ESPACES PRÉHILBERTIENS RÉELS.

Dans tout ce chapitre, le corps de base est $\mathbb{R}$.

I. Produit scalaire

I.1. Définitions - Exemples

Déf 1:

Une forme bilinéaire sur un $\mathbb{R}$-espace vectoriel E est une application φ de $E \times E$ dans $\mathbb{R}$ linéaire par rapport à chacune des deux variables, c'est-à-dire telle que :

a) Pour tout $y \in E$, l'application $x \mapsto \varphi(x,y)$ est une forme linéaire sur E.

b) Pour tout $x \in E$, l'application $y \mapsto \varphi(x,y)$ est une forme linéaire sur E.

Pour vérifier qu'une application φ de E^2 dans $\mathbb{R}$ est une forme bilinéaire, il faut donc vérifier :

$$\forall (x_1, x_2, y) \in E^3,\ \forall \lambda \in \mathbb{R},\ \varphi(\lambda x_1 + x_2, y) = \lambda \varphi(x_1, y) + \varphi(x_2, y)$$

(linéarité par rapport à la 1ère variable), et :

$$\forall (x, y_1, y_2) \in E^3,\ \forall \lambda \in \mathbb{R},\ \varphi(x, \lambda y_1 + y_2) = \lambda \varphi(x, y_1) + \varphi(x, y_2)$$

(linéarité par rapport à la 2ème variable).

Déf 2:

Une application $\varphi : E^2 \to \mathbb{R}$ est dite symétrique lorsque pour tout $(x,y) \in E^2$, $\varphi(x,y) = \varphi(y,x)$.

Rem: Si φ est une application de E^2 dans $\mathbb{R}$, symétrique et linéaire d'un côté, elle est bilinéaire.

Déf 3:

On appelle produit scalaire sur un $\mathbb{R}$-espace vectoriel E une forme bilinéaire symétrique φ qui est de plus :

- positive, c'est-à-dire : $\forall x \in E\ ,\ \varphi(x,x) \geqslant 0$.
- et définie, c'est-à-dire : $\forall x \in E\ ,\ \varphi(x,x) = 0 \Longrightarrow x = 0_E$.

Dire que φ est définie positive peut aussi s'écrire : $\forall x \in E \setminus \{0\}\ ,\ \varphi(x,x) > 0$.

Déf 4:

Si φ est un produit scalaire sur un $\mathbb{R}$-espace vectoriel E, le couple (E, φ) s'appelle un espace préhilbertien réel.

Le produit scalaire de deux vecteurs x et y se note en général : $\varphi(x,y) = (x|y)$ ou $\langle x\,|\,y\rangle$ ou $\langle x\,,\,y\rangle$.

Exemples :

1. Dans $\mathbb{R}^n$, pour $x = (x_1, \ldots, x_n)$ et $y = (y_1, \ldots, y_n)$: $\quad \langle x\,|\,y\rangle = x_1y_1 + \cdots + x_ny_n$.
   Il s'agit du produit scalaire canonique sur $\mathbb{R}^n$; la démonstration du fait qu'il s'agit bien d'un produit scalaire est simple :

   ☞ *Démonstration:*

   - L'application $\varphi \colon (x,y) \mapsto \sum\limits_{i=1}^{n} x_i y_i$ est bien définie sur $(\mathbb{R}^n)^2$, et à valeurs dans $\mathbb{R}$.
   - Il est clair que pour tout $(x,y) \in (\mathbb{R}^n)^2$, $\varphi(x,y) = \varphi(y,x)$: φ est symétrique.
   - À y fixé, chaque application $x \mapsto x_i$ est une forme linéaire sur $\mathbb{R}^n$ (forme linéaire coordonnée) donc l'application partielle $x \mapsto \varphi(x,y)$ est linéaire comme combinaison linéaire de formes linéaires.
     Puisqu'elle est symétrique, φ est donc bilinéaire.
   - Pour tout $x \in \mathbb{R}^n$, $\varphi(x,x) = \sum\limits_{i=1}^{n} x_i^2$ est un nombre réel positif, donc φ est positive.
   - Enfin, φ est définie puisque, si $\varphi(x,x) = \sum\limits_{i=1}^{n} x_i^2 = 0$ alors tous les x_i sont nuls (ce sont des réels !), donc x est nul.

Rem: En notant $X \in \mathcal{M}_{n,1}(\mathbb{R})$ le vecteur colonne $\begin{pmatrix} x_1 & \dots & x_n \end{pmatrix}^\top$ et $Y \in \mathcal{M}_{n,1}(\mathbb{R})$ le vecteur colonne $\begin{pmatrix} y_1 & \dots & y_n \end{pmatrix}^\top$, on a aussi : $\boxed{\langle x \,|\, y \rangle = X^\top Y = Y^\top X\,.}$

2. De façon plus générale, si E est un $\mathbb{R}$-espace vectoriel rapporté à une base $\mathscr{B} = (e_1, \dots, e_n)$, et si x et y sont deux vecteurs de E de coordonnées respectives $(x_1, \dots, x_n)$ et $(y_1, \dots, y_n)$ dans cette base, on définit un produit scalaire sur E en posant : $\langle x \,|\, y \rangle = \sum\limits_{i=1}^{n} x_i y_i$.
3. Dans $\mathcal{M}_n(\mathbb{R})$, si $A = (a_{i,j})_{1 \leqslant i,j \leqslant n}$ et $B = (b_{i,j})_{1 \leqslant i,j \leqslant n}$, on peut poser :
$$\langle A \,|\, B \rangle = \operatorname{tr}(A^\top B) = \sum_{1 \leqslant i,j \leqslant n} a_{i,j} b_{i,j}\,.$$
Il s'agit du produit scalaire canonique sur $\mathcal{M}_n(\mathbb{R})$.

Démonstration:

- Déjà on a, en utilisant la formule du produit matriciel :
$$\forall\, (i,j) \in [\![1\,;n]\!]^2,\ (A^\top B)_{ij} = \sum_{k=1}^{n} (A^\top)_{ik} B_{kj} = \sum_{k=1}^{n} a_{ki} b_{kj}$$
d'où
$$\operatorname{tr}(A^\top B) = \sum_{i=1}^{n} (A^\top B)_{ii} = \sum_{i=1}^{n} \sum_{k=1}^{n} a_{ki} b_{ki}$$
ce qui après changement d'indices donne la formule annoncée.
- Pour démontrer qu'il s'agit bien d'un produit scalaire, on pourrait vérifier une à une les propriétés, mais il est plus rapide de dire que, si l'on munit $\mathcal{M}_n(\mathbb{R})$ de sa base canonique (E_{ij}), la définition ci-dessus est la même que celle de l'exemple 2. (avec ici un espace vectoriel de dimension n^2).

4. Dans $\mathbb{R}[X]$, pour $P = \sum\limits_{k=0}^{n} a_k X^k$ et $Q = \sum\limits_{k=0}^{p} b_k X^k$: $\quad \langle P \,|\, Q \rangle = \sum\limits_{k=0}^{\min(n,p)} a_k b_k$.
Il s'agit du produit scalaire canonique sur $\mathbb{R}[X]$.
5. Dans l'ensemble $\mathscr{C}([a\,;b], \mathbb{R})$ des fonctions continues sur le segment $[a\,;b]$ on peut poser : $\langle f \,|\, g \rangle = \displaystyle\int_a^b fg$.

Démonstration:

Le caractère bilinéaire symétrique ne pose pas de problème.
Le caractère positif non plus.
Enfin, si f est une application continue telle que $\displaystyle\int_a^b f^2 = 0$ alors f est nulle sur $[a\,;b]$, puisque l'intégrale d'une fonction positive *continue* et non identiquement nulle est strictement positive (cf. théorème dans un chapitre d'intégration).

I.2. Norme associée à un produit scalaire

Déf 5:

Si E est un espace préhilbertien réel, où le produit scalaire est noté $\langle \cdot \,|\, \cdot \rangle$, on appelle norme (euclidienne) associée l'application :
$$N : x \in E \mapsto \|x\| = \sqrt{\langle x \,|\, x \rangle}\,.$$
(cela a bien un sens car $\langle x \,|\, x \rangle \geqslant 0$).

Exemples :

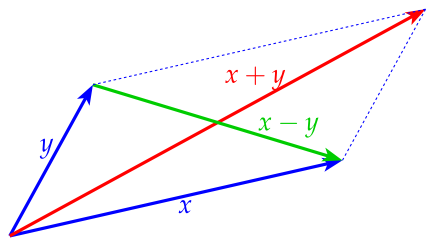
1. Dans $\mathbb{R}^n$, muni du produit scalaire canonique, si $x = (x_1, \dots, x_n)$ on a : $\|x\| = \sqrt{\sum\limits_{i=1}^{n} x_i^2}$.
2. Dans $\mathcal{M}_n(\mathbb{R})$, muni du produit scalaire canonique, si $A = (a_{i,j})_{1 \leqslant i,j \leqslant n}$ on a ; $\|A\| = \sqrt{\sum\limits_{1 \leqslant i,j \leqslant n} a_{i,j}^2}$.
3. Dans l'ensemble $\mathscr{C}([a\,;b], \mathbb{R})$ des fonctions continues sur le segment $[a\,;b]$ muni du produit scalaire vu ci-dessus, on aura : $\|f\| = \sqrt{\int\limits_a^b f(t)^2 \,\mathrm{d}t}$.

De la définition d'un produit scalaire découlent facilement les propriétés suivantes :

Propriétés :

1. $\forall\, x \in E\,, \|x\| = 0_{\mathbb{R}} \iff x = 0_E\,.$
2. $\forall\, x \in E\,,\ \forall\, \lambda \in \mathbb{R}\,,\ \|\lambda x\| = |\lambda|\,\|x\|\,.$
3. Un vecteur x est dit <u>unitaire</u> si $\|x\| = 1$.
   Si x est un vecteur non nul, alors $\dfrac{x}{\|x\|}$ est unitaire.
4. $\forall\,(x,y) \in E^2\,,\ \|x+y\|^2 = \|x\|^2 + 2\,\langle x\,|\,y\rangle + \|y\|^2\,.$
5. *Identité de polarisation :* $\forall\,(x,y) \in E^2\,,\ \langle x\,|\,y\rangle = \dfrac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2\right)\,.$
6. $\forall\,(x,y) \in E^2\,,\ \langle x+y\,|\,x-y\rangle = \|x\|^2 - \|y\|^2\,.$
7. *Identité du parallélogramme :*

$$\forall\,(x,y) \in E^2\,,\ \|x+y\|^2 + \|x-y\|^2 = 2\left(\|x\|^2 + \|y\|^2\right)\,.$$



Théorème 1: Inégalité de Cauchy-Schwarz

Soit E un espace préhilbertien réel. Pour tous x, y dans E on a :

$$|\langle x\,|\,y\rangle| \leqslant \|x\|\,.\,\|y\|$$

et il y a égalité si et seulement si le système $\{x,y\}$ est lié.

Démonstration:

- *1ère démonstration*
  - Si $y = 0$, le résultat est immédiat.
  - Sinon, on écrit que, pour tout $t \in \mathbb{R}$ le réel $\|x+ty\|^2$ est positif.
    Or $\|x+ty\|^2 = t^2\,\|y\|^2 + 2t\,\langle x\,|\,y\rangle + \|x\|^2$. Il s'agit là d'un trinôme en t (car le coefficient de t^2, $\|y\|^2$ est ici non nul); puisque ce trinôme est de signe constant sur $\mathbb{R}$, son discriminant est négatif; cela donne : $\langle x\,|\,y\rangle^2 - \|x\|^2\,\|y\|^2 \leqslant 0$, d'où l'inégalité cherchée.
  - Supposons qu'il y ait égalité :
    - si $y = 0$, la famille $\{x,y\}$ est bien liée.
    - sinon, cela signifie que le trinôme précédent a un discriminant nul, donc possède une racine double t_0 ; on aura alors $\|x+t_0 y\|^2 = 0$, d'où $x + t_0 y = 0$ et x est bien colinéaire à y.
  - Pour terminer, il est facile de vérifier directement que, si la famille $\{x,y\}$ est liée, il y a bien égalité.
- *2ème démonstration (abrégée)*

  Dans le cas non trivial où x et y sont non nuls, on écrit que : $\left\|\frac{x}{\|x\|} \pm \frac{y}{\|y\|}\right\|^2 \geqslant 0$ et en développant on trouve les deux inégalités voulues.

<u>Corollaire 1.1:</u> Inégalité triangulaire

Soit E un espace préhilbertien réel. Pour tous x, y dans E on a :

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

et il y a égalité si et seulement si la famille $\{x,y\}$ est positivement liée (c'est-à-dire qu'il existe un réel λ *positif* tel que $x = \lambda y$ ou $y = \lambda x$).

Démonstration:

- $\|x+y\|^2 = \|x\|^2 + 2\langle x\,|\,y\rangle + \|y\|^2 \underbrace{\leqslant}_{\text{C.S}} \|x\|^2 + 2\|x\|\cdot\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2$
- Il y a égalité si et seulement si $\langle x\,|\,y\rangle = \|x\|\cdot\|y\|$, ce qui équivaut à dire qu'il y a égalité dans l'inégalité de Cauchy-Schwarz *et que* le produit scalaire $\langle x\,|\,y\rangle$ est positif, d'où le résultat.

Corollaire 1.2:

Soit E un espace préhilbertien réel. Pour tous x, y dans E on a :

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

Démonstration:

En appliquant l'inégalité triangulaire à $x+y$ et $-y$ on obtient :

$$\|x\| \leqslant \|x+y\| + \|-y\| = \|x+y\| + \|y\| \ ,$$

soit :

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

De la même façon, en échangeant les rôles de x et y :

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

La juxtaposition de ces deux résultats donne bien l'inégalité voulue.

Exemples :

1. Dans $\mathbb{R}^n$, muni du produit scalaire canonique, si $x = (x_1, \ldots, x_n)$ et $y = (y_1, \ldots, y_n$, l'inégalité de Cauchy-Schwarz s'écrit :
$$\left|\sum_{i=1}^{n} x_i y_i\right| \leqslant \sqrt{\sum_{i=1}^{n} x_i^2}\ \sqrt{\sum_{i=1}^{n} y_i^2}\,.$$
2. Dans $\mathcal{M}_n(\mathbb{R})$, muni du produit scalaire canonique, si $A,\ B \in \mathcal{M}_n(\mathbb{R})$, l'inégalité de Cauchy-Schwarz s'écrit :
$$\left|\operatorname{tr}(A^\top B\right| \leqslant \sqrt{\operatorname{tr}(A^\top A)}\ \sqrt{\operatorname{tr}(B^\top B)}\,.$$
3. Dans l'espace vectoriel $E = \mathscr{C}([a\,;b], \mathbb{R})$ des fonctions continues sur le segment $[a\,;b]$ muni du produit scalaire usuel, si $f,\ g \in E$, l'inégalité de Cauchy-Schwarz s'écrit :
$$\left|\int_a^b f(t)g(t)\,\mathrm{d}t\right| \leqslant \sqrt{\int_a^b f(t)^2\,\mathrm{d}t}\ \sqrt{\int_a^b g(t)^2\,\mathrm{d}t}\,.$$

II. Orthogonalité

Dans ce paragraphe, $(E, \langle\cdot\,|\,\cdot\rangle)$ désigne un espace préhilbertien réel.

II.1. Orthogonal d'une partie

Déf 6:

On dit que deux vecteurs x et y de E sont orthogonaux, et on note $x \perp y$, lorsque $\langle x\,|\,y\rangle = 0$.

Rem: La relation $\langle x\,|\,y\rangle = 0$ est équivalente à $\langle y\,|\,x\rangle = 0$. La relation d'orthogonalité est donc symétrique.

Déf 7:

Soit A une partie non vide d'un espace préhilbertien E.
Un vecteur x de E est dit orthogonal à A si il est orthogonal à tout vecteur de A.
L'ensemble des vecteurs orthogonaux à A s'appelle l'orthogonal de A, noté $A^\perp$. Ainsi :

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

Théorème 2:

Si x est un vecteur non nul de E, alors $\{x\}^{\perp}$ est un hyperplan de E.
Un supplémentaire en est la droite vectorielle $\mathbb{K}.x$.

Démonstration:

- L'application $\varphi_x : \begin{array}{rcl} E & \longrightarrow & \mathbb{R} \\ y & \longmapsto & \langle x \mid y \rangle \end{array}$ est une forme linéaire (par définition du produit scalaire euclidien ou hermitien), non nulle (car $\varphi_x(x) = \|x\|^2 \neq 0$).
  $\{x\}^{\perp}$ en est le noyau, c'est donc un hyperplan.
- La droite vectorielle $\mathbb{R}.x$ n'est pas incluse dans $\{x\}^{\perp}$ (car x orthogonal à x implique $x = 0$), donc en est un supplémentaire d'après le cours sur les hyperplans.

Prop 1:

Si A est une partie non vide de E, alors $A^{\perp}$ est un sous-espace vectoriel de E.

Démonstration:

- $0_E \in A^{\perp}$ puisque $\langle 0 \mid a \rangle = 0$ pour tout $a \in A$;
- Si x et y appartiennent à $A^{\perp}$, alors pour tout $a \in A$, $\langle x \mid a \rangle = \langle y \mid a \rangle = 0$ donc pour tout $\lambda \in \mathbb{R}$, $\langle \lambda x + y \mid a \rangle = 0$ d'où $\lambda x + y \in A^{\perp}$.

On pouvait aussi écrire : $A^{\perp} = \bigcap_{a \in A} \{a\}^{\perp}$ et utiliser le théorème précédent puisque l'intersection d'une famille quelconque de sous-espaces vectoriels est un sous-espace vectoriel.

Prop 2:

Propriétés de l'orthogonal :

1. $E^{\perp} = \{0\}$ et $\{0\}^{\perp} = E$;
2. $\forall (A, B) \in \mathscr{P}(E)^2$, $A \subset B \Rightarrow B^{\perp} \subset A^{\perp}$;
3. $\forall A \in \mathscr{P}(E)$, $A^{\perp} = \mathrm{Vect}(A)^{\perp}$;
   il en résulte que, si F est un sous-espace vectoriel de E et si $(f_i)_{i \in I}$ est une base de F, un vecteur x appartient à $F^{\perp}$ si et seulement si pour tout $i \in I$ on a $\langle x \mid f_i \rangle = 0$;
4. $\forall A \in \mathscr{P}(E)$, $A \subset \left(A^{\perp}\right)^{\perp}$;
5. Si F et G sont des sous-espaces vectoriels de E : $(F + G)^{\perp} = F^{\perp} \cap G^{\perp}$.

Démonstration:

1. Soit $y \in E^{\perp}$; alors pour tout $x \in E$ on a $\langle y \mid x \rangle = 0$; en particulier, on a $\langle y \mid y \rangle = 0$ d'où $y = 0$; cela montre l'inclusion $E^{\perp} \subset \{0\}$, l'inclusion réciproque étant évidente.
   Tout vecteur x est orthogonal à 0, donc $E \subset \{0\}^{\perp} \subset E$, d'où la deuxième égalité.
2. Soient A et B deux parties de E telles que $A \subset B$. Soit alors $x \in B^{\perp}$. Pour tout $b \in B$, $\langle x \mid b \rangle = 0$. En particulier, pour tout $a \in A$, on a $a \in B$ donc $\langle x \mid a \rangle = 0$. D'où $x \in A^{\perp}$ et l'inclusion s'ensuit.
3. Soit $F = \mathrm{Vect}(A)$. $A \subset F$ donc d'après la propriété n°2, $F^{\perp} \subset A^{\perp}$.
   Réciproquement, soit $x \in A^{\perp}$ et $y \in F$. Puisque $F = \mathrm{Vect}(A)$, on peut écrire $y = \sum_{i=1}^{p} \lambda_i a_i$ avec $\lambda_i \in \mathbb{R}$ et $a_i \in A$. On a alors $\langle x \mid y \rangle = \left\langle x \,\middle|\, \sum_{i=1}^{p} \lambda_i a_i \right\rangle = \sum_{i=1}^{p} \lambda_i \langle x \mid a_i \rangle$. Or $\langle x \mid a_i \rangle = 0$ pour tout i puisque $a_i \in A$ et $x \in A^{\perp}$. Donc $\langle x \mid y \rangle = 0$ et ce pour tout $y \in F$. Ainsi, $x \in F^{\perp}$, , ce qui prouve l'inclusion $A^{\perp} \subset F^{\perp}$.
4. Soit $a \in A$. Pour tout $b \in A^{\perp}$, $\langle a \mid b \rangle = 0$ donc $a \perp b$ donc $a \in \left(A^{\perp}\right)^{\perp}$.
5. Déjà, les inclusions $F \subset F + G$ et $G \subset F + G$ impliquent d'après 2. $(F + G)^{\perp} \subset F^{\perp}$ et $(F + G)^{\perp} \subset G^{\perp}$ donc $(F + G)^{\perp} \subset F^{\perp} \cap G^{\perp}$.
   Ensuite, si $x \in F^{\perp} \cap G^{\perp}$, alors x est orthogonal à tout vecteur de F et de G, donc par linéarité il est orthogonal à tout vecteur de $F + G$, ce qui donne l'inclusion réciproque.

Déf 8:

Soient F et G deux sous-espaces vectoriels d'un espace préhilbertien E.
On dit que F et G sont des sous-espaces vectoriels orthogonaux, et on note $F \perp G$ lorsque :

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

autrement dit lorsque tout vecteur de F est orthogonal à tout vecteur de G.
Cela équivaut à : $F \subset G^{\perp}$ et/ou $G \subset F^{\perp}$.

Prop 3:

Soient F et G deux sous-espaces vectoriels d'un espace préhilbertien E. On suppose que l'on dispose d'une base $(f_i)_{i\in I}$ de F et d'une base $(g_j)_{j\in J}$ de G.
Alors F et G sont orthogonaux si et seulement si pour tout $(i,j) \in I \times J$, $f_i \perp g_j$.

Démonstration:

- Si $F \perp G$ alors tout vecteur de F est orthogonal à tout vecteur de G, donc en particulier les f_i sont orthogonaux aux g_j.
- Réciproquement, si les f_i sont orthogonaux aux g_j, alors pour toutes familles de réels (λ_i) et (μ_j) à support fini on a :

$$\left\langle \sum_{i\in I} \lambda_i f_i \,\middle|\, \sum_{j\in J} \mu_j g_j \right\rangle = \sum_{i\in I}\sum_{j\in J} \lambda_i \mu_j \langle f_i \mid g_j \rangle = 0$$

donc tout vecteur de $F = \mathrm{Vect}\left(\{f_i,\ i \in I\}\right)$ est orthogonal à tout vecteur de $G = \mathrm{Vect}\left(\{g_j,\ j \in J\}\right)$.

Exemples

1. Dans $\mathbb{R}^3$, muni du produit scalaire canonique, soit la droite $F = \mathbb{R}.e_3$ et le plan $G = \mathrm{Vect}(e_1 - e_2, e_1 + 2e_2)$.
   Alors $F \perp G$. En fait, ici, $F^{\perp} = G$.
   Si $H = \mathbb{R}.e_1$, on a aussi $F \perp H$, mais seulement $F^{\perp} \supset H$..
2. Dans $\mathbb{R}^3$, soit $F = \mathrm{Vect}(e_1, e_2)$ et $G = \mathrm{Vect}(e_3, e_2 + e_3)$. On n'a pas $F \perp G$ mais par contre $F^{\perp} \perp G^{\perp}$.
   Les deux plans F et G sont alors dits *perpendiculaires*.

Rem : Les exemples précédents montrent qu'il faut faire attention à la terminologie employée.
Ainsi, dire que F et G sont orthogonaux ne veut PAS dire que l'un est l'orthogonal de l'autre !

Prop 4:

Si F et G sont deux sous-espaces vectoriels orthogonaux de E, alors $F \cap G = \{0\}$.

Démonstration:

Soit $x \in F \cap G$. Alors $x \in F$ est orthogonal à $x \in G$, donc $\langle x \mid x \rangle = 0$, donc $x = 0$.

II.2. Familles orthogonales, orthonormées (ou orthonormales)

Déf 9:

Une famille $(x_i)_{i\in I}$ de vecteurs d'un espace préhilbertien réel E est dite orthogonale si : $\forall (i,j) \in I^2$, $i \neq j \Longrightarrow \langle x_i \mid x_j \rangle = 0$.
Elle est dite orthonormée (ou orthonormale) si, *en plus*, $\|x_i\| = 1$ pour tout $i \in I$, c'est-à-dire si

$$\forall (i,j) \in I^2, \langle x_i \mid x_j \rangle = \delta_{i,j}.$$

Rem: Si $(x_i)_{i\in I}$ est une famille orthogonale de vecteurs *non nuls*, alors la famille des vecteurs $\left(\dfrac{x_i}{\|x_i\|}\right)_{i\in I}$ est orthonormée.

Prop 5: Relation de Pythagore

Si $(x_1, \ldots, x_p)$ est une famille orthogonale d'un espace préhilbertien réel, alors :

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

Démonstration:

En effet : $\left\|\sum_{i=1}^{p} x_i\right\|^2 = \left\langle \sum_{i=1}^{p} x_i \,\middle|\, \sum_{i=1}^{p} x_i \right\rangle = \sum_{1\leqslant i,j\leqslant p} \langle x_i \,|\, x_j\rangle = \sum_{i=1}^{p} \langle x_i \,|\, x_i\rangle$, puisque $\langle x_i \,|\, x_j\rangle =)0$ pour $i \neq j$.

Cas particulier :

Théorème 3: de Pythagore

Soit $(E, \langle . \,|\, . \rangle)$ un espace préhilbertien réel, et $x, y \in E$. Alors :

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

Démonstration:

Découle directement de la relation : $\|x+y\|^2 = \|x\|^2 + 2\langle x \,|\, y\rangle + \|y\|^2$.

Théorème 4:

Si $(x_i)_{i\in I}$ est une famille orthogonale de vecteurs *non nuls*, alors cette famille est libre.

Démonstration:

En effet, si l'on a $\sum_{i\in I} \lambda_i x_i = 0$ (avec λ_i scalaires tous nuls sauf un nombre fini), alors, pour tout $j \in I$, $\left\langle x_j \,\middle|\, \sum_{i\in I} \lambda_i x_i \right\rangle = 0$, c'est-à-dire $\sum_{i\in I} \lambda_i \langle x_i \,|\, x_j\rangle = 0$ donc $\lambda_j \langle x_j \,|\, x_j\rangle = 0$ d'où $\lambda_j = 0$ puisque $\|x_j\| \neq 0$.

Corollaire 4.1:

Toute famille orthonormée est libre.

Exemple

Dans $E = \mathscr{C}([0;\pi], \mathbb{R})$ muni du produit scalaire : $\langle f \,|\, g\rangle = \dfrac{2}{\pi}\displaystyle\int_0^{\pi} f(t)g(t)\,\mathrm{d}t$, la famille de fonctions $f_n : t \mapsto \cos nt$, avec $n \in \mathbb{N}^*$, est orthonormée.

II.3. Orthogonalité en dimension finie

Déf 10:

On appelle espace vectoriel euclidien un espace préhilbertien réel de dimension finie.
Si E est un espace euclidien, on appelle base orthonormée (ou base orthonormale) de E toute base de E qui est aussi une famille orthonormée.

Rem: En vertu du corollaire 4.1, pour qu'une famille $(x_1, \ldots, x_n)$ de vecteurs d'un espace vectoriel euclidien forme une base orthonormée, il faut et il suffit que :

- la famille soit orthonormée, c'est-à-dire : $\forall (i,j) \in [\![1;n]\!]^2$, $\langle x_i \,|\, x_j\rangle = \delta_{i,j}$.
- et que $n = \dim E$.

Théorème 5:

Tout espace euclidien (non réduit à $\{0\}$) possède une base orthonormée.

Démonstration:

Par récurrence sur $n = \dim E$:

- si $n = 1$, E est une droite vectorielle, et si $\{x\}$ en est une base, alors $\left\{\dfrac{x}{\|x\|}\right\}$ en sera une base orthonormée
- Supposons le résultat acquis pour tout espace préhilbertien de dimension $n-1$, et soit E préhilbertien de dimension n. Il existe dans E un vecteur unitaire e_n ; d'après le théorème 2, le sous-espace vectoriel $H = \{e_n\}^{\perp}$ est un hyperplan de E. D'après l'hypothèse de récurrence, H possède une base orthonormée $(e_1, \ldots, e_{n-1})$. Il est clair alors que la famille $(e_1, \ldots, e_n)$ sera orthonormée, donc sera une base orthonormée de E, ce qui achève la récurrence.

Exemples

1. Dans $\mathbb{R}^n$ muni du produit scalaire canonique, défini par :
$$\text{si } x = (x_1, \ldots, x_n) \text{ et } y = (y_1, \ldots, y_n)\ ,\ \langle x \,|\, y \rangle = \sum_{i=1}^{n} x_i y_i$$
la base canonique est une base orthonormée.
2. Dans $\mathcal{M}_n(\mathbb{R})$ muni du produit scalaire canonique défini par :
$$\text{si } A = (a_{i,j})_{1\leqslant i,j\leqslant n} \text{ et } B = (b_{i,j})_{1\leqslant i,j\leqslant n},\ \langle A \,|\, B \rangle = \operatorname{tr}(A^\top B) = \sum_{1\leqslant i,j\leqslant n} a_{i,j} b_{i,j}$$
la base canonique $(E_{i,j})_{1\leqslant i,j\leqslant n}$ est une base orthonormée.

II.4. Expression analytique du produit scalaire dans une base orthonormée

Prop 6:

Soit E un espace euclidien et $\mathscr{B} = (e_1, \ldots, e_n)$ une base *orthonormée* de E.

Les coordonnées $(x_1, \ldots, x_n)$ d'un vecteur x dans cette base sont données par : $\boxed{x_i = \langle e_i \,|\, x \rangle\ .}$

Ainsi : $\boxed{\forall\, x \in E\ ,\ x = \sum_{i=1}^{n} \langle e_i \,|\, x \rangle\, e_i\ .}$

Démonstration:

On a par définition des coordonnées : $x = \sum_{k=1}^{n} x_k e_k$. Donc pour tout $i \in [\![1\,;n]\!]$, compte tenu de la linéarité à droite du produit scalaire : $\langle e_i \,|\, x \rangle = \left\langle e_i \,\middle|\, \sum_{k=1}^{n} x_k e_k \right\rangle = \sum_{k=1}^{n} x_k \underbrace{\langle e_i \,|\, e_k \rangle}_{=\delta_{i,k}} = x_i$.

Soit E un espace euclidien, muni d'une base *orthonormée* $\mathscr{B} = (e_1, \ldots, e_n)$. Soient x et y deux vecteurs de E :
$$x = \sum_{i=1}^{n} x_i e_i \quad \text{et } y = \sum_{i=1}^{n} y_i e_i \quad \text{avec } (x_1, \ldots, x_n) \text{ et } (y_1, \ldots, y_n) \in \mathbb{R}^n\ .$$

On notera aussi $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ et $Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ les matrices colonnes formées des coordonnées dans $\mathscr{B}$ des vecteurs x et y.

Par bilinéarité du produit scalaire : $\langle x \,|\, y \rangle = \left\langle \sum_{i=1}^{n} x_i e_i \,\middle|\, \sum_{j=1}^{n} y_j e_j \right\rangle = \sum_{1\leqslant i,j\leqslant n} x_i y_j \underbrace{\langle e_i \,|\, e_j \rangle}_{=\delta_{i,j}}$

donc :

$$\boxed{\langle x \,|\, y \rangle = \sum_{i=1}^{n} x_i y_i = X^\top Y = Y^\top X} \quad \text{et} \quad \boxed{\|x\| = \sqrt{\sum_{i=1}^{n} x_i^2}\quad ,\quad \|x\|^2 = X^\top X\,.}$$

III. Projections orthogonales

III.1. Supplémentaire orthogonal

Théorème 6:

Soit F un sous-espace vectoriel *de dimension finie* d'un espace préhilbertien E. Alors :

1. $E = F \oplus F^\perp$.
2. $\left(F^\perp\right)^\perp = F$.

Démonstration:

1. Le résultat étant évident si $F = \{0\}$, on supposera F de dimension $n \geqslant 1$.
F possède alors une base orthonormée $(e_1, \dots, e_n)$. Soit $x \in E$. On cherche à écrire $x = y + z$ avec $y \in F$ et $z \in F^\perp$, donc on cherche $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ et $z \in F^\perp$ tels que $x = \sum_{i=1}^{n} \lambda_i e_i + z$.

$$z \in F^\perp \iff \forall j \in [\![1\,;n]\!],\ \langle e_j \,|\, z\rangle = 0 \iff \forall j \in [\![1\,;n]\!],\ \left\langle e_j \,\middle|\, x - \sum_{i=1}^{n} \lambda_i e_i \right\rangle = 0$$

$$\iff \forall j \in [\![1\,;n]\!],\ \langle e_j \,|\, x\rangle = \sum_{i=1}^{n} \lambda_i \underbrace{\langle e_j \,|\, e_i\rangle}_{=\delta_{i,j}} = \lambda_j$$

Ce qui précède démontre l'existence et l'unicité de la décomposition cherchée :

$$x = \underbrace{\sum_{i=1}^{n} \langle e_i \,|\, x\rangle\, e_i}_{\in F} + \underbrace{\left(x - \sum_{i=1}^{n} \langle e_i \,|\, x\rangle\, e_i\right)}_{\in F^\perp}.$$

2. On a déjà vu l'inclusion : $F \subset \left(F^\perp\right)^\perp$.
Réciproquement, soit $x \in \left(F^\perp\right)^\perp$. D'après le résultat précédent, x s'écrit de façon unique sous la forme $x = y + z$ avec $y \in F$ et $z \in F^\perp$. Puisque $x \in \left(F^\perp\right)^\perp$ et $z \in F^\perp$, on a $\langle x \,|\, z\rangle = 0$ soit $\langle y \,|\, z\rangle + \langle z \,|\, z\rangle = 0$.
Mais $\langle z \,|\, y\rangle = 0$ puisque $y \in F$ et $z \in F^\perp$ donc on en déduit $\langle z \,|\, z\rangle = 0$, d'où $z = 0$ et $x = y$ appartient à F,ce qui démontre l'inclusion $\left(F^\perp\right)^\perp \subset F$.

Déf 11:

Si F est un sous-espace vectoriel de dimension finie d'un espace préhilbertien E, on a : $E = F \oplus F^\perp$.
$F^\perp$ s'appelle le supplémentaire orthogonal de F.

Corollaire 6.1:

Si F et G sont deux sous-espaces vectoriels d'un espace préhilbertien de dimension finie E on a :

$$(F \cap G)^\perp = F^\perp + G^\perp.$$

Démonstration:

On a vu (proposition 2) que, si F' et G' sont des sous-espaces vectoriels de E, on a $(F' + G')^\perp = F'^\perp \cap G'^\perp$. Il suffit d'appliquer ce résultat avec $F' = F^\perp$ et $G' = G^\perp$, en utilisant le deuxième résultat du théorème précédent.

Théorème 7: Théorème de la base orthonormée incomplète

Si E est un espace préhilbertien de dimension finie, toute famille orthonormée de vecteurs de E peut être complétée en une base orthonormée.

Démonstration:

Si $(e_1, \dots, e_p)$ est une famille orthonormée dans E, soit $F = \mathrm{Vect}(e_1, \dots, e_p)$. Il suffit alors de choisir une base orthonormée dans $F^\perp$ (il en existe d'après le théorème 5).

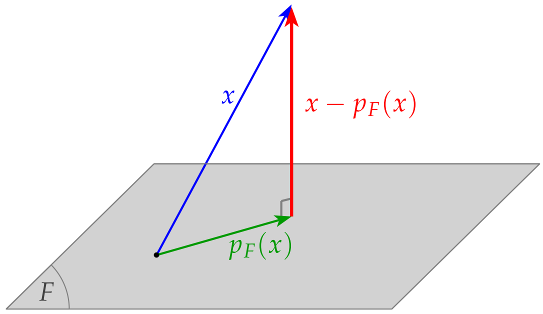
III.2. Projection orthogonale

Déf 12:

Soit F un sous-espace vectoriel de dimension finie d'un espace préhilbertien E. On appelle :

- projection orthogonale sur F, notée p_F, la projection sur F de direction $F^\perp$.
- symétrie orthogonale par rapport à F, notée s_F, la symétrie par rapport à F et de direction $F^\perp$.
  Rappel : $s_F = 2p_F - \mathrm{Id}_E$.

- On a, pour tout $x \in E$: $p_F(x) \perp x - p_F(x)$.
- Par définition d'une projection, $y = p_F(x)$ est l'unique vecteur de E vérifiant $y \in F$ et $x - y \in F^{\perp}$.



Directement d'après la démonstration du théorème 6, on obtient :

Théorème 8:

Si $(e_1, \dots, e_n)$ est une *base orthonormée* d'un sous-espace vectoriel F de dimension finie de E, on a, pour tout $x \in E$:

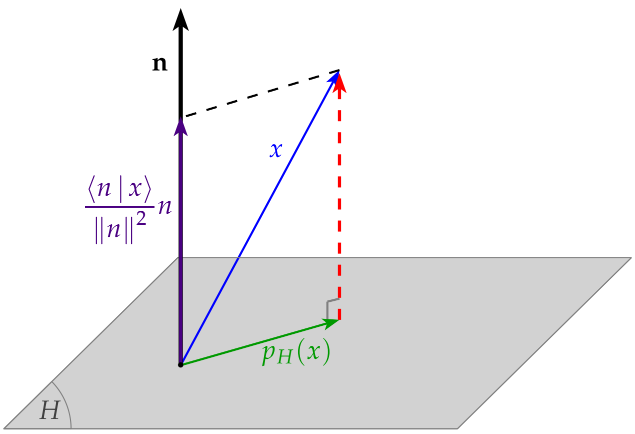
$$p_F(x) = \sum_{i=1}^{n} \langle e_i \,|\, x \rangle \, e_i \,.$$

Rem: Si l'on connaît seulement une base *orthogonale* $(e_1, \dots, e_n)$ de F, on a, pour tout $x \in E$:

$$p_F(x) = \sum_{i=1}^{n} \frac{\langle e_i \,|\, x \rangle}{\|e_i\|^2} \, e_i \,.$$

Rem: Dans le cas particulier où H est un hyperplan, et si n est un vecteur orthogonal à H, la formule précédente donne le projeté orthogonal de x sur la droite $\mathbb{K}.n$: $p_{H^{\perp}}(x) = \dfrac{\langle n \,|\, x \rangle}{\|n\|^2} \, n$ d'où :

$$p_H(x) = x - \frac{\langle n \,|\, x \rangle}{\|n\|^2} \, n \,.$$



Exercices

1. Dans $\mathbb{R}^3$ muni de sa structure euclidienne canonique, déterminer la matrice dans la base canonique de la projection orthogonale sur le plan P d'équation : $x + y + z = 0$.

Solution:

Soit p la projection orthogonale sur le plan P et $q = \mathrm{Id}_{\mathbb{R}^3} - p$ celle sur la droite D orthogonale à P. D a pour base $n = (1, 1, 1)$.
Soit $u = (x, y, z) \in \mathbb{R}^3$ et $q(u) = (x', y', z')$ son image par q. Trouver la matrice de q revient à trouver son expression analytique, c'est-à-dire à exprimer x', y' et z' en fonction de x, y et z.

D'après les formules ci-dessus :

$$q(u) = \frac{\langle n \mid u \rangle}{\|n\|^2} n = \frac{x+y+z}{3} n$$

donc on a

$$\begin{cases} x' = \dfrac{x+y+z}{3} \\ y' = \dfrac{x+y+z}{3} \\ z' = \dfrac{x+y+z}{3} \end{cases}$$

et par suite la matrice de q est : $M(q) = \frac{1}{3}\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$, puis celle de p est $M(p) = I_3 - M(q) = \frac{1}{3}\begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}$.

2. Dans $\mathbb{R}^4$ muni de sa structure euclidienne canonique, déterminer la matrice dans la base canonique de la symétrie orthogonale par rapport au sous-espace vectoriel F donné par les équations :

$$\begin{cases} x_1 + x_2 + x_3 + x_4 = 0 \\ x_1 + 2x_2 + 3x_3 + 4x_4 = 0 \end{cases}.$$

Solution:

• On remarque déjà que F est bien un sous-espace vectoriel de $\mathbb{R}^4$ (intersection d'hyperplans), et que c'est un plan (car le système d'équations est de rang 2, voir cours sur les formes linéaires).

• Si s est la symétrie orthogonale par rapport au plan F, on a $s = 2p - \mathrm{Id}_{\mathbb{R}^4}$ où p est la projection orthogonale sur F. Il suffit donc de trouver la matrice de p.

En soustrayant les deux lignes, on trouve que le système définissant F est équivalent à $\begin{cases} x_1 = x_3 + 2x_4 \\ x_2 = -2x_3 - 3x_4 \end{cases}$. En choisissant $(x_3, x_4) = (1,0)$ on trouve qu'un vecteur de F est $u_1 = (1,-2,1,0)$. Pour trouver une base orthogonale de F il reste à trouver un vecteur $u_2 = (x_1, x_2, x_3, x_4)$ de F qui est de plus orthogonal à u_1 donc qui vérifie les deux équations précédentes ainsi que l'équation $x_1 - 2x_2 + x_3 = 0$. Cela conduit à $6x_3 + 8x_4 = 0$. En prenant $x_4 = 3$ et $x_3 = -4$ on obtient $u_2 = (2,-1,-4,3)$.

On a ainsi trouvé une base orthogonale du plan F. Si $x = (x_1, x_2, x_3, x_4)$ est un vecteur quelconque de $\mathbb{R}^4$ et si l'on note $p(x) = (x_1', x_2', x_3', x_4')$, les formules précédentes donnent ;

$$\begin{aligned} p(x) &= \frac{\langle u_1 \mid x \rangle}{\|u_1\|^2} u_1 + \frac{\langle u_2 \mid x \rangle}{\|u_2\|^2} u_2 = \frac{x_1 - 2x_2 + x_3}{6} u_1 + \frac{2x_1 - x_2 - 4x_3 + 3x_4}{30} u_2 \\ &= \frac{x_1 - 2x_2 + x_3}{6}(1,-2,1,0) + \frac{2x_1 - x_2 - 4x_3 + 3x_4}{30}(2,-1,-4,3). \end{aligned}$$

On en tire les coordonnées de $p(x)$:

$$\begin{cases} x_1' = \dfrac{x_1 - 2x_2 + x_3}{6} \cdot 1 + \dfrac{2x_1 - x_2 - 4x_3 + 3x_4}{30} \cdot 2 & = \dfrac{9x_1 - 12x_2 - 3x_3 + 6x_4}{30} \\ x_2' = \dfrac{x_1 - 2x_2 + x_3}{6} \cdot (-2) + \dfrac{2x_1 - x_2 - 4x_3 + 3x_4}{30} \cdot (-1) & = \dfrac{-12x_1 + 21x_2 - 6x_3 - 3x_4}{30} \\ x_3' = \dfrac{x_1 - 2x_2 + x_3}{6} \cdot 1 + \dfrac{2x_1 - x_2 - 4x_3 + 3x_4}{30} \cdot (-4) & = \dfrac{-3x_1 - 6x_2 + 21x_3 - 12x_4}{30} \\ x_4' = \dfrac{x_1 - 2x_2 + x_3}{6} \cdot 0 + \dfrac{2x_1 - x_2 - 4x_3 + 3x_4}{30} \cdot 3 & = \dfrac{2x_1 - x_2 - 4x_3 + 3x_4}{10} \end{cases}$$

d'où la matrice de p :

$$M(p) = \frac{1}{10}\begin{pmatrix} 3 & -4 & -1 & 2 \\ -4 & 7 & -2 & -1 \\ -1 & -2 & 7 & -4 \\ 2 & -1 & -4 & 3 \end{pmatrix}$$

et enfin celle de s :

$$M(s) = 2M(p) - I_4 = \frac{1}{5}\begin{pmatrix} -2 & -4 & -1 & 2 \\ -4 & 2 & -2 & -1 \\ -1 & -2 & 2 & -4 \\ 2 & -1 & -4 & -2 \end{pmatrix}.$$

III.3. Distance à un sous-espace vectoriel

Déf 13:

Soit E un espace préhilbertien réel.
On appelle distance entre deux vecteurs x et y de E le réel

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

Déf 14:

Soit A une partie non vide d'un espace préhilbertien E.
si x est un vecteur de E, on appelle <u>distance de x à A</u> le réel :

$$d(x, A) = \inf\{d(x, a)\ ,\ a \in A\} = \inf\{\|x - a\|\ ,\ a \in A\}$$

(cette borne inférieure existe bien car c'est celle d'une partie non vide et minorée (par 0) de $\mathbb{R}$).

Dans le cas où A est un sous-espace vectoriel de dimension finie d'un espace préhilbertien, on a le résultat suivant :

Théorème 9:

Soit F est un sous-espace vectoriel de dimension finie d'un espace préhilbertien E.
Pour tout $x \in E$ il existe un et un seul vecteur $y \in F$ tel que

$$d(x, F) = d(x, y)\,.$$

Ce vecteur y est le projeté orthogonal de x sur F.

Démonstration:

• Déjà, cela a bien un sens de parler de la projection orthogonale p_F sur F : puisque F est de dimension finie, F et $F^\perp$ sont supplémentaires.

• Soit z un vecteur quelconque de F. Puisque $x - p_F(x)$ appartient à $F^\perp$ (par définition de p_F) et que $p_F(x) - z$ appartient à F (car F sous-espace vectoriel), ces deux vecteurs sont orthogonaux donc d'après la relation de Pythagore :

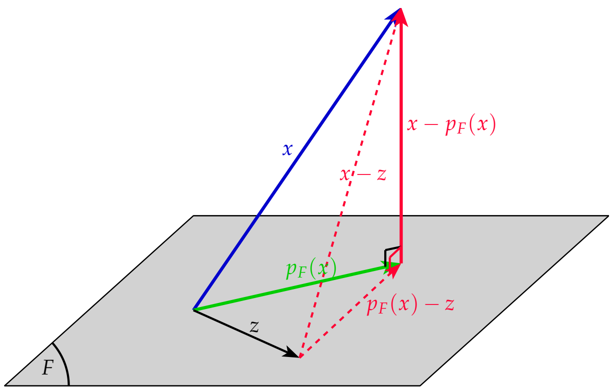
$$\|x - z\|^2 = \|x - p_F(x)\|^2 + \|p_F(x) - z\|^2 \quad (*)$$

On en déduit :

$$\forall z \in F\,,\ \|x - z\| \geqslant \|x - p_F(x)\|$$

donc

$$\|x - p_F(x)\| \leqslant \inf\{\|x - z\|\ ,\ z \in F\} = d(x, F)\,.$$



Mais comme $p_F(x)$ est aussi élément de F, cela implique $\|x - p_F(x)\| = \inf\{\|x - z\|\ ,\ z \in F\} = d(x, F)$.

• *Unicité :* S'il existe $z \in F$ tel que $d(x, z) = d(x, F) = \|x - p_F(x)\|$, alors la relation $(*)$ implique $\|p_F(x) - z\|^2 = 0$ d'où $z = p_F(x)$.

Remarques

1. En reprenant les mêmes notations, si $(e_1, \ldots, e_n)$ est une base orthonormée de F, on sait que $p_F(x) = \sum_{i=1}^{n} \langle e_i \,|\, x\rangle\, e_i$.

   On a donc $\|p_F(x)\|^2 = \sum_{i=1}^{n} \langle e_i \,|\, x\rangle^2$ puis, d'après la relation de Pythagore (les vecteurs $x - p_F(x)$ et $p_F(x)$ étant orthogonaux),

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

   ce qui donne la formule :

   $$\boxed{d(x, F)^2 = \|x\|^2 - \sum_{i=1}^{n} \langle e_i \,|\, x\rangle^2\,.}$$

2. Dans le cas particulier où F est un hyperplan H, et si n est un vecteur orthogonal à H, on a vu que le projeté orthogonal de x sur la droite $\mathbb{K}.n$ est : $p_{H^\perp}(x) = \dfrac{\langle n \,|\, x\rangle}{\|n\|^2}\, n$ d'où :

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

   (voir figure page 10).

Exercices

1. Dans $\mathbb{R}^4$ muni de sa structure euclidienne canonique, soit le vecteur $u = (1,1,1,3)$ et P le plan dont un système d'équations est : $\begin{cases} x+y+z+t=0 \\ x-y+z-t=0 \end{cases}$.

   Déterminer la distance de u à P.

   Solution:

   Le système d'équations définissant P est équivalent à $\begin{cases} x+z=0 \\ y+t=0 \end{cases}$ donc facilement, une base orthonormée de P est formée des vecteurs $e_1 = \frac{1}{\sqrt{2}}(-1,0,-1,0)$ et $e_2 = \frac{1}{\sqrt{2}}(0,1,0,-1)$.
   Inutile ici de chercher la matrice de la projection orthogonale sur P ; on se contente d'appliquer bêtement la formule précédente :

$$d(u,P) = \sqrt{\|u\|^2 - \langle e_1 \,|\, u\rangle^2 - \langle e_2 \,|\, u\rangle^2} = \sqrt{12-1-1} = \sqrt{10}.$$

2. Déterminer $(a,b,c) \in \mathbb{R}^3$ tels que $\displaystyle\int_0^1 x^2(\ln x - ax^2 - bx - c)^2 \,\mathrm{d}x$ soit minimum, et calculer la valeur de ce minimum.

   Solution:

   On se place dans l'espace vectoriel $E = \mathscr{C}([0;1],\mathbb{R})$, muni du produit scalaire usuel défini par :

$$\forall (f,g) \in E^2,\ \langle f \,|\, g\rangle = \int_0^1 f(x)g(x)\,\mathrm{d}x.$$

   Notons f l'application : $x \in\,]0;1] \mapsto x\ln x$, prolongée par continuité en 0 en posant $f(0)=0$, et $f_{a,b,c}$ l'application : $x \in [0;1] \mapsto ax^3+bx^2+cx$.

   Alors $\displaystyle\int_0^1 x^2(\ln x - ax^2 - bx - c)^2\,\mathrm{d}x = \|f - f_{a,b,c}\|^2$. Le minimum cherché est donc la distance au carré de la fonction f au sous-espace vectoriel F des fonctions de la forme $x \mapsto ax^3+bx^2+cx$ (c'est bien un sous-espace vectoriel de E de dimension finie, car engendré par les trois fonctions $x \mapsto x^k$ pour $k \in \{1,2,3\}$). D'après le théorème 9, ce minimum existe et est atteint lorsque $f_{a,b,c}$ est le projeté orthogonal de la fonction f sur F.
   Il ne reste plus qu'à chercher ce projeté. Nous n'allons pas ici utiliser la formule du cours, qui exige de connaître une base orthonormée de F. Il est plus rapide de revenir à la définition de la projection orthogonale, c'est-à-dire d'écrire que $f - f_{a,b,c}$ est orthogonal à F, donc orthogonal aux trois fonctions $x \mapsto x^k$ pour $k \in \{1,2,3\}$. Cela conduit aux équations :

$$\int_0^1 (x\ln x - ax^3 - bx^2 - cx)x^3\,\mathrm{d}x = \int_0^1 (x\ln x - ax^3 - bx^2 - cx)x^2\,\mathrm{d}x = \int_0^1 (x\ln x - ax^3 - bx^2 - cx)x\,\mathrm{d}x = 0.$$

   On calcule d'abord, en intégrant par parties et par une récurrence simple :

$$\forall k \in \mathbb{N},\ \int_0^1 x^k \ln x\,\mathrm{d}x = -\frac{1}{(k+1)^2}\,.$$

   Le système des trois équations précédentes s'écrit alors :

$$\frac{a}{7}+\frac{b}{6}+\frac{c}{5} = -\frac{1}{25} \quad,\quad \frac{a}{6}+\frac{b}{5}+\frac{c}{4} = -\frac{1}{16} \quad,\quad \frac{a}{5}+\frac{b}{4}+\frac{c}{3} = -\frac{1}{9}\,.$$

   On trouve après quelques calculs : $a = -\frac{7}{4},\ b = 4,\ c = -\frac{137}{60}$.
   Le minimum cherché vaut alors δ^2 où

$$\begin{aligned}\delta^2 &= \|f - f_{a,b,c}\|^2 = \langle f - f_{a,b,c} \,|\, f - f_{a,b,c}\rangle \underbrace{=}_{\text{astuce !}} \langle f \,|\, f - f_{a,b,c}\rangle \ (\text{car } f - f_{a,b,c} \in F^\perp \text{ donc est } \perp \text{ à } f_{a,b,c}) \\ &= \int_0^1 x\ln x(x\ln x - ax^3 - bx^2 + cx)\,\mathrm{d}x = \int_0^1 x^2(\ln x)^2\,\mathrm{d}x - a\int_0^1 x^4\ln x\,\mathrm{d}x - b\int_0^1 x^3\ln x\,\mathrm{d}x - c\int_0^1 x^2\ln x\,\mathrm{d}x \\ &= \frac{2}{27}+\frac{a}{25}+\frac{b}{16}+\frac{c}{9} = \frac{1}{2700}\,.\end{aligned}$$

III.4. Procédé d'orthonormalisation de Gram-Schmidt

Nous venons de voir dans le paragraphe précédent qu'il est important de pouvoir connaître une base orthonormée d'un espace euclidien. Le théorème ci-dessous en fournit un procédé de construction.

Théorème 10: Procédé d'orthonormalisation de Gram-Schmidt

Soit E un espace préhilbertien et $(a_k)_{k\in I}$ une famille *libre* de vecteurs de E, où I désigne $\mathbb{N}$ ou un intervalle d'entiers de la forme $[\![1\,;n]\!]$ avec $n\in\mathbb{N}^*$.

Alors il existe une et une seule famille *orthonormée* $(e_k)_{k\in I}$ telle que :

- Pour tout $k\in I$, $\mathrm{Vect}\{e_1,\dots,e_k\}=\mathrm{Vect}\{a_1,\dots,a_k\}$;
- $\forall\, k\in I$, $\langle e_k\,|\,a_k\rangle>0$ (cette condition assure l'unicité).

Démonstration:

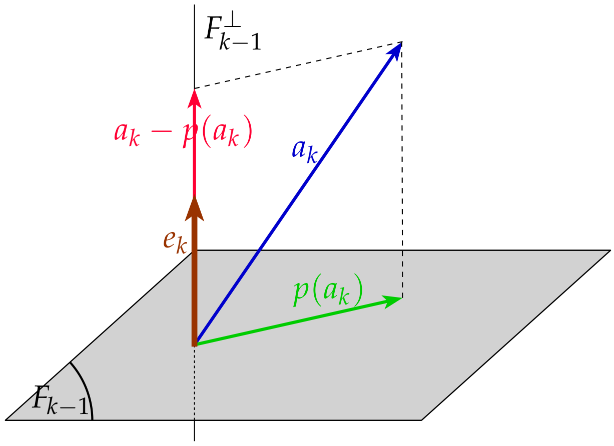
On construit la suite $(e_k)_{k\in I}$ par récurrence sur k. Pour cela, on notera pour $k\in I$, $F_k=\mathrm{Vect}\{a_1,\dots,a_k\}$. On remarque que, puisque la famille (a_k) est libre, on a $\dim F_k=k$.

- La propriété est trivialement réalisée pour $k=1$: e_1 est nécessairement égal à $\dfrac{a_1}{\|a_1\|}$ (ce qui est possible car $a_1\neq 0$).
- Supposons construits $e_1,\dots,e_{k-1}$, famille orthonormée vérifiant les conditions du théorème jusqu'au rand $k-1$.
  - La condition « $\{e_1,\dots,e_k\}$ est une famille orthogonale » est équivalente à e_k orthogonal à $e_1,\dots,e_{k-1}$ puisque l'on sait déjà $\{e_1,\dots,e_{k-1}\}$ orthonormée, c'est-à-dire à $e_k\in F_{k-1}^{\perp}$.
  - La condition « $\mathrm{Vect}\{e_1,\dots,e_k\}=\mathrm{Vect}\{a_1,\dots,a_k\}=F_k$ » est équivalente à $e_k\in F_k$ puisque l'on sait déjà que $\mathrm{Vect}\{e_1,\dots,e_{k-1}\}=\mathrm{Vect}\{a_1,\dots,a_{k-1}\}$.

  e_k appartient donc à l'orthogonal de F_{k-1} dans F_k ; cet orthogonal est une droite, celle de base le vecteur $a_k-p(a_k)$, où $p(a_k)$ est le projeté orthogonal de a_k sur F_{k-1}.
  $(e_1,\dots,e_{k-1})$ étant une base orthonormée de F_{k-1}, la formule du théorème 8 donne

  $$p(a_k)=\sum_{i=1}^{k-1}\langle e_i\,|\,a_k\rangle\,e_i\,.$$

  En notant $b_k=a_k-p(a_k)=a_k-\sum_{i=1}^{k-1}\langle e_i\,|\,a_k\rangle\,e_i$, on doit donc avoir e_k colinéaire à b_k.



  - La condition « e_k unitaire » conduit ensuite à $e_k=\varepsilon\dfrac{b_k}{\|b_k\|}$ avec $\varepsilon=\pm1$ (le vecteur b_k ne peut pas être nul, sinon on aurait $a_k=p(a_k)\in F_{k-1}$ ce qui est impossible puisque la famille $(a_1,\dots,a_k)$ est libre).
  - Enfin, la condition $\langle e_k\,|\,a_k\rangle>0$ permet de choisir ε, c'est-à-dire détermine e_k de façon unique.
    En effet, $\langle e_k\,|\,a_k\rangle=\langle e_k\,|\,a_k-p(a_k)\rangle+\underbrace{\langle e_k\,|\,p(a_k)\rangle}_{=0}=\langle e_k\,|\,b_k\rangle=\varepsilon\,\|b_k\|$, donc il faut $\varepsilon=+1$.

Remarques

1. Les formules $b_k=a_k-\sum_{i=1}^{k-1}\langle e_i\,|\,a_k\rangle\,e_i$ puis $e_k=\dfrac{b_k}{\|b_k\|}$, qui permettent de construire la famille orthonormée (e_k) de proche en proche à partir des a_k, sont à retenir (mais il est tout aussi facile de retenir la façon dont elles ont été obtenues...)
2. Si l'on veut seulement construire une famille *orthogonale* à partir des a_k, il suffit de ne pas normer les b_k à la fin du calcul, donc d'utiliser la formule
$$e_k=a_k-\sum_{i=1}^{k-1}\frac{\langle e_i\,|\,a_k\rangle}{\|e_i\|^2}e_i\,.$$
3. Si E est un espace euclidien, $\mathscr{B}=(a_1,\dots,a_n)$ une base de E et si $\mathscr{B}'=(e_1,\dots,e_n)$ est la base orthonormée de E obtenue par le procédé de Schmidt, la matrice de passage de la base $\mathscr{B}$ à la base $\mathscr{B}'$ est triangulaire supérieure à éléments diagonaux positifs. En effet :
   - pour tout $k\in[\![1\,;n]\!]$ on a $\mathrm{Vect}(a_1,\dots,a_k)=\mathrm{Vect}(e_1,\dots,e_k)$ donc la matrice est triangulaire supérieure ;
   - si l'on considère la matrice inverse, c'est-à-dire la matrice de passage $\mathrm{P}_{\mathscr{B}'}^{\mathscr{B}}$, le i-ème coefficient diagonal est la i-ème coordonnée de a_i sur e_i, donc est égal à $\langle a_i\,|\,e_i\rangle$ puisque la base (e_i) est orthonormée, et ce coefficient est positif par construction. Les coefficients diagonaux de la matrice inverse sont donc aussi positifs.

Exercice On munit $\mathbb{R}_3[X]$ du produit scalaire défini par :

$$\forall P,Q \in \mathbb{R}_3[X]\ ,\ \langle P\,|\,Q\rangle = \int_{-1}^{1} P(t)Q(t)\,\mathrm{d}t\,.$$

Déterminer la base orthonormée obtenue à partir de la base canonique par le procédé de Schmidt.

Solution:

Notons (a_0,a_1,a_2,a_3) la base canonique de $\mathbb{R}_3[X]$: $a_k = X^k$. On déroule alors les calculs, pour déterminer d'abord une base *orthogonale* (e_0,e_1,e_2,e_3).

- On prend $e_0 = a_0 = 1$;
- puis $e_1 = a_1 - \dfrac{\langle e_0\,|\,a_1\rangle}{\|e_0\|^2}e_0$. Or $\langle e_0\,|\,a_1\rangle = \displaystyle\int_{-1}^{1} e_0(t)a_1(t)\,\mathrm{d}t = \int_{-1}^{1} t\,\mathrm{d}t = 0$, donc $e_1 = a_1 = X$.
- puis $e_2 = a_2 - \dfrac{\langle e_0\,|\,a_2\rangle}{\|e_0\|^2}e_0 - \dfrac{\langle e_1\,|\,a_2\rangle}{\|e_1\|^2}e_1$. Or $\langle e_1\,|\,a_2\rangle = \langle a_1\,|\,a_2\rangle = \displaystyle\int_{-1}^{1} t^3\,\mathrm{d}t = 0$ et $\langle e_0\,|\,a_2\rangle = \displaystyle\int_{-1}^{1} t^2\,\mathrm{d}t = \frac{2}{3}$ et $\|e_0\|^2 = 2$ donc $e_2 = X^2 - \dfrac{1}{3}$.
- enfin, $e_3 = a_3 - \dfrac{\langle e_0\,|\,a_3\rangle}{\|e_0\|^2}e_0 - \dfrac{\langle e_1\,|\,a_3\rangle}{\|e_1\|^2}e_1 - \dfrac{\langle e_2\,|\,a_3\rangle}{\|e_2\|^2}e_2$. On a $\langle e_0\,|\,a_3\rangle = \langle e_2\,|\,a_3\rangle = 0$ (intégrale d'une fonction impaire sur $[-1\,;1]$), et $\langle e_1\,|\,a_3\rangle = \displaystyle\int_{-1}^{1} t^4\,\mathrm{d}t = \frac{2}{5}$ et $\|e_1\|^2 = \displaystyle\int_{-1}^{1} t^2\,\mathrm{d}t = \frac{2}{3}$, donc $e_3 = X^3 - \dfrac{3}{5}X$.
- Il reste à normer la base orthogonale obtenue. Or $\|e_0\| = \sqrt{2}$, $\|e_1\| = \sqrt{\frac{2}{3}}$, $\|e_2\| = \sqrt{\int_{-1}^{1}\left(t^2-\frac{1}{3}\right)^2\mathrm{d}t} = \frac{2}{3}\sqrt{\frac{2}{5}}$ et $\|e_3\| = \sqrt{\int_{-1}^{1}\left(t^3-\frac{3}{5}t\right)^2\mathrm{d}t} = \frac{2}{5}\sqrt{\frac{2}{7}}$, et l'on a finalement la base orthonormée :

$$\frac{\sqrt{2}}{2}\quad,\quad \sqrt{\frac{3}{2}}X\quad,\quad \frac{3}{2}\sqrt{\frac{5}{2}}\left(X^2-\frac{1}{3}\right)\quad,\quad \frac{5}{2}\sqrt{\frac{7}{2}}\left(X^3-\frac{3}{5}X\right).$$

III.5. Représentation des formes linéaires d'un espace euclidien

Théorème 11:

Soit φ une forme linéaire sur un espace vectoriel euclidien E ;
Alors il existe un et un seul vecteur $a \in E$ tel que

$$\forall x \in E\ ,\ \varphi(x) = \langle a\,|\,x\rangle\ .$$

Démonstration:

Soit $\mathscr{B} = (e_1,\dots,e_n)$ une base orthonormée de E. Alors pour tout $x = \sum_{i=1}^{n} x_i e_i$ dans E on a

$$\varphi(x) = \sum_{i=1}^{n} x_i\varphi(e_i) = \sum_{i=1}^{n} a_i x_i = \langle a\,|\,x\rangle$$

où a est le vecteur de coordonnées $a_i = \varphi(e_i)$ dans $\mathscr{B}$ (il s'agit de l'expression analytique de la forme linéaire φ).
Cela prouve l'existence de a. Enfin, a est unique car si l'on a $\langle a\,|\,x\rangle = \langle b\,|\,x\rangle$ pour tout $x \in E$, alors $a - b \in E^{\perp}$ donc $a = b$.

Rem: Lorsque φ est non nulle, son noyau $\operatorname{Ker}\varphi$ est un hyperplan H. D'après le théorème précédent, il existe donc $a \in E$ tel que

$$H = \{x \in E \text{ tq } \langle a\,|\,x\rangle = 0\} = \{a\}^{\perp}\ .$$

$\mathbb{R}.a$ est la droite orthogonale à H ; ses vecteurs sont les vecteurs <u>normaux</u> à H.