbb{R})$  qui sont naturellement isomorphes.

(2) Les produits scalaires présentés ci-avant sont canoniques, y compris celui sur  $\mathcal{M}_n(\mathbb{R})$  en considérant l'isomorphisme  $\mathcal{M}_n(\mathbb{R}) \simeq \mathbb{R}^{n^2}$ .

**Proposition 4.** Soient  $a, b$  des réels avec  $a < b$ .

1. L'espace  $E = \mathcal{C}^0([a; b], \mathbb{R})$  muni de  $\langle f, g \rangle = \int_a^b f(t)g(t) dt$  est préhilbertien réel.
2. L'espace  $E = \mathbb{R}[X]$  muni de  $\langle P, Q \rangle = \int_0^1 P(t)Q(t) dt$  est préhilbertien réel.
3. L'espace  $E = \mathbb{R}[X]$  muni de  $\langle P, Q \rangle = \int_0^{+\infty} P(t)Q(t)e^{-t} dt$  est préhilbertien réel.

*Démonstration.* 1. Symétrie immédiate, linéarité en la première variable par bilinéarité du produit à droite et de l'intégrale, positive par positivité de l'intégrale.

Si  $\langle f, f \rangle = \int_0^1 f^2(t) dt = 0$ , comme  $f^2$  est continue positive, alors d'après la propriété de séparation de l'intégrale, on a  $f = 0$ .

2. Presque identique au cas précédent. Comme  $P(t) = 0$  pour tout  $t \in [0; 1]$ , alors  $P$  admet une infinité de racines et est donc le polynôme nul.

3. On justifie la convergence de l'intégrale puis quasi identique au précédent.  $\square$

### 3 Norme euclidienne

**Définition 3.** Soit  $E$  espace préhilbertien réel. On appelle norme euclidienne associée au produit scalaire  $\langle \cdot, \cdot \rangle$  l'application notée  $\|\cdot\|$  définie par

$$\|\cdot\| : \begin{cases} E \rightarrow \mathbb{R}_+ \\ x \mapsto \|x\| = \sqrt{\langle x, x \rangle} \end{cases}$$

**Remarque :** La norme euclidienne associée à un produit scalaire est bien définie par positivité du produit scalaire.

**Proposition 5.** Soit  $E$  espace préhilbertien réel et  $(x, y) \in E^2$ . On a

1.  $\|x + y\|^2 = \|x\|^2 + 2\langle x, y \rangle + \|y\|^2$
2.  $\|x - y\|^2 = \|x\|^2 - 2\langle x, y \rangle + \|y\|^2$
3.  $\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2)$  (identité du parallélogramme)

*Démonstration.* Immédiate par définition de la norme et bilinéarité du produit scalaire.  $\square$

**Remarques :** (1) Vérifier l'identité du parallélogramme est caractéristique d'une norme euclidienne (exercice difficile). Par exemple, la norme  $\|\cdot\|_1$  sur  $\mathbb{R}^n$  n'est pas une norme euclidienne (considérer  $x = (1, 0, \dots)$  et  $y = (0, 1, 0, \dots)$ ).

(2) L'identité du parallélogramme peut aussi être reformulée en *identité de la médiane* :

$$\forall (x, y) \in E^2 \quad \|x\|^2 + \|y\|^2 = 2\left\|\frac{x+y}{2}\right\|^2 + \frac{1}{2}\|x-y\|^2$$

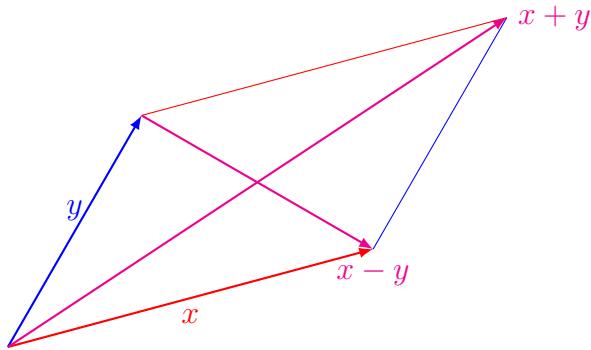


FIGURE 1 – Identité du parallélogramme

**Proposition 6 (Identités de polarisation).** Soit  $E$  espace préhilbertien réel et  $\|\cdot\|$  la norme euclidienne associée. On a

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

*Démonstration.* Immédiate avec le résultat de la proposition précédente.  $\square$

**Remarque :** Les identités de polarisation permettent de reconstruire le produit scalaire d'un espace préhilbertien réel à partir de la connaissance de la norme euclidienne.

**Théorème 1 (Inégalité de Cauchy-Schwarz).** Soit  $E$  espace préhilbertien réel et  $\|\cdot\|$  la norme euclidienne associée. On a

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

De plus  $|\langle x, y \rangle| = \|x\| \|y\| \iff (x, y)$  famille liée

*Démonstration.* Soit  $(x, y) \in E^2$ . On pose

$$\forall t \in \mathbb{R} \quad P(t) = \|x+ty\|^2 = t^2\|y\|^2 + 2t\langle x, y \rangle + \|x\|^2$$

- Si  $\|y\|^2 > 0$ , la fonction  $t \mapsto P(t)$  est polynomiale de degré deux positive et admet donc au plus une racine réelle d'où son discriminant  $\Delta = 4[\langle x, y \rangle^2 - \|x\|^2 \times \|y\|^2] \leq 0$ .
- Si  $\|y\|^2 = 0$ , la fonction  $P$  est une fonction affine de signe constant positif ce qui impose  $\langle x, y \rangle = 0$  et l'égalité donc l'inégalité a lieu.

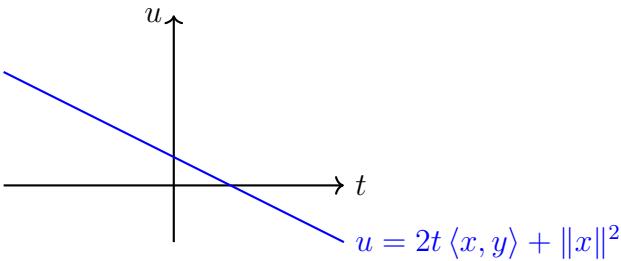


FIGURE 2 – Graphe pour  $\langle x, y \rangle < 0$

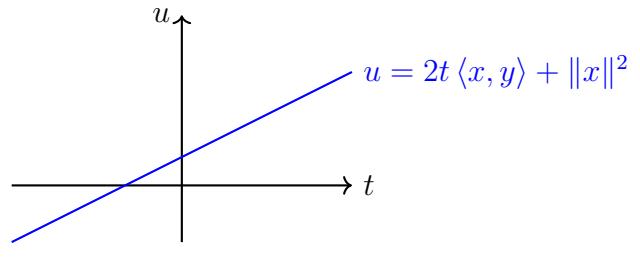


FIGURE 3 – Graphe pour  $\langle x, y \rangle > 0$

L'équivalence a clairement lieu pour  $y = 0_E$ . Pour  $y \neq 0_E$  ce qui équivaut à  $\|y\|^2 > 0$ , on a

$$|\langle x, y \rangle| = \|x\| \|y\| \iff \Delta = 0 \iff \exists! t_0 \in \mathbb{R} \mid P(t_0) = 0 \iff \exists! t_0 \in \mathbb{R} \mid x = -t_0 y$$

Cette dernière assertion équivaut au caractère lié de  $(x, y)$  car  $y \neq 0_E$  d'où l'existence de  $\lambda$  réel tel que  $x = \lambda y$  et ce  $\lambda$  est unique car  $(y)$  est libre.  $\square$

**Remarques :** (1) Dans  $\mathbb{R}^2$ , résultat bien connu car

$$|\vec{x} \cdot \vec{y}| = \|\vec{x}\| \|\vec{y}\| |\cos \alpha| \leq \|\vec{x}\| \|\vec{y}\|$$

(2) La preuve de l'inégalité sans le cas d'égalité n'utilise pas le caractère défini du produit scalaire (remarque qui servira plus tard dans le cours **Probabilités Discrètes**).

**Corollaire 1.** Soit  $E$  espace préhilbertien réel. La norme euclidienne associée est une norme, plus précisément elle vérifie les propriétés suivantes :

1.  $\forall (x, \lambda) \in E \times \mathbb{R} \quad \|\lambda x\| = |\lambda| \|x\|.$
2.  $\forall x \in E \quad \|x\| = 0 \iff x = 0_E$
3.  $\forall (x, y) \in E^2 \quad \|x + y\| \leq \|x\| + \|y\|$  avec égalité si et seulement si la famille  $(x, y)$  est positivement liée.

*Démonstration.* 1. et 2. Immédiates.

3. D'après l'inégalité de Cauchy-Schwarz, on a

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

et le résultat suit par croissance de  $\sqrt{\phantom{x}}$  sur  $\mathbb{R}_+$ . Le cas d'égalité résulte du cas d'égalité dans l'inégalité de Cauchy-Schwarz.  $\square$

## II Orthogonalité

Dans tout ce qui suit, l'ensemble  $E$  est un espace préhilbertien réel.

### 1 Vecteurs orthogonaux

**Définition 4.** Deux vecteurs  $x$  et  $y$  de  $E$  sont dits orthogonaux si  $\langle x, y \rangle = 0$ . On note  $x \perp y$ .

**Définition 5.** Une famille  $(u_i)_{i \in I}$  de vecteurs de  $E$  est dite orthogonale si elle constituée de vecteurs deux à deux orthogonaux, i.e.

$$\forall (i, j) \in I^2 \text{ avec } i \neq j \quad \langle u_i, u_j \rangle = 0$$

**Proposition 7.** Soit  $(u_i)_{i \in I}$  une famille orthogonale de vecteurs non nuls. Alors  $(u_i)_{i \in I}$  est une famille libre.

*Démonstration.* Soit  $(\alpha_i)_{i \in I}$  famille de réels presque nulle telle que  $\sum_{i \in I} \alpha_i u_i = 0_E$ . On a

$$\forall j \in I \quad \left\langle \sum_{i \in I} \alpha_i u_i, u_j \right\rangle = \alpha_j \|u_j\|^2 = 0 \implies \alpha_j = 0$$

d'où le résultat.  $\square$

**Théorème 2 (Théorème de Pythagore).** Soit  $(u_i)_{i \in \llbracket 1; p \rrbracket} \in E^p$  une famille orthogonale. Alors

$$\left\| \sum_{i=1}^p u_i \right\|^2 = \sum_{i=1}^p \|u_i\|^2$$

*Démonstration.* On a  $\left\| \sum_{i=1}^p u_i \right\|^2 = \left\langle \sum_{i=1}^p u_i, \sum_{j=1}^p u_j \right\rangle = \sum_{(i,j) \in \llbracket 1; p \rrbracket^2} \langle u_i, u_j \rangle = \sum_{i=1}^p \|u_i\|^2$

$\square$

**Proposition 8.** Soit  $(x, y) \in E^2$ . On a

$$(x, y) \text{ orthogonale} \iff \|x + y\|^2 = \|x\|^2 + \|y\|^2$$

*Démonstration.* On a

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

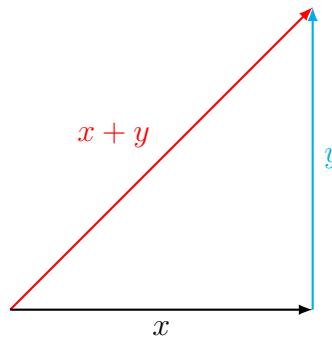


FIGURE 4 – Pythagore, le grand retour

$\square$

**Remarque :** La réciproque du théorème de Pythagore est fausse pour une famille de plus de deux vecteurs. Par exemple, dans  $E = \mathbb{R}^3$  avec  $a = e_1 + e_2$ ,  $b = e_1 - e_2$  et  $c = e_2 + e_3$ . On a

$$\langle a, b \rangle + \langle a, c \rangle + \langle b, c \rangle = 0 \implies \|a + b + c\|^2 = \|a\|^2 + \|b\|^2 + \|c\|^2$$

mais les vecteurs  $a, b, c$  ne sont pas deux à deux orthogonaux (et on peut même choisir une telle famille libre).

## 2 Orthogonal d'une partie

**Définition 6.** Soit  $A$  une partie non vide de  $E$ . On appelle orthogonal de  $A$  l'ensemble noté  $A^\perp$  défini par

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

**Proposition 9.** On a  $E^\perp = \{0_E\}$  et  $\{0_E\}^\perp = E$ .

*Démonstration.* La deuxième égalité est immédiate. Pour la première, soit  $x \in E^\perp$ . On a donc  $\langle x, y \rangle = 0$  pour tout  $y \in E$  et en particulier pour  $y = x$  d'où  $\|x\| = 0$  et par suite  $x = 0_E$ .  $\square$

**Proposition 10.** Soient  $A, B$  des parties non vides de  $E$ . On a

1.  $A^\perp$  est un sev de  $E$  ;
2.  $A \subset B \implies B^\perp \subset A^\perp$  ;
3.  $A^\perp = \text{Vect}(A)^\perp$  ;
4.  $A \subset (A^\perp)^\perp$ .

*Démonstration.* 1. Vérification naïve immédiate.

Variante :  $A^\perp = \bigcap_{a \in A} \text{Ker } (x \mapsto \langle x, a \rangle)$ .

2. Soit  $x \in B^\perp$ . Alors, pour tout  $a \in A \subset B$ , on a  $\langle x, a \rangle = 0$  d'où  $x \in A^\perp$ .

3. Comme  $A \subset \text{Vect}(A)$ , on a  $\text{Vect}(A)^\perp \subset A^\perp$  d'après le 2.

Réiproquement, soit  $x \in A^\perp$ . Soit  $y \in \text{Vect}(A)$ , c'est-à-dire, notant  $A = (a_i)_{i \in I}$ , qu'il existe  $(\alpha_i)_{i \in I}$  famille presque nulle telle que  $y = \sum_{i \in I} \alpha_i a_i$  et

$$\langle x, y \rangle = \sum_{i \in I} \alpha_i \langle x, a_i \rangle = 0$$

et le résultat suit.

4. Soit  $x \in A$ . Alors pour tout  $y \in A^\perp$ , on a  $\langle x, y \rangle = 0$ .  $\square$

**Remarque :** Si  $A$  n'est pas un sev, l'inclusion  $A \subset (A^\perp)^\perp$  est stricte puisque  $(A^\perp)^\perp$  est un sev contenant  $A$  donc le contenant strictement. Si  $A$  est un sev, c'est moins évident. Un contre-exemple est donné ultérieurement.

**Définition 7.** Les sev  $F$  et  $G$  de  $E$  sont dits orthogonaux si

$$\forall (x, y) \in F \times G \quad \langle x, y \rangle = 0$$

On note  $F \perp G$ .

**Proposition 11.** Soient  $F, G$  des sev de  $E$ .

1.  $F \perp G \implies F \cap G = \{0_E\}$
2.  $F \perp G \iff F \subset G^\perp \iff G \subset F^\perp$
3.  $F \perp F^\perp$  et  $F \cap F^\perp = \{0_E\}$

*Démonstration.* 1. Soit  $x \in F \cap G$ . Alors  $\langle x, x \rangle = 0$  d'où  $x = 0_E$ .

2. On a  $F \perp G \iff \forall x \in F \quad \forall y \in G \quad \langle x, y \rangle = 0 \iff F \subset G^\perp$

et l'autre équivalence s'en déduit par symétrie des rôles.

3. Immédiat.  $\square$

**Théorème 3 (Orthogonal d'une famille génératrice).**

1. Soit  $F = \text{Vect}(u_i)_{i \in I}$ . On a

$$x \in F^\perp \iff \forall i \in I \quad x \perp u_i$$

2. Soit  $F = \text{Vect}(u_i)_{i \in I}$  et  $G = \text{Vect}(v_j)_{j \in J}$ . On a

$$F \perp G \iff \forall (i, j) \in I \times J \quad u_i \perp v_j$$

*Démonstration.* 1. D'après la proposition 10, on a  $F^\perp = \{u_i, i \in I\}^\perp$  et le résultat suit.  
2. Le sens direct est immédiat. Réciproquement, on a

$$\forall (i, j) \in I \times J \quad u_i \perp v_j \iff \forall i \in I \quad u_i \in \text{Vect}(v_j, j \in J)^\perp = G^\perp \implies \text{Vect}(u_i, i \in I) \subset G^\perp$$

et le résultat suit.  $\square$

**Proposition 12.** Soient  $F, G$  des sev de  $E$  tels que  $E = F + G$ . Dans ce cas,

$$F \perp G \iff F = G^\perp \iff G = F^\perp$$

*Démonstration.* D'après la proposition 11, on sait que

$$F \perp G \iff F \subset G^\perp \iff G \subset F^\perp$$

Montrons simplement que sous l'hypothèse  $E = F + G$ , on a

$$F \perp G \iff F = G^\perp$$

On suppose  $F \perp G$ . On sait déjà  $F \subset G^\perp$ . Soit  $x \in G^\perp$ . Il existe  $(u, v) \in F \times G$  tel que  $x = u + v$ .

On a  $\langle x, v \rangle = 0 = \underbrace{\langle u, v \rangle}_{=0} + \langle v, v \rangle = \|v\|^2 \implies x = u \in F$

Les rôles joués par  $F$  et  $G$  étant symétriques, on en déduit le résultat attendu.  $\square$

**Notation :** Si  $E = F + G$  avec  $F \perp G$ , alors on a  $E = F \oplus G$  et on note  $E = F \overset{\perp}{\oplus} G$ .

**Remarque importante :** Une conséquence de la proposition 12 est que si un sev admet un supplémentaire orthogonal, alors celui-ci est unique. En effet, si  $E = F \overset{\perp}{\oplus} G$ , alors on a  $G = F^\perp$ .

### 3 Famille orthonormale

**Définition 8.** Une famille de vecteurs de  $E$  est dite orthonormale (ou orthonormée) si elle est orthogonale et constituée de vecteurs unitaires (on dit aussi normés), i.e. de norme égale à 1.

**Proposition 13.** Toute famille orthonormale de  $E$  est libre.

*Démonstration.* Conséquence de la proposition 7.  $\square$

**Théorème 4 (Orthonormalisation de Gram-Schmidt).** Soit  $(u_1, \dots, u_p)$  une famille libre de  $E$ . Il existe  $(v_1, \dots, v_p)$  une famille orthonormale de  $E$  telle que

$$\forall k \in \llbracket 1 ; p \rrbracket \quad \text{Vect}(u_1, \dots, u_k) = \text{Vect}(v_1, \dots, v_k)$$

*Démonstration.* On met en œuvre l’algorithme dit d’*orthonormalisation* de Gram-Schmidt.

- On initialise avec  $v_1 = \frac{u_1}{\|u_1\|}$ .
- Pour  $k \in \llbracket 2 ; p \rrbracket$ , on pose  $z_k = u_k - \sum_{i=1}^{k-1} \langle v_i, u_k \rangle v_i$  puis  $v_k = \frac{z_k}{\|z_k\|}$ .

Montrons par récurrence que la famille  $(v_1, \dots, v_p)$  vérifie les propriétés voulues. Soit

$$\mathcal{P}(k) : "(v_1, \dots, v_k) \text{ orthonormé et } \text{Vect}(v_1, \dots, v_k) = \text{Vect}(u_1, \dots, u_k)"$$

- $\mathcal{P}(1)$  : L’initialisation est immédiate.

- $\mathcal{P}(k) \implies \mathcal{P}(k+1)$  : Supposons  $\mathcal{P}(k)$  vraie pour  $k \in \llbracket 1 ; p-1 \rrbracket$  fixé (on exclut  $k=p$  sinon on suppose le résultat final). On a

$$z_{k+1} = u_{k+1} - \sum_{i=1}^k \langle v_i, u_{k+1} \rangle v_i$$

Alors, pour tout  $j \in \llbracket 1 ; k \rrbracket$ ,

$$\begin{aligned} \langle z_{k+1}, v_j \rangle &= \left\langle u_{k+1} - \sum_{i=1}^k \langle v_i, u_{k+1} \rangle v_i, v_j \right\rangle = \langle u_{k+1}, v_j \rangle - \sum_{i=1}^k \langle v_i, u_{k+1} \rangle \underbrace{\langle v_i, v_j \rangle}_{=\delta_{i,j}} \\ &= \langle u_{k+1}, v_j \rangle - \langle v_j, u_{k+1} \rangle = 0 \end{aligned}$$

Le vecteur  $z_{k+1}$  est non nul. Sinon, on aurait

$$z_{k+1} = 0_E \iff u_{k+1} = \sum_{i=1}^k \langle v_i, u_{k+1} \rangle v_i \in \text{Vect}(v_1, \dots, v_k) = \text{Vect}(u_1, \dots, u_k)$$

Dans ce cas, la famille  $(u_1, \dots, u_{k+1})$  serait liée ce qui est absurde puisqu’elle est extraite de la famille libre  $(u_1, \dots, u_p)$ . Par suite, la norme  $\|z_{k+1}\|$  est non nulle et le vecteur  $v_{k+1} = \frac{z_{k+1}}{\|z_{k+1}\|}$  est normé. Ainsi, la famille  $(v_1, \dots, v_{k+1})$  est orthonormée. On a

$$\text{Vect}(v_1, \dots, v_k) = \text{Vect}(u_1, \dots, u_k) \subset \text{Vect}(u_1, \dots, u_{k+1})$$

et

$$v_{k+1} \in \text{Vect}(u_1, \dots, u_{k+1})$$

d’où

$$\text{Vect}(v_1, \dots, v_{k+1}) \subset \text{Vect}(u_1, \dots, u_{k+1})$$

Enfin, la famille  $(v_1, \dots, v_{k+1})$  est libre car orthonormée et  $(u_1, \dots, u_{k+1})$  est libre car extraite de la famille libre  $(u_1, \dots, u_p)$ . Par suite

$$\dim \text{Vect}(v_1, \dots, v_{k+1}) = \dim \text{Vect}(u_1, \dots, u_{k+1}) = k+1$$

Par inclusion et égalité des dimensions, il vient

$$\text{Vect}(v_1, \dots, v_{k+1}) = \text{Vect}(u_1, \dots, u_{k+1})$$

ce qui clôture la récurrence.

Ainsi,  $\mathcal{P}(k)$  est vraie pour tout  $k \in \llbracket 1 ; p \rrbracket$  d’où  $(v_1, \dots, v_p)$  orthonormée et

$$\forall k \in \llbracket 1 ; p \rrbracket \quad \text{Vect}(u_1, \dots, u_k) = \text{Vect}(v_1, \dots, v_k)$$

□

**Remarque :** Si on impose une condition additionnelle d’orientation à la famille  $(v_1, \dots, v_p)$ , alors celle-ci est unique.

**Théorème 5.** Soit  $(u_1, \dots, u_p)$  une famille libre de  $E$ . Il existe une unique famille orthonormée  $(v_1, \dots, v_p)$  de  $E$  telle que

$$\forall k \in \llbracket 1 ; p \rrbracket \quad \text{Vect}(u_1, \dots, u_k) = \text{Vect}(v_1, \dots, v_k) \quad \text{et} \quad \langle u_k, v_k \rangle > 0$$

*Démonstration.* Soit  $k \in \llbracket 1 ; p \rrbracket$ . Notons  $\pi_k = \sum_{i=1}^{k-1} \langle u_k, v_i \rangle v_i$  et  $z_k = u_k - \pi_k$ . Par construction, on a  $z_k \perp v_i$  pour  $i \in \llbracket 1 ; k-1 \rrbracket$  d'où

$$\langle z_k, u_k \rangle = \langle z_k, z_k + \pi_k \rangle = \|z_k\|^2 + \underbrace{\langle z_k, \pi_k \rangle}_{=0} = \|z_k\|^2 > 0$$

et par suite

$$\langle v_k, u_k \rangle = \frac{1}{\|z_k\|} \langle z_k, u_k \rangle > 0$$

Ainsi, la base construite par l'algorithme de Gram-Schmidt vérifie la condition d'orientation. Soit  $(w_1, \dots, w_p)$  une famille orthonormée qui vérifie les mêmes propriétés que  $(v_1, \dots, v_p)$ . Alors, pour  $k \in \llbracket 1 ; p \rrbracket$ , on a

$$w_k \in \text{Vect}(w_1, \dots, w_k) \cap \text{Vect}(w_1, \dots, w_{k-1})^\perp = \text{Vect}(v_1, \dots, v_k) \cap \text{Vect}(v_1, \dots, v_{k-1})^\perp$$

autrement dit  $w_k \in \text{Vect}(v_k)$ . Comme  $w_k$  est normé et vérifie la même condition d'orientation que  $v_k$ , on conclut  $w_k = v_k$ .  $\square$

**Remarques :** (1) Les résultats et démonstrations d'orthonormalisation s'étendent au cas d'une famille libre dénombrable.

(2) On a établi  $\|z_k\|^2 = \langle z_k, u_k \rangle$  pour  $k \in \llbracket 1 ; p \rrbracket$ . Cette observation peut être mise à profit lors de la mise en œuvre pratique du procédé d'orthonormalisation, le calcul de  $\langle z_k, u_k \rangle$  s'avérant souvent plus simple que le calcul direct de  $\|z_k\|^2$  avec  $k \in \llbracket 1 ; p \rrbracket$ .

### Application : Un cas particulier du théorème de Riesz

Soit  $E$  préhilbertien réel de dimension infinie. Alors, la sphère  $S(0_E, 1)$  n'est pas compacte. En effet, soit  $(u_n)_n$  une famille libre de vecteurs de  $E$ . On note  $(v_n)_n$  la famille obtenue par orthonormalisation de Gram-Schmidt. Pour  $p, q$  entiers distincts, on a

$$\|v_p - v_q\|^2 = \|v_p\|^2 - 2 \langle v_p, v_q \rangle + \|v_q\|^2 = 2$$

Ainsi, la suite  $(v_n)_n$  est à valeurs dans  $S(0_E, 1)$  avec  $\|v_p - v_q\| = \sqrt{2}$  pour tout  $p \neq q$  ce qui interdit l'existence d'une valeur d'adhérence.

## 4 Bases orthonormales

**Définition 9.** Soit  $E$  euclidien. On appelle base orthonormale (ou base orthonormée) notée BON en abrégé une base de  $E$  qui est également une famille orthonormale.

**Corollaire 2.** Soit  $F$  sev de dimension finie de  $E$ . Alors  $F$  possède une base orthonormée.

*Démonstration.* Le sev  $F$  admet une base que l'on orthonormalise avec l'algorithme de Schmidt pour obtenir une base orthonormée de  $F$ .  $\square$

**Corollaire 3.** Soit  $E$  un espace euclidien. Alors

1. l'espace  $E$  possède une base orthonormée ;
2. Toute famille orthonormale de  $E$  peut être complétée en une base orthonormée de  $E$ .

*Démonstration.* 1. C'est le corollaire précédent avec  $F = E$ .

2. Soit  $(u_1, \dots, u_p)$  une famille orthonormée de  $E$  avec  $p \leq n = \dim E$ . D'après le théorème de la base incomplète, on peut compléter  $(u_1, \dots, u_p)$  en  $(u_1, \dots, u_n)$  base de  $E$ . On fixe ensuite  $v_i = u_i$  pour tout  $i \in \llbracket 1; p \rrbracket$  puis on poursuit la construction des  $v_i$  en suivant l'algorithme de Schmidt. On obtient alors une base orthonormée de  $E$  qui est une complétion de la famille initiale.  $\square$

**Proposition 14.** Soit  $E$  euclidien muni d'une base orthonormée  $\mathcal{B} = (e_1, \dots, e_n)$ . Soit  $(x, y) \in E^2$  avec

$$x = \sum_{i=1}^n x_i e_i \quad \text{et} \quad y = \sum_{i=1}^n y_i e_i$$

On a

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

*Démonstration.* On a

$$\langle x, y \rangle = \sum_{(i,j) \in \llbracket 1; n \rrbracket^2} x_i y_j \underbrace{\langle e_i, e_j \rangle}_{=\delta_{i,j}} = \sum_{i=1}^n x_i y_i$$

Le reste suit.  $\square$

**Proposition 15.** Soit  $E$  euclidien muni d'une base orthonormée  $\mathcal{B} = (e_1, \dots, e_n)$ . Soit  $(x, y) \in E^2$  avec et  $X = \text{mat}_{\mathcal{B}} x$ ,  $Y = \text{mat}_{\mathcal{B}} y$ . On a

$$\langle x, y \rangle = X^\top Y, \quad \|x\|^2 = X^\top X$$

*Démonstration.* Immédiate.  $\square$

**Proposition 16.** Soit  $E$  euclidien muni d'une base orthonormée  $\mathcal{B} = (e_1, \dots, e_n)$ ,  $f \in \mathcal{L}(E)$  et  $A = (a_{i,j}) = \text{mat}_{\mathcal{B}} f$ . On a

$$\forall (i, j) \in \llbracket 1; n \rrbracket^2 \quad a_{i,j} = \langle f(e_j), e_i \rangle \quad \text{et} \quad \text{Tr}(f) = \sum_{i=1}^n \langle f(e_i), e_i \rangle$$

*Démonstration.* Soit  $(i, j) \in \llbracket 1; n \rrbracket^2$ . On a  $f(e_j) = \sum_{k=1}^n a_{k,j} e_k$  puis

$$\langle f(e_j), e_i \rangle = \sum_{k=1}^n a_{k,j} \underbrace{\langle e_k, e_i \rangle}_{=\delta_{k,i}} = a_{i,j}$$

Le reste suit.  $\square$

### III Projection orthogonale

Dans tout ce qui suit, l'ensemble  $E$  est un espace préhilbertien réel.

#### 1 Supplémentaire orthogonal

**Théorème 6.** Soit  $F$  sev de dimension finie de  $E$ . Alors on a  $E = F \overset{\perp}{\oplus} F^\perp$ .

*Démonstration.* On suppose  $F \neq \{0_E\}$  sans quoi le résultat est trivial. Soit  $(e_1, \dots, e_p)$  une base orthonormée de  $F$  et  $x \in E$ . On procède par analyse/synthèse.

Analyse : Soit  $(\alpha_1, \dots, \alpha_p) \in \mathbb{R}^p$  tel que

$$x = \sum_{i=1}^p \alpha_i e_i + x - \sum_{i=1}^p \alpha_i e_i \quad \text{avec} \quad x - \sum_{i=1}^p \alpha_i e_i \in F^\perp$$

On a  $\forall k \in \llbracket 1 ; p \rrbracket \quad \left\langle x - \sum_{i=1}^p \alpha_i e_i, e_k \right\rangle = \langle x, e_k \rangle - \alpha_k = 0$

Ainsi les  $\alpha_k$  sont déterminés (cela prouve aussi l'unicité de la décomposition dans  $F \oplus F^\perp$  sous réserve d'existence mais on l'a déjà en réalité car on sait  $F \perp F^\perp$ ).

Synthèse : On a

$$x = \underbrace{\sum_{i=1}^p \langle x, e_i \rangle e_i}_{\in F} + x - \sum_{i=1}^p \langle x, e_i \rangle e_i$$

Puis  $\forall k \in \llbracket 1 ; p \rrbracket \quad \left\langle x - \sum_{i=1}^p \langle x, e_i \rangle e_i, e_k \right\rangle = \langle x, e_k \rangle - \langle x, e_k \rangle \|e_k\|^2 = 0$

donc, d'après le théorème 3, on a  $x - \sum_{i=1}^p \langle x, e_i \rangle e_i \in F^\perp$  ce qui prouve le résultat attendu.  $\square$

**Remarque :** Le résultat est faux si  $F$  est de dimension infinie. Considérons par exemple  $E = \mathcal{C}^0([0;1], \mathbb{R})$  muni du produit scalaire  $\langle f, g \rangle = \int_0^1 f(t)g(t) dt$  pour  $(f, g) \in E^2$ . On pose  $F = \{f \in E \mid f(0) = 0\}$ . Soit  $f \in F^\perp$ . Notant  $g : t \mapsto tf(t)$ , on a  $g \in F$  d'où

$$\langle f, g \rangle = \int_0^1 tf(t)^2 dt = 0$$

Comme  $t \mapsto tf(t)^2$  est continue positive sur  $[0;1]$ , on en déduit  $tf^2(t) = 0$  pour  $t \in [0;1]$  d'où  $f(t) = 0$  pour  $t \in ]0;1]$ . Par continuité de  $f$  en 0, on a également  $f(0) = 0$  d'où  $F^\perp = \{0_E\}$  et par conséquent

$$F \oplus F^\perp \neq E$$

**Corollaire 4.** Soit  $F$  un sev de dimension finie de  $E$ . On a  $(F^\perp)^\perp = F$ .

*Démonstration.* On a  $E = F \oplus F^\perp$  donc, d'après la proposition 12, comme  $F \perp F^\perp$ , on a  $(F^\perp)^\perp = F$ .  $\square$

**Remarque :** Le résultat est faux en dimension infinie. On reprend le contre-exemple précédent. On a trouvé  $F^\perp = \{0_E\}$  d'où  $(F^\perp)^\perp = E$  et par conséquent

$$F \not\subseteq (F^\perp)^\perp$$

## 2 Projection orthogonale

**Définition 10.** Soit  $F$  un sev de dimension finie de  $E$ . On appelle projection orthogonale sur  $F$  notée  $p_F$  la projection sur  $F$  parallèlement à  $F^\perp$ .

**Remarque :** (1) Comme  $F$  et  $F^\perp$  sont supplémentaires, on peut définir une telle projection.  
(2) Si  $E$  est euclidien, alors  $F$  et  $F^\perp$  sont de dimension finie. Par suite, le projecteur  $\text{id} - p_F$  associé à  $p_F$  peut se noter  $p_{F^\perp}$  (notation licite car  $F^\perp$  de dimension finie).

**Théorème 7 (Caractérisation géométrique du projeté orthogonal).** Soit  $F$  un sev de  $E$  de dimension finie et  $x \in E$ . Pour  $y \in E$ , on a

$$y = p_F(x) \iff \begin{cases} y \in F \\ x - y \in F^\perp \end{cases}$$

*Démonstration.* Le sens direct est immédiat par définition de  $p_F$  et le sens indirect résulte de l'unicité de la décomposition dans  $F \oplus F^\perp$  :

$$\begin{aligned} x &= y + (x - y) \quad \text{avec } (y, x - y) \in F \times F^\perp \\ &= p_F(x) + x - p_F(x) \quad \text{avec } (p_F(x), x - p_F(x)) \in F \times F^\perp \end{aligned}$$

et on identifie terme à terme.  $\square$

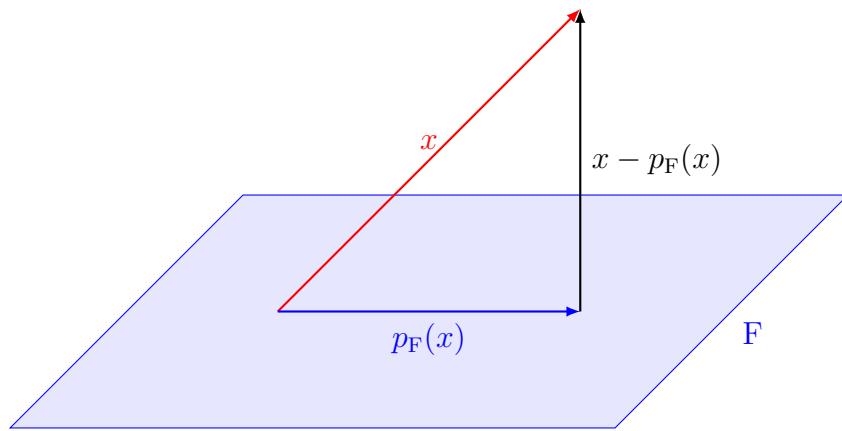


FIGURE 5 – Décomposition d'une projection orthogonale

**Exemple :** Soit  $E = \mathbb{R}_2[X]$  muni de  $\langle P, Q \rangle = \int_0^{+\infty} P(t)Q(t)e^{-t} dt$ .

Déterminer la projection de  $1$  sur  $F = \text{Vect}(X, X^2)$ . On note  $p_F(1) = aX + bX^2$ . On a

$$\begin{cases} \langle 1 - (aX + bX^2), X \rangle = 0 \\ \langle 1 - (aX + bX^2), X^2 \rangle = 0 \end{cases} \iff \begin{cases} 2a + 6b = 1 \\ 6a + 24b = 2 \end{cases} \iff (a, b) = \left(1, -\frac{1}{6}\right)$$

Ainsi

$$p_F(1) = X - \frac{X^2}{6}$$

**Théorème 8.** Soit  $F$  un sev de dimension finie de  $E$  et  $(e_1, \dots, e_p)$  une base orthonormée de  $F$ . On a

$$\forall x \in E \quad p_F(x) = \sum_{i=1}^p \langle x, e_i \rangle e_i$$

*Démonstration.* Dans la preuve du théorème 6, on a établi que  $E = F \oplus F^\perp$  avec  $(e_1, \dots, e_p)$  une base orthonormée de  $F$  et  $x \in E$  se décompose de manière unique en

$$x = \underbrace{\sum_{i=1}^p \langle x, e_i \rangle e_i}_{\in F} + \underbrace{x - \sum_{i=1}^p \langle x, e_i \rangle e_i}_{\in F^\perp}$$

On en déduit

$$\forall x \in E \quad p_F(x) = \sum_{i=1}^p \langle x, e_i \rangle e_i$$

□

**Remarque :** Dans l'algorithme de Gram-Schmidt, l'étape itérative consiste à construire  $z_k = u_k - p_{F_k}(u_k)$  où  $F_k = \text{Vect}(u_1, \dots, u_{k-1}) = \text{Vect}(v_1, \dots, v_{k-1})$ . Ainsi, on a

$$v_k = \frac{z_k}{\|z_k\|} \in F_k^\perp \quad \text{d'où} \quad v_k \perp v_i \quad \forall i \in \llbracket 1 ; k-1 \rrbracket$$

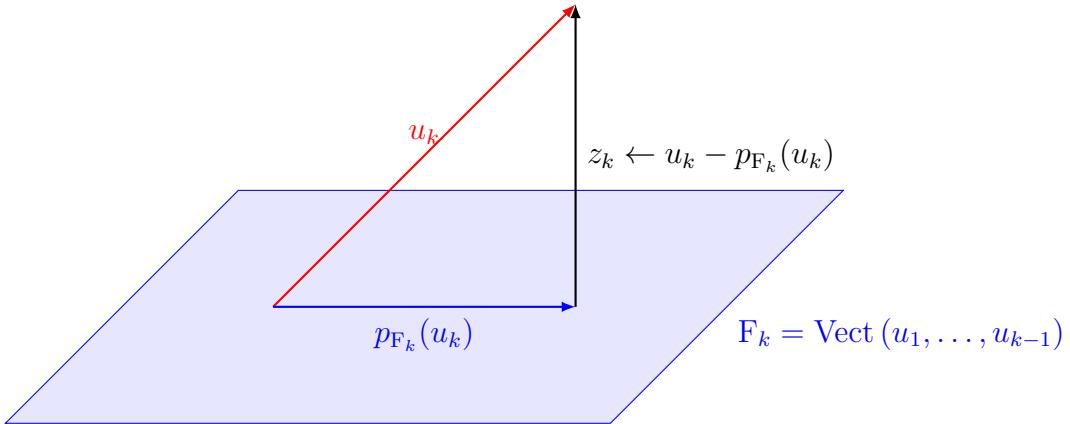


FIGURE 6 – Étape itérative de l'algorithme d'orthonormalisation

**Proposition 17 (Inégalité de Bessel).** Soit  $(e_1, \dots, e_p)$  une famille orthonormée de  $E$ . On a

$$\forall x \in E \quad \sum_{i=1}^p \langle x, e_i \rangle^2 \leq \|x\|^2$$

*Démonstration.* On note  $F = \text{Vect}(e_1, \dots, e_p)$ . Soit  $x \in E$ . On décompose

$$x = \underbrace{\sum_{i=1}^p \langle x, e_i \rangle e_i}_{\in F} + \underbrace{x - \sum_{i=1}^p \langle x, e_i \rangle e_i}_{\in F^\perp}$$

D'après le théorème de Pythagore, il vient

$$\|x\|^2 = \left\| \sum_{i=1}^p \langle x, e_i \rangle e_i \right\|^2 + \left\| x - \sum_{i=1}^p \langle x, e_i \rangle e_i \right\|^2 \geq \left\| \sum_{i=1}^p \langle x, e_i \rangle e_i \right\|^2 = \sum_{i=1}^p \langle x, e_i \rangle^2$$

□

**Proposition 18 (à refaire).** Soit  $a$  un vecteur normé de  $E$  euclidien. On a

$$\forall x \in E \quad p_{\text{Vect}(a)^\perp}(x) = x - \langle x, a \rangle a$$

*Démonstration.* Pour  $x \in E$ , comme  $(a)$  est une base orthonormée de  $\text{Vect}(a)$ , on a

$$p_{\text{Vect}(a)^\perp}(x) = (\text{id} - p_{\text{Vect}(a)})(x) = x - \langle x, a \rangle a$$

□

**Définition 11.** Soit  $p$  un projecteur de  $E$ . Le projecteur  $p$  est dit orthogonal si  $\text{Im } p \perp \text{Ker } p$ .

**Remarque :** Cette définition étend la définition 10 puisque  $\text{Im } p$  n'est pas nécessairement de dimension finie. Les deux définitions sont cohérentes entre elles : si  $F$  est de dimension finie de  $E$ , alors  $p_F$  est un projecteur orthogonal au sens de la définition 11. Le projecteur associé  $\text{id} - p_F$  est également un projecteur orthogonal puisqu'il est le projecteur sur  $F^\perp$  parallèlement à  $F$ .

**Proposition 19 (à refaire).** Soit  $p$  un projecteur de  $E$ . On a

$$p \text{ projecteur orthogonal} \iff \forall x \in E \quad \|p(x)\| \leq \|x\|$$

*Démonstration.* Supposons  $p$  projecteur orthogonal. Soit  $x \in E$ . On a

$$x = \underbrace{p(x)}_{\in \text{Im } p} + \underbrace{x - p(x)}_{\in \text{Ker } p}$$

D'après le théorème de Pythagore, il vient

$$\|x\|^2 = \|p(x)\|^2 + \|x - p(x)\|^2 \geq \|p(x)\|^2$$

Le résultat suit. Réciproquement, soit  $(a, b) \in \text{Im } p \times \text{Ker } p$ . on a

$$\forall t \in \mathbb{R} \quad \|p(ta + b)\|^2 \leq \|ta + b\|^2 \iff \forall t \in \mathbb{R} \quad 0 \leq 2t \langle a, b \rangle + \|b\|^2$$

Dans cette dernière assertion, une fonction affine prend des valeurs toujours positives.

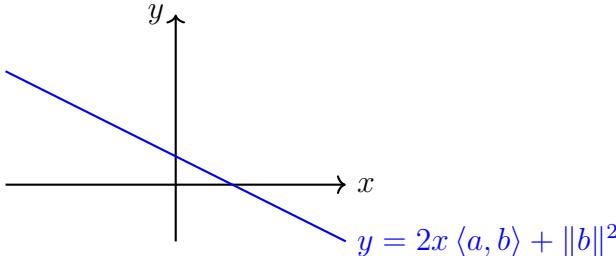


FIGURE 7 – Graphe pour  $\langle a, b \rangle < 0$

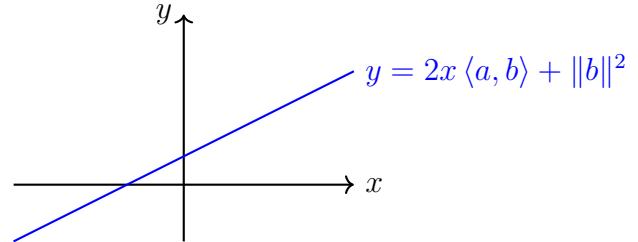


FIGURE 8 – Graphe pour  $\langle a, b \rangle > 0$

Ceci n'est possible que si la fonction est constante d'où  $\langle a, b \rangle = 0$  pour  $(a, b) \in \text{Im } p \times \text{Ker } p$  ce qui prouve que  $p$  est orthogonal. □

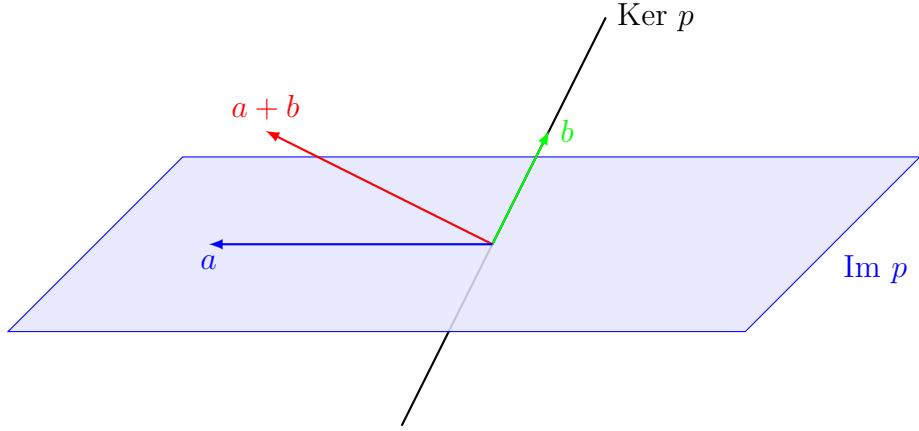


FIGURE 9 – Décomposition d'une projection non orthogonale

Sur la figure ci-avant, on observe que pour un projecteur  $p$  non orthogonal, on peut trouver  $(a, b) \in \text{Im } p \times \text{Ker } p$  tel que

$$\|a\| > \|a + b\|$$

**Remarque :** On peut interpréter le résultatat de la proposition 19 ainsi : pour  $p$  un projecteur, on a

$$p \text{ orthogonal} \iff p \in \mathcal{L}_c(E) \text{ et } \|p\|_{\text{op}} \leq 1$$

### 3 Distance à un sev de dimension finie

**Théorème 9 (Caractérisation métrique du projeté orthogonal).** Soit  $F$  un sev de dimension finie et  $x \in E$ . Alors, on a  $\|x - p_F(x)\| = d(x, F)$  et  $p_F(x)$  est l'unique élément de  $F$  qui réalise cette égalité, i.e. pour  $y \in E$

$$\begin{cases} y \in F \\ \|x - y\| = d(x, F) \end{cases} \iff y = p_F(x)$$

*Démonstration.* Soit  $x \in E$ . Pour  $y \in F$ , on a

$$x - y = \underbrace{x - p_F(x)}_{\in F^\perp} + \underbrace{p_F(x) - y}_{\in F} \stackrel{\text{Pythagore}}{\implies} \|x - y\|^2 = \|x - p_F(x)\|^2 + \|p_F(x) - y\|^2 \geq \|x - p_F(x)\|^2$$

d'où

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

Si  $y \in F$  tel que  $\|x - y\| = d(x, F)$ , alors

$$\|p_F(x) - y\|^2 = \|x - y\|^2 - \|x - p_F(x)\|^2 = 0 \implies y = p_F(x)$$

□

**Corollaire 5.** Soit  $F$  un sev de dimension finie et  $x \in E$ . On a

$$d(x, F)^2 = \langle x - p_F(x), x \rangle$$

*Démonstration.* On a  $d(x, F)^2 = \|x - p_F(x)\|^2 = \langle x - p_F(x), x \rangle - \underbrace{\langle x - p_F(x), p_F(x) \rangle}_{\in F^\perp}$

□

**Remarque :** En pratique, ce résultat permet de beaucoup simplifier certains calculs.

**Exemple :** Soit  $E = \mathbb{R}_2[X]$  muni de  $\langle P, Q \rangle = \int_0^{+\infty} P(t)Q(t)e^{-t} dt$  et  $F = \text{Vect}(X, X^2)$ . Déterminer  $d(1, F)$ . On a

$$d(1, F)^2 = \langle 1 - p_F(1), 1 \rangle = \int_0^{+\infty} \left[ 1 - t + \frac{t^2}{6} \right] e^{-t} dt = \frac{1}{3}$$

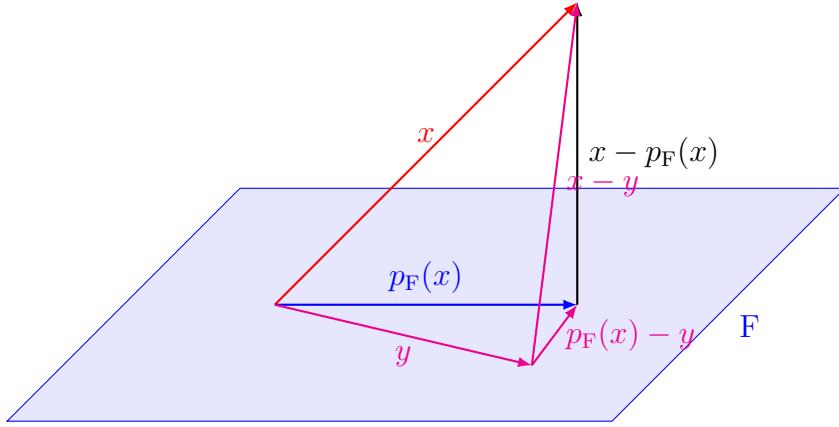


FIGURE 10 – Distance à un sous-espace de dimension finie

## IV Exemples importants

### 1 Distance à un hyperplan

**Proposition 20.** Soit  $E$  un espace euclidien et  $H$  un hyperplan de  $E$  avec  $n \in H^\perp$  et  $n \neq 0_E$ .

On a

$$\forall x \in E \quad d(x, H) = \frac{|\langle x, n \rangle|}{\|n\|}$$

*Démonstration.* On a  $E = H^\perp \oplus \text{Vect}(n)$ . Comme  $\text{Vect}(n) = H^\perp$  est de dimension finie, on a  $\text{id} - p_H = p_{H^\perp}$ . La famille  $(n/\|n\|)$  est une base orthonormée de  $H^\perp$  et on obtient

$$d(x, H) = \|(\text{id} - p_H)(x)\| = \|p_{H^\perp}(x)\| = \frac{|\langle x, n \rangle|}{\|n\|}$$

□

**Remarque :** Le résultat vaut aussi avec  $E$  préhilbertien réel mais c'est un travail plus délicat.

**Exemple :** Soit  $E = \mathbb{R}^n$  et  $H : \sum_{i=1}^n x_i = 0$ . Notons  $u = (1, \dots, 1) \in E$ . On a

$$H = \{x \in E \mid x \perp u\} = \text{Vect}(u)^\perp$$

d'où

$$\forall x \in E \quad d(x, H) = \frac{|\langle x, u \rangle|}{\|u\|} = \frac{\left| \sum_{i=1}^n x_i \right|}{\sqrt{n}}$$

## 2 Droite des moindres carrés

**Proposition 21.** Soit  $(M_i)_{1 \leq i \leq n}$  une suite de points ( $n \geq 2$ ) de coordonnées  $(x_i, y_i)$  telle que toutes les abscisses ne soient pas égales. Étant donnée une droite  $\mathcal{D}$  d'équation  $y = ax + b$ , on note  $H_i$  le projeté de  $M_i$  sur  $\mathcal{D}$  parallèlement à l'axe  $(Oy)$  avec  $i \in \llbracket 1 ; n \rrbracket$ . Alors il existe une unique droite  $\mathcal{D}^*$  telle que la somme des distances au carrés  $\sum_{i=1}^n M_i H_i^2$  soit minimale. Cette droite s'appelle la droite des moindres carrés et est décrite par l'équation

$$y = a^*x + b^* \quad \text{avec} \quad a^* = \frac{\sigma_{x,y}}{\sigma_x^2} \quad b^* = \bar{y} - a^*\bar{x}$$

où  $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$      $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$      $\sigma_x^2 = \frac{1}{n} \sum_{i=1}^n x_i^2 - \bar{x}^2$      $\sigma_{x,y} = \frac{1}{n} \sum_{i=1}^n x_i y_i - \bar{x} \bar{y}$

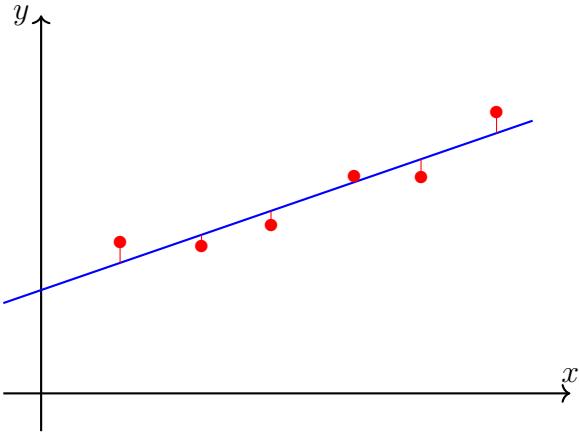


FIGURE 11 – Droite des moindres carrés

La droite des moindres carrés d'équation  $y = ax + b$  est choisie telle que la quantité

$$\sum_{i=1}^n M_i H_i^2 = \sum_{i=1}^n [y_i - (ax_i + b)]^2$$

soit minimale. C'est la droite qui passe « au mieux au plus près » des points  $M_1, \dots, M_n$ . Le problème consiste donc en la minimisation en  $(a, b) \in \mathbb{R}^2$  de la somme des carrés  $\sum_{i=1}^n [y_i - (ax_i + b)]^2$ .

**Vocabulaire :** La suite de points  $M_1, \dots, M_n$  s'appelle un échantillon, en général obtenu par une suite de mesures. La quantité  $\bar{x}$  est la moyenne des abscisses,  $\bar{y}$  la moyenne des ordonnées,  $\sigma_x^2$  la variance des abscisses,  $\sigma_x$  l'écart-type des abscisses et  $\sigma_{x,y}$  la covarience entre les abscisses et ordonnées.

*Démonstration.* On a

$$\sum_{i=1}^n M_i H_i^2 = \sum_{i=1}^n [y_i - (ax_i + b)]^2$$

L'ensemble

$$\Lambda = \left\{ \sum_{i=1}^n [y_i - ax_i - b]^2, (a, b) \in \mathbb{R}^2 \right\}$$

est une partie non vide minorée de  $\mathbb{R}$  et admet donc une borne inférieure finie. Munissons  $\mathbb{R}^n$  du produit scalaire canonique. Ainsi, notant  $y = (y_1, \dots, y_n)$ ,  $x = (x_1, \dots, x_n)$  et  $u = (1, \dots, 1)$ , on a

$$\sum_{i=1}^n M_i H_i^2 = \sum_{i=1}^n [y_i - (ax_i + b)]^2 = \|y - (ax + bu)\|^2$$

Par croissance et continuité de  $t \mapsto t^2$  sur  $\mathbb{R}_+$ , on a

$$\inf_{(a,b) \in \mathbb{R}^2} \sum_{i=1}^n M_i H_i^2 = \inf_{(a,b) \in \mathbb{R}^2} \|y - (ax + bu)\|^2 = \left( \inf_{(a,b) \in \mathbb{R}^2} \|y - (ax + bu)\| \right)^2$$

Notons  $F = \text{Vect}(u, x)$ . D'après la caractérisation métrique du projeté orthogonal, on sait que

$$\inf \{\|y - (ax + bu)\|, (a, b) \in \mathbb{R}^2\} = d(y, F) = \|y - p_F(y)\|$$

et le projeté  $p_F(y)$  est l'unique vecteur de  $F$  qui réalise cette égalité. La famille  $(x, u)$  est libre car les  $x_i$  ne sont pas toutes égales. On a donc reformulé le problème de la droite des moindres carrés comme recherche de distance à un plan vectoriel. On construit une base  $(v_1, v_2)$  orthonormée de  $F$  avec l'algorithme d'orthonormalisation :

$$v_1 = \frac{u}{\sqrt{n}} \quad v_2 = \frac{1}{\sqrt{n}\sigma_x} (x - \bar{x}u)$$

Puis

$$\begin{aligned} p_F(y) &= \langle y, v_1 \rangle v_1 + \langle y, v_2 \rangle v_2 \\ &= \bar{y}u + \frac{1}{n\sigma_x^2} (\langle y, x \rangle - \langle y, u \rangle \bar{x}) (x - \bar{x}u) \\ p_F(y) &= \bar{y}u + \frac{n\sigma_{x,y}}{n\sigma_x^2} (x - \bar{x}u) = \frac{\sigma_{x,y}}{\sigma_x^2} x + \left( \bar{y} - \frac{\sigma_{x,y}}{\sigma_x^2} \bar{x} \right) u \end{aligned}$$

Par liberté de  $(x, u)$ , on identifie les coefficients  $a^*$  et  $b^*$  attendus.

*Variante* : On peut opter pour une approche un peu différente, moins géométrique et plus probabiliste, pour obtenir le résultat. On a

$$\sum_{i=1}^n M_i H_i^2 = \|y - ax - bu\|^2 = \|y - ax - (\bar{y} - a\bar{x})u + (\bar{y} - a\bar{x})u - bu\|^2$$

Après calcul, on vérifie  $\langle y - ax - (\bar{y} - a\bar{x})u, u \rangle = 0$

Ainsi, on obtient

$$\sum_{i=1}^n M_i H_i^2 = \|y - \bar{y}u - a(x - \bar{x}u)\|^2 + \|(\bar{y} - a\bar{x} - b)u\|^2$$

Le premier terme est un trinôme en  $a$  qu'on minimise avec le choix

$$a = \frac{\langle x - \bar{x}u, y - \bar{y}u \rangle}{\|x - \bar{x}u\|^2}$$

puis, on choisit  $b = \bar{y} - a\bar{x}$  qui annule le second terme et on donc minimisé la somme des carrés des écarts. Cette démarche est inspirée des probabilités. Soit  $(\Omega, \mathcal{A}, \mathbb{P})$  espace probabilisé. Si on suppose que le couple  $(X, Y)$  suit une loi uniforme sur  $(x_i, y_i)_{1 \leq i \leq n}$ , on a par transfert

$$n\mathbb{E}((Y - aX - b)^2) = \|y - ax - bu\|^2$$

et la décomposition proposée consiste à écrire

$$\mathbb{E}((Y - aX - b)^2) = \mathbb{V}(Y - aX - b) + (\mathbb{E}(Y - aX - b))^2$$

et on observe  $\mathbb{V}(Y - aX - b) = \mathbb{V}(Y - aX)$

Ce terme ne dépend donc que de  $a$ . On le minimise puis on choisit  $b$ . □

**Applications** : Innombrables !!

## Annexe

### Borne inférieure et distance

Soit  $A$  partie non vide minorée de  $\mathbb{R}$  et  $f$  définie sur  $A \cup \inf A$ . Si  $f$  est croissante et continue en  $\inf A$ , on a

$$\inf f(A) = f(\inf A)$$

Par croissance de  $f$ , on a

$$\forall x \in A \quad x \geq \inf A \implies f(x) \geq f(\inf A)$$

d'où

$$\inf f(A) \geq f(\inf A)$$

Puis, on a l'existence d'une suite  $(a_n)$  à valeurs dans  $A$  avec  $a_n \xrightarrow{n \rightarrow \infty} \inf A$ . Or

$$\forall n \in \mathbb{N} \quad f(a_n) \geq \inf f(A)$$

Comme  $f$  est continue en  $\inf A$ , faisant tendre  $n \rightarrow +\infty$ , il vient

$$f(\inf A) = \lim_{n \rightarrow +\infty} f(a_n) \geq \inf f(A)$$

On conclut

$$\boxed{f(\inf A) = \inf f(A)}$$

Si  $f$  n'est pas croissante, le résultat est faux : prendre  $f(x) = x^2$  et  $A = [-1 ; 1]$  par exemple.  
On a

$$f(\inf A) = f(-1) = 1 \neq 0 = \inf \{x^2, x \in [-1 ; 1]\}$$

Si  $f$  n'est pas continue en  $\inf A$ , le résultat est faux également : prendre  $f(x) = \begin{cases} x & \text{si } x > 0 \\ -1 & \text{si } x = 0 \end{cases}$   
et  $A = [0 ; +\infty[$ . On a

$$f(\inf A) = f(0) = -1 \neq 0 = \inf \{x, x > 0\}$$

**Application :** Situation typique dans  $E$  préhilbertien avec  $(x, y, z) \in E^3$

$$\boxed{\inf_{(a,b) \in \mathbb{R}^2} \|x + ay + bz\|^2 = \left( \inf_{(a,b) \in \mathbb{R}^2} \|x + ay + bz\| \right)^2 = d(x, \text{Vect}(y, z))^2}$$

La fonction  $t \mapsto t^2$  est croissante continue sur  $\mathbb{R}_+$  d'où l'application du résultat qui précède.