Chapitre 16

Projection orthogonale.

Compétences attendues.

- ✓ Déterminer l'orthogonal d'un sous-espace.
- ✓ Déterminer le projeté orthogonal d'un vecteur sur un sous-espace vectoriel.
- ✓ Utiliser une projection orthogonale pour minimiser une quantité.

Mathieu Mansuy - Professeur de Mathématiques en ECS2 au Lycée Louis Pergaud (Besançon)
`mansuy.mathieu@hotmail.fr`

1 Supplémentaire orthogonal

Dans tout ce chapitre, E désigne un espace euclidien de dimension $n \geq 1$, dont on note $\langle \cdot, \cdot \rangle$ le produit scalaire, et $\|\cdot\|$ la norme euclidienne associée.

1.1 Définition

Définition.

Soit F un sous-espace vectoriel de E. On appelle *orthogonal de* F, et on note $F^\perp$, l'ensemble des vecteurs orthogonaux à F, c'est-à-dire :

$$F^\perp = \{x \in E,\ \forall y \in F,\ \langle x, y \rangle = 0\}.$$

Ainsi on a : $x \in F^\perp \Leftrightarrow \forall y \in F,\ \langle x, y \rangle = 0.$

Propriété 1

$F^\perp$ est un sous-espace vectoriel de E orthogonal à F.

Preuve.

$\square$

Exemple. $\{0_E\}^\perp = E$ et $E^\perp = \{0_E\}$.

Propriété 2

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

(1) Si G est orthogonal à F, alors $G \subset F^\perp$. (2) $F \subset (F^\perp)^\perp$.

Preuve.

$\square$

Propriété 3

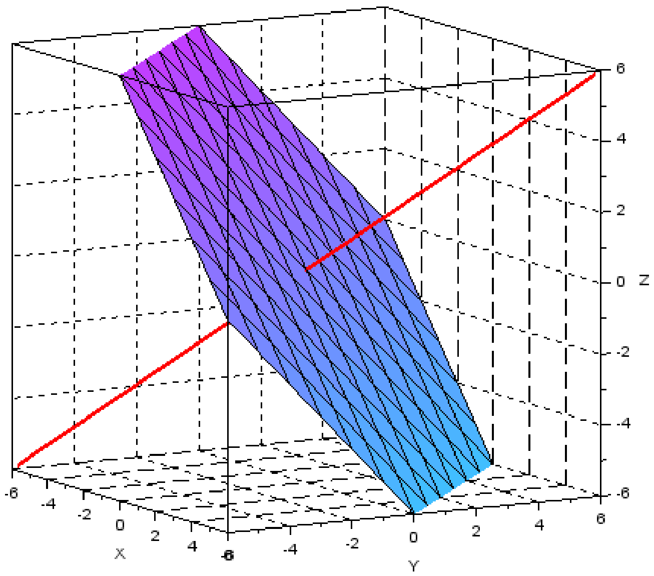
Soit $F = \text{Vect}(e_1, \dots, e_k)$ un sous-espace vectoriel de E. On a :

$$x \in F^\perp \quad \Leftrightarrow \quad \forall\, i \in [\![1, k]\!], \quad \langle x, e_i \rangle = 0.$$

Preuve.

□

Exercice. Soit $F = \{(x, y, z),\ x + y + z = 0\}$ sous-espace vectoriel de $\mathbb{R}^3$ muni du produit scalaire canonique. Déterminer $F^\perp$.



Représentation de F et $F^\perp$.

1.2 Supplémentaire orthogonal

Théorème 4

Soit F un sous-espace vectoriel de E. Alors $\dim(F^\perp) = \dim(E) - \dim(F)$, et on a :

$$E = F \oplus F^\perp.$$

On dit que $F^\perp$ est **le** *supplémentaire orthogonal* de F.

Preuve.

□

Corollaire 5

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

(2) La concaténation d'une base orthonormée $(e_1, \dots, e_p)$ de F et d'une base orthonormée $(e_{p+1}, \dots, e_n)$ de $F^\perp$ est une base orthonormée de E.

Preuve.

□

Remarque. On peut définir de même l'orthogonal d'un sous-espace vectoriel F en dimension infinie. Cependant les propriétés énoncées dans ce paragraphe ne sont alors plus vraies : on n'a pas nécessairement $E = F \oplus F^\perp$ ou encore $F = \left(F^\perp\right)^\perp$ lorsque E est de dimension infinie.

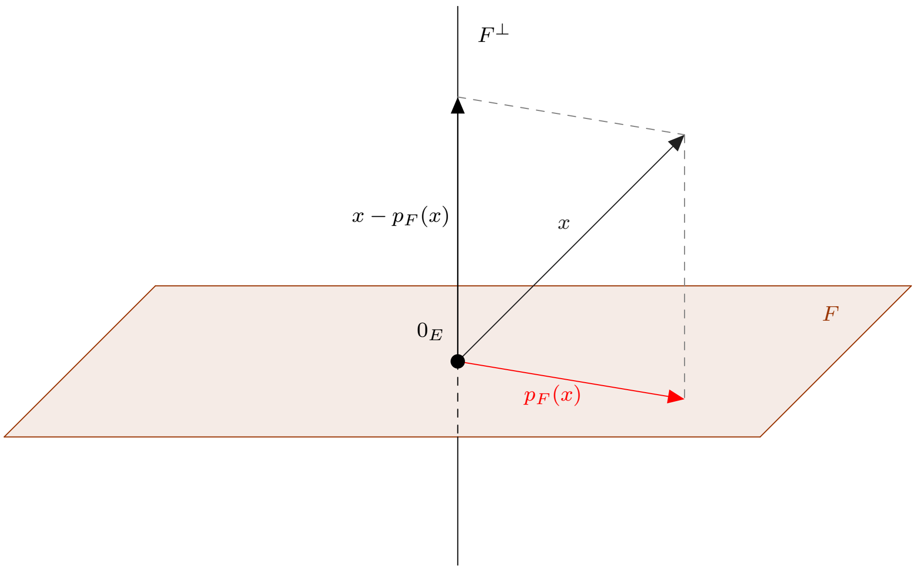
2 Projection orthogonale sur un sous-espace vectoriel de dimension finie

2.1 Projeté orthogonal

Soit F un sous-espace vectoriel de E. On vient de voir que $E = F \oplus F^{\perp}$.

Définition.

On appelle *projection orthogonale sur* F, notée p_F, la projection sur F dans la direction de $F^{\perp}$.
Pour tout $x \in E$, $p_F(x)$ est appelé *le projeté orthogonal de* x *sur* F.



Projeté orthogonal de x *sur* F.

Remarques.

- Soit p un projecteur de E. p est un projecteur orthogonal si et seulement si $\mathrm{Im}(p) \perp \mathrm{Ker}(p)$.
- Le projeté orthogonal $p_F(x)$ d'un vecteur $x \in E$ sur F est caractérisé par $\begin{cases} p_F(x) \in F \\ x - p_F(x) \in F^{\perp} \end{cases}$.

Méthode.

Pour déterminer le projeté orthogonal de $x \in E$ *sur* F*, on peut procéder ainsi :*

(i) *On détermine une base* $(u_1, \dots, u_p)$ *de* F *;*

(ii) $p_F(x) \in F$ *: il existe donc* $\lambda_1, \dots, \lambda_p \in \mathbb{R}$ *tels que* $p_F(x) = \sum_{i=1}^{p} \lambda_i u_i$ *;*

(iii) $x - p_F(x) \in F^{\perp}$ *: on a donc* $\begin{cases} \langle x - p_F(x), u_1 \rangle = 0 \\ \quad\vdots \\ \langle x - p_F(x), u_p \rangle = 0 \end{cases}$ *;*

(iv) *On résout alors ce système linéaire pour obtenir* $\lambda_1, \dots, \lambda_p$*, et* $p_F(x)$.

Exercice. Dans $\mathbb{R}^3$ muni du produit scalaire canonique, déterminer le projeté orthogonal de $u = (2, 2, 2)$ sur le sous-espace vectoriel $F = \left\{(x, y, z) \in \mathbb{R}^3,\ x - \dfrac{y}{2} + z = 0\right\}$.

2.2 Expression dans une base orthonormée de F

Propriété 6

Soit $(e_1, \ldots, e_p)$ une **base orthonormée** de F. Alors on a :

$$\forall x \in E, \quad p_F(x) = \sum_{i=1}^{p} \langle x, e_i \rangle e_i.$$

Preuve.

$\square$

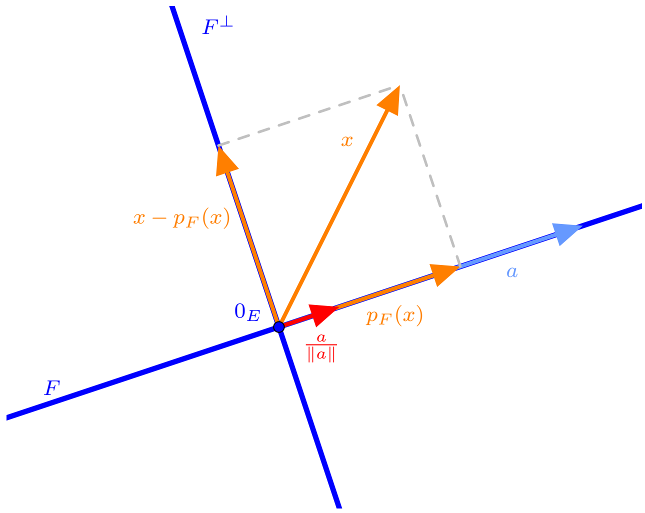
Méthode.

*Lorsqu'on dispose d'une **base orthonormée** $(e_1, \ldots, e_p)$ de F (et uniquement dans ce cas), on utilisera la formule précédente pour obtenir le projeté orthogonal de x sur F :*

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

Exercice. Obtenir le projeté orthogonal de $u = (2, 2, 2)$ sur le sous-espace $F = \mathrm{Vect}((1, 1, -1), (0, 1, 1))$.

Exercice. Soit F une droite vectorielle dirigée par $a \neq 0_E$. Exprimer le projeté orthogonal d'un vecteur x de E sur F, puis celui de x sur $F^\perp$.



Remarque. Si on ne dispose pas d'une base orthonormée de F, on pourrait procéder ainsi :

(i) on détermine une base $(x_1, \dots, x_p)$ de F ;

(ii) on l'orthonormalise par Gram-Schmidt : on obtient une base orthonormée $(e_1, \dots, e_p)$ de F ;

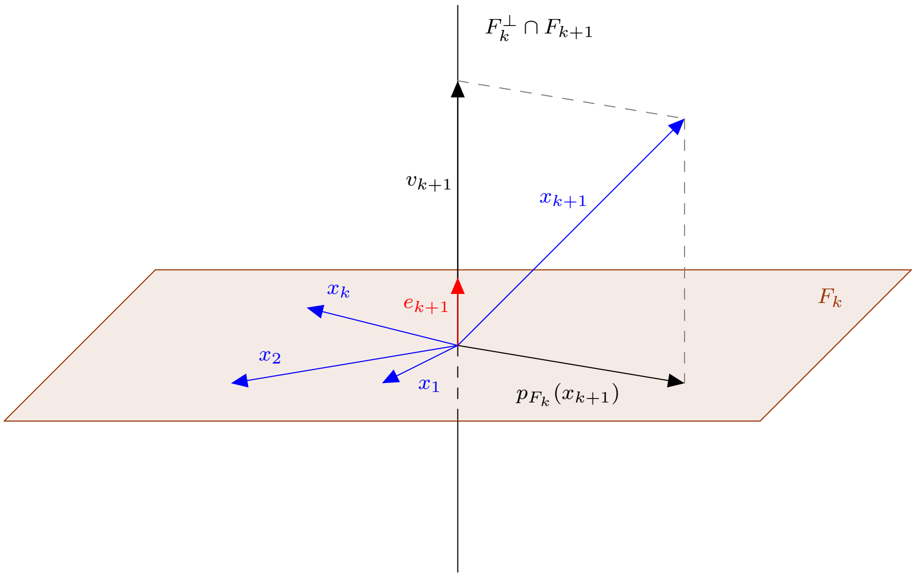
(iii) on utilise la formule précédente : $p_F(x) = \displaystyle\sum_{i=1}^{p} \langle x, e_i \rangle e_i$.

Cependant, cette méthode est en général plus longue que la méthode décrite à la section 2.1. On utilisera donc la propriété précédente **uniquement** si on dispose déjà d'une base orthonormée de F.

Remarque. Retour sur le procédé d'orthonormalisation de Gram-Schmidt.
Soit $(x_1, \dots, x_n)$ une famille libre de vecteurs de E, et notons $F_k = \mathrm{Vect}(x_1, \dots, x_k)$ pour $1 \leq k \leq n-1$. On peut réécrire le procédé d'orthonormalisation de Gram-Schmidt de la manière suivante :

- poser $e_1 = \dfrac{x_1}{||x_1||}$;
- une fois les vecteurs $e_1, \dots, e_k$ construits,
  - poser $v_{k+1} = x_{k+1} - p_{F_k}(x_{k+1})$;
  - poser $e_{k+1} = \dfrac{v_{k+1}}{||v_{k+1}||}$.

Pour « redresser » x_{k+1} en un vecteur v_{k+1} orthogonal à F_k, on lui soustrait donc sa projection orthogonale sur F_k. Reste ensuite à le normaliser pour obtenir e_{k+1}.



Algorithme de Gram-Schmidt.

Propriété 7

Soit $\mathscr{B}$ une **base orthonormée** de E, et soit $(u_1, \ldots, u_p)$ une **base orthonormée** de F. Soient $U_1, \ldots, U_p$ les vecteurs colonnes des coordonnées de $u_1, \ldots, u_p$ dans la base $\mathscr{B}$.

La matrice de p_F dans la base $\mathscr{B}$ est donnée par :

$$M_{\mathscr{B}}(p_F) = \sum_{k=1}^{p} U_k\ {}^tU_k.$$

Preuve. Puisque $(u_1, \ldots, u_p)$ est une base orthonormée de F, on a :

$$\forall x \in E, \quad p_F(x) = \sum_{k=1}^{p} \langle u_k, x\rangle u_k.$$

Puisque $\mathscr{B} = (e_1, \ldots, e_n)$ est une base orthonormée de E, on a :

$$\forall x \in E, \quad x = \sum_{j=1}^{n} \langle x, e_j\rangle e_j. \tag{$*$}$$

Notons $M_{\mathscr{B}}(p_F) = (m_{i,j})$. Pour tout $(i,j) \in [\![1,n]\!]^2$, $m_{i,j}$ est la j-ème coordonnée de $p_F(e_i)$ dans la base $\mathscr{B}$, c'est-à-dire $\langle p_F(e_i), e_j\rangle$ d'après $(*)$ pour $x = p_F(e_i)$. On a donc :

$$m_{i,j} = \langle p_F(e_i), e_j\rangle = \langle \sum_{k=1}^{p} \langle u_k, e_i\rangle u_k, e_j\rangle \underset{\text{lin. à gauche}}{=} \sum_{k=1}^{p} \langle u_k, e_i\rangle\langle u_k, e_j\rangle \underset{(*) \text{ avec } x=u_k}{=} \sum_{k=1}^{p} (U_k)_i (U_k)_j$$

Or on a : $\quad (U_k\ {}^tU_k) = \begin{pmatrix} (U_k)_1 \\ \vdots \\ (U_k)_n \end{pmatrix} \begin{pmatrix} (U_k)_1 & \ldots & (U_k)_n \end{pmatrix} = \begin{pmatrix} (U_k)_1(U_k)_1 & \ldots & (U_k)_1(U_k)_n \\ \vdots & & \vdots \\ (U_k)_n(U_k)_1 & \ldots & (U_k)_n(U_k)_n \end{pmatrix}.$

Donc $(U_k\ {}^tU_k)_{i,j} = (U_k)_i(U_k)_j$ et on obtient :

$$m_{i,j} = \sum_{k=1}^{p} (U_k)_i(U_k)_j = \sum_{k=1}^{p} \left(U_k\ {}^tU_k\right)_{i,j} = \left(\sum_{k=1}^{p} U_k\ {}^tU_k\right)_{i,j}.$$

D'où le résultat. $\square$

Exercice. Soit $E = \mathbb{R}^3$ muni du produit scalaire canonique, et soit $F = \text{Vect}((1,0,0),(0,1,1))$. Déterminer la matrice de p_F dans la base canonique de $\mathbb{R}^3$.

2.3 Distance d'un vecteur à un sous-espace

Théorème 8

Soit F un sous-espace vectoriel de E, et $x \in E$.
L'ensemble $\{\|x-y\|, y \in F\}$ admet un minimum, et ce minimum est atteint en un unique vecteur $v \in F$ qui est $v = p_F(x)$ la projection orthogonale de x sur F.

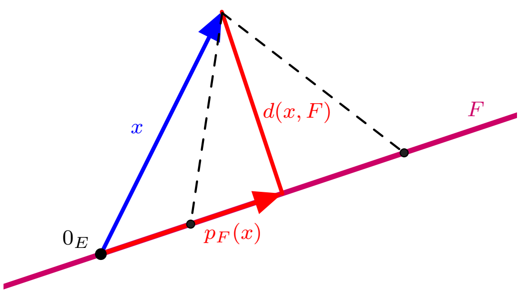
Preuve.

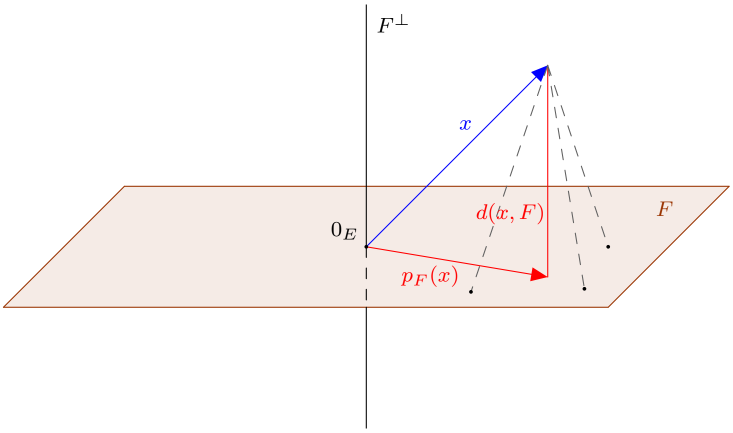
□

Définition.

On dit que $p_F(x)$ est la *meilleur approximation de x dans F au sens des moindres carrés*, c'est-à-dire pour la norme euclidienne, et on appelle *distance de x à F* le réel :

$$d(x,F) = \min\{\|x-y\|, y \in F\} = \|x - p_F(x)\|.$$





Distance de x à une droite ou un plan.

Méthode.

Afin de minimiser une quantité à l'aide d'un projeté orthogonal, on procèdera comme suit :

(i) s'il n'est pas donné dans l'énoncé, on identifie l'espace euclidien E et la norme euclidienne $\|\cdot\|$ en jeu ;

(ii) on écrit la quantité à minimiser sous la forme $\|x-u\|^2$ en identifiant x un vecteur de E fixé et u un vecteur de E qui varie dans un sous-espace F de E ;

(iii) on sait qu'un tel minimum existe, et est atteint en un unique point $u = p_F(x)$. On calcule donc le projeté orthogonal $p_F(x)$ de x sur F ;

(iv) on calcule $\|x - p_F(x)\|^2 \left(\underset{Pythagore}{=} \|x\|^2 - \|p_F(x)\|^2 \right)$ qui minimise cette quantité.

Exercice. Montrer que $\displaystyle\inf_{(a,b)\in\mathbb{R}^2} \int_{-1}^{1} (t^2 - at - b)^2 \,\mathrm{d}t$ existe et le calculer.

2.4 Pseudo-solutions d'un système linéaire

Considérons l'espace vectoriel $\mathscr{M}_{n,1}(\mathbb{R})$ muni du produit scalaire canonique :

$$\forall X, Y \in \mathscr{M}_{n,1}(\mathbb{R}), \quad \langle X, Y \rangle = {}^tXY.$$

Soit $1 \leq p \leq n$. On considère le système linéaire suivant :

$$AX = B \tag{S}$$

où $A \in \mathscr{M}_{n,p}(\mathbb{R})$ est la matrice du système qu'on supposera de rang p, $B \in \mathscr{M}_{n,1}(\mathbb{R})$ est le second membre et $X \in \mathscr{M}_{p,1}(\mathbb{R})$ le vecteur-colonne des inconnues.

On cherche à trouver l'ensemble des solutions de (S). Or ce système a plus d'équations que d'inconnues ($n \geq p$), et cet ensemble de solutions est possiblement vide lorsque $B \notin \mathrm{Im}(A)$.

Dans le cas où (S) n'admet pas de solution, on souhaite trouver un vecteur $X_0 \in \mathscr{M}_{p,1}(\mathbb{R})$ qui s'en « approche le plus », c'est-à-dire tel que $AX_0 - B$ soit « le plus proche possible » de $0_{n,1}$. On cherche ainsi X_0 satisfaisant :

$$\|AX_0 - B\| = \min_{X \in \mathscr{M}_{p,1}(\mathbb{R})} \|AX - B\|.$$

Propriété 9

Soient $1 \leq p \leq n$ et soit $A \in \mathscr{M}_{n,p}(\mathbb{R})$ de rang p, et $B \in \mathscr{M}_{n,1}(\mathbb{R})$.
Il existe un unique vecteur $X_0 \in \mathscr{M}_{p,1}(\mathbb{R})$ minimisant $\{\|AX - B\|\,,\ X \in \mathscr{M}_{p,1}(\mathbb{R})\}$.

Preuve. Soit $f : \begin{cases} \mathscr{M}_{p,1}(\mathbb{R}) & \to \mathscr{M}_{n,1}(\mathbb{R}) \\ X & \mapsto AX \end{cases}$ l'application canoniquement associée à A, et soit $F = \mathrm{Im}(f)$. Alors l'ensemble

$$\{\|Y - B\|,\ Y \in F\}$$

possède un minimum atteint en un unique point $Y_0 \in \mathrm{Im}(f)$ qui est le projeté orthogonal de B sur $\mathrm{Im}(f)$.
Par le théorème du rang, on a $\dim(\mathrm{Ker}(f)) = \dim(\mathscr{M}_{p,1}(\mathbb{R})) - \mathrm{rg}(f) = p - \mathrm{rg}(A) = p - p = 0$. Donc f est injective, et il existe un unique $X_0 \in \mathscr{M}_{p,1}(\mathbb{R})$ tel que $Y_0 = AX_0$. $\square$

Remarques.

- Le vecteur X_0 du théorème précédent est ce qui « se rapproche le plus » d'une solution de $AX = B$ au sens des moindres carrés (i.e. de la norme euclidienne). On parle de *pseudo-solution* du système linéaire. Pour en apprendre plus à ce sujet, on pourra consulter l'épreuve d'ESSEC 2012.
- On peut montrer que $X_0 = ({}^tAA)^{-1}AB$ (formule non exigible qu'on démontrera en TP). Entre autres, ces résultats nous permettront en TP d'approximer un nuage de points par une droite « s'en approchant le plus », appelée *droite des moindres carrés*.