Chapitre 32

Espaces préhilbertiens

1 Produit scalaire et norme euclidienne **2**
1.1 Produit scalaire 2
1.2 Inégalité de Cauchy-Schwarz 3
1.3 Norme associée à un produit scalaire . . 4

2 Orthogonalité **5**
2.1 Vecteurs et sous-espaces orthogonaux . . 5
2.2 Familles orthogonales, orthonormales . . 7
2.3 Orthonormalisation de Gram-Schmidt . . 9

3 Bases orthonormées d'un espace euclidien **10**
3.1 Existence de bases orthonormées 10
3.2 Calculs dans une base orthonormée . . . 11

4 Projection orthogonale **12**
4.1 Orthogonal d'une partie 12
4.2 Supplémentaire orthogonal 13
4.3 Projection orthogonale 13
4.4 Expression dans une base orthonormée . 15
4.5 Cas particulier des hyperplans 16
4.6 Distance d'un vecteur à un sous-espace . 17
4.7 Pseudo-solution d'un système linéaire . . 18

Compétences attendues.

- ✓ Montrer qu'une application $\varphi : E \times E \to \mathbb{R}$ est un produit scalaire.
- ✓ Savoir reconnaitre et appliquer l'inégalité de Cauchy-Schwartz.
- ✓ Montrer qu'une famille de vecteurs est orthogonale ou orthonormale.
- ✓ Construire une base orthonormée à l'aide de l'algorithme de Gram-Schmidt.
- ✓ Calculer les coordonnées d'un vecteur et sa norme dans une base orthonormale.
- ✓ Déterminer l'orthogonal d'un sous-espace.
- ✓ Déterminer le projeté orthogonal d'un vecteur sur un sous-espace vectoriel.
- ✓ Utiliser un projeté orthogonal pour minimiser une quantité.

Mathieu Mansuy - Professeur en MP2I au Lycée Carnot (Dijon)
Page personnelle : `mathieu-mansuy.fr/`
E-mail : `mathieu.mansuy@ac-dijon.fr`

1 Produit scalaire et norme euclidienne

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

1.1 Produit scalaire

Définition.

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

- φ est *bilinéaire* : $\forall\, x, x', y, y' \in E,\ \forall\, \lambda, \in \mathbb{R},$

$$\varphi(\lambda x + x', y) = \lambda\varphi(x, y) + \varphi(x', y) \quad \text{et} \quad \varphi(x, \lambda y + y') = \lambda\varphi(x, y) + \varphi(x, y').$$

- φ est *symétrique* : $\forall\, x, y \in E, \quad \varphi(y, x) = \varphi(x, y).$
- φ est *positive* : $\forall\, x \in E, \quad \varphi(x, x) \geq 0.$
- φ est *définie positive* : $\forall\, x \in E, \quad \varphi(x, x) = 0 \quad \Leftrightarrow \quad x = 0_E.$

✎ **Notation.**

On note alors $\varphi(x, y) = \langle x, y\rangle$ (ou parfois $(x|y)$, ou tout simplement $x \cdot y$).

Remarques.

- Pour montrer la bilinéarité, on se contentera de prouver la linéarité par rapport à la première variable et la symétrie. En effet, on obtient alors directement la linéarité par rapport à la deuxième variable puisque pour tout $(x, y, y') \in E^3$, pour tout $\lambda \in \mathbb{R}$:

$$\varphi(x, \lambda y + \mu y') \underbrace{=}_{\text{symétrie}} \varphi(\lambda y + \mu y', x) \underbrace{=}_{\text{lin. } 1^{ere} \text{ var.}} \lambda\varphi(y, x) + \mu\varphi(y', x) \underbrace{=}_{\text{symétrie}} \lambda\varphi(x, y) + \mu\varphi(x, y').$$

- On montrera en détails le caractère défini positif, souvent le point non trivial. Notons que l'implication $x = 0_E \Rightarrow \varphi(x, x) = 0$ est toujours satisfaite pour une forme bilinéaire, il est donc inutile de s'y attarder.

Exemples classiques.

- *Produit scalaire sur* $\mathbb{R}^2$. Pour tout $x = (x_1, x_2), y = (y_1, y_2) \in \mathbb{R}^2$, on pose :

$$\langle x, y\rangle = x_1y_1 + x_2y_2.$$

- *Produit scalaire canonique sur* $\mathbb{R}^n$. Pour tout $x = (x_1, \ldots, x_n), y = (y_1, \ldots, y_n) \in \mathbb{R}^n$, on pose :

$$\langle x, y\rangle = \sum_{i=1}^{n} x_iy_i.$$

- *Produit scalaire sur* $\mathscr{M}_{n,1}(\mathbb{R})$. Pour tout $X, Y \in \mathscr{M}_{n,1}(\mathbb{R})$, on pose :

$$\langle X, Y\rangle = X^\top \times Y.$$

Exercice 1 Montrer que les applications suivantes définissent un produit scalaire sur E :

- $E = \mathscr{C}([a,b],\mathbb{R})$, $\langle f, g\rangle = \displaystyle\int_a^b f(t)g(t)\,\mathrm{d}t$;
- $E = \mathbb{R}_n[X]$, $\langle P, Q\rangle = \displaystyle\sum_{i=0}^{n} P(i)Q(i)$;
- $E = \mathscr{M}_n(\mathbb{R})$, $\langle A, B\rangle = \mathrm{tr}(A^\top B)$.

Exemple. L'ensemble des variables aléatoires discrètes sur $(\Omega, \mathscr{A}, P)$ admettant une variance est un espace vectoriel E, sur lequel $\mathrm{Cov} : (X,Y) \mapsto \mathrm{Cov}(X,Y)$ est une forme bilinéaire symétrique et positive. Elle n'est cependant pas définie positive car :

$$\mathrm{Cov}(X,X) = 0 \quad \Leftrightarrow \quad V(X) = 0 \quad \Leftrightarrow \quad X \text{ variable certaine.}$$

Ainsi X n'est pas nécessairement la variable aléatoire nulle.

Définition.

Un $\mathbb{R}$-espace vectoriel E muni d'un produit scalaire $\langle\cdot,\cdot\rangle$ est appelé *espace préhilbertien réel*, et noté $(E, \langle\cdot,\cdot\rangle)$.

Le saviez-vous ?

Le terme préhilbertien a été donné en hommage au mathématicien David Hilbert (1862 - 1943). Il fut en effet le premier à s'intéresser aux espaces qu'on appelle aujourd'hui espaces de Hilbert ou espaces hilbertiens, et qui sont « le bon cadre » pour formuler la mécanique quantique. Il s'agit d'espaces vectoriels munis d'un produit scalaire et vérifiant une hypothèse supplémentaire (appelée complétude). Si on enlève cette dernière hypothèse, on obtient donc les espaces pré-hilbertiens.

Hilbert a eu une grande influence sur les mathématiques du 20ème siècle, notamment à travers la liste des 23 problèmes (aujourd'hui connus sous le nom de problèmes de Hilbert) qu'il présenta en 1900 au congrès international des mathématiciens, et dont une dizaine ne sont toujours pas complètement résolus 120 ans plus tard.

David Hilbert (1862 - 1943).

1.2 Inégalité de Cauchy-Schwarz

Théorème 1 (*Inégalité de Cauchy* (1789 - 1857) - *Schwarz* (1843 - 1921))

Soit $(E, \langle\cdot,\cdot\rangle)$ un espace préhilbertien réel. Alors :

$$\forall\, x, y \in E, \quad |\langle x, y\rangle| \leq \sqrt{\langle x, x\rangle}\sqrt{\langle y, y\rangle}.$$

De plus, il y a égalité si, et seulement si, x et y sont colinéaires.

Remarque. L'inégalité de Cauchy-Schwarz prend des formes variées selon l'espace préhilbertien $(E, \langle\cdot,\cdot\rangle)$ dans lequel on travaille. Par exemple,

- dans $E = \mathbb{R}^n$ muni du produit scalaire canonique :

$$\forall (x_1, \dots, x_n), (y_1, \dots, y_n) \in \mathbb{R}^n, \quad \left| \sum_{i=1}^n x_i y_i \right| \leq \sqrt{\sum_{i=1}^n x_i^2} \sqrt{\sum_{i=1}^n y_i^2}.$$

- dans $E = \mathscr{C}([a,b], \mathbb{R})$ muni du produit scalaire défini plus haut :

$$\forall f, g \in \mathscr{C}([a,b], \mathbb{R}), \quad \left| \int_a^b f(t)g(t)\,\mathrm{d}t \right| \leq \sqrt{\int_a^b f(t)^2\,\mathrm{d}t} \sqrt{\int_a^b g(t)^2\,\mathrm{d}t}.$$

1.3 Norme associée à un produit scalaire

Définition.

Soit $(E, \langle \cdot, \cdot \rangle)$ un espace préhilbertien réel. On appelle *norme associée au produit scalaire* $\langle \cdot, \cdot \rangle$ l'application $\| \cdot \| : E \to \mathbb{R}_+$ définie par :

$$\forall\, x \in E, \quad \|x\| = \sqrt{\langle x, x \rangle}.$$

On appelle *norme du vecteur* $x \in E$ le réel positif $\|x\|$, et on dira que x est *unitaire* si $\|x\| = 1$.

Exemples.

- Dans $\mathbb{R}^2$ muni du produit scalaire usuel : $\|(2,1)\| = \sqrt{2^2 + 1^2} = \sqrt{5}$.
- Dans $\mathscr{C}([0,1])$ muni du produit scalaire $\langle f, g \rangle = \displaystyle\int_0^1 f(t)g(t)\,\mathrm{d}t$:

$$\|\exp\|^2 = \int_0^1 e^t \times e^t\,\mathrm{d}t = \int_0^1 e^{2t}\,\mathrm{d}t = \frac{e^2 - 1}{2} \quad \text{et donc} \quad \|\exp\| = \sqrt{\frac{e^2-1}{2}}.$$

Remarque. L'inégalité de Cauchy-Schwarz se récrit à l'aide de normes de la façon suivante :

$$\forall (x, y) \in E^2, \quad |\langle x, y \rangle| \leq \|x\| \cdot \|y\|.$$

Propriété 2

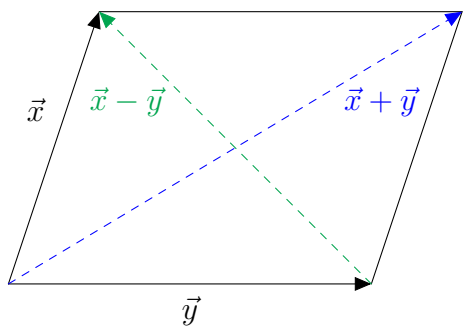
(1) *Identité remarquable :* $\forall x, y \in E,\ \|x+y\|^2 = \|x\|^2 + 2\langle x, y \rangle + \|y\|^2$.

(2) *Identité de polarisation :* $\forall x, y \in E,\ \langle x, y \rangle = \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2\right)$.

(3) *Identité du parallélogramme :* $\forall x, y \in E,\ \|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2)$.

Remarques.

- Lorsqu'une norme est associée à un produit scalaire, connaître la norme permet de retrouver le produit scalaire grâce à l'identité de polarisation. Cela peut être utile en exercice : si on dispose d'une propriété satisfaite par la norme, on pourra en déduire une propriété analogue pour le produit scalaire associé.
- D'après l'identité du parallélogramme, la somme des carrés des diagonales d'un parallélogramme est égale à la somme des carrés des côtés.



Propriété 3

La norme associée à $\langle \cdot, \cdot \rangle$ satisfait les trois propriétés suivantes :

(1) *séparation :* $\forall x \in E,\ \|x\| \geq 0$ et $\|x\| = 0 \Leftrightarrow x = 0_E$;

(2) *homogénéité :* $\forall \lambda \in \mathbb{R}, \forall x \in E,\ \|\lambda \cdot x\| = |\lambda| \cdot \|x\|$;

(3) *inégalité triangulaire :* $\forall x, y \in E,\ \|x + y\| \leq \|x\| + \|y\|$.

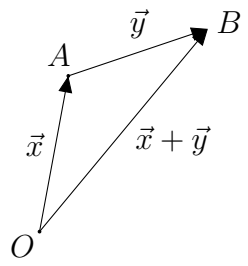
Pour aller plus loin.

Plus généralement, les propriétés évoquées à la proposition précédente définissent ce qu'on appelle une *norme* sur l'espace vectoriel E, notion que vous étudierez davantage l'an prochain.

Remarques.

- Comme conséquence de l'homogénéité, si $x \in E$ est non nul, alors $\dfrac{x}{\|x\|}$ est unitaire. On dit qu'on a *normalisé* le vecteur x.
- L'inégalité triangulaire dans $\mathbb{R}^2$ donne l'inégalité entre distances :

$$OB \leq OA + AB.$$



Propriété 4

- *Cas d'égalité :* $\|x + y\| = \|x\| + \|y\| \Leftrightarrow (x = 0_E)$ ou $(\exists\, \alpha \geq 0,\ y = \alpha \cdot x)$.
- *Deuxième inégalité triangulaire :* $\forall x, y \in E,\ |\|x\| - \|y\|| \leq \|x + y\|$.

Remarque. Sur $(E, \langle \cdot, \cdot \rangle)$ muni de la norme euclidienne, on définit la *distance* entre deux éléments x et y de E par le réel positif $d(x, y) = \|x - y\|$. On vérifie alors aisément les propriétés suivantes :

- *séparation :* $\forall x, y \in E,\quad d(x, y) = 0 \Leftrightarrow x = y$;
- *symétrie :* $\forall x, y \in E,\quad d(x, y) = d(y, x)$;
- *inégalité triangulaire :* $\forall x, y, z \in E,\quad d(x, z) \leq d(x, y) + d(y, z)$.

2 Orthogonalité

2.1 Vecteurs orthogonaux, sous-espaces orthogonaux

Définition.

Soit $(E, \langle \cdot, \cdot \rangle)$ un espace préhilbertien réel. On dit que les vecteurs $x, y \in E$ sont *orthogonaux* si $\langle x, y \rangle = 0$. On note parfois $x \perp y$.

Remarque. Dans $\mathbb{R}^2$, vous avez vu au lycée que le produit scalaire canonique de deux vecteurs non nuls x et y s'écrit sous la forme :

$$\langle x, y\rangle = \|x\|\,\|y\| \cos\widehat{(x,y)}$$

où $\widehat{(x,y)} \in [0,\pi]$ désigne l'angle non orienté entre les vecteurs x et y. Ainsi :

$$x \perp y \quad \Leftrightarrow \quad x = 0_{\mathbb{R}^2} \ \text{ ou } \ y = 0_{\mathbb{R}^2} \ \text{ ou } \ \widehat{(x,y)} = \frac{\pi}{2}.$$

Exemples.

- Le vecteur nul est orthogonal à tout vecteur de E.
- Dans $\mathbb{R}^3$ muni du produit scalaire canonique, $u = (1,0,1)$ et $v = (-1,-5,1)$ sont orthogonaux car :

$$\langle u, v\rangle = 1 \times (-1) + 0 \times (-5) + 1 \times 1 = 0.$$

- Dans $\mathscr{C}([-\pi,\pi])$ muni du produit scalaire $\langle f, g\rangle = \displaystyle\int_{-\pi}^{\pi} f(t)g(t)\,\mathrm{d}t$, $f = \cos$ et $g = \sin$ sont orthogonaux car :

$$\langle f, g\rangle = \int_{-\pi}^{\pi} \cos(t)\sin(t)\,\mathrm{d}t = 0$$

  puisque $f \times g : [-\pi,\pi] \to \mathbb{R}$ est une fonction impaire.

Propriété 5

- Si $x \in E$ est orthogonal à tout vecteur de E, alors $x = 0_E$.
- Pour tout $x, y \in E$:

$$(\forall a \in E,\ \langle a, x\rangle = \langle a, y\rangle) \quad \Rightarrow \quad (x = y).$$

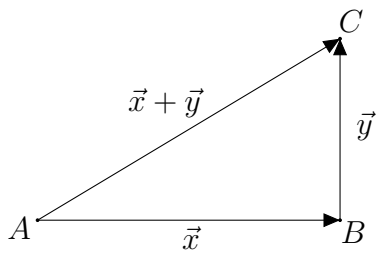
Propriété 6 (*Théorème de Pythagore* (580 av. J.-C. 495 av. J.-C.))

Soit $(E, \langle\cdot,\cdot\rangle)$ un espace pré-hilbertien réel. Alors :

$$x \text{ et } y \text{ sont orthogonaux} \quad \Leftrightarrow \quad \|x+y\|^2 = \|x\|^2 + \|y\|^2$$

Illustration graphique. On retrouve comme cas particulier le théorème de Pythagore dans le plan :

$$AC^2 = AB^2 + BC^2.$$

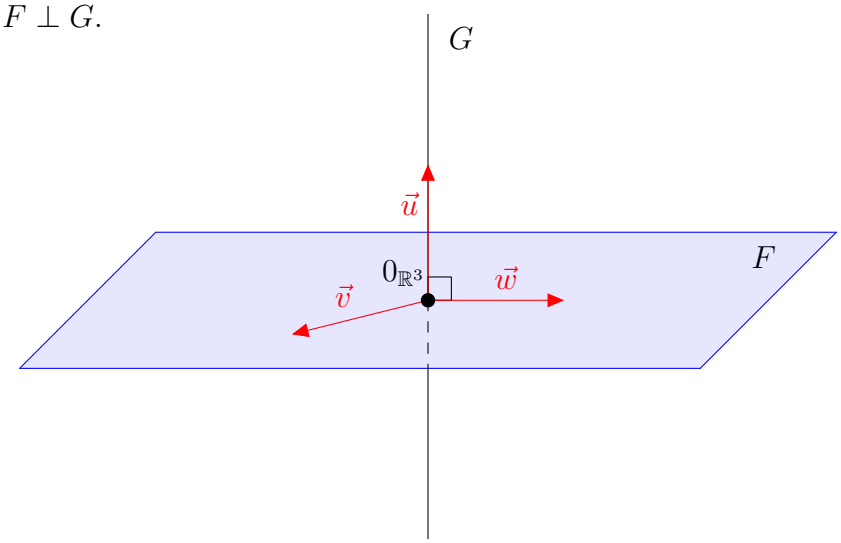


Définition.

Soient F et G deux sous-espaces vectoriels de E. On dit que F et G sont *orthogonaux* si :

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

Autrement dit, F et G sont orthogonaux si tout vecteur de F est orthogonal à tout vecteur de G. On note alors parfois $F \perp G$.



Deux sous-espaces orthogonaux de $\mathbb{R}^3$.

Propriété 7

On suppose que $F = \mathrm{Vect}(x_1, \dots, x_p)$ et $G = \mathrm{Vect}(y_1, \dots, y_q)$. Alors :

$$F \perp G \quad \Leftrightarrow \quad \forall (i,j) \in [\![1,p]\!] \times [\![1,q]\!], \ \langle x_i, y_j \rangle = 0.$$

Propriété 8

Si F et G sont orthogonaux, alors F et G sont en somme directe (i.e. $F \cap G = \{0_E\}$).

2.2 Familles orthogonales, familles orthonormales

Définition.

Soit $(E, \langle \cdot, \cdot \rangle)$ un espace préhilbertien réel. Considérons une famille $(x_i)_{i \in I}$ de vecteurs de E.

- On dit que $(x_i)_{i \in I}$ est *orthogonale* si :

$$\forall (i,j) \in I^2, \quad i \neq j \ \Rightarrow \ \langle x_i, x_j \rangle = 0.$$

- On dit que $(x_i)_{i \in I}$ est *orthonormale* ou *orthonormée* si elle est orthogonale et formée de vecteurs unitaires. Autrement dit, $(x_i)_{i \in I}$ est une famille orthonormée si, et seulement si :

$$\forall (i,j) \in I^2, \quad \langle x_i, x_j \rangle = \delta_{i,j} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases}.$$

Remarque. Si $(x_1, \dots, x_k)$ est une famille orthogonale ne contenant pas le vecteur nul, alors $\left(\frac{x_1}{\|x_1\|}, \cdots, \frac{x_k}{\|x_k\|}\right)$ est une famille orthonormale.

Exemples.

- La base canonique de $\mathbb{R}^n$ est orthonormée pour le produit scalaire canonique.

- Soit E un espace vectoriel de dimension finie, et soit $\mathscr{B} = (e_1, \dots, e_n)$ une base de E. On vérifie que l'application $\varphi_{\mathscr{B}}$ définie par :
$$\varphi_{\mathscr{B}}\left(\sum_{i=1}^{n} x_i e_i, \sum_{i=1}^{n} y_i e_i\right) = \sum_{i=1}^{n} x_i y_i$$
est un produit scalaire sur E, et que $\mathscr{B}$ est orthonormée pour $\varphi_{\mathscr{B}}$.

Exercice 2 On considère $E = \mathscr{C}([0, 2\pi])$ muni du produit scalaire $\langle f, g\rangle = \int_0^{2\pi} f(t)g(t)\,\mathrm{d}t$. Pour tout $k \in \mathbb{N}$, on note $f_k : \begin{array}{ccc} [0, 2\pi] & \to & \mathbb{R} \\ x & \mapsto & \cos(kx) \end{array}$. Montrer que $(f_k)_{k\in\mathbb{N}}$ est une famille orthogonale de $(E, \langle\cdot,\cdot\rangle)$. En déduire une famille orthonormée de $(E, \langle\cdot,\cdot\rangle)$.

Propriété 9 (Généralisation du théorème de Pythagore à n vecteurs orthogonaux)

Si $(x_1, \dots, x_k)$ est une famille orthogonale de vecteurs de E, alors :
$$\left\|\sum_{i=1}^{k} x_i\right\|^2 = \sum_{i=1}^{k} \|x_i\|^2 .$$

Mise en garde.

La réciproque du théorème de Pythagore est vraie lorsque $k = 2$, mais fausse en général lorsque $k \geq 3$.

Propriété 10

Toute famille orthogonale de vecteurs **non nuls** de E est libre. En particulier, toute famille orthonormale est libre.

2.3 Orthonormalisation de Gram-Schmidt

Propriété 11 (*Orthonormalisation de Gram* (1850 - 1916) *-Schmidt* (1876 - 1959))

Soit $(x_1, \dots, x_n)$ une famille libre de vecteurs de E. On définit par récurrence une famille $(e_1, \dots, e_n)$ comme suit :

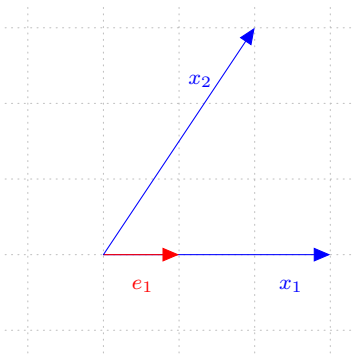
- **Étape 1.** On pose $e_1 = \dfrac{x_1}{\|x_1\|}$.
- **Étape $k+1$.** Une fois les vecteurs $e_1, \dots, e_k$ construits :
  (a) on pose $v_{k+1} = x_{k+1} - \langle e_1, x_{k+1}\rangle e_1 - \dots - \langle e_k, x_{k+1}\rangle e_k$;
  (b) on pose $e_{k+1} = \dfrac{v_{k+1}}{\|v_{k+1}\|}$.

Alors :

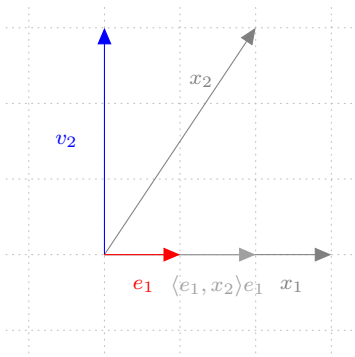
- $(e_1, \dots, e_n)$ est une famille orthonormée ;
- $\mathrm{Vect}(e_1, \dots, e_k) = \mathrm{Vect}(x_1, \dots, x_k)$ pour tout $1 \leq k \leq n$.

Illustration graphique. Procédé d'orthonormalisation de Gram-Schmidt dans le plan.

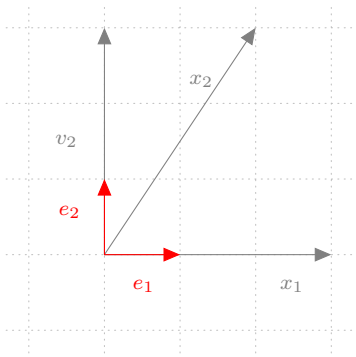
Étape 1 :
Normalisation de x_1 en e_1.



Étape 2.(a) :
« Redressement » de x_2 en v_2.



Étape 2.(b) :
Normalisation de v_2 en e_2.



Orthonormalisation de Gram-Schmidt de la famille (x_1, x_2).

Le procédé d'orthonormalisation de Gram-Schmidt consiste donc à chaque étape :

- tout d'abord à « redresser » le vecteur x_{i+1} en un vecteur v_{i+1} orthogonal aux $x_1, \dots, x_i$;
- ensuite à normaliser v_{i+1} pour obtenir e_{i+1}.

Remarques.

- Si $(x_1, \dots, x_k)$ est déjà orthonormale, alors $(e_1, \dots, e_k) = (x_1, \dots, x_k)$.
- Si $(x_1, \dots, x_k)$ est déjà orthogonale, alors $(e_1, \dots, e_k) = \left(\frac{x_1}{\|x_1\|}, \dots, \frac{x_k}{\|x_k\|}\right)$.

Exercice 3 On considère $\mathbb{R}_2[X]$ muni du produit scalaire $\langle P, Q\rangle = \sum_{k=0}^{2} P(k)Q(k)$. Orthonormaliser la base canonique par le procédé de Gram-Schmidt.

3 Bases orthonormées d'un espace euclidien

3.1 Existence de bases orthonormées d'un espace euclidien

Définition.

On appelle *espace euclidien* tout espace préhilbertien réel de dimension finie.

Définition.

Une famille $(e_1, \ldots, e_n)$ est une *base orthonormée* (ou *orthonormale*) d'un espace euclidien E s'il s'agit d'une famille orthonormée et d'une base de E. Autrement dit, c'est une base de E satisfaisant :

$$\forall (i,j) \in [\![1,n]\!]^2, \quad \langle e_i, e_j\rangle = \delta_{i,j} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{sinon} \end{cases}.$$

Exemples.

- La base canonique de $\mathbb{R}^n$ est une base orthonormée pour le produit scalaire usuel.
- La famille $\left(\frac{1}{\sqrt{3}}, \frac{X-1}{\sqrt{2}}, \sqrt{\frac{3}{2}}\left(X^2 - 2X + \frac{1}{3}\right)\right)$ est une base orthonormée de $\mathbb{R}_2[X]$ muni du produit scalaire $\langle P, Q\rangle = \sum_{k=0}^{2} P(k)Q(k)$.

Remarque. Une famille orthonormée de cardinal $\dim(E)$ est automatiquement une base orthonormée, puisqu'elle est libre (car orthogonale ne contenant pas le vecteur nul, ses vecteurs étant tous unitaires) et de cardinal $\dim(E)$.

Propriété 12

Soit $(E, \langle\cdot,\cdot\rangle)$ un espace vectoriel euclidien. Alors il existe une base orthonormée de E.

Remarque. Il n'y a pas unicité d'une base orthonormée pour un espace euclidien. Par exemple pour $E = \mathbb{R}_2[X]$ muni du produit scalaire $\langle P, Q\rangle = \sum_{k=0}^{2} P(k)Q(k)$, la famille des polynômes de Lagrange (L_0, L_1, L_2) associés à $0, 1, 2$ est aussi une base orthonormée de E.

Propriété 13

Soit $(E, \langle\cdot,\cdot\rangle)$ un espace vectoriel euclidien. Toute famille orthonormée de E peut être complétée en une base orthonormée de E.

3.2 Calculs dans une base orthonormée

L'intérêt des bases orthonormales résulte des propriétés suivantes.

Propriété 14

Soit $(E, \langle \cdot, \cdot \rangle)$ un espace euclidien rapporté à une **base orthonormée** $\mathscr{B} = (e_1, \ldots, e_n)$.

(1) Pour tout $x \in E$:

$$x = \langle e_1, x \rangle \, e_1 + \ldots + \langle e_n, x \rangle \, e_n.$$

Autrement dit, $M_{\mathscr{B}}(x) = \begin{pmatrix} \langle x, e_1 \rangle \\ \vdots \\ \langle x, e_n \rangle \end{pmatrix}$.

(2) Pour tout $x = x_1 e_1 + \ldots + x_n e_n$ et $y = y_1 e_1 + \ldots + y_n e_n$ de E :

$$\langle x, y \rangle = x_1 y_1 + \ldots + x_n y_n = \langle x, e_1 \rangle \langle y, e_1 \rangle + \cdots + \langle x, e_n \rangle \langle y, e_n \rangle.$$

(3) Pour tout $x = x_1 e_1 + \ldots + x_n e_n$ de E :

$$\|x\|^2 = x_1^2 + \ldots + x_n^2 = \langle x, e_1 \rangle^2 + \cdots + \langle x, e_n \rangle^2.$$

Mise en garde.

Ces formules sont valables uniquement lorsque la base $\mathscr{B}$ considérée est **orthonormée**.

Propriété 15

Soient $x, y \in E$, et notons X, Y les vecteurs coordonnées de x et y dans une base **orthonormée** $\mathscr{B}$ de E. Alors :

- $\langle x, y \rangle = {}^t XY$.
- $\|x\|^2 = {}^t XX$.

Remarque. Cette proposition justifie l'intérêt de travailler en base orthonormée : tous les calculs de produits scalaires (et donc ceux de normes également) se font comme dans $\mathscr{M}_{n,1}(\mathbb{R})$ muni de son produit scalaire canonique (et donc comme dans $\mathbb{R}^n$ muni de son produit scalaire canonique).

4 Projection orthogonale sur un sous-espace de dimension finie

4.1 Orthogonal d'une partie

Définition.

Soit E un espace préhilbertien, et soit A une partie de E.

On appelle *orthogonal de* A, et on note $A^\perp$, l'ensemble des vecteurs orthogonaux à tous les vecteurs de A, c'est-à-dire :

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

Ainsi : $\quad x \in A^\perp \ \Leftrightarrow\ \forall y \in A,\ \langle x, y\rangle = 0.$

Exemples.

- $\{0_E\}^\perp = E$ car tout vecteur est orthogonal à 0_E.
- $E^\perp = \{0_E\}$ car le seul vecteur orthogonal à tout vecteur de E est 0_E.

Propriété 16

Soient A et B deux parties d'un espace préhilbertien E. Alors :

(1) $A^\perp$ est un sous-espace vectoriel de E.

(2) Si $A \subset B$, alors $B^\perp \subset A^\perp$.

(3) $A \subset \left(A^\perp\right)^\perp$.

(4) $A^\perp = (\mathrm{Vect}(A))^\perp$.

Propriété 17

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

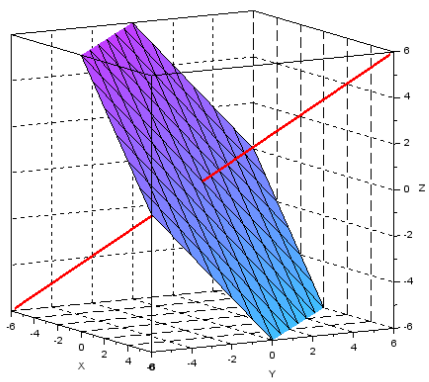
- F et G sont orthogonaux si, et seulement si, $F \subset G^\perp$.
- Si $F = \mathrm{Vect}(e_1, \ldots, e_k)$, alors :

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

Exercice 4 On considère le sous-espace vectoriel

$$F = \{(x, y, z) \mid x + y + z = 0\}$$

de $\mathbb{R}^3$ muni du produit scalaire canonique. Déterminer $F^\perp$.



Représentation de F et $F^\perp$.

4.2 Supplémentaire orthogonal d'un sous-espace de dimension finie

Théorème 18

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

- $F^\perp$ est un supplémentaire de F dans E ;
- si G est un supplémentaire de F orthogonal à F, alors $G = F^\perp$;
- $\left(F^\perp\right)^\perp = F$.

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

Danger.

Ce résultat n'est plus vrai lorsque F est de dimension infinie : on n'a pas $E = F \oplus F^\perp$ en général, ni $F = \left(F^\perp\right)^\perp$.

Propriété 19

Soient E un espace euclidien (donc de dimension finie), et F un sous-espace vectoriel de E. Alors :

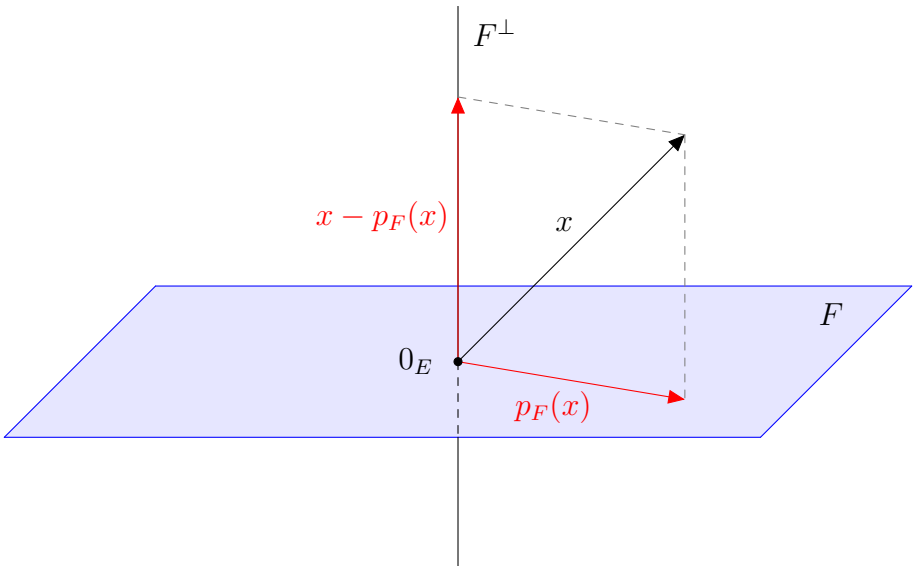
- $\dim(F^\perp) = \dim(E) - \dim(F)$;
- la concaténation d'une base orthonormée de F et d'une base orthonormée de $F^\perp$ est une base orthonormée de E.

4.3 Projection orthogonale

Soit F un sous-espace vectoriel de dimension finie d'un espace préhilbertien $(E, \langle \cdot, \cdot \rangle)$. On vient de voir qu'alors $E = F \oplus F^\perp$.

Définition.

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



Projeté orthogonal de x sur F.

Remarque. Un projecteur p de E est orthogonal si, et seulement si, $\mathrm{Im}(p) \perp \mathrm{Ker}(p)$.

Propriété 20

Soit F un sous-espace vectoriel de dimension finie de $(E, \langle \cdot, \cdot \rangle)$, et soit $x \in E$.
Pour tout $y \in E$:

$$y = p_F(x) \quad \Leftrightarrow \quad \begin{cases} y \in F \\ x - y \in F^\perp \end{cases} .$$

Méthode. Calcul du projeté orthogonal à partir d'une base de F.
Pour déterminer le projeté orthogonal de $x \in E$ sur F, on peut procéder ainsi :

(i) on détermine une base $(u_1, \ldots, u_p)$ de F ;

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

(iii) puisque $x - p_F(x) \in F^\perp$, il satisfait $\begin{cases} \langle x - p_F(x), u_1 \rangle = 0 \\ \vdots \\ \langle x - p_F(x), u_p \rangle = 0 \end{cases}$;

(iv) on résout alors ce système linéaire pour obtenir $\lambda_1, \ldots, \lambda_p$, et donc $p_F(x)$.

Exercice 5 Dans $\mathbb{R}^3$ muni du produit scalaire canonique, déterminer le projeté orthogonal de $x = (2, 2, 2)$ sur le sous-espace vectoriel $F = \left\{(x, y, z) \in \mathbb{R}^3 \mid x - \frac{y}{2} + z = 0\right\}$.

4.4 Expression dans une base orthonormée de F

Propriété 21

Soit F un sous-espace vectoriel de dimension finie de $(E, \langle \cdot, \cdot \rangle)$, et soit $(e_1, \dots, e_p)$ une **base orthonormée** de F. Alors :

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

Méthode. Calcul du projeté orthogonal à partir d'une B.O.N. de F.
Lorsqu'on dispose d'une ***base orthonormée*** $(e_1, \dots, e_p)$ *de F, on utilisera la formule précédente pour obtenir le projeté orthogonal de x sur F :*

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

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

Cependant, cette méthode est en général plus longue que celle décrite à la section précédente.

Exercice 6 Déterminer le projeté orthogonal de $x = (2, 2, 2)$ sur le sous-espace $F = \mathrm{Vect}((1, 1, -1), (0, 1, 1))$.

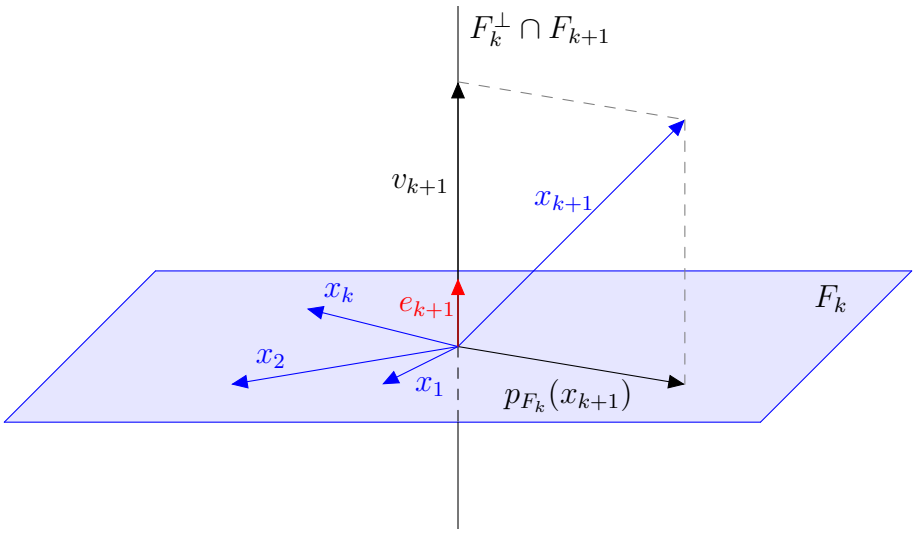
Remarque. Retour sur le procédé d'orthonormalisation de Gram-Schmidt.
Soit $(x_1, \dots, x_n)$ une famille libre de vecteurs de E, et notons $F_k = \mathrm{Vect}(x_1, \dots, x_k)$ pour $1 \leq k \leq n-1$. On peut récrire le procédé d'orthonormalisation de Gram-Schmidt de la manière suivante :

- **Étape 1.** Poser $e_1 = \dfrac{x_1}{||x_1||}$;
- **Étape $k+1$.** Une fois les vecteurs $e_1, \dots, e_k$ construits,

  (a) poser $v_{k+1} = x_{k+1} - p_{F_k}(x_{k+1})$;

  (b) 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.

4.5 Cas particulier des hyperplans

Définition.

Soit $(E, \langle \cdot, \cdot \rangle)$ un espace euclidien, et soit H un hyperplan de E.
On appelle *vecteur normal à* H tout vecteur directeur de la droite vectoriel $H^\perp$.

Remarques.

- Si a est un vecteur normal à l'hyperplan H, alors $H = \mathrm{Vect}(a)^\perp$, et H est le noyau de la forme linéaire non nulle $\varphi : x \in E \mapsto \langle x, a \rangle$.

- Si on se donne une base orthonormée $\mathscr{B} = (e_1, \ldots, e_n)$ de E, alors l'équation cartésienne de H dans la base $\mathscr{B}$ s'écrit :
$$0 = \langle x, a \rangle = x_1 a_1 + \ldots, x_n a_n$$
où $(a_1, \ldots, a_n)$ sont les coordonnées de a dans la base $\mathscr{B}$.

- Dans le cas d'un hyperplan de $\mathbb{R}^n$ muni de son produit scalaire canonique, il est très facile d'obtenir un vecteur normal à partir d'une équation de l'hyperplan. En effet, supposons qu'une équation de H soit $a_1x_1 + a_2x_2 + \cdots + a_nx_n = 0$, où les a_i ne sont pas tous nuls. On peut alors encore écrire
$$H = \{(x_1, \ldots, x_n) \in \mathbf{R}^n \mid \langle (x_1, \ldots, x_n), (a_1, \ldots, a_n) \rangle = 0\} = \{(a_1, \ldots, a_n)\}^\perp$$
de sorte que $(a_1, \ldots, a_n)$ est un vecteur normal à H. Pour reprendre un exemple traité précédemment, un vecteur normal à $H = \{(x, y, z) \in \mathbb{R}^3 \mid x + y + z = 0\}$ est $(1, 1, 1)$, et donc $H^\perp = \mathrm{Vect}(1, 1, 1)$.

Propriété 22 (Expression du projeté orthogonal sur un hyperplan)

Soient $(E, \langle \cdot, \cdot \rangle)$ un espace euclidien, H un hyperplan de E et a un vecteur normal à H.
Pour tout $x \in E$:
$$p_F(x) = x - \frac{\langle x, a \rangle}{\|a\|^2} a.$$

4.6 Distance d'un vecteur à un sous-espace

Définition.

Soit $(E, \langle\cdot,\cdot\rangle)$ un espace préhilbertien réel. Considérons $x \in E$ et A une partie non vide de E.
On appelle *distance de x à A*, et on note $d(x, A)$, le réel positif défini par :

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

Remarques.

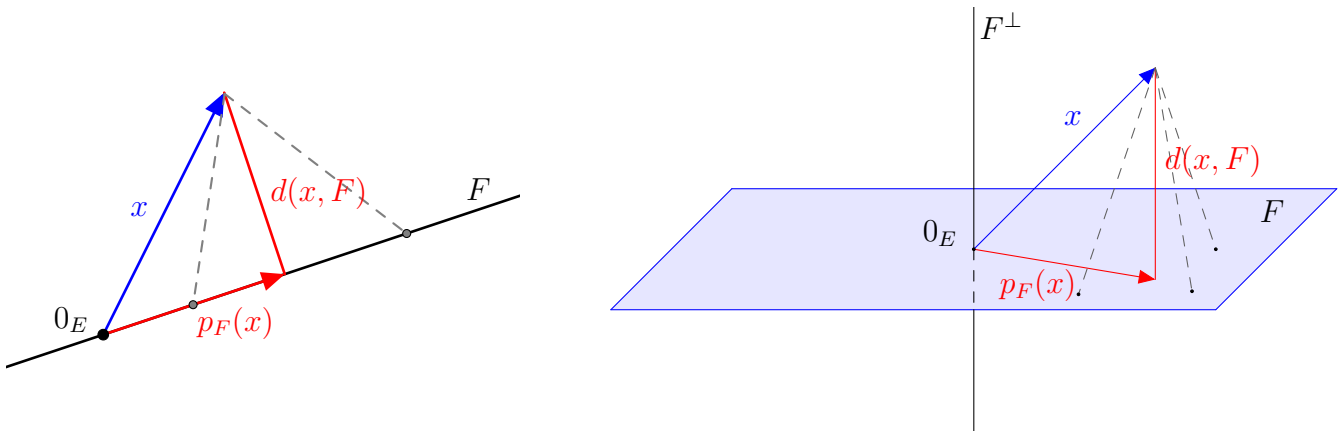
- Cette borne inférieure existe bien car $\{\|x - a\|,\ a \in A\}$ est une partie non vide et minorée (par 0) de $\mathbb{R}$.
- Cette borne inférieure n'est pas forcément atteinte, et si elle l'est, cela peut être en plusieurs points. Par exemple dans $E = \mathbb{R}$ muni du produit scalaire canonique pour lequel $d(x, a) = |x - a|$:
  - $d(\sqrt{2}, \mathbb{Q}) = 0$ (par densité de $\mathbb{Q}$ dans $\mathbb{R}$) mais cette distance n'est pas atteinte puisque $\sqrt{2} \notin \mathbb{Q}$;
  - $d(1/2, \mathbb{Z}) = \frac{1}{2}$, atteint en les deux points 0 et 1 de $\mathbb{Z}$.

Dans le cas où A est un sous-espace vectoriel de dimension finie, la distance d'un point à A est facile à exprimer à l'aide d'un projeté orthogonal.

Théorème 23 (de la distance à un sous-espace)

Soient F un sous-espace vectoriel de dimension finie d'un espace préhilbertien E, et $x \in E$.
La distance de x à F est atteinte en un unique point de F, qui est la projection orthogonale $p_F(x)$ de x sur F. Ainsi :

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



Distance de x à une droite ou un plan.

Exercice 7 Déterminer la distance de $x = (2, 2, 2)$ au sous-espace $F = \{(x, y, z) \in \mathbb{R}^3,\ x - \frac{y}{2} + z = 0\}$.

Méthode. Minimisation d'une quantité à l'aide d'un projeté orthogonal.
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 8 Montrer que $\inf\limits_{(a,b)\in\mathbb{R}^2} \int_{-1}^{1} (t^2 - at - b)^2 \, \mathrm{d}t$ existe et le calculer.

4.7 Pseudo-solution d'un système linéaire

Cette section est hors programme. Elle constitue cependant une belle application des outils développés précédemment.

On considère 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 :

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

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). Ce système ayant plus d'équations que d'inconnues ($n \geq p$), cet ensemble 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 mieux », c'est-à-dire tel que $AX_0 - B$ soit « le plus proche possible » de $0_{n,1}$. On cherche ainsi X_0 satisfaisant :

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

Théorème 24 (Pseudo-solution d'un système linéaire)

Il existe un unique vecteur $X_0 \in \mathscr{M}_{p,1}(\mathbb{R})$ minimisant $\{\|AX - B\|, \ X \in \mathscr{M}_{p,1}(\mathbb{R})\}$, donné par :

$$X_0 = (A^\top \times A)^{-1} AB.$$

Le vecteur X_0 est appelé *pseudo-solution du système (S).*

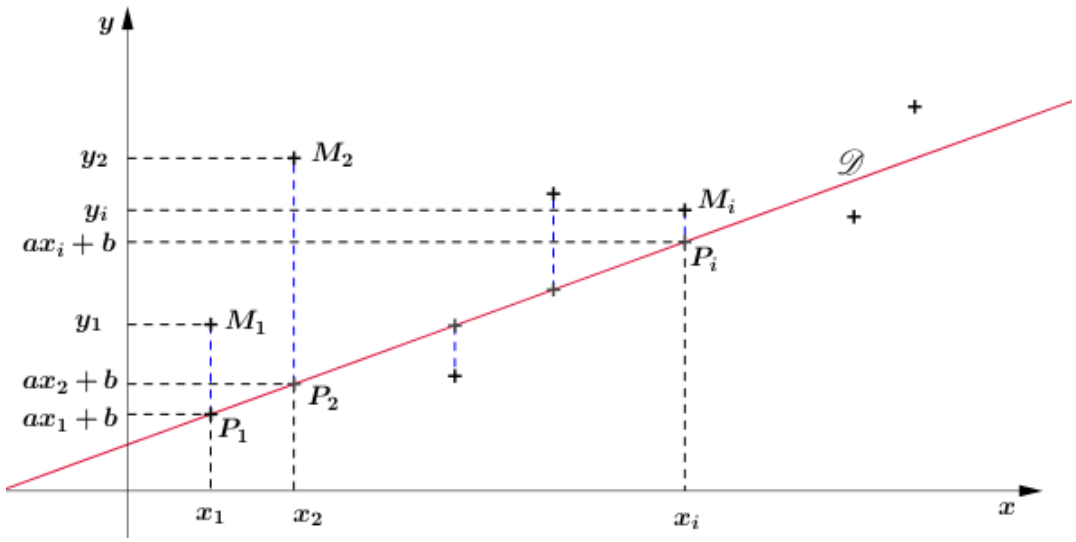
Application à la droite des moindres carrés. On étudie l'existence d'une corrélation linéaire pour une série statistique double $((x_i, y_i))_{i=1,\dots,n}$: on souhaite « placer » tous les points M_i de coordonnées (x_i, y_i) sur une même droite $\mathscr{D}$. On cherche donc $(a, b) \in \mathbb{R}^2$ tel que :

$$\forall i = 1, \dots, n, \quad y_i = ax_i + b.$$

Ce problème ce récrit matriciellement sous la forme :

$$Y = AU$$

où $Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$, $A = \begin{pmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{pmatrix}$ et $U = \begin{pmatrix} a \\ b \end{pmatrix} \in \mathscr{M}_{2,1}(\mathbb{R})$. Seulement, il y a très peu de chance qu'une telle droite existe, et donc que ce système ait une solution, nos points n'étant très probablement pas alignés.



Par ce qui précède, il existe un unique vecteur $U = \begin{pmatrix} a \\ b \end{pmatrix}$ qui minimise

$$\|Y - AU\|^2 = \sum_{i=1}^{n} (y_i - (ax_i + b))^2 = \sum_{i=1}^{n} P_i M_i^2.$$

De plus, $U = (A^\top \times A)^{-1} AB$. La droite $y = ax + b$ ainsi obtenue est appelée *droite des moindre carrés*.

Le saviez-vous ?

Cérès est la plus petite planète naine du système solaire. Avec un diamètre d'environ 950 kilomètres, il s'agit de l'objet le plus grand et le plus massif de la ceinture d'astéroïdes située entre les orbites des planètes Mars et Jupiter. Elle fut découverte le 1er janvier 1801 par Giuseppe Piazzi, astronome italien, qui pu suivre sa trajectoire jusqu'au 14 février 1801, date à laquelle l'astéroïde s'approcha trop près du Soleil pour continuer à être observé. Difficile alors pour les astronomes de l'époque de prédire la position exacte de Cérès sur la base des seules observations de Piazzi afin de confirmer sa découverte.

Afin de retrouver l'astéroïde, Carl Friedrich Gauss (1777 - 1855) développa la méthode des moindres carrés permettant de comparer les données expérimentales de Piazzi au modèle mathématique censé décrire ces données. Il obtint ainsi un résultat suffisamment précis pour permettre à Zach, un astronome allemand, de localiser à nouveau Cérès à la fin de l'année 1801.