Révisions sur les espaces préhilbertiens

(1) Définition d'un produit scalaire. — Un produit scalaire sur un $\mathbf{R}$-espace vectoriel E est une application :

$$\langle \cdot\,,\,\cdot \rangle \left| \begin{array}{ccc} E\times E & \longrightarrow & \mathbf{R} \\ (x,y) & \longmapsto & \langle x\,,\,y\rangle \end{array} \right.$$

qui vérifie les axiomes suivants.

(a) $\langle \cdot\,,\,\cdot \rangle$ est *bilinéaire* :

$$\forall\,(\lambda,\mu)\in\mathbf{R}^2 \quad \forall,(x,y,z)\in E^3 \quad \begin{cases} \langle \lambda\,x+\mu\,y\,,\,z\rangle &=& \lambda\,\langle x\,,\,z\rangle+\mu\,\langle y\,,\,z\rangle \\ \langle x\,,\,\lambda\,y+\mu\,z\rangle &=& \lambda\,\langle x\,,\,y\rangle+\mu\,\langle x\,,\,z\rangle \end{cases}$$

(b) $\langle \cdot\,,\,\cdot \rangle$ est *symétrique* :

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

(c) $\langle \cdot\,,\,\cdot \rangle$ est *positive* :

$$\forall\,x\in E \quad \langle x\,,\,x\rangle\geqslant 0$$

(d) $\langle \cdot\,,\,\cdot \rangle$ est *définie* :

$$\forall\,x\in E \quad \langle x\,,\,x\rangle=0 \implies x=0_E$$

(2) Produits scalaires usuels. — Il convient de savoir démontrer que chacune des applications suivantes est un produit scalaire.

(a) $$\langle \cdot\,,\,\cdot \rangle \left| \begin{array}{ccc} \mathbf{R}^n\times\mathbf{R}^n & \longrightarrow & \mathbf{R} \\ ((x_1,\dots,x_n),(y_1,\dots,y_n)) & \longmapsto & \displaystyle\sum_{i=1}^{n} x_i\,y_i \end{array} \right.$$

(b) $$\langle \cdot\,,\,\cdot \rangle \left| \begin{array}{ccc} \mathcal{M}_{n,1}(\mathbf{R})\times\mathcal{M}_{n,1}(\mathbf{R}) & \longrightarrow & \mathbf{R} \\ (X,Y) & \longmapsto & X^\top\times Y \end{array} \right.$$

(c) $$\langle \cdot\,,\,\cdot \rangle \left| \begin{array}{ccc} \mathbf{R}[X]\times\mathbf{R}[X] & \longrightarrow & \mathbf{R} \\ (A,B) & \longmapsto & \displaystyle\sum_{i=0}^{+\infty}[A]_i\,[B]_i \end{array} \right.$$

(d) $$\langle \cdot\,,\,\cdot \rangle \left| \begin{array}{ccc} \mathcal{C}^0([0,1],\mathbf{R})\times\mathcal{C}^0([0,1],\mathbf{R}) & \longrightarrow & \mathbf{R} \\ (f,g) & \longmapsto & \displaystyle\int_0^1 f(t)\,g(t)\ \mathrm{d}t \end{array} \right.$$

(e) $$\langle \cdot\,,\,\cdot \rangle \left| \begin{array}{ccc} \mathcal{M}_{n,p}(\mathbf{R})\times\mathcal{M}_{n,p}(\mathbf{R}) & \longrightarrow & \mathbf{R} \\ (A,B) & \longmapsto & \mathrm{tr}\left(A^\top\times B\right)=\displaystyle\sum_{i=1}^{n}\sum_{j=1}^{p}[A]_{i,j}\,[B]_{i,j} \end{array} \right.$$

(3) Trois identités remarquables. — Fixons un espace préhilbertien réel $(E,\langle \cdot\,,\,\cdot \rangle)$ pour la suite. On pose :

$$\forall\,x\in E \quad ||\,x\,||:=\sqrt{\langle x\,,\,x\rangle} \qquad \text{[norme du vecteur } x\text{]}$$

On dipose de trois identités remarquables.

(a) *Développement du carré de la norme d'une somme de deux vecteurs.*

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

(b) *Développement du carré de la norme d'une somme d'un nombre fini de vecteurs.*

$$\forall\,(x_1,\dots,x_p)\in E^p \quad \left|\left|\sum_{i=1}^{p}x_i\right|\right|^2=\sum_{i=1}^{p}||\,x_i\,||^2+2\sum_{1\leqslant i<j\leqslant p}\langle x_i\,,\,x_j\rangle$$

(c) *Identité de polarisation.*

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

En conséquence, l'application $||\cdot||:E\longrightarrow\mathbf{R}_+$ détermine le produit scalaire $\langle \cdot\,,\,\cdot \rangle:E\times E\longrightarrow\mathbf{R}$.

(4) Inégalité de Cauchy-Schwarz et son cas d'égalité. — Le premier résultat saisissant est *l'inégalité de Cauchy-Schwarz et son cas d'égalité.*

(a) Pour tout $(x,y) \in E^2$, $|\langle x\,,\,y \rangle| \leqslant ||\,x\,||\;||\,y\,||$.
(b) Pour tout $(x,y) \in E^2$, $|\langle x\,,\,y \rangle| = ||\,x\,||\;||\,y\,||$ si et seulement si la famille de vecteurs (x,y) est liée.

La démonstration de ce résultat est à connaître. Dans le cas où $x \in E$ et $y \in E \setminus \{0_E\}$, on peut remarquer que l'application :

$$f \left| \begin{array}{ccc} \mathbf{R} & \longrightarrow & \mathbf{R} \\ t & \longmapsto & ||\,x + t\,y\,||^2 = ||\,y\,||^2\, t^2 + 2\,\langle x\,,\,y \rangle\, t + \left|\left|\, x^2 \,\right|\right| \end{array} \right.$$

est polynomiale de degré 2 et ne prend que des valeurs positives ou nulles, pour en déduire que :

$$\Delta(f) := 4\,\langle x\,,\,y \rangle^2 - 4\,||\,x\,||^2\;||\,y\,||^2 \leqslant 0$$

Ce résultat est un outil précieux dans des problèmes d'optimisation (e.g. recherche d'un maximum).

(5) Inégalité de Minkowski et son cas d'égalité. — De l'inégalité de Cauchy-Schwarz et de son cas d'égalité, on déduit *l'inégalité de Minkowski et son cas d'égalité.*

(a) Pour tout $(x,y) \in E^2$, $||\,x + y\,|| \leqslant ||\,x\,|| + ||\,y\,||$.
(b) Pour tout $(x,y) \in E^2$, $||\,x + y\,|| = ||\,x\,|| + ||\,y\,||$ si et seulement si $x = 0_E$ ou il existe $\lambda \in \mathbf{R}_+$ tel que $y = \lambda\, x$.

(6) Un espace préhilbertien réel est naturellement muni d'une norme. — L'application :

$$||\cdot|| \left| \begin{array}{ccc} E & \longrightarrow & \mathbf{R}_+ \\ x & \longmapsto & ||\,x\,|| := \sqrt{\langle x\,,\,x \rangle} \end{array} \right.$$

est une norme sur E.

(a) *Homogénéité.*

$$\forall\,(\lambda, x) \in \mathbf{R} \times E \quad ||\,\lambda\, x\,|| = |\lambda|\;||\,x\,||$$

(b) *Séparation.*

$$\forall\, x \in E \quad ||\,x\,|| = 0 \implies x = 0_E$$

(c) *Inégalité triangulaire (inégalité de Minkowski).*

$$\forall\,(x,y) \in E^2 \quad ||\,x + y\,|| \leqslant ||\,x\,|| + ||\,y\,||$$

En conséquence, la fonction distance :

$$\mathrm{d} \left| \begin{array}{ccc} E \times E & \longrightarrow & \mathbf{R}_+ \\ (x,y) & \longmapsto & ||\,x - y\,|| \end{array} \right.$$

est 1-lipschitzienne, i.e. :

$$\forall\,(x,y) \in E^2 \quad \big|\, ||\,x\,|| - ||\,y\,||\, \big| \leqslant ||\,x - y\,|| \qquad \text{[seconde inégalité triagulaire]}$$

(7) Vecteurs orthogonaux. — Deux vecteurs x et y de E sont dits *orthogonaux* si $\langle x\,,\,y \rangle = 0$. Dans ce cas, on note $x \perp y$.

(8) Famille orthogonale/orthonormale de vecteurs. — Soit $(x_i)_{i \in I}$ une famille de vecteurs de E.

(a) La famille $(x_i)_{i \in I}$ est dite *orthogonale* si :

$$\forall\,(i,j) \in I^2 \quad i \neq j \implies x_i \perp x_j$$

(b) La famille $(x_i)_{i \in I}$ est dite *orthonormale* ou *orthonormée* si :

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

(9) Orthogonalité et liberté. — Une famille orthogonale de vecteurs non nuls de E est libre.

En conséquence, si E est de dimension finie, une famille orthogonale de vecteurs non nuls comportant $\dim(E)$ vecteurs est une base de E.

(10) Théorème de Pythagore. — Le *théorème de Pythagore* stipule que, pour toute famille orthogonale $(x_1, \ldots, x_n)$ de vecteurs de E :

$$\left|\left|\sum_{k=1}^{n} x_k\right|\right|^2 = \sum_{k=1}^{n} ||\, x_k \,||^2$$

(11) Vecteur normalisé d'un vecteur non nul. — Le normalisé d'un vecteur non nul $x \in E$ est le vecteur $\dfrac{x}{||\, x\,||}$.

(12) Algorithme d'orthonormalisation de Gram-Schmidt. — Soit $(e_1, \ldots, e_n)$ une famille libre de vecteurs de E. On construit une famille $(\varepsilon_1', \ldots, \varepsilon_n')$ de vecteurs de E de proche en proche comme suit.

(i) ε_1' est le vecteur normalisé de $\varepsilon_1 := e_1$.

(ii) Pour tout $i \in [\![1, n-1]\!]$:

$$\varepsilon_{i+1}' \text{ est le vecteur normalisé de } \varepsilon_{i+1} := e_{i+1} - \sum_{k=1}^{i} \langle\, e_{i+1} \,,\, \varepsilon_k' \,\rangle\, \varepsilon_k'$$

Alors :

(a) la famille $(\varepsilon_1', \ldots, \varepsilon_n')$ est bien définie et orthonormale ;

(b) pour tout $i \in [\![1, n]\!]$, $\mathrm{Vect}\,(e_1, \ldots, e_i) = \mathrm{Vect}\,(\varepsilon_1', \ldots, \varepsilon_i')$.

La démonstration des propriétés (a) et (b) est à connaître (récurrence forte). De plus, on doit savoir appliquer cet algorithme, dans les cas $n = 2$ et $n = 3$, dans un espace préhilbertien réel « concret ».

(13) Matrice de passage d'une base à son orthonormalisée par l'algorithme de Gram-Schmidt. — Supposons que l'espace E est de dimension finie. En appliquant l'algorithme de Gram-Schmidt à une base $\mathcal{B} = (e_1, \ldots, e_n)$ de E, on obtient une base $\mathcal{C} = (\varepsilon_1', \ldots, \varepsilon_n')$ de E et la matrice de passage :

$$P_{\mathcal{B}\to\mathcal{C}} := \mathrm{Mat}_{\mathcal{C},\mathcal{B}}(\mathrm{id})$$

est triangulaire supérieure, avec des coefficients diagonaux strictement positifs.

(14) Espaces euclidiens et bases orthonormées. — Supposons que le $\mathbf{R}$-espace vectoriel E, muni d'un produit scalaire $\langle \cdot\,,\,\cdot \rangle$, est de dimension finie $n \geqslant 1$. On dit alors que $(E, \langle \cdot\,,\,\cdot \rangle)$ est un *espace euclidien.*

L'espace E possède une base orthonormée (conséquence de l'algorithme de Gram-Schmidt déjà expliquée).

(15) Théorème de la base orthonormée incomplète. — Si $(e_1, \ldots, e_p)$ est une famille orthonormée (donc libre) de vecteurs de E $(p < n)$, alors on peut la compléter en une famille orthonormée $(e_1, \ldots, e_p, e_{p+1}, \ldots, e_n)$ de vecteurs de E, qui est une base orthonormée de E.

(16) Coordonnées dans une base orthonormée, produit scalaire et norme. — Les trois propriétés suivantes mettent en lumière qu'il est aisé de travailler dans un espace euclidien muni d'une base orthornormée.

Munissons E d'une base orthonormée $\mathcal{B} = (e_1, \ldots, e_n)$ et considérons deux vecteurs x, y de E.

(a) *Coordonnées dans une BON.*

$$x = \sum_{i=1}^{n} \langle\, x \,,\, e_i \,\rangle\, e_i$$

(b) *Produit scalaire et norme dans une BON.* Soient $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ et $Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ les coordonnées des vecteurs x et y dans la base $\mathcal{B}$. Alors :

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

(17) Sous-espaces vectoriels orthogonaux. — On ne suppose plus le $\mathbf{R}$-espace vectoriel E de dimension finie.

Deux sous-espaces vectoriels F et G de E sont dits *orthogonaux* si :

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

(18) Sous-espaces vectoriels orthogonaux et somme directe. — Si $F_1, \ldots, F_p$ sont des sous-espaces vectoriels de E deux à deux orthogonaux, alors la somme $F_1 \oplus \ldots \oplus F_p$ est directe.

(19) Orthogonal d'une partie. — Soit A une partie de E. *L'orthogonal de* A est la partie $A^\perp$ de E définie par :

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

L'orthorgonal de A est un sous-espace vectoriel de E.

(20) Calcul de l'orthogonal d'un sous-espace vectoriel engendré. — Si $x_1, \ldots, x_p$ sont des vecteurs de E :

$$\mathrm{Vect}\,(x_1, \ldots, x_p)^\perp = \{y \in E \ : \ \langle\, y \,,\, x_1 \,\rangle = \ldots = \langle\, y \,,\, x_p \,\rangle = 0\}$$

Pour déterminer l'orthogonal du sous-espace vectoriel $\mathrm{Vect}\,(x_1, \ldots, x_p)$, on est ainsi ramené à résoudre le système linéaire :

$$\left\{\begin{array}{rcl} \langle\, y \,,\, x_1 \,\rangle & = & 0 \\ & \vdots & \\ \langle\, y \,,\, x_n \,\rangle & = & 0 \end{array}\right.$$

d'inconnue $y \in E$.

(21) Somme d'un sous-espace et de son orthogonal, orthogonal de son orthogonal. — Soit F un sous-espace vectoriel de E. Alors :

(a) la somme $F \oplus F^\perp$ est directe ;

(b) $F \subset \left(F^\perp\right)^\perp$.

Si le sous-espace vectoriel F est quelconque, il peut arriver que $F \oplus F^\perp \neq E$ et $F \neq \left(F^\perp\right)^\perp$.

(22) Projeté orthogonal d'un vecteur sur un sous-espace vectoriel de dimension finie. — Le résultat suivant est fondamental et « sa » démonstration est à connaître.

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

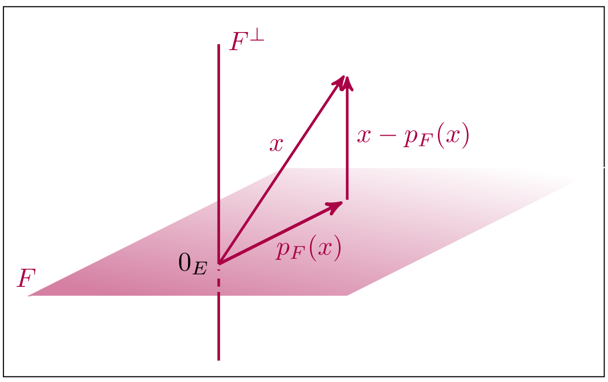
(a) Il existe un unique vecteur y de E tel que :

$$y \in F \quad \text{et} \quad x - y \in F^\perp$$

Ce vecteur y est noté $p_F(x)$ et est appelé projeté orthogonal de x sur F.

(b) Si $(f_1, \ldots, f_p)$ est une base orthonormale de F :

$$p_F(x) = \sum_{i=1}^{p} \langle\, x \,,\, f_i \,\rangle\, f_i$$



Pour calculer $p_F(x)$, on dispose de deux grandes méthodes :

- déterminer une base orthonormée de F (e.g. en appliquant l'algorithme de Gram-Schmidt à une base de F), puis appliquer la formule (b) ;
- déterminer un vecteur y tel que :

$$y \in F \qquad \text{et} \qquad x - y \in F^\perp$$

(23) Définition d'un projecteur orthogonal. — Un projecteur p de E est dit *orthogonal* si $\mathrm{Ker}\,(p) \perp \mathrm{Im}\,(p)$.

(24) Projection orthogonale sur un sous-espace vectoriel de dimension finie. — Soit F un sous-espace vectoriel de dimension finie de E.

(a) *Projection orthogonale de E sur F.* L'application :

$$p_F \left| \begin{array}{ccc} E & \longrightarrow & E \\ x & \longmapsto & p_F(x) \text{ qui est l'unique vecteur } y \in F \text{ tel que } x - y \in F^\perp \end{array} \right.$$

est un projecteur orthogonal, appelé projection orthogonale de E sur F.

(b) *Noyau et image de la projection orthogonale de E sur F.*

$$\mathrm{Ker}\,(p_F) = F^\perp \quad \text{et} \quad \mathrm{Im}\,(p_F) = \mathrm{Ker}\,(p_F - \mathrm{id}_E) = F$$

(c) *Expression de la projection orthogonale de E sur F à l'aide d'une base orthonormale de F.* Si $(f_1, \ldots, f_p)$ est une base orthonormale de F, alors :

$$\forall x \in E \quad \sum_{i=1}^{p} \langle\, x\,,\, f_i \,\rangle\, f_i$$

(25) Rappel sur la distance d'un vecteur à une partie non vide. — Soit $x \in E$ et A une partie non vide de E. La *distance de x à A* est définie par :

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

Nous savons démontrer que l'application :

$$\mathrm{d}(\,\cdot\,, A) \left| \begin{array}{ccc} E & \longrightarrow & \mathbf{R}_+ \\ x & \longmapsto & \mathrm{d}(x, A) := \inf \{ ||\, x - a \,|| \;:\; a \in A \} \end{array} \right.$$

est 1-lipschitzienne (avec des passages à la borne inférieure soignés), donc continue. En outre, nous savons établir que :

$$\forall x \in E \quad \mathrm{d}(x, A) = 0 \iff x \in \overline{A} \qquad \left[\overline{A} \text{ désigne l'adhérence de } A \right]$$

(26) Somme d'un s.e.v. de dimension finie et de son orthogonal, orthogonal de son orthogonal. — Soit F un sous-espace vectoriel de dimension finie de E. Alors :

(a) $F \oplus F^\perp = E$

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

Pour établir ces deux résultats, on peut prendre appui sur la projection orthogonale de E sur le sous-espace vectoriel F de dimension finie.

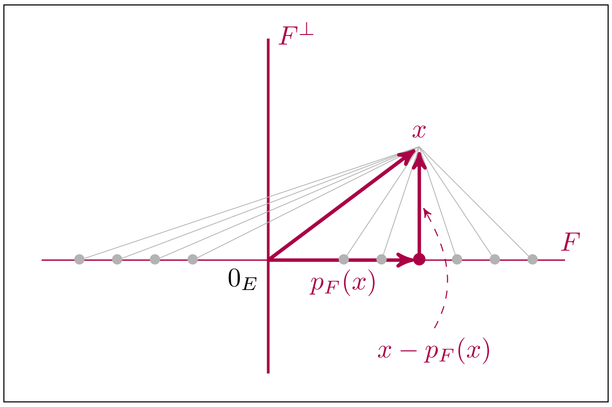
(27) Distance d'un vecteur à un sous-espace vectoriel F de dimension finie. — Soient F un sous-espace vectoriel de dimension finie de E et $x \in E$. On note $p_F : E \longrightarrow E$ la projection orthogonale de E sur F.

(a) $\mathrm{d}(x, F) = ||\, x - p_F(x) \,||$

(b) $\forall y \in F \quad y \neq p_F(x) \Longrightarrow ||\, x - y \,|| > \mathrm{d}(x, F)$

(c) $||\, x \,||^2 = ||\, p_F(x) \,||^2 + \mathrm{d}(x, F)^2$

La distance de x à F est donc atteinte en un unique point de F, qui est le projeté orthogonal $p_F(x)$ du vecteur x sur F.



Ce résultat est un outil précieux dans des problèmes d'optimisation (e.g. recherche d'un minimum).

(28) Projection sur un hyperplan d'un espace euclidien, distance d'un vecteur à cet hyperplan. — Supposons E de dimension finie (E est donc euclidien). Considérons un hyperplan H de E et v un vecteur non nul normal à H (i.e. un vecteur de base de la droite vectorielle $H^\perp$). Alors, pour tout vecteur x de E :

(a) le projeté orthogonal de x sur H est :

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

(b) la distance de x à H est :

$$\mathrm{d}(x, H) = \frac{|\langle\, x\,,\, v \,\rangle|}{||\, v \,||}$$