

12

Espaces préhilbertiens réels

« C'est par la logique que nous démontrons, mais c'est par l'intuition que nous découvrons; sans elle, le géomètre serait comme un écrivain qui serait ferré sur la grammaire, mais qui n'aurait pas d'idées. »

Henri Poincaré (1854–1912)

Plan de cours

I	Généralités	1
II	Orthogonalité	5
III	Suites totales (*)	11
IV	Méthode des moindres carrés (complément)	13

I | Généralités

Dans tout ce chapitre, E désigne un \mathbb{R} -espace vectoriel de dimension finie ou non.

A – Produit scalaire

Définition 12.1 : Produit scalaire

On appelle produit scalaire sur E toute forme bilinéaire symétrique définie positive, c'est-à-dire toute application $\varphi : E \times E \rightarrow \mathbb{R}$ telle que :

- φ est bilinéaire :
 $\forall x_1, x_2, y \in E, \forall \lambda \in \mathbb{R}, \varphi(\lambda x_1 + x_2, y) = \lambda \varphi(x_1, y) + \varphi(x_2, y)$.
 $\forall x, y_1, y_2 \in E, \forall \lambda \in \mathbb{R}, \varphi(x, \lambda y_1 + y_2) = \lambda \varphi(x, y_1) + \varphi(x, y_2)$.
- φ est symétrique : $\forall x, y \in E, \varphi(x, y) = \varphi(y, x)$.
- φ est définie positive : $\forall x \in E, \varphi(x, x) \geq 0$ et $\varphi(x, x) = 0$ si et seulement si $x = 0_E$.

On note généralement le produit scalaire $(\cdot | \cdot)$, $\langle \cdot | \cdot \rangle$ ou $\langle \cdot, \cdot \rangle$.

Il suffit de vérifier la linéarité à gauche et la symétrie pour justifier la bilinéarité.

Définition 12.2 : Espaces préhilbertiens réels

- On appelle espace préhilbertien réel tout \mathbb{R} -espace vectoriel muni d'un produit scalaire. Notation usuelle : $(E, (\cdot | \cdot))$.
- Un espace préhilbertien réel de dimension finie est appelé espace euclidien.

Voici quatre exemples fondamentaux d'espaces préhilbertiens réels, à connaître sur le bout des doigts.

Exemple 1 – \mathbb{R}^n muni du produit scalaire canonique

Le produit scalaire canonique est défini par :

$$\forall x, y \in \mathbb{R}^n \quad (x | y) = \sum_{i=1}^n x_i y_i \text{ en notant } x = (x_1, \dots, x_n), y = (y_1, \dots, y_n)$$

Si on pose $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ et $Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$, on a $(x | y) = {}^t X Y = X^T Y$.

On vérifie aisément que l'application ainsi définie est bilinéaire et symétrique.

De plus, l'application $(\cdot|\cdot)$ est définie positive car quel que soit $x \in \mathbb{R}^n$:

$$(x|x) = \sum_{i=1}^n x_i^2 \geq 0 \quad \text{et} \quad (x|x) = \sum_{i=1}^n x_i^2 = 0 \iff \forall i \in [1, n] \quad x_i = 0 \iff x = \mathbf{0}_{\mathbb{R}^n}$$

Exemple 2 – $E = \mathcal{C}([a, b], \mathbb{R})$ muni de $(f, g) \mapsto \int_a^b f(t)g(t) dt$

Si $f, g \in \mathcal{C}([a, b], \mathbb{R})$, l'intégrale existe car $f \times g$ est continue sur le segment $[a, b]$.

$(\cdot|\cdot)$ est bien à valeurs dans \mathbb{R} . De plus,

- $(\cdot|\cdot)$ est bilinéaire. Soient $f, g, h \in \mathcal{C}([a, b], \mathbb{R})$ et $\lambda \in \mathbb{R}$.

$$(\lambda f + g|h) = \int_a^b (\lambda f(t) + g(t))h(t) dt = \lambda \int_a^b f(t)h(t) dt + \int_a^b g(t)h(t) dt = \lambda(f|h) + (g|h)$$

par linéarité de l'intégrale; ce qui justifie la linéarité à gauche. On obtient la linéarité à droite par symétrie.

- $(\cdot|\cdot)$ est symétrique. Soient $f, g \in \mathcal{C}([a, b], \mathbb{R})$.

$$(f|g) = \int_a^b f(t)g(t) dt = \int_a^b g(t)f(t) dt = (g|f)$$

- $(\cdot|\cdot)$ est définie positive. Soit $f \in \mathcal{C}([a, b], \mathbb{R})$. $(f|f) = \int_a^b f^2(t) dt \geq 0$ par positivité de l'intégrale et :

$$(f|f) = 0 \iff \int_a^b f^2(t) dt = 0 \iff \begin{array}{c} \iff \forall t \in [a, b] \quad f^2(t) = 0 \\ \text{\small } f^2 \text{ est continue} \\ \text{\small et positive sur } [a, b] \end{array} \iff f \text{ est nulle sur } [a, b]$$

Exemple 3 – $E = \mathbb{R}[X]$ muni de $(P, Q) \mapsto \int_0^1 P(t)Q(t) dt$

Quels que soient les polynômes P et Q , l'intégrale existe car la fonction polynomiale $t \mapsto P(t)Q(t)$ est continue sur le segment $[0, 1]$. $(\cdot|\cdot)$ est bien à valeurs dans \mathbb{R} . De plus,

- $(\cdot|\cdot)$ est bilinéaire. Soient $P, Q, R \in \mathbb{R}[X]$ et $\lambda \in \mathbb{R}$.

$$(\lambda P + Q|R) = \int_0^1 (\lambda P(t) + Q(t))R(t) dt = \lambda \int_0^1 P(t)R(t) dt + \int_0^1 Q(t)R(t) dt = \lambda(P|R) + (Q|R)$$

par linéarité de l'intégrale; ce qui justifie la linéarité à gauche.

On obtient la linéarité à droite par symétrie.

- $(\cdot|\cdot)$ est symétrique. Soient $P, Q \in \mathbb{R}[X]$.

$$(P|Q) = \int_0^1 P(t)Q(t) dt = \int_0^1 Q(t)P(t) dt = (Q|P)$$

- $(\cdot|\cdot)$ est définie positive. Soit $P \in \mathbb{R}[X]$. $(P|P) = \int_0^1 P^2(t) dt \geq 0$ par positivité de l'intégrale et :

$$(P|P) = 0 \iff \int_0^1 P^2(t) dt = 0 \iff \begin{array}{c} \iff \forall t \in [0, 1] \quad P^2(t) = 0 \\ \text{\small } P^2 \text{ est continue} \\ \text{\small et positive sur } [0, 1] \\ \iff P \text{ admet une} \\ \text{\small infinité de racines} \end{array} \iff P = \mathbf{0}_{\mathbb{R}[X]}$$

Exemple 4 – $E = \mathcal{M}_n(\mathbb{R})$ muni de $(A, B) \mapsto \text{Tr}({}^t AB) = \text{Tr}(A^T B)$

Rappelons tout d'abord que :

$$\forall (i, j) \in \llbracket 1, n \rrbracket^2 \quad (AB)_{ij} = \sum_{k=1}^n a_{ik} b_{kj} \quad \text{et} \quad \text{Tr}(AB) = \sum_{i=1}^n \sum_{k=1}^n a_{ik} b_{ki}$$

- $(\cdot | \cdot)$ est bilinéaire. Soient $A, B, C \in \mathcal{M}_n(\mathbb{R})$ et $\lambda \in \mathbb{R}$.

$$(\lambda A + B | C) = \text{Tr}({}^t(\lambda A + B)C) = \lambda \text{Tr}({}^t AC) + \text{Tr}({}^t BC) = \lambda(A | C) + (B | C)$$

par linéarité de la trace; ce qui justifie la linéarité à gauche.

On obtient la linéarité à droite par symétrie.

- $(\cdot | \cdot)$ est symétrique. Soient $A, B \in \mathcal{M}_n(\mathbb{R})$.

$$(A | B) = \text{Tr}({}^t AB) \underset{\text{Tr}({}^t M) = \text{Tr}(M)}{=} \text{Tr}({}^t({}^t AB)) = \text{Tr}({}^t BA) = (B | A)$$

- $(\cdot | \cdot)$ est définie positive. Soit $A \in \mathcal{M}_n(\mathbb{R})$. $(A | A) = \text{Tr}({}^t AA) = \sum_{i=1}^n \sum_{k=1}^n a_{ik}^2 \geq 0$ et :

$$(A | A) = 0 \iff \sum_{i=1}^n \sum_{k=1}^n a_{ik}^2 = 0 \iff \forall (i, k) \in \llbracket 1, n \rrbracket^2 \quad a_{ik}^2 = 0 \iff A = 0_{\mathcal{M}_n(\mathbb{R})}$$

Parmi ces différents espaces, lesquels sont euclidiens? Préciser alors la dimension.

Exercice 1

Montrer que pour $n \in \mathbb{N}^*$, $(P, Q) \mapsto \sum_{k=0}^n P(k)Q(k)$ définit un produit scalaire sur $\mathbb{R}_n[X]$.

Exercice 2

Soit $\mathcal{L}^2(I)$ l'ensemble des fonctions définies sur un intervalle I , à valeurs dans \mathbb{R} et de carré intégrable.

1. Soient $f, g \in \mathcal{L}^2(I)$. En utilisant l'ingéualité, $(f(t) - g(t))^2 \geq 0$, montrer que $(f + g)^2$ est intégrable sur I .
2. En déduire que $\mathcal{L}^2(I)$ possède une structure d'espace vectoriel.
3. Montrer que l'application $(f, g) \mapsto \int_I f(t)g(t) dt$ définit un produit scalaire sur $\mathcal{L}^2(I)$.

Exercice 3

Faire de même avec $\ell^2(\mathbb{N}) = \{(u_n) \in \mathbb{R}^{\mathbb{N}} \mid \sum u_n^2 \text{ converge}\}$ muni de $(u, v) \mapsto \sum_{n=0}^{+\infty} u_n v_n$.

B – Norme euclidienne**Définition 12.3 : Norme euclidienne et distance**

Soit $(E, (\cdot | \cdot))$ un espace préhilbertien réel.

On appelle norme (euclidienne) sur E l'application $\|\cdot\| : E \rightarrow \mathbb{R}_+$ définie par :

$$\forall x \in E \quad \|x\| = \sqrt{(x | x)}$$

On appelle alors distance de x à y le réel positif $d(x, y) = \|x - y\|$ pour $x, y \in E$.

Si $x \neq 0_E$, $\frac{x}{\|x\|}$ est de norme 1, il est dit unitaire.

Identités remarquables vérifiées par la norme euclidienne :

Quels que soient $x, y \in E$,

- $\|x + y\|^2 = \|x\|^2 + \|y\|^2 + 2(x|y)$;
- $\|x - y\|^2 = \|x\|^2 + \|y\|^2 - 2(x|y)$;
- Identité du parallélogramme : $\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2$;
- Identité de polarisation : $(x|y) = \frac{1}{4}(\|x + y\|^2 - \|x - y\|^2)$.

Théorème 12.4 : Inégalité de Cauchy-Schwarz

Soit $(E, (\cdot|\cdot))$ un espace préhilbertien réel. On a alors :

$$\forall x, y \in E \quad |(x|y)| \leq \|x\| \cdot \|y\|$$

Il y a égalité si et seulement si x et y sont colinéaires.

Démonstration 1

Soient $x, y \in E$ et $\lambda \in \mathbb{R}$.

- Si $x = 0_E$, le résultat est immédiat, y compris le cas d'égalité. Supposons désormais $x \neq 0_E$.
- $(\lambda x + y|\lambda x + y) \geq 0$ et $(\lambda x + y|\lambda x + y) = \lambda^2(x|x) + 2\lambda(x|y) + (y|y)$.
C'est un trinôme en λ de signe constant donc son discriminant Δ est négatif ou nul.

$$\Delta = (2(x|y))^2 - 4(x|x)(y|y) = 4((x|y)^2 - (x|x)(y|y)) \leq 0$$

Ainsi, $|(x|y)| \leq \sqrt{(x|x)}\sqrt{(y|y)} = \|x\| \cdot \|y\|$.

- Cas d'égalité :

$\Delta = 0$ donc il existe une racine double notée λ_0 vérifiant :

$$(\lambda_0 x + y|\lambda_0 x + y) = 0$$

On a donc $\lambda_0 x + y = 0$ ce qui signifie $y = -\lambda_0 x$. ■

Démonstration 2

Soient $x, y \in E$.

- Si l'un des deux vecteurs est nul, le résultat est trivial, y compris le cas d'égalité.
- Sinon, posons $\varepsilon = 1$ si $(x|y) \geq 0$, -1 sinon.

$$0 \leq \left\| \frac{x}{\|x\|} - \varepsilon \frac{y}{\|y\|} \right\|^2 = \left\| \frac{x}{\|x\|} \right\|^2 + \left\| \frac{y}{\|y\|} \right\|^2 - 2 \frac{\varepsilon \times (x, y)}{\|x\| \cdot \|y\|} = 2 \left(1 - \frac{|(x, y)|}{\|x\| \cdot \|y\|} \right)$$

On retrouve bien $|(x|y)| \leq \|x\| \cdot \|y\|$.

- Cas d'égalité : $\frac{x}{\|x\|} = \varepsilon \frac{y}{\|y\|}$. ■

Exemple

Soit $P \in \mathbb{R}[X]$. Montrer que $\int_0^1 P(t) dt \leq \sqrt{\int_0^1 P^2(t) dt}$.

On pose $E = \mathbb{R}[X]$ et $\langle P|Q \rangle = \int_0^1 P(t)Q(t) dt$. On applique l'inégalité de Cauchy-Schwarz avec $Q = 1$.

Exercice 4

Soient $a, b \in \mathbb{R}$ avec $a < b$. Montrer que si $f \in \mathcal{C}^1([a, b]; \mathbb{R})$, alors :

$$\left(\int_a^b f^2(t) dt \right) \cdot \left(\int_a^b f'^2(t) dt \right) \geq \left(\frac{f(b)^2 - f(a)^2}{2} \right)^2$$

Théorème 12.5 : Norme

L'application $\|\cdot\|$ est une norme sur E .

Démonstration

Soient $x, y \in E$ et $\lambda \in \mathbb{R}$.

- $\|x\| = (x|x) \geq 0$ donc $\|\cdot\|$ est bien à valeurs dans \mathbb{R}_+ .
- $\|x\| = 0 \iff (x|x) = 0 \iff x = 0$ et $\|\lambda x\| = \sqrt{(\lambda x|\lambda x)} = \sqrt{\lambda^2(x|x)} = |\lambda| \cdot \|x\|$.
- D'après l'inégalité de Cauchy-Schwarz, $|(x|y)| \leq \|x\| \cdot \|y\|$. Donc :

$$\|x + y\|^2 = (x + y|x + y) = \|x\|^2 + \|y\|^2 + 2(x|y) \leq \|x\|^2 + \|y\|^2 + 2 \cdot \|x\| \cdot \|y\| = (\|x\| + \|y\|)^2$$

D'où $\|x + y\| \leq \|x\| + \|y\|$. ■

II | Orthogonalité

On considère un espace préhilbertien réel $(E, (\cdot|\cdot))$.

A – Vecteurs orthogonaux**Définition 12.6**

Deux vecteurs x et y de E sont dits orthogonaux si $(x|y) = 0$.

Exemples

- $E = \mathbb{R}^3, (x|y) = \sum_{i=1}^3 x_i y_i$. Les vecteurs $x = (1, 0, 2)$ et $y = (2, 1, -1)$ sont orthogonaux.
- $E = \mathcal{C}([0, 2\pi], \mathbb{R}), \langle f, g \rangle = \frac{1}{2\pi} \int_0^{2\pi} f(t)g(t) dt$.
Comme $\langle \cos, \sin \rangle = 0$, les vecteurs \cos et \sin sont orthogonaux.

Théorème 12.7 : Pythagore

Soient $x, y \in E$. $\|x + y\|^2 = \|x\|^2 + \|y\|^2 \iff (x|y) = 0$.

Démonstration

$\|x + y\|^2 = \|x\|^2 + \|y\|^2 + 2(x|y)$ donc $\|x + y\|^2 = \|x\|^2 + \|y\|^2 \iff (x|y) = 0$. ■

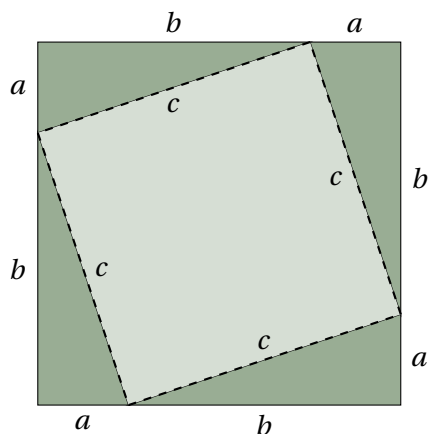


Illustration du théorème de Pythagore

L'aire du grand carré est égale à la somme de l'aire du petit carré et de l'aire des quatre triangles rectangles. Ainsi,

$$(a + b)^2 = c^2 + 4 \times \frac{ab}{2}$$

On trouve donc après simplification :

$$a^2 + b^2 = c^2$$

Théorème 12.8

Le vecteur nul est le seul vecteur orthogonal à tous les autres.

Démonstration

Considérons un vecteur x orthogonal à tous les autres, c'est-à-dire que :

$$\forall y \in E \quad (x|y) = 0$$

Il est en particulier orthogonal à lui-même, donc $(x|x) = \|x\|^2 = 0$. On a bien $x = 0_E$. ■

B – Familles orthogonales et orthonormales

Soit I un ensemble d'indices fini ou infini.

Définition 12.9 : Familles orthogonales et orthonormales

- Une famille de vecteurs $(e_i)_{i \in I}$ de E est dite orthogonale si :

$$\forall (i, j) \in I^2, i \neq j \implies (e_i|e_j) = 0.$$

- Elle est dite orthonormale si elle vérifie de plus : $\forall i \in I, \|e_i\| = 1$.

Proposition 12.10 : Pythagore « généralisé »

Soit (x_1, \dots, x_n) une famille orthogonale de vecteurs de E . Alors, $\left\| \sum_{i=1}^n x_i \right\|^2 = \sum_{i=1}^n \|x_i\|^2$.

Théorème 12.11

Une famille orthogonale constituée de vecteurs non nuls est libre. En particulier, toute famille orthonormale est libre.

Démonstration

Démontrons ce résultat dans le cas d'une famille finie (e_1, \dots, e_n) de vecteurs.

- Soient $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ tels que $\lambda_1 e_1 + \dots + \lambda_n e_n = 0_E$. Ainsi, quel que soit $j \in \llbracket 1, n \rrbracket$,

$$\left(\sum_{i=1}^n \lambda_i e_i | e_j \right) = \sum_{i=1}^n \lambda_i (e_i | e_j) = \lambda_j \|e_j\|^2 = 0$$

Le vecteur e_j étant non nul, $\lambda_j = 0$. Et ceci, pour tout $j \in \llbracket 1, n \rrbracket$. La famille est bien libre.

- Une famille orthonormale est orthogonale et ses vecteurs sont unitaires donc non nuls. ■

Une famille orthonormale contient donc au plus $\dim(E)$ vecteurs si E est de dimension finie. Si elle en contient précisément $\dim(E)$, c'est une base. On la qualifie de *base orthonormale* ou de *base orthonormée*.

Théorème 12.12 : Décomposition dans une base orthonormée

Soient E un espace euclidien de dimension $n \in \mathbb{N}^*$ et (e_1, \dots, e_n) une base orthonormée de E .

$$\forall x \in E, x = (x|e_1)e_1 + \dots + (x|e_n)e_n = \sum_{i=1}^n (x|e_i)e_i$$

Autrement dit, les coordonnées de x dans la base (e_1, \dots, e_n) sont $((x|e_i))_{1 \leq i \leq n}$.

Démonstration

Soient $\mathcal{B} = (e_1, \dots, e_n)$ une base orthonormale de E et $x \in E$.

Il existe $x_1, \dots, x_n \in \mathbb{R}$ tels que $x = \sum_{i=1}^n x_i e_i$ donc pour tout $j \in \llbracket 1, n \rrbracket$,

$$(x|e_j) = \left(\sum_{i=1}^n x_i e_i | e_j \right) = \sum_{i=1}^n x_i (e_i | e_j) = \sum_{i=1}^n x_i \delta_{ij} = x_j$$

Ainsi, $x = (x|e_1)e_1 + \dots + (x|e_n)e_n = \sum_{i=1}^n (x|e_i)e_i$. ■

Proposition 12.13

Soient $\mathcal{B} = (e_1, \dots, e_n)$ une base orthonormale de E . On considère $x, y \in E$ de coordonnées respectives $X = (x_1, \dots, x_n)$ et $Y = (y_1, \dots, y_n)$. On a alors :

$$(x|y) = \sum_{i=1}^n x_i y_i = \sum_{i=1}^n (x|e_i)(y|e_i) = {}^t X Y \quad \text{et} \quad \|x\|^2 = \sum_{i=1}^n x_i^2 = \sum_{i=1}^n (x|e_i)^2 = {}^t X X$$

Ce dernier résultat montre que finalement, tous les produits scalaires se ramènent au produit scalaire canonique de \mathbb{R}^n via le choix d'une base orthonormale. Mais tout espace euclidien possède-t-il une base orthonormale ?

Réponse : oui ! Tout espace vectoriel de dimension finie (donc tout espace euclidien) possède une base. Dans le cas d'un espace euclidien, on peut même construire une base orthonormée à l'aide du *procédé* ou *algorithme d'orthonormalisation de Gram-Schmidt*. Ceci nous assure l'existence d'une base orthonormale.

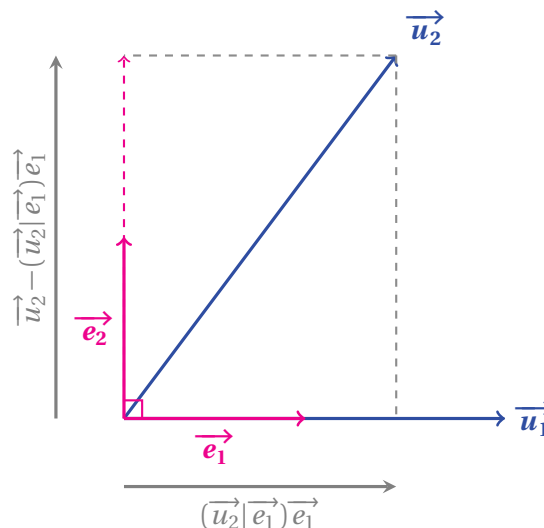
Le procédé d'orthonormalisation repose sur l'idée fondamentale suivante :

On considère une famille libre (\vec{u}_1, \vec{u}_2) de \mathbb{R}^2 .

- Commençons par poser $\vec{e}_1 = \frac{\vec{u}_1}{\|\vec{u}_1\|}$ pour obtenir un vecteur unitaire.
- On retranche ensuite à \vec{u}_2 sa composante suivant \vec{e}_1 .
On obtient alors un vecteur $\vec{e}_2' = \vec{u}_2 - (\vec{u}_2 | \vec{e}_1) \vec{e}_1$ orthogonal à \vec{e}_1 . Il ne reste plus qu'à le diviser par sa norme pour obtenir un vecteur unitaire :

$$\vec{e}_2 = \frac{\vec{u}_2 - (\vec{u}_2 | \vec{e}_1) \vec{e}_1}{\|\vec{u}_2 - (\vec{u}_2 | \vec{e}_1) \vec{e}_1\|}$$

- La famille (\vec{e}_1, \vec{e}_2) obtenue est orthonormale.

**Théorème 12.14**

Soit $n \in \mathbb{N}^*$ et (u_1, \dots, u_n) une famille libre de vecteurs de E . Il existe alors une famille orthonormale (e_1, \dots, e_n) de E telle que :

$$\text{Vect}(e_1, \dots, e_n) = \text{Vect}(u_1, \dots, u_n)$$

Démonstration

Démontrons ce résultat par récurrence sur n .

- **Initialisation** – La famille (u_1) étant libre, u_1 est non nul. On pose alors $e_1 = \frac{u_1}{\|u_1\|}$.

- **Hérédité** – Supposons la propriété vraie au rang n et montrons qu'elle l'est encore au rang $n + 1$.
 Considérons pour cela la famille $(u_1, \dots, u_n, u_{n+1})$ que l'on suppose libre.
 La famille (u_1, \dots, u_n) étant libre, il existe une famille orthonormale (e_1, \dots, e_n) telle que $\text{Vect}(u_1, \dots, u_n) = \text{Vect}(e_1, \dots, e_n)$. On pose alors :

$$e'_{n+1} = u_{n+1} - \sum_{i=1}^n \lambda_i e_i$$

- (i) On souhaite que la famille (e_1, \dots, e'_{n+1}) soit orthogonale.

$$\forall j \in \llbracket 1, n \rrbracket \quad (e'_{n+1} | e_j) = 0 \iff \forall j \in \llbracket 1, n \rrbracket \quad (u_{n+1} | e_j) - \sum_{i=1}^n \lambda_i (e_i | e_j) = (u_{n+1} | e_j) - \lambda_j = 0$$

On pose donc $\lambda_j = (u_{n+1} | e_j)$ pour tout $j \in \llbracket 1, n \rrbracket$.

- (ii) e'_{n+1} est non nul. Dans le cas contraire, u_{n+1} serait combinaison linéaire de (e_1, \dots, e_n) donc de (u_1, \dots, u_n) . La famille (u_1, \dots, u_{n+1}) ne pourrait être libre ! On peut donc poser $e_{n+1} = \frac{e'_{n+1}}{\|e'_{n+1}\|}$.

La famille (e_1, \dots, e_{n+1}) est alors orthonormale.

- (iii) Enfin, puisque $\text{Vect}(u_1, \dots, u_n) = \text{Vect}(e_1, \dots, e_n)$ et que e_{n+1} est combinaison linéaire de e_1, \dots, e_n et de u_{n+1} , $\text{Vect}(u_1, \dots, u_n, u_{n+1}) = \text{Vect}(e_1, \dots, e_n, e_{n+1})$.

Ceci achève la récurrence. ■

Quelques remarques :

- Une telle famille (e_1, \dots, e_n) est unique à condition que $(u_k, e_k) > 0$ pour tout $k \in \llbracket 1, n \rrbracket$.
- La matrice de passage de la base (u_1, \dots, u_n) à (e_1, \dots, e_n) est triangulaire supérieure.
- On peut normaliser les vecteurs e'_k à chaque étape ou bien normaliser la famille (e'_1, \dots, e'_n) une fois construite.

Théorème 12.15

Tout espace euclidien admet une base orthonormale.

Exercice 5

Montrer que la famille $\mathcal{B} = ((2, 1, 0), (0, 1, 1), (1, 2, 1))$ est une base de \mathbb{R}^3 puis construire une base orthonormée de \mathbb{R}^3 pour le produit scalaire usuel à l'aide du procédé vu précédemment.

C – Orthogonal d'un sous-espace vectoriel

E désigne toujours un espace préhilbertien réel de dimension quelconque.

Définition 12.16 : Orthogonal

Soit F un sous-espace vectoriel de E . On appelle orthogonal de F l'ensemble :

$$F^\perp = \{x \in E \mid \forall y \in F \ (x|y) = 0\}$$

Proposition 12.17

Soient F et G deux sous-espaces vectoriels de E .

- F^\perp est un sous-espace vectoriel de E .
- Si $F \subset G$ alors $G^\perp \subset F^\perp$.

Démonstration

- (i) • Tout d'abord, F^\perp est non vide car il contient le vecteur nul.
 • Soient $x_1, x_2 \in F^\perp$ et $\lambda \in \mathbb{R}$. Alors,

$$\forall y \in F \quad (\lambda x_1 + x_2 | y) = \lambda \underbrace{(x_1 | y)}_{=0} + \underbrace{(x_2 | y)}_{=0} = 0$$

Donc F^\perp est stable par combinaison linéaire.

C'est bien un sous-espace vectoriel de E .

- (ii) Soit $x \in G^\perp$. Pour tout $y \in F$, $(x | y) = 0$ car $y \in G$. Ainsi, $x \in F^\perp$. On a bien $G^\perp \subset F^\perp$. ■

Exercice 6

| Montrer que si E est un espace préhilbertien réel, $\{0_E\}^\perp = E$ et $E^\perp = \{0_E\}$.

Exercice 7

| Soit $F = \text{Vect}((a, b, c)) \subset \mathbb{R}^3$. Déterminer F^\perp et en donner une équation cartésienne.

Attention, dire que deux sous-espaces vectoriels sont orthogonaux ne signifie pas que l'un est l'orthogonal de l'autre. Penser à l'exemple de deux droites orthogonales dans l'espace.

Proposition 12.18

Soient F un sous-espace vectoriel de E et $u \in E$.

$u \in F^\perp$ si et seulement si u est orthogonal aux vecteurs d'une base quelconque de F .

Démonstration

L'implication est immédiate, montrons simplement la réciproque dans le cas d'un espace de dimension finie. Supposons u orthogonal aux vecteurs d'une base (e_1, \dots, e_p) de F . Soit $y \in F$. Il existe donc $(\lambda_1, \dots, \lambda_p) \in \mathbb{R}^p$ tel que $y = \sum_{i=1}^p \lambda_i e_i$. Ainsi, $(u | y) = \left(u \mid \sum_{i=1}^p \lambda_i e_i \right) = \sum_{i=1}^p \lambda_i (u | e_i) = 0$. Donc $u \in F^\perp$. ■

Théorème 12.19

Soit F un sous-espace vectoriel de dimension finie de E . On a $E = F \oplus F^\perp$.

Démonstration

Soit $x \in E$ et (e_1, \dots, e_p) une base orthonormée de F . Raisonnons par analyse/synthèse.

• *Analyse*

On suppose que $x = x_F + x_{F^\perp}$ avec $x_F \in F$ et $x_{F^\perp} \in F^\perp$.

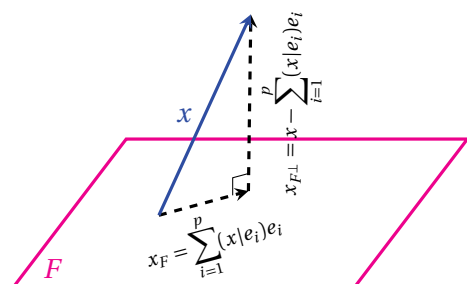
$x_F \in F$ donc $x_F = \sum_{i=1}^p (x_F | e_i) e_i$ et $x_{F^\perp} = x - x_F \in F^\perp$ donc :

$$\forall i \in \llbracket 1, p \rrbracket \quad (x - x_F | e_i) = 0 \quad \text{c'est-à-dire} \quad x_F = \sum_{i=1}^p (x | e_i) e_i$$

• *Synthèse*

On peut écrire :

$$x = \underbrace{\sum_{i=1}^p (x | e_i) e_i}_{\in F} + \left(x - \sum_{i=1}^p (x | e_i) e_i \right)$$



Il reste à montrer que $x - \sum_{i=1}^p (x|e_i)e_i \in F^\perp$, ce qui est bien le cas car :

$$\forall j \in \llbracket 1, p \rrbracket \quad \left(x - \sum_{i=1}^p (x|e_i)e_i \mid e_j \right) = (x|e_j) - \sum_{i=1}^p (x|e_i)(e_i|e_j) = 0$$

Corollaire 12.20 : Inégalité de Bessel

Soient (e_1, \dots, e_p) une famille orthonormale de E et $x \in E$. Alors $\sum_{i=1}^p (x|e_i)^2 \leq \|x\|^2$.

Il y a égalité si et seulement si $x \in \text{Vect}(e_1, \dots, e_p)$.

Corollaire 12.21

Soient E un espace euclidien (donc de dimension finie) et F un sous-espace vectoriel de E . F^\perp est un espace vectoriel de dimension finie et $\dim(F^\perp) = \dim(E) - \dim(F)$. De plus, $(F^\perp)^\perp = F$.

Démonstration

- Si $E = F \oplus G$ et que E est de dimension finie, alors $\dim(E) = \dim(F) + \dim(G)$.
- Il suffit de montrer que $F \subset (F^\perp)^\perp$ puis on conclut par égalité des dimensions.

L'orthogonal d'une droite vectorielle est donc un hyperplan de E .

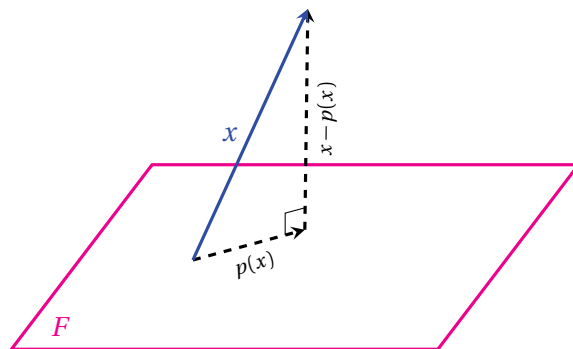
D – Projection orthogonale et distance

E désigne toujours un espace préhilbertien réel.

Définition 12.22 : Projecteur orthogonal

Soit F un sous-espace vectoriel de E de dimension finie. On a $E = F \oplus F^\perp$.

On appelle projection orthogonale sur F la projection sur F parallèlement à F^\perp .



Représentation du projeté orthogonal de x sur F

Théorème 12.23

En notant p la projection orthogonale sur F , sous-espace vectoriel de E de dimension finie n ,

- Si $x \in E$, $p(x)$ est entièrement caractérisé par : $p(x) \in F$ et $x - p(x) \in F^\perp$.
- Si (e_1, \dots, e_n) est une base orthonormale de F alors $p(x) = (x|e_1)e_1 + \dots + (x|e_n)e_n$.

Démonstration

Redémontrons rapidement le deuxième point.

$$\forall i \in \llbracket 1, n \rrbracket \quad (x|e_i) = (p(x) + (x - p(x))|e_i) = (p(x)|e_i) + (x - p(x)|e_i) = (p(x)|e_i)$$

On retrouve donc le fait que $p(x) = \sum_{i=1}^n (x|e_i)e_i$. ■

Exercice 8

| Donner la matrice dans la base canonique de la projection orthogonale sur le plan d'équation $x + y + z = 0$.

Exercice 9

Soient E un espace préhilbertien réel et p la projection orthogonale sur une droite vectorielle D de E .

- Pour $x \in E$, exprimer $p(x)$ en fonction de x .
- Même question lorsque p est la projection orthogonale sur un hyperplan H de E .

Définition 12.24 : Distance

Soient $x \in E$ et F un sous-espace vectoriel de E de dimension finie.

On appelle distance de x à F le réel $d(x, F) = \inf_{u \in F} d(x, u) = \inf_{u \in F} \|x - u\|$.

Intuitivement, la distance de x à F est la plus petite des distances entre x et les vecteurs de F . Cependant, rien ne nous garantit l'existence d'une *distance minimale*. Noter que la définition a bien un sens car $\{\|x - u\| \mid u \in F\}$ est une partie de \mathbb{R} non vide et minorée; elle admet une borne inférieure.

Théorème 12.25

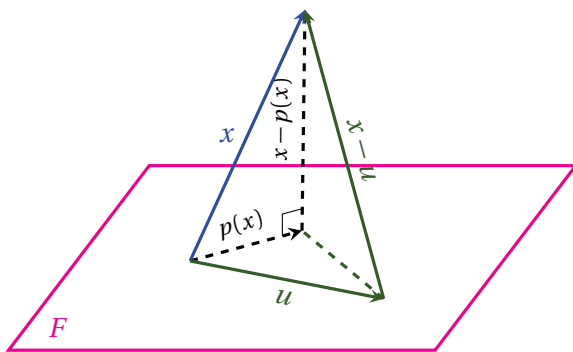
Soient $x \in E$ et F un sous-espace vectoriel de E de dimension finie.

$d(x, F) = \inf_{u \in F} \|x - u\| = \|x - p(x)\|$ où p est la projection orthogonale sur F .

La distance est donc un minimum qui est atteint pour $u = p(x)$.

Démonstration

On peut à nouveau s'appuyer sur une simple figure.



Soit $u \in F$.

D'après le théorème de Pythagore,

$$\|x - u\|^2 = \underbrace{\|x - p(x)\|^2}_{\in F^\perp} + \underbrace{\|u - p(x)\|^2}_{\in F}$$

On a donc $\|x - u\| \geq \|x - p(x)\|$ et la borne inférieure est atteinte pour $u = p(x) \in F$.

La borne inférieure est un minimum. ■

Exercice 10

| Déterminer la distance du vecteur $u = (1, 2, 3)$ au plan de \mathbb{R}^3 d'équation $x + y + z = 0$.

III | Suites totales (*)

$(E, (\cdot|\cdot))$ désigne toujours un espace préhilbertien réel. Étant donné une famille orthonormale de vecteurs $(e_i)_{i \in \mathbb{N}}$, on cherche à généraliser l'expression $x = \sum_{i=1}^n (x|e_i)e_i$ valable en dimension finie.

Définition 12.26 : Suite totale

On dit qu'une suite de vecteurs $(e_i)_{i \in \mathbb{N}}$ de E est totale si pour tout $x \in E$, on a :

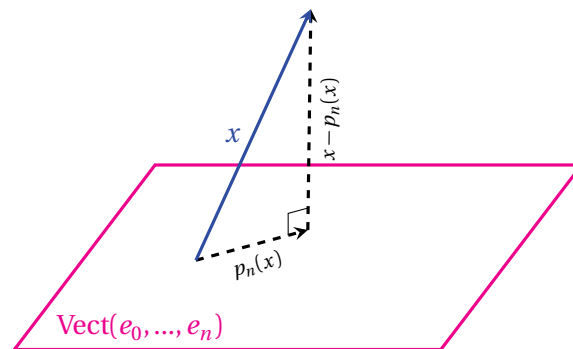
$$\forall \varepsilon > 0, \quad \exists y \in \text{Vect}(e_i)_{i \in \mathbb{N}}, \quad \|x - y\| < \varepsilon$$

Par caractérisation séquentielle de la limite, cela revient à dire qu'il existe une suite $(x_n)_{n \in \mathbb{N}}$ d'éléments de $\text{Vect}(e_i)_{i \in \mathbb{N}}$ tel que $(x_n)_{n \in \mathbb{N}}$ converge vers x .

Autrement dit, comme nous le verrons dans un prochain chapitre, $\text{Vect}(e_i)_{i \in \mathbb{N}}$ est *dense* dans E .

Théorème 12.27

Soient $(e_i)_{i \in \mathbb{N}}$ une suite orthonormale totale d'éléments de E et pour tout $n \in \mathbb{N}$, p_n le projecteur orthogonal sur $\text{Vect}(e_0, \dots, e_n)$. Alors, pour tout x de E , $(p_n(x))_{n \in \mathbb{N}}$ converge vers x .



Représentation du projeté orthogonal de x sur $\text{Vect}(e_0, \dots, e_n)$

Démonstration

Soient $x \in E$ et $(e_i)_{i \in \mathbb{N}}$ une suite orthonormale totale d'éléments de E .

On pose pour $n \in \mathbb{N}$, $F_n = \text{Vect}(e_i)_{0 \leq i \leq n}$ et l'on cherche à montrer que $\|x - p_n(x)\| \xrightarrow{n \rightarrow +\infty} 0$. et l'on fixe $\varepsilon > 0$.

Par définition d'une famille totale, il existe $y \in \text{Vect}(e_i)_{i \in \mathbb{N}}$ tel que $\|x - y\| < \varepsilon$.

Le vecteur y étant combinaison linéaire d'un nombre nécessairement fini de vecteurs e_i , il existe $n_0 \in \mathbb{N}$ tel que $y \in F_{n_0}$. Remarquons qu'alors, pour tout $n \geq n_0$, $y \in F_n$ et donc,

$$\|x - p_n(x)\| = \inf_{u \in F_n} \|x - u\| \leq \|x - y\| \leq \varepsilon$$

On vient finalement d'établir que pour tout $x \in E$, $x = \lim_{n \rightarrow +\infty} p_n(x) = \lim_{n \rightarrow +\infty} \sum_{i=0}^n (x|e_i)e_i = \sum_{i=0}^{+\infty} (x|e_i)e_i$.

On obtient alors la version « complète » de l'inégalité de Bessel.

Corollaire 12.28 : Égalité de Parseval

Si $(e_i)_{i \in \mathbb{N}}$ une suite orthonormale totale d'éléments de E , alors :

$$\forall x \in E, \quad \|x\|^2 = \sum_{i=0}^{+\infty} (x|e_i)^2$$

Démonstration

Le théorème de Pythagore va encore une fois venir à notre rescousse. Il suffit d'écrire pour tout $x \in E$:

$$\|x\|^2 = \|p_n(x)\|^2 + \|x - p_n(x)\|^2 = \sum_{i=1}^n (x|e_i)^2 + \|x - p_n(x)\|^2 \xrightarrow{n \rightarrow +\infty} \sum_{i=0}^{+\infty} (x|e_i)^2$$

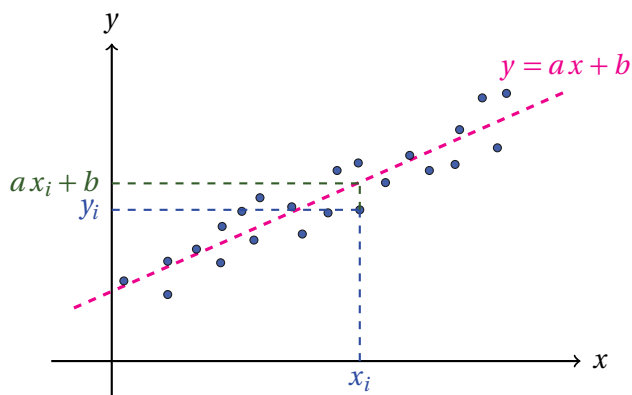
Des applications de ces résultats seront traitées en TD.

IV | Méthode des moindres carrés (complément)

Il est courant, en physique-chimie, en sciences industrielles, ou plus généralement dans toute discipline expérimentale (biologie, chimie, économie, ...), d'avoir à comparer des données expérimentales et de conjecturer une éventuelle dépendance linéaire entre deux paramètres donnés (par exemple entre l'allongement d'un ressort et la force de traction exercée sur celui-ci).

Supposons que l'on dispose d'une série de n mesures de la forme (x_i, y_i) avec $i \in \llbracket 1, n \rrbracket$. On cherche à trouver « la droite de meilleure approximation » de nos mesures, c'est-à-dire la droite qui décrit au mieux la tendance du nuage observé. C'est le principe de régression linéaire.

Mais quel sens donner à cette fameuse « droite de meilleure approximation » ?



Si la droite recherchée a pour équation d'équation $y = ax + b$, l'écart ponctuel entre la mesure obtenue (x_i, y_i) et la mesure attendue $(x_i, ax_i + b)$ vaut $|y_i - ax_i - b|$. On peut dès lors chercher à minimiser l'écart global entre les points et la droite, écart qui peut être défini de différentes façons. Par exemple,

$$\max_{1 \leq i \leq n} |y_i - ax_i - b|; \quad \sum_{i=1}^n |y_i - ax_i - b|; \quad \sum_{i=1}^n (y_i - ax_i - b)^2$$

C'est cette dernière quantité que l'on souhaite minimiser dans la méthode dite des moindres carrés.

On peut déterminer $\inf_{(a,b) \in \mathbb{R}^2} \sum_{i=1}^n (y_i - ax_i - b)^2$ en l'interprétant comme la distance d'un vecteur à un certain sous-espace vectoriel de \mathbb{R}^n .

Posons $X = (x_1, \dots, x_n)$, $Y = (y_1, \dots, y_n)$, $Z = aX + b = (ax_1 + b, \dots, ax_n + b)$ et $F = \text{Vect}(X, (1, \dots, 1))$. On a donc

$$\inf_{(a,b) \in \mathbb{R}^2} \sum_{i=1}^n (y_i - ax_i - b)^2 = \inf_{(a,b) \in \mathbb{R}^2} \|Y - Z\|^2 = \inf_{Z \in F} \|Y - Z\|^2 = d^2(Y, F).$$

D'après ce qui précède, $d^2(Y, F)$ vaut $\|Y - p(Y)\|^2$ où p est la projection orthogonale sur F .

On a donc $Y - Z = Y - p(Y) \in F^\perp$ c'est-à-dire :

$$\begin{cases} (Y - Z | X) = 0 \\ (Y - Z | (1, \dots, 1)) = 0 \end{cases}$$

Cela nous conduit à résoudre le système suivant d'inconnues a et b :

$$\begin{cases} \sum_{i=1}^n (y_i - ax_i - b)x_i = 0 \\ \sum_{i=1}^n (y_i - ax_i - b) = 0 \end{cases}$$

On obtient après simplification le système linéaire 2×2 suivant :

$$\begin{cases} a \sum_{i=1}^n x_i^2 + b \sum_{i=1}^n x_i = \sum_{i=1}^n x_i y_i \\ a \sum_{i=1}^n x_i + nb = \sum_{i=1}^n y_i \end{cases}$$

On obtient ainsi les coefficients a et b recherchés.

Une petite mise en garde cependant, rien ne nous garantit que la loi étudiée est linéaire !