CHAPITRE 9
Espaces préhilbertiens réels

Plan du chapitre

I Généralités 2
A - Produit scalaire 2
B - Norme associée à un produit scalaire 2

II Orthogonalité 3
A - Vecteurs et sous-espaces orthogonaux 3
B - Familles orthogonales et orthonormées 4
C - Algorithme d'orthonormalisation de Gram-Schmidt 4

III Projection orthogonale 5
A - Projection orthogonale sur un sous-espace 5
B - Distance d'un vecteur à un sous-espace 5

IV Méthode : L'algorithme de Gram-Schmidt 6

Introduction

Un espace préhilbertien réel est un espace vectoriel réel dans lequel on dispose d'une opération algébrique supplémentaire permettant d'associer un nombre réel à chaque couple de vecteurs : un produit scalaire. Cette nouvelle application permet d'étendre les notions traditionnelles de géométrie euclidienne du plan et de l'espace à un espace vectoriel réel quelconque : les longueurs, les angles, l'orthogonalité et la projection orthogonale. Le préfixe « pré » dans le mot « préhilbertien » fait référence à l'absence d'une hypothèse particulière : la complétude. Lorsque cette hypothèse supplémentaire est vérifiée, l'espace porte le nom d'espace hilbertien ou d'espace de Hilbert en référence au mathématicien allemand David Hilbert.

Bien que la géométrie euclidienne remonte à l'antiquité, les produits scalaires n'apparaissent qu'au XIX^{e} siècle au fur et à mesure qu'émerge et se formalise la notion générale d'espace vectoriel. L'attribution de sa paternité n'est pas évidente, mais on peut citer les mathématiciens Grassmann, Peano et Clifford.

Au début du XX^{e} siècle, les mathématiciens Hilbert, Schmidt, Riesz et Fréchet commencent à s'intéresser à certains espaces préhilbertiens de dimension infinie dont les éléments sont des fonctions afin d'étudier des équations intégrales : c'est la naissance de l'analyse fonctionnelle. Cette dernière est omniprésente de nos jours dans de nombreux problèmes issus de la physique.

Les espaces préhilbertiens interviennent dans de nombreux domaines de la physique. Ils sont par exemple au cœur de la théorie de la mécanique quantique. D'autre part, ils permettent d'introduire de nombreux outils pour étudier les équations aux dérivées partielles qui sont omniprésentes dans les problèmes de la physique. En mathématiques, la généralisation des notions de produit scalaire et de projection orthogonale nous permettra de construire la théorie des séries de Fourier dans un futur chapitre.

Dans ce chapitre, nous commencerons par définir la notion générale de produit scalaire sur un espace vectoriel réel. Dans un second temps, nous étudierons l'orthogonalité dans ce nouveau cadre, puis finalement nous généraliserons la notion de projection orthogonale avec ses propriétés classiques.

Dans tout le chapitre, on fixe un espace vectoriel E sur $\mathbb{R}$.

I - Généralités

I.A - Produit scalaire

Définition (Produit scalaire) : Une application $\varphi : E \times E \to \mathbb{R}$ est un produit scalaire si elle vérifie les conditions suivantes.

(i) Linéaire à gauche : pour tout $y \in E$, l'application $x \mapsto \varphi(x,y)$ est linéaire,

(ii) Linéaire à droite : pour tout $x \in E$, l'application $y \mapsto \varphi(x,y)$ est linéaire,

(iii) Symétrique : pour tout $(x,y) \in E^2$, on a $\varphi(x,y) = \varphi(y,x)$,

(iv) Positive : pour tout $x \in E$, on a $\varphi(x,x) \geqslant 0$,

(v) Définie : si $x \in E$ vérifie $\varphi(x,x) = 0$, alors $x = 0$.

Remarque 1 : Si la condition (iii) est vérifiée, alors les conditions (i) et (ii) sont équivalentes.

Exemples 1 :

a) Le produit scalaire euclidien canonique sur $E = \mathbb{R}^n$ est défini par

$$\forall (x,y) \in \mathbb{R}^n \times \mathbb{R}^n, \quad \varphi(x,y) = \sum_{i=1}^{n} x_i y_i.$$

b) Si $E = \mathscr{C}^0([0,1],\mathbb{R})$, l'application $\varphi : E \times E \to \mathbb{R}$ définie par

$$\forall (f,g) \in \mathscr{C}^0([0,1],\mathbb{R})^2, \quad \varphi(f,g) = \int_0^1 f(t)g(t)\,\mathrm{d}t$$

est un produit scalaire sur $\mathscr{C}^0([0,1],\mathbb{R})$.

c) Si $E = \mathscr{M}_n(\mathbb{R})$, l'application $\varphi : E \times E \to \mathbb{R}$ définie par

$$\forall (A,B) \in \mathscr{M}_n(\mathbb{R})^2, \quad \varphi(A,B) = \mathrm{Tr}\left(A^{\mathrm{T}}B\right)$$

est un produit scalaire sur $\mathscr{M}_n(\mathbb{R})$.

Notation : Le produit scalaire de $(x,y) \in E^2$ se note $(x \mid y)$, $\langle x \mid y \rangle$ ou $x \cdot y$.

Définition (Espace préhilbertien réel) : Un espace préhilbertien réel est un espace vectoriel réel muni d'un produit scalaire.

Définition (Espace euclidien) : Un espace euclidien est un espace vectoriel réel de dimension finie muni d'un produit scalaire.

I.B - Norme associée à un produit scalaire

Jusqu'à la fin du chapitre, on fixe un produit scalaire $(\cdot \mid \cdot)$ sur E.

Définition (Norme associée à un produit scalaire) : La norme associée au produit scalaire $(\cdot \mid \cdot)$ est l'application $\|\cdot\| : E \to \mathbb{R}$ définie par

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

Exemple 2 : La norme associée au produit scalaire canonique sur $\mathbb{R}^n$ est

$$\forall x \in \mathbb{R}^n, \quad \|x\| = \sqrt{x_1^2 + \cdots + x_n^2}.$$

Proposition 1 : La norme $\|\cdot\| : E \to \mathbb{R}$ vérifie les propriétés suivantes.

(i) Positivité : pour tout $x \in E$, on a $\|x\| \geqslant 0$,

(ii) Séparation : pour tout $x \in E$, on a $\|x\| = 0$ si et seulement si $x = 0$,

(iii) Homogénéité : pour tout $x \in E$ et tout $\lambda \in \mathbb{R}$, on a $\|\lambda \cdot x\| = |\lambda| \cdot \|x\|$.

Remarque 2 : Pour tout $(x,y) \in E^2$, on a la relation $\|x \pm y\|^2 = \|x\|^2 \pm 2(x \mid y) + \|y\|^2$.

Proposition (Identité de polarisation) : Pour tout $(x,y)^2 \in E$, on a l'identité

$$(x \mid y) = \frac{1}{2}\left(\|x+y\|^2 - \|x\|^2 - \|y\|^2\right).$$

Proposition (Inégalité de Cauchy-Schwarz) : Pour tout $(x,y)\in E^2$, on a l'inégalité

$$|(x\mid y)|\leqslant \|x\|\|y\|.$$

De plus, on a égalité si et seulement si les vecteurs x et y sont colinéaires.

Proposition (Inégalité triangulaire) : Pour tout $(x,y)\in E^2$, on a l'inégalité

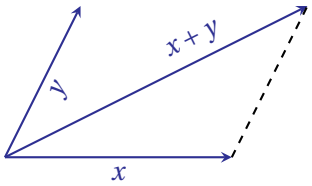
$$\|x+y\|\leqslant \|x\|+\|y\|.$$

De plus, on a égalité si et seulement si $x=0$ ou s'il existe $\lambda\geqslant 0$ tel que $y=\lambda x$.

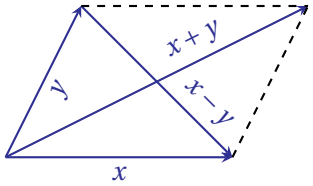
Proposition (Identité du parallélogramme) : Pour tout $(x,y)\in E^2$, on a l'identité

$$\|x+y\|^2+\|x-y\|^2=2\|x\|^2+2\|y\|^2.$$

Illustration : Les deux figures ci-dessous illustrent les propositions précédentes.



L'inégalité triangulaire



L'identité du parallélogramme

Définition (Distance associée à un produit scalaire) : La distance associée à un produit scalaire sur E est l'application $d : E\times E\to\mathbb{R}$ définie par

$$\forall (x,y)\in E^2,\quad d(x,y)=\|x-y\|.$$

II - Orthogonalité

II.A - Vecteurs et sous-espaces orthogonaux

Définition (Vecteurs orthogonaux) : Deux vecteurs $x\in E$ et $y\in E$ sont dits orthogonaux si $(x\mid y)=0$.

Exemples 3 :
En reprenant les espaces préhilbertiens de l'exemple 1,

a) Les vecteurs $(1,-1,0,\dots,0)$ et $(1,1,0,\dots,0)$ sont orthogonaux.

b) Les fonctions $t\mapsto 1$ et $t\mapsto t-1/2$ sont orthogonales.

c) Si A est une matrice triangulaire strictement supérieur et si B est une matrice triangulaire inférieur, alors A et B sont orthogonaux.

Théorème de Pythagore : Deux vecteurs $x\in E$ et $y\in E$ sont orthogonaux si et seulement si

$$\|x+y\|^2=\|x\|^2+\|y\|^2.$$

Définition (Sous-espaces orthogonaux) : Deux sous-espaces vectoriels F et G de E sont dits orthogonaux si tous les vecteurs de F sont orthogonaux à tous les vecteurs de G, i.e.

$$\forall (x,y)\in F\times G,\ (x\mid y)=0.$$

Exemple 4 : Si $E=\mathscr{M}_n(\mathbb{R})$ muni du produit scalaire $(A\mid B)=\mathrm{Tr}\left(A^{\mathrm{T}}B\right)$, alors les sous-espaces vectoriels $\mathscr{S}_n(\mathbb{R})$ et $\mathscr{A}_n(\mathbb{R})$ sont orthogonaux.

Définition (Orthogonal d'un sous-espace vectoriel) : Soit F un sous-espace vectoriel de E. L'orthogonal de F est l'ensemble des vecteurs de E orthogonaux à tous les vecteurs de F, i.e.

$$F^{\perp}=\left\{x\in E\mid \forall y\in F,\ (x\mid y)=0\right\}.$$

Proposition 2 : Soit F un sous-espace vectoriel de E. L'ensemble $F^\perp$ est un sous-espace vectoriel de E.

Exemple 5 : Dans $E=\mathbb{R}^3$ muni du produit scalaire euclidien canonique, l'orthogonal de

$$H=\{(x,y,z)\in\mathbb{R}^3 \mid x+y+z=0\}$$

est la droite $H^\perp=\text{Vect}((1,1,1))$.

Théorème 1 : Si E est un espace euclidien et si F est un sous-espace vectoriel de E, alors

$$\dim(E)=\dim(F)+\dim\left(F^\perp\right).$$

Exemple 6 : Dans l'exemple précédent, on a $\dim(H)=2$, donc $\dim\left(H^\perp\right)=1$.

II.B - Familles orthogonales et orthonormées

Définition (Famille orthogonale) : Une famille $(u_i)_{i\in I}$ de vecteurs de E est dite orthogonale si ses vecteurs sont orthogonaux deux à deux, i.e.

$$\forall (i,j)\in I^2,\ \ i\neq j \ \Rightarrow\ (u_i\mid u_j)=0.$$

Définition (Famille orthonormée) : Une famille $(u_i)_{i\in I}$ de vecteurs de E est dite orthonormée (ou orthonormale) si elle est orthogonale et que $\|u_i\|=1$ pour tout indice $i\in I$.

Exemple 7 : La base canonique de $\mathbb{R}^n$ est une base orthonormée pour le produit scalaire canonique sur $\mathbb{R}^n$.

Proposition 3 : Toute famille orthogonale de vecteurs non nuls est libre.

Théorème 2 : On suppose que E est un espace euclidien dont $\mathscr{B}=(e_1,\dots,e_n)$ est une base orthonormée. Soient $x,y\in E$ deux vecteurs de E qui se décompose dans $\mathscr{B}$ sous la forme

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

où $(x_1,\dots,x_n)\in\mathbb{R}^n$ et $(y_1,\dots,y_n)\in\mathbb{R}^n$. On a alors pour tout $i\in[\![1,n]\!]$

$$x_i=(x\mid e_i),\quad \|x\|^2=\sum_{i=1}^n x_i^2,\quad (x\mid y)=\sum_{i=1}^n x_i y_i.$$

Remarque 3 : Autrement dit, une fois que l'on a fixé une base orthonormée, les règles de calculs sont les mêmes que dans $\mathbb{R}^n$ avec le produit scalaire usuel.

II.C - Algorithme d'orthonormalisation de Gram-Schmidt

Théorème de Gram-Schmidt : Soit $(v_1,\dots,v_p)$ une famille libre de vecteurs de E. Il existe une unique famille orthonormale $(e_1,\dots,e_p)$ vérifiant

$$\forall k\in[\![1,p]\!],\quad \text{Vect}(v_1,\dots,v_k)=\text{Vect}(e_1,\dots,e_k)\quad\text{et}\quad (e_k\mid v_k)>0.$$

Remarque 4 : La démonstration du théorème ci-dessus est basée sur l'algorithme de Gram-Schmidt qui est expliqué dans la dernière partie.

Corollaire 1 : Si E est un espace euclidien, alors il existe une base orthonormée de E.

III - Projection orthogonale

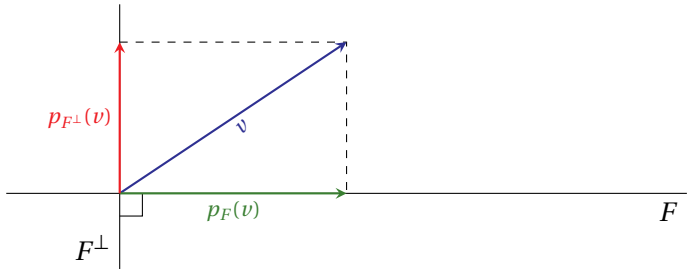
III.A - Projection orthogonale sur un sous-espace

Proposition 4 : Si F un sous-espace vectoriel de dimension finie de E, alors on a la décomposition $E = F \oplus F^\perp$.

ATTENTION : L'hypothèse de dimension finie est importante : en général, on a toujours $F \cap F^\perp = \{0\}$, mais on a pas $E = F + F^\perp$.

Définition (Projection orthogonale) : Soit F un sous-espace vectoriel de dimension finie de E. La projection orthogonale sur F est le projecteur $p_F : E \to E$ sur F parallèlement à $F^\perp$.

Illustration : On peut se représenter la projection orthogonale d'un vecteur avec le schéma ci-dessous.



Remarque 5 : On a la relation $p_F + p_{F^\perp} = \mathrm{Id}_E$.

Exemple 8 : La projection orthogonale sur le sous-espace vectoriel

$$H = \{(x,y,z) \in \mathbb{R}^3 \mid x+y+z = 0\}$$

de $\mathbb{R}^3$ muni du produit scalaire canonique est

$$p_H(x,y,z) = \left(\frac{2x-y-z}{3}, \frac{-x+2y-z}{3}, \frac{-x-y+2z}{3}\right).$$

Proposition (Formule de projection) : Soit F un sous-espace vectoriel de dimension finie de E. Si $(e_1, \dots, e_p)$ est une base orthonormée de F, alors

$$\forall v \in E, \quad p_F(v) = \sum_{i=1}^{p} (v \mid e_i) \cdot e_i.$$

III.B - Distance d'un vecteur à un sous-espace

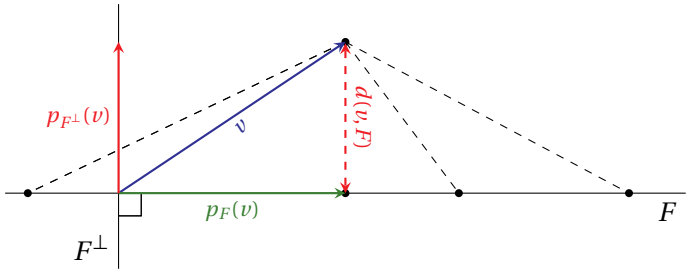
Définition (Distance d'un vecteur à un sous-espace vectoriel) : La distance d'un vecteur $v \in E$ à un sous espace vectoriel F de E est le nombre

$$d(v,F) = \inf_{f \in F} \|v - f\| = \inf_{f \in F} d(v,f).$$

Théorème 3 : Soit $v \in E$. Si F un sous-espace vectoriel de dimension finie de E, alors

$$d(v,F) = \left\|v - p_F(v)\right\| = \|p_{F^\perp}(v)\|.$$

Illustration : On peut visualiser le théorème sur le schéma ci-dessous.



Exemple 9 : On souhaite calculer la distance du vecteur $(1,0,0)$ au sous-espace

$$H = \{(x,y,z) \in \mathbb{R}^3 \mid x+y+z = 0\}$$

de $\mathbb{R}^3$ muni du produit scalaire canonique. En reprenant le formule pour la projection calculer dans l'exemple précédent, on a

$$d((1,0,0),H) = \left\|(1,0,0) - p_H(1,0,0)\right\| = \left\|(1,0,0) - \left(\frac{2}{3}, -\frac{1}{3}, -\frac{1}{3}\right)\right\| = \frac{1}{\sqrt{3}}.$$

IV - Méthode : L'algorithme de Gram-Schmidt

L'algorithme de Gram-Schmidt permet de construire la famille $(e_1,\dots,e_p)$ du théorème de Gram-Schmidt en partant d'une famille libre $(v_1,\dots,v_p)$ de vecteurs de E.

Méthode : L'algorithme de Gram-Schmidt
On commence par construire une famille orthogonale $(u_1,\dots,u_p)$ vérifiant les conditions du théorème.

1) On pose $u_1 = v_1$.
2) On cherche u_2 sous la forme

$$u_2 = v_2 + a u_1 \quad \text{avec} \quad a \in \mathbb{R}.$$

On souhaite que u_2 et u_1 soit orthogonaux, i.e.

$$(u_2 \mid u_1) = 0 \iff (v_2 \mid u_1) + a(u_1 \mid u_1) = 0 \iff a = -\frac{(v_2 \mid u_1)}{(u_1 \mid u_1)}.$$

On trouve ainsi u_2.
3) On itère le procédé. Si on a construit $(u_1,\dots,u_{k-1})$, on cherche u_k sous la forme

$$u_k = v_k + \lambda_1 u_1 + \cdots + \lambda_{k-1} u_{k-1} \quad \text{avec} \quad \lambda_1,\dots,\lambda_{k-1} \in \mathbb{R}.$$

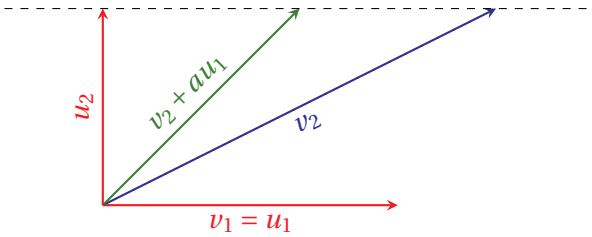
Pour $i \in [\![1, k-1]\!]$, on souhaite que u_k soit orthogonale à u_i, i.e.

$$(u_k \mid u_i) = 0 \iff (v_k \mid u_i) + \lambda_i (u_i \mid u_i) = 0 \iff \lambda_i = -\frac{(v_k \mid u_i)}{(u_i \mid u_i)}.$$

4) A la fin du procédé, on a construit une famille orthogonale $(u_1,\dots,u_p)$ de vecteurs non nuls. On en déduit la famille orthonormée $(e_1,\dots,e_p)$ vérifiant les conditions du théorème en normalisant les vecteurs u_i

$$\forall i \in [\![1,p]\!], \quad e_i = \frac{u_i}{\|u_i\|}.$$

Illustration : Le schéma ci-dessous représente la seconde étape dans l'algorithme de Gram-Schmidt.



Exemple 10 : On cherche une base orthonormée de l'espace vectoriel

$$H = \text{Vect}(v_1, v_2) \quad \text{où} \quad v_1 = (1,0,1) \text{ et } v_2 = (1,1,0)$$

muni de la restriction du produit scalaire canonique de $\mathbb{R}^3$. On applique l'algorithme de Gram-Schmidt à la base (v_1, v_2) de H.

1) On pose $u_1 = v_1 = (1,0,1)$.
2) On cherche u_2 sous la forme

$$u_2 = v_2 + a u_1 \quad \text{avec} \quad a \in \mathbb{R}.$$

On souhaite que u_2 et u_1 soit orthogonaux, i.e.

$$(u_2 \mid u_1) = 0 \iff (v_2 \mid u_1) + a(u_1 \mid u_1) = 0 \iff a = -\frac{(v_2 \mid u_1)}{(u_1 \mid u_1)} = -\frac{1}{2}.$$

On a donc

$$u_2 = v_2 - \frac{1}{2} u_1 = \left(\frac{1}{2}, 1, -\frac{1}{2}\right).$$

3) Ainsi la famille (u_1, u_2) est une base orthogonale de H et on en déduit une base orthonormée (e_1, e_2) de H avec

$$e_1 = \frac{u_1}{\|u_1\|} = \left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right), \quad e_2 = \frac{u_2}{\|u_2\|} = \left(\frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}, -\frac{1}{\sqrt{6}}\right).$$

Exemple 11 : On considère $E = \mathbb{R}_2[X]$ muni du produit scalaire définie par

$$\forall (P,Q) \in \mathbb{R}_2[X]^2, \quad (P \mid Q) = \int_{-1}^{1} P(t)Q(t)\,\mathrm{d}t.$$

Nous allons construire une base orthonormée de $\mathbb{R}_2[X]$ en appliquant l'algorithme de Gram-Schmidt à la base $(V_0, V_1, V_2) = (1, X, X^2)$.

1) On pose $U_0 = V_0 = 1$.
2) On cherche U_1 sous la forme

$$U_1 = V_1 + aU_0 \quad \text{avec} \quad a \in \mathbb{R}.$$

On souhaite que U_0 et U_1 soit orthogonaux, i.e.

$$(U_1 \mid U_0) = 0 \;\Leftrightarrow\; (V_1 \mid U_0) + a(U_0 \mid U_0) = 0 \;\Leftrightarrow\; a = -\frac{(V_1 \mid U_0)}{(U_0 \mid U_0)} = 0.$$

On a donc $U_1 = X$.

3) On cherche U_2 sous la forme

$$U_2 = V_2 + bU_1 + cU_0 \quad \text{avec} \quad b, c \in \mathbb{R}.$$

On souhaite que U_2 soit orthogonal à U_0 et U_1, donc

$$(U_2 \mid U_0) = 0 \;\Leftrightarrow\; (V_2 \mid U_0) + c(U_0 \mid U_0) = 0 \;\Leftrightarrow\; c = -\frac{(V_2 \mid U_0)}{(U_0 \mid U_0)} = -\frac{1}{3}.$$

$$(U_2 \mid U_1) = 0 \;\Leftrightarrow\; (V_2 \mid U_1) + b(U_1 \mid U_1) = 0 \;\Leftrightarrow\; b = -\frac{(V_2 \mid U_1)}{(U_1 \mid U_1)} = 0.$$

On a donc $U_2 = X^2 - 1/3$.

4) Ainsi la famille (U_0, U_1, U_2) est une base orthogonale de $\mathbb{R}_2[X]$ et on en déduit une base orthonormée (E_0, E_1, E_2) de $\mathbb{R}_2[X]$ avec

$$E_0 = \frac{U_0}{\|U_0\|} = \frac{1}{\sqrt{2}}, \quad E_1 = \frac{U_1}{\|U_1\|} = \sqrt{\frac{3}{2}}X, \quad E_2 = \frac{U_2}{\|U_2\|} = \frac{\sqrt{5}}{2\sqrt{2}}\left(3X^2 - 1\right).$$