14 Espaces préhilbertiens réels

« Ce qui est simple est toujours faux. Ce qui ne l'est pas est inutilisable. »
Paul Valéry (1871 – 1945)

Plan de cours

Dans tout ce chapitre, E désigne un $\mathbb{R}$-espace vectoriel de dimension quelconque.

I | Généralités

A – Produit scalaire

Définition 14.1 : Produit scalaire

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

- φ est *bilinéaire* : pour tous $x_1, x_2, y_1, y_2 \in E$ et $\lambda \in \mathbb{R}$,

$$\varphi(\lambda x_1 + x_2, y_1) = \lambda\varphi(x_1, y_1) + \varphi(x_2, y_1) \quad \text{et} \quad \varphi(x_1, \lambda y_1 + y_2) = \lambda\varphi(x_1, y_1) + \varphi(x_1, y_2)$$

- φ est *symétrique* : pour tous $x, y \in E$, $\varphi(x, y) = \varphi(y, x)$.
- φ est *définie positive* : pour tout $x \in E$, $\varphi(x, x) \geqslant 0$ 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 14.2 : 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 de son produit scalaire canonique

Le produit scalaire canonique est défini par :

$$\forall X, Y \in \mathbb{R}^n, \quad (X|Y) = X^\top Y = \sum_{i=1}^{n} x_i y_i \text{ en notant } X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \text{ et } Y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$$

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

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

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

Si $f,g\in\mathscr{C}([a,b],\mathbb{R})$, l'intégrale existe par continuité de fg sur le segment $[a,b]$. $(\cdot|\cdot)$ est à valeurs dans $\mathbb{R}$.

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

$$(\lambda f+g|h)=\int_a^b(\lambda f(t)+g(t))h(t)\,\mathrm{d}t=\lambda\int_a^b f(t)h(t)\,\mathrm{d}t+\int_a^b g(t)h(t)\,\mathrm{d}t=\lambda(f|h)+(g|h)$$

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

$$(f|f)=0\iff\int_a^b f^2(t)\,\mathrm{d}t=0\underset{\substack{f^2\text{ est continue}\\\text{et positive sur }[a,b]}}{\iff}\forall t\in[a,b],\quad f^2(t)=0\quad\iff f\text{ est nulle sur }[a,b]$$

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

Là aussi, l'intégrale est bien définie. $(\cdot|\cdot)$ est à valeurs dans $\mathbb{R}$ et de plus,

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

$$(\lambda P+Q|R)=\int_0^1(\lambda P(t)+Q(t))R(t)\,\mathrm{d}t=\lambda\int_0^1 P(t)R(t)\,\mathrm{d}t+\int_0^1 Q(t)R(t)\,\mathrm{d}t=\lambda(P|R)+(Q|R)$$

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

$$(P|P)=0\iff\int_0^1 P^2(t)\,\mathrm{d}t=0\underset{\substack{P^2\text{ est continue}\\\text{et positive sur }[0,1]}}{\iff}\forall t\in[0,1],\quad P^2(t)=0\underset{\substack{P\text{ admet une}\\\text{infinité de racines}}}{\iff}P=0_{\mathbb{R}[X]}$$

Exemple 4 – $E=\mathscr{M}_n(\mathbb{R})$ muni de $(A,B)\mapsto\mathrm{Tr}(A^\top B)$

Rappelons tout d'abord que : $\quad\forall(i,j)\in[\![1,n]\!]^2,\quad(AB)_{ij}=\displaystyle\sum_{k=1}^n a_{ik}b_{kj}\quad$ et $\quad\mathrm{Tr}(AB)=\displaystyle\sum_{i=1}^n\sum_{k=1}^n a_{ik}b_{ki}$.

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

$$(\lambda A+B|C)=\mathrm{Tr}((\lambda A+B)^\top C)=\lambda\,\mathrm{Tr}(A^\top C)+\mathrm{Tr}(B^\top C)=\lambda(A|C)+(B|C)$$

  par linéarité de la trace ; ce qui justifie la linéarité à gauche. La linéarité à droite est obtenue par symétrie.
- $(\cdot|\cdot)$ est symétrique. Soient $A,B\in\mathscr{M}_n(\mathbb{R})$.

$$(A|B)=\mathrm{Tr}(A^\top B)\underset{\mathrm{Tr}(M^\top)=\mathrm{Tr}(M)}{=}\mathrm{Tr}((A^\top B)^\top)=\mathrm{Tr}(B^\top A)=(B|A)$$

- $(\cdot|\cdot)$ est définie positive. Soit $A\in\mathscr{M}_n(\mathbb{R})$. $(A|A)=\mathrm{Tr}(A^\top A)=\displaystyle\sum_{i=1}^n\sum_{k=1}^n a_{ik}^2\geqslant 0$ et :

$$(A|A)=0\iff\sum_{i=1}^n\sum_{k=1}^n a_{ik}^2=0\iff\forall(i,k)\in[\![1,n]\!]^2,\ a_{ik}^2=0\iff A=0_{\mathscr{M}_n(\mathbb{R})}$$

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

1. Montrer que $\mathscr{L}^2(I)$ possède une structure d'espace vectoriel.
2. Montrer que l'application $(f,g) \mapsto \int_I f(t)g(t)\,\mathrm{d}t$ définit un produit scalaire sur $\mathscr{L}^2(I)$.

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

B – Norme euclidienne

Définition 14.3 : Norme euclidienne et distance

Soit $(E,(\cdot|\cdot))$ un espace préhilbertien réel.
On appelle norme (euclidienne) sur E l'application $\|\cdot\| : E \to \mathbb{R}_+$ 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 $x \neq 0_E$, $\dfrac{x}{\|x\|}$ est de norme 1, il est dit unitaire.

Proposition 14.4 : Identités remarquables

Soit $(E,(\cdot|\cdot))$ un espace préhilbertien réel. Pour tous $x,y \in E$,

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

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

(iii) Identité du parallélogramme : $\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2$;

(iv) Identité de polarisation : $(x|y) = \dfrac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2\right)$.

Théorème 14.5 : Inégalité de Cauchy-Schwarz

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

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

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 $x \neq 0_E$.
- $(\lambda x + y|\lambda x + y) \geqslant 0$ 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.

  $$\Delta = (2(x|y))^2 - 4(x|x)(y|y) = 4\left((x|y)^2 - (x|x)(y|y)\right) \leqslant 0$$

  Ainsi, $|(x|y)| \leqslant \sqrt{(x|x)}\sqrt{(y|y)} = \|x\| \cdot \|y\|$.
- Cas d'égalité : $\Delta = 0$ donc il existe une racine double notée λ_0 vérifiant $(\lambda_0 x + y|\lambda_0 x + y) = 0$.
  D'où $\lambda_0 x + y = 0$, soit $y = -\lambda_0 x$. ■

Démonstration 2

Soient $x, y \in E$.

- Si l'un des deux vecteurs est nul, le résultat est immédiat, y compris le cas d'égalité.
- Sinon, posons $x' = \dfrac{x}{\|x\|}$, $y' = \dfrac{y}{\|y\|}$ et enfin, $\varepsilon = \begin{cases} 1 \text{ si } (x|y) \geqslant 0 \\ -1 \text{ sinon} \end{cases}$

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

On retrouve bien $|(x|y)| \leqslant \|x\| \cdot \|y\|$. Le cas d'égalité est clair : $\dfrac{x}{\|x\|} = \varepsilon \dfrac{y}{\|y\|}$. ■

Exemple

Soit $P \in \mathbb{R}[X]$. Montrer que $\displaystyle\int_0^1 P(t)\,\mathrm{d}t \leqslant \sqrt{\int_0^1 P^2(t)\,\mathrm{d}t}$.

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

Exercice 3

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

$$\left(\int_a^b f^2(t)\,\mathrm{d}t\right) \cdot \left(\int_a^b f'^2(t)\,\mathrm{d}t\right) \geqslant \left(\frac{f(b)^2 - f(a)^2}{2}\right)^2$$

Théorème 14.6 : Norme

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

Démonstration

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

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

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

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

II | Orthogonalité

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

A – Vecteurs orthogonaux

Définition 14.7

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

Exemples

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

Théorème 14.8 : 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 $\|x+y\|^2 = \|x\|^2 + \|y\|^2 \iff (x|y) = 0$. ■

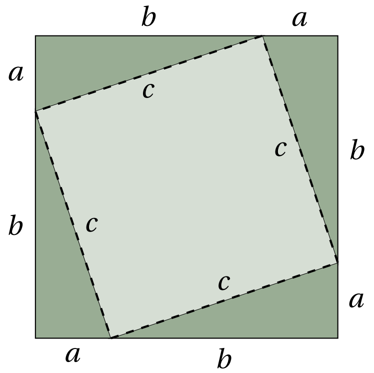


Illustration du théorème de Pythagore

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 14.9

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 $(x|x) = \|x\|^2 = 0$. Ainsi, $x = 0_E$. ■

B – Familles orthogonales et orthonormales

Définition 14.10 : Familles orthogonales et orthonormales

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

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

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

Proposition 14.11 : Pythagore « généralisé »

Soit $(e_1, \dots, e_n)$ une famille orthogonale de vecteurs de E. Alors, $\left\| \sum_{i=1}^{n} e_i \right\|^2 = \sum_{i=1}^{n} \|e_i\|^2$.

Théorème 14.12

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 $j \in [\![1, n]\!]$,

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

  Le vecteur e_j étant non nul, $\lambda_j = 0$. Et ceci, pour tout $j \in [\![1, n]\!]$. 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 14.13 : 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.

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

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

Démonstration

Soient $\mathscr{B} = (e_1, \dots, e_n)$ 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 $j \in [\![1, n]\!]$,

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

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

Proposition 14.14

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

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

Ce dernier résultat montre qu'en dimension finie, 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.

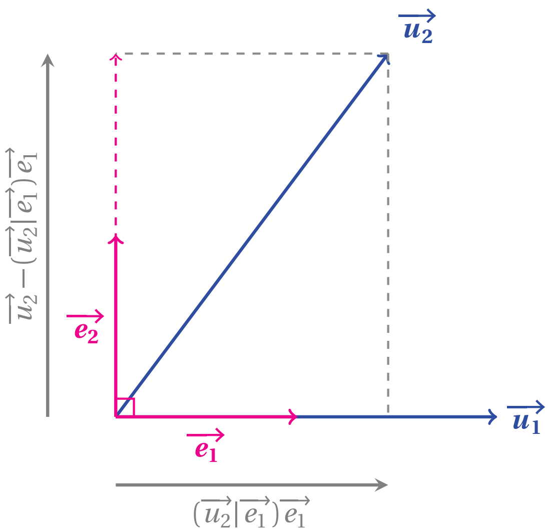
Le procédé d'orthonormalisation repose sur l'idée fondamentale suivante :

On considère une famille libre $(\overrightarrow{u_1}, \overrightarrow{u_2})$ de $\mathbb{R}^2$.

- Commençons par poser $\overrightarrow{e_1} = \dfrac{\overrightarrow{u_1}}{\|\overrightarrow{u_1}\|}$ pour obtenir un vecteur unitaire.
- On retranche ensuite à $\overrightarrow{u_2}$ sa composante suivant $\overrightarrow{e_1}$. On obtient alors un vecteur $\overrightarrow{e_2'} = \overrightarrow{u_2} - (\overrightarrow{u_2}|\overrightarrow{e_1})\overrightarrow{e_1}$ orthogonal à $\overrightarrow{e_1}$. Il ne reste plus qu'à le diviser par sa norme pour obtenir un vecteur unitaire :

$$\overrightarrow{e_2} = \frac{\overrightarrow{u_2} - (\overrightarrow{u_2}|\overrightarrow{e_1})\overrightarrow{e_1}}{\|\overrightarrow{u_2} - (\overrightarrow{u_2}|\overrightarrow{e_1})\overrightarrow{e_1}\|}$$

La famille $(\overrightarrow{e_1}, \overrightarrow{e_2})$ obtenue est orthonormale.



Théorème 14.15

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

$$\mathrm{Vect}(e_1, \ldots, e_n) = \mathrm{Vect}(u_1, \ldots, u_n)$$

Démonstration

Démontrons ce résultat par récurrence sur n.

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

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

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

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

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

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

(iii) Enfin, puisque $\mathrm{Vect}(u_1, \ldots, u_n) = \mathrm{Vect}(e_1, \ldots, e_n)$ et que e_{n+1} est combinaison linéaire de $e_1, \ldots, e_n$ et de u_{n+1}, $\mathrm{Vect}(u_1, \ldots, u_n, u_{n+1}) = \mathrm{Vect}(e_1, \ldots, e_n, e_{n+1})$. Ceci achève la récurrence. ■

Quelques remarques :

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

Théorème 14.16

Tout espace euclidien admet une base orthonormale.

Exercice 4

Montrer que la famille $\mathscr{B} = ((2,1,0), (0,1,1), (1,2,1))$ 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.

Exercice 5

Soit E un espace vectoriel muni d'une base $(e_1, \ldots, e_n)$. Montrer qu'il existe un unique produit scalaire sur E tel que cette base est orthonormale.

C – Orthogonal d'un sous-espace vectoriel

E désigne toujours un espace préhilbertien réel de dimension quelconque.

Définition 14.17 : 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 14.18

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

(i) $F^\perp$ est un sous-espace vectoriel de E.

(ii) Si $F \subset G$, alors $G^\perp \subset F^\perp$. De plus, $F \subset (F^\perp)^\perp$.

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

Donc $F^\perp$ est stable par combinaison linéaire.

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

Assez facilement, $\{0_E\}^\perp = E$ et $E^\perp = \{0_E\}$. 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.

Exercice 6

Soit $u \in E$ non nul. Montrer que $\mathrm{Vect}(u)^\perp$ est un hyperplan et en donner une équation.

Proposition 14.19

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, $(u|y) = \left(u \middle| \sum_{i=1}^{p} \lambda_i e_i\right) = \sum_{i=1}^{p} \lambda_i (u|e_i) = 0$. Donc $u \in F^\perp$. ■

Théorème 14.20

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

Démonstration

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

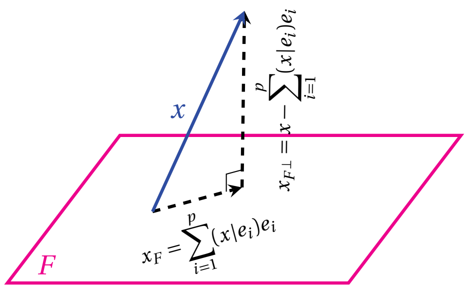
• *Analyse* – On suppose que $x = x_F + x_{F^\perp}$ avec $x_F \in F$ et $x_{F^\perp} \in F^\perp$.

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

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

• *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 :

$$\forall j \in [\![1,p]\!], \quad \left(x - \sum_{i=1}^{p} (x|e_i) e_i \middle| e_j\right) = (x|e_j) - \sum_{i=1}^{p} (x|e_i)(e_i|e_j) = 0$$

■

Corollaire 14.21 : Inégalité de Bessel

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

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

Corollaire 14.22

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, $\left(F^\perp\right)^\perp = F$.

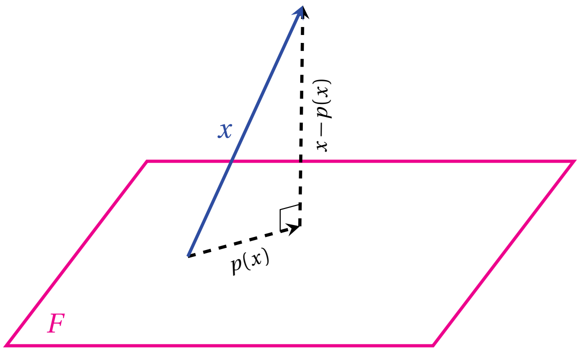
Démonstration

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

D – Projection orthogonale et distance

Définition 14.23 : Projecteur orthogonale

Soit F un sous-espace vectoriel de E de dimension finie. Alors, $E = F \oplus F^\perp$.
On appelle projection orthogonale sur F la projection sur F parallèlement à $F^\perp$.



Représentation du projeté orthogonal de x sur F

Théorème 14.24

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 $p(x) = (x|e_1)e_1 + \cdots + (x|e_n)e_n$.

Démonstration

Redémontrons rapidement le deuxième point.

$$\forall i \in [\![1,n]\!],\quad (x|e_i) = (p(x) + (x - p(x))|e_i) = (p(x)|e_i) + (x - p(x)|e_i) = (p(x)|e_i)$$

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

Exercice 7

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

Exercice 8

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.

Exercice 9

On munit $\mathbb{R}^n$ de sa structure euclidienne canonique et on note x un vecteur unitaire de $\mathbb{R}^n$.

- Déterminer la matrice dans la base canonique de la projection orthogonale sur $\text{Vect}(x)$.
- Soit M la matrice dans la base canonique d'une projection orthogonale.

  Montrer qu'il existe des vecteurs unitaires $X_1, \ldots, X_r$ de $\mathbb{R}^n$ tels que $M = \displaystyle\sum_{k=1}^{r} X_k X_k^\top$.

Définition 14.25 : 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 $d(x,F) = \inf\limits_{u \in F} d(x,u) = \inf\limits_{u \in F} \|x-u\|$.

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.

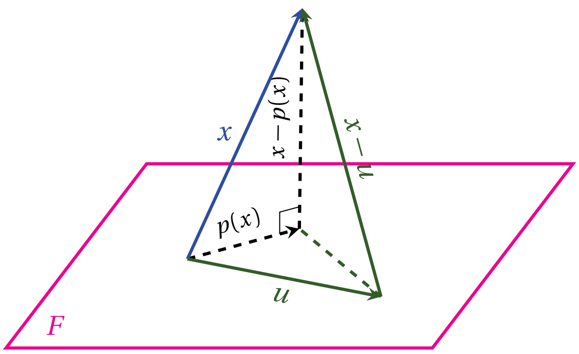
Théorème 14.26

Soient $x \in E$ et F un sous-espace vectoriel de E de dimension finie.
$d(x,F) = \inf\limits_{u \in F} \|x-u\| = \|x-p(x)\|$ 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,

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

Ainsi, $\|x-u\| \geqslant \|x-p(x)\|$ 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 (pour votre culture, HP)

$(E, (\cdot|\cdot))$ 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 $x = \displaystyle\sum_{i=1}^{n} (x|e_i)e_i$ valable en dimension finie.

Définition 14.27 : Suite totale

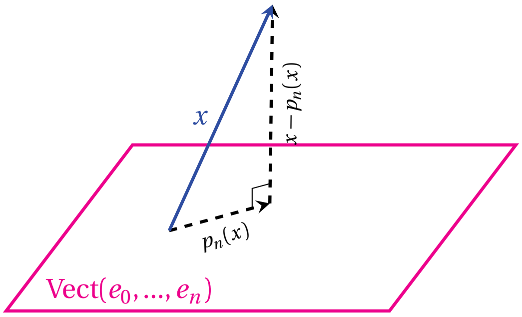
On dit qu'une suite de vecteurs $(e_i)_{i\in\mathbb{N}}$ de E est totale si pour tout $x \in E$, :

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

Par caractérisation séquentielle de la limite, cela revient à dire qu'il existe une suite $(x_n)_{n\in\mathbb{N}}$ d'éléments de $\underset{i\in\mathbb{N}}{\text{Vect}}(e_i)$ qui converge vers x. En d'autres termes, $\underset{i\in\mathbb{N}}{\text{Vect}}(e_i)$ est *dense* dans E.

Théorème 14.28

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 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 $n\in\mathbb{N}$, $F_n=\underset{0\leqslant i\leqslant n}{\text{Vect}}(e_i)$ et l'on cherche à montrer que $\|x-p_n(x)\|\xrightarrow[n\to+\infty]{}0$. On fixe $\varepsilon>0$.
Par définition d'une famille totale, il existe $y\in\underset{i\in\mathbb{N}}{\text{Vect}}(e_i)$ 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 $n\geqslant n_0$, $y\in F_n$ et donc,

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

On vient finalement d'établir que pour tout $x\in E$, $x=\lim\limits_{n\to+\infty}p_n(x)=\lim\limits_{n\to+\infty}\sum\limits_{i=0}^{n}(x|e_i)e_i=\sum\limits_{i=0}^{+\infty}(x|e_i)e_i$.
On obtient alors la version « complète » de l'inégalité de Bessel.

Corollaire 14.29 : Égalité de Parseval

Si $(e_i)_{i\in\mathbb{N}}$ une suite orthonormale totale d'éléments de E, alors, pour tout $x\in E$, $\|x\|^2=\sum\limits_{i=0}^{+\infty}(x|e_i)^2$.

Démonstration

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

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

Exercice 11

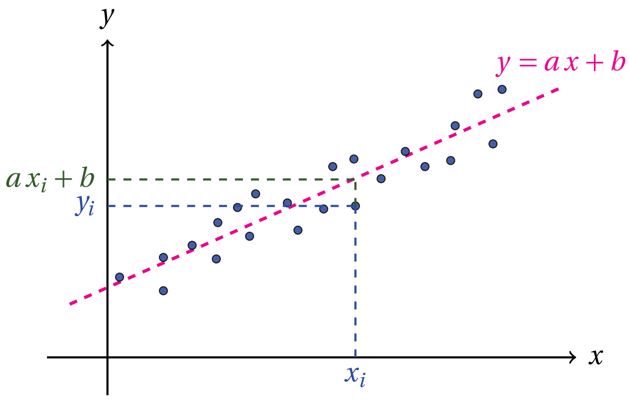
Montrer que la famille $(x\longmapsto x^n)_{n\in\mathbb{N}}$ est une suite totale de l'espace $\mathscr{C}([0,1])$ pour un produit scalaire à préciser.

IV | Méthode des moindres carrés (pour votre culture, HP)

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 $i \in [\![1,n]\!]$. 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 $y = ax+b$, 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,

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

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

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

Posons $X = (x_1, \ldots, x_n)$, $Y = (y_1, \ldots, y_n)$, $Z = aX + b = (ax_1 + b, \ldots, ax_n + b)$ et $F = \mathrm{Vect}(X, (1, \ldots, 1))$.

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

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

Alors, $Y - Z = Y - p(Y) \in F^{\perp}$ c'est-à-dire :

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

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

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

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

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

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 !