ALG 14 | Espaces préhilbertiens réels

On généralise à des espaces vectoriels réels quelconque la notion de produit scalaire et de norme euclidienne. L'objectif essentiel est d'arriver à la méthode d'approximation par les moindres carrés.

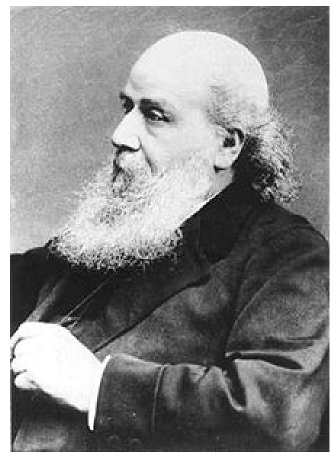
Personification de la Géométrie, Anonyme

La géométrie s'est essentiellement développée depuis les travaux d'**Euclide** selon trois grands axes : consolidation des *fondements axiomatiques*, généralisation des dimensions deux et trois à un entier $n \geqslant 4$ quelconque (voire à la dimension infinie !) et découverte de nouvelles géométries, différentes du cadre euclidien, comme par exemple la géométrie hyperbolique. De nos jours encore, le cadre euclidien est celui de la physique, à l'exception notoire de la théorie générale de la relativité et donc de l'astronomie.

Ce fut en effet la physique et ses nombreux développements qui incitèrent les mathématiciens à étudier des espaces de dimensions supérieures. Par exemple, la position d'un solide dans l'espace nécessite six dimensions réelles : trois dimensions pour décrire la position de son centre de gravité G et trois autres correspondant aux angles d'Euler décrivant les trois degrés de libertés de rotation du solide « autour » du point G.

Felix Christian Klein

La généralisation de la géométrie à ces espaces s'est opérée en deux temps. Il fallut d'abord décrire un cadre algébrique formalisant la notion de vecteur : il s'agit de l'algèbre linéaire (théorie des espaces vectoriels, matrices, déterminants, etc). Mais ce cadre s'avèra insuffisant car purement descriptif... Il fallut alors introduire une notion de *distance* (de *norme*) permettant de mesurer des vecteurs et de définir des angles. Le cadre des espaces euclidiens que nous connaissons aujourd'hui fut forgé par **Felix Klein** à la fin du XIX[e] siècle à partir de la notion générale de produit scalaire définie par **James Joseph Sylvester**.

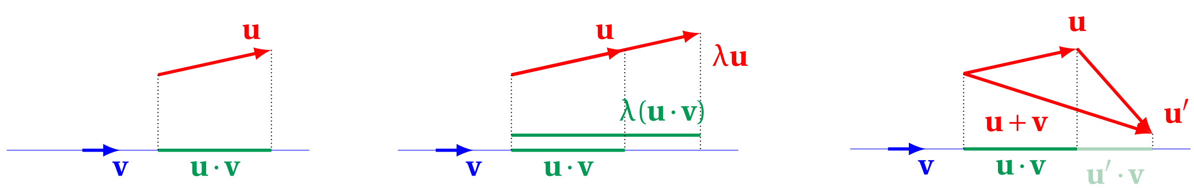


James Joseph Sylvester

Ces applications sont nombreuses et son efficacité incontestable. C'est dans ce cadre que l'on peut résoudre rigoureusement l'équation de chaleur en suivant les idées de Fourier.

1. Généraliser la notion de produit scalaire et de norme

Le principe de départ est de dégager les propriétés algébriques d'un produit scalaire afin de généraliser cet outil. Intuitivement, pour $\mathbf{v}$ non nul, $\mathbf{u}\cdot\mathbf{v}$ est la mesure algébrique de la projection de $\mathbf{u}$ sur la droite $\text{Vect}(\mathbf{v})$ orientée par $\mathbf{v}$, l'unité étant $\|v\|$.



On déduit de ces figures que le produit scalaire vérifie les propriétés suivantes[1] :

⇛ *Bilinéarité :* $\forall(\mathbf{u},\mathbf{u}',\mathbf{v}) \in \mathbf{P}^3 \quad (\mathbf{u}+\mathbf{u}')\cdot\mathbf{v} = \mathbf{u}\cdot\mathbf{v}+\mathbf{u}'\cdot\mathbf{v}$ et $(\lambda\mathbf{u})\cdot\mathbf{v} = \lambda(\mathbf{u}\cdot\mathbf{v})$.

⇛ *Symétrie :* $\forall(\mathbf{u},\mathbf{v}) \in \mathbf{P}^2 \quad \mathbf{u}\cdot\mathbf{v} = \mathbf{v}\cdot\mathbf{u}$.

1. La propriété illustrée par la deuxième figure découle du théorème de Thalès.

⇛ *Positivité* : $\forall \mathbf{u} \in \mathbf{P}, \quad \mathbf{u}\cdot\mathbf{u} \geqslant 0$.

⇛ *Caractère défini* : $\forall \mathbf{u} \in \mathbf{P}, \quad \mathbf{u}\cdot\mathbf{u} = 0 \iff \mathbf{u} = \mathbf{0}$.

2. Produits scalaires et normes euclidiennes

On rappelle qu'une application est une forme bilinéaire[2] sur un $\mathbb{R}$-ev E si :

$$\begin{cases} \forall (u,v,w,\lambda) \in \mathbf{E}^3 \times \mathbb{R}, \ \psi(u+\lambda v, w) = \psi(u,w) + \lambda\psi(v,w) & \textit{(linéarité à gauche)} \\ \forall (u,v,w,\lambda) \in \mathbf{E}^3 \times \mathbb{R}, \ \psi(w, u+\lambda v) = \psi(w,u) + \lambda\psi(w,v) & \textit{(linéarité à droite)} \end{cases}$$

Définition 14.0. Produit scalaire

Soit E un $\mathbb{R}$-ev de dimension quelconque. On appelle produit scalaire sur E toute forme bilinéaire ψ sur E vérifiant les axiomes suivants :

⇛ ψ est symétrique : pour tous u et v dans E, $\psi(u,v) = \psi(v,u)$.

⇛ ψ est positive : pour tous u dans E, on a $\psi(u,u) \geqslant 0$.

⇛ ψ est définie : pour tout u dans E, on a $\psi(u,u) = 0 \iff u = 0$.

En résumé, un produit scalaire sur E est une forme bilinéaire symétrique définie positive.

Montrer qu'une application bilinéaire est un produit scalaire

Dans la pratique on vérifiera en premier la symétrie de ψ ce qui permettra de se limiter à la vérification de la linéarité à gauche.

Notation 14.1. Produits scalaires

Plutôt que la notation fonctionnelle $\psi(u,v)$, on emploie couramment les écritures $\langle u|v\rangle$ ou encore $(u|v)$ pour le produit scalaire de deux vecteurs u et v.

✘ Le produit scalaire canonique sur $E = \mathbb{R}^n$ défini par $\langle X|Y\rangle = X^T Y = \sum_{i=1}^{n} X_i Y_i$;

- ✔ On écrit les éléments de $\mathbb{R}^n$ sous forme de matrices-colonnes et on identifie $X^T Y$, élément de $\mathscr{M}_1(\mathbb{R})$, avec son seul coefficient.
- ✔ Les linéarités à gauche et à droite découlent de la distributivité du produit matriciel. La symétrie est immédiate par commutativité du produit sur $\mathbb{R}$. Soit $X \in \mathbb{R}^n$. On a $\langle X|X\rangle = \sum_{i=1}^n X_i^2 \geqslant 0$ et cette somme est nulle *si et seulement si* tous les X_i sont nuls, ie $X = 0$. On en déduit le caractère défini positif.

✘ Le produit scalaire canonique sur $E = \mathscr{M}_n(\mathbb{R})$ défini par $\langle A|B\rangle = \operatorname{tr}\left(A^T B\right)$;

- ✔ La symétrie est immédiate en utilisant l'expression suivante :

$$\forall (A,B) \in \mathscr{M}_n(\mathbb{R})^2, \ \langle A|B\rangle = \sum_{1\leqslant i,j\leqslant n} A_{i,j} B_{i,j} \quad (\star)$$

2. Plus généralement, nous avons défini dans ALG 11 la notion de forme n-linéaire sur un $\mathbb{K}$-ev.

- ✔ Les linéarités à droite découle de la distributivité du produit matriciel et de la linéarité de la trace. Par symétrie, la linéarité à gauche est acquise. Soit $A \in \mathscr{M}_n(\mathbb{R})$. Puisque les coefficients de E sont réels, on a $\langle A|A\rangle = \sum_{1\leqslant i,j\leqslant n}^{n} A_{i,j}^2 \geqslant 0$ et cette somme est nulle *si et seulement si* tous les $A_{i,j}$ sont nuls, ie $A = 0$. On en déduit le caractère défini positif.

✘ Sur $E = \mathbb{R}_n[X]$, $\langle P|Q\rangle := \int_0^1 P(t)Q(t)\mathrm{d}t$;

- ✔ Les linéarités à gauche et à droite découlent de la distributivité du produit dans $\mathbb{R}[X]$ et de la linéarité de l'intégrale. La symétrie est immédiate par commutativité du produit sur $\mathbb{R}$.
- ✔ Soit P dans $\mathbb{R}_n[X]$. Puisque P est à coefficients réels, on a $\langle P|P\rangle = \int_0^1 P(t)^2\mathrm{d}t \geqslant 0$ par positivité de l'intégrale et, puisque la fonction polynomiale P^2 est positive et continue, cette intégrale est nulle *si et seulement si* $\forall t \in [0,1]$, $P(t) = 0$, ce qui équivaut à $P = 0$ (le seul polynôme admettant une infinité de racines est le polynôme nul). On en déduit le caractère défini positif.

✘ Le lecteur poursuivra par le test **(⚡ 14.0)**.

Définition 14.2. Espaces préhilbetiens réels, espaces euclidiens

Dans ce chapitre, tous les ev sont des $\mathbb{R}$-ev.

- ⇛ Un espace préhilbertien réel est un $\mathbb{R}$-ev E muni d'un produit scalaire $\langle|\rangle$.
- ⇛ On écrira ainsi en abrégé que $(E, \langle|\rangle)$ est un espace préhilbertien réel.
- ⇛ On appelle espace euclidien tout espace préhilbertien réel de dimension finie.

Définition 14.3. Norme euclidienne associée à un produit scalaire, vecteur unitaire

Soit $(E, \langle|\rangle)$ un espace préhilbertien réel.

- ⇛ Pour tout $u \in E$, on pose $\|u\| = \sqrt{\langle u|u\rangle}$. Ce réel est appelé norme de u.
- ⇛ Un vecteur u est dit unitaire si $\|u\| = 1$.

Pour deux vecteurs u et v, on a :

$$\begin{aligned}\|u+v\|^2 &= \langle u+v|u+v\rangle = \langle u|u\rangle + \langle u|v\rangle + \langle v|u\rangle + \langle v|v\rangle \quad \text{par bilinéarité}\\ &= \langle u|u\rangle + 2\langle u|v\rangle + \langle v|v\rangle \quad \text{par symétrie}\end{aligned}$$

Plus généralement, par bilinéarité et symétrie, on obtient le développement de la norme d'une somme :

$$\left\|\sum_{i=1}^m u_i\right\|^2 = \left\langle \sum_{i=1}^m u_i \,\middle|\, \sum_{i=1}^m u_i \right\rangle = \sum_{1\leqslant i,j\leqslant m} \langle u_i|u_j\rangle = \sum_{i=1}^m \|u_i\|^2 + 2\sum_{1\leqslant i<j\leqslant m} \langle u_i|v_i\rangle$$

Le cas de deux vecteurs est particulièrement important :

Proposition 14.4. Identités remarquables

Soit $(E, \langle|\rangle)$ un espace préhilbertien réel. Pour tout $(u,v) \in E^2$,

a. Développement du carré de la norme d'une somme : $\|u \pm v\|^2 = \|u\|^2 \pm 2\langle u|v\rangle + \|v\|^2$.

b. *Identité de polarisation (1)* : $\langle u|v\rangle = \dfrac{\|u+v\|^2 - \|u\|^2 - \|v\|^2}{2} = \dfrac{-\|u-v\|^2 + \|u\|^2 + \|v\|^2}{2}$.

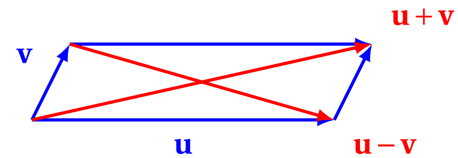
c. *Identité de polarisation (2)* : $\langle u|v\rangle = \dfrac{\|u+v\|^2 - \|u-v\|^2}{4}$.

d. *Identité du parallélogramme* : $\|u+v\|^2 + \|u-v\|^2 = 2\|u\|^2 + 2\|v\|^2$.

Les identités de polarisation et l'identité du parallélogramme

Les relations b) et c) de la proposition précédentes sont connues sous le nom *d'identités de polarisation* : elles permettent « *d'inverser* » l'ordre naturel (la norme est définie au moyen du produit scalaire) en exprimant le produit scalaire au moyen de la norme ; connaissant l'application norme $u \mapsto \|u\|$, on peut *reconstruire* le produit scalaire $(u,v) \mapsto \langle u|v\rangle$.

L'identité du parallélogramme est la généralisation d'une propriété des parallélogramme du plan et de l'espace : la somme des carrés des longueurs des côtés est égale à la somme des carrés des longueurs des diagonales.



La proposition suivante unifie les inégalités établies pour les éléments de $\mathbb{R}^n$ et les intégrales.

Proposition 14.5. Inégalité de Cauchy-Schwarz

Pour tous u et v dans E, on a $|\langle u|v\rangle| \leqslant \|u\|\|v\|$. Il a égalité si et seulement si (u,v) est liée.

Le lecteur trouvera une autre preuve de cette inégalité dans le test **(14.1)**.
On a plus précisément que $\langle u|v\rangle = \|u\|\|v\|$ si et seulement si u et v sont colinéaires de même sens (ie $u=0$ ou $\exists\lambda \geqslant 0$ tel que $v = \lambda u$, on dit aussi que u et v sont *positivement liés*). On a également $\langle u|v\rangle = -\|u\|\|v\|$ si et seulement si u et v sont colinéaires de sens contraires.

- Pour tout $(x_1,\dots,x_n,y_1,\dots,y_n) \in \mathbb{R}^{2n}$, $\quad \left|\sum_{i=1}^n x_i y_i\right| \leqslant \sqrt{\sum_{i=1}^n x_i^2}\sqrt{\sum_{i=1}^n y_i^2}$.
- Pour $a<b$ et tout $(f,g) \in \mathscr{C}^0([a,b],\mathbb{R})^2$, $\quad \left|\int_a^b f(t)g(t)\mathrm{d}t\right| \leqslant \sqrt{\int_a^b f(t)^2\mathrm{d}t}\sqrt{\int_a^b g(t)^2\mathrm{d}t}$.
- Le lecteur trouvera un autre exemple dans le test **(14.2)**.

Proposition 14.6. Norme euclidienne

Soit $(\mathrm{E}, \langle|\rangle)$ une espace préhilbertien réel.

a. *Séparation* : pour tout u dans E, on a $\|u\| = 0 \iff u = 0$.

b. *Homogénéité* : pour tous u dans E et λ dans $\mathbb{R}$, on a $\|\lambda u\| = |\lambda|\,\|u\|$.

c. *Inégalité triangulaire* : pour tous u et v dans E, on a $\|u+v\| \leqslant \|u\| + \|v\|$.

d. *Cas d'égalité* : $\forall (u,v) \in \mathrm{E}^2$, on a $\|u+v\| = \|u\| + \|v\| \iff u$ et v sont positivement liés.

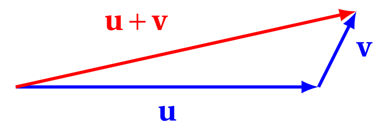
Interprétations géométriques

On revient à la proposition précédente. L'axiome de séparation permet de montrer qu'un vecteur est nul en calculant sa norme.

La propriété d'homogénéité permet de « *normer* » des vecteurs :

$$\forall u \in \mathrm{E}\setminus\{0\},\quad \left\|\frac{u}{\|u\|}\right\| = 1,\quad \text{ie } \frac{u}{\|u\|} \text{ est unitaire}$$

L'inégalité triangulaire généralise aux espaces préhilbertiens réels, la propriété bien connue dans le plan et l'espace : *le chemin le plus court entre deux points est la ligne droite.*



3. Orthogonalité

On généralise d'abord la notion de couple de vecteurs orthogonaux puis on l'étend à des familles quelconque. On définit ensuite la notion d'orthogonal d'une partie.

3.1. Familles orthogonales

Par rapport à la géométrie euclidienne apprise dans les « *petites classes* », l'approche est inversée : l'orthogonalité est définie par le produit scalaire[3].

Définition 14.7. Vecteurs orthogonaux, familles orthogonales, familles orthonormales

Soit $(\mathrm{E}, \langle\cdot|\cdot\rangle)$ un espace préhilbertien réel.

- Deux vecteurs u et v de E sont dits orthogonaux si $\langle u|v\rangle = 0$. On note alors $u \perp v$.
- Une famille $(u_i)_{i\in\mathrm{I}}$ de vecteurs de E est dite orthogonale si

$$\forall (i,j)\in\mathrm{I}^2,\ \ i\neq j \implies u_i \perp u_j$$

- Une famille $(u_i)_{i\in\mathrm{I}}$ de vecteurs de E est dite orthonormale ou encore orthonormée si elle est orthogonale et si tous ses éléments sont unitaires. Ceci équivaut à $\forall (i,j)\in\mathrm{I}^2,\ \langle u_i|u_j\rangle = \delta_{i,j}$.

- Dans $\mathscr{C}^0([0,\pi],\mathbb{R})$ muni du produit scalaire $\langle u|v\rangle := \int_0^\pi fg$ la famille de fonctions $(f_n)_{n\in\mathbb{Z}}$ définie par $f_n : t\mapsto \sin(nt)$ est orthogonale :

$$\forall (n,m)\in\mathbb{N}^2,\, n\neq m \implies \langle f_n|f_m\rangle = \left[\frac{\sin(n-m)t}{n-m} - \frac{\sin(n+m)t}{n+m}\right]_0^\pi = 0$$

- Pour $(u,v)\in\left(\mathbb{R}^3\right)^2$ tels que $u\perp v$, la famille $(u, v, u\wedge v)$ est orthogonale.
- Dans $\mathscr{M}_n(\mathbb{R})$ muni de $\langle \mathrm{A}|\mathrm{B}\rangle := \mathrm{tr}\left(\mathrm{A}^\mathrm{T}\mathrm{B}\right)$, toute famille de la forme (A,S), avec A antisymétrique et S symétrique, est orthogonale. En effet, pour $(\mathrm{A},\mathrm{S})\in\mathscr{A}_n(\mathbb{R})\times\mathscr{S}_n(\mathbb{R})$, on a

$$\langle \mathrm{A}|\mathrm{S}\rangle = \mathrm{tr}(-\mathrm{AS}) = -\mathrm{tr}(\mathrm{AS}) = -\mathrm{tr}(\mathrm{SA}) = -\langle \mathrm{S}|\mathrm{A}\rangle$$

  d'où $\langle \mathrm{A}|\mathrm{S}\rangle = 0$ par symétrie du produit scalaire.

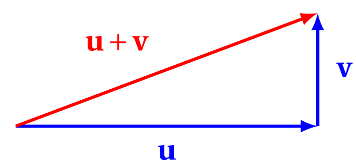
3. La notion d'angle (ou plutôt *les notions d'angle*) n'est pas au programme. La construction moderne s'appuie sur la notion de rotation.

En normant une famille orthogonale de vecteurs non nuls, on obtient une famille orthonormale.

- La base canonique de l'espace euclidien canonique $\mathbb{R}^n$ est orthonormée (immédiat).
- La base canonique de l'espace euclidien canonique $\mathscr{M}_n(\mathbb{R})$ est orthonormée (immédiat en utilisant la relation ($\star$) de l'exemple **??** à la page **??**).
- Le lecteur poursuivra avec profit par le test (↯ **14.3**).

Et maintenant, une veille connaissance : Pythagore.

- Il se généralise aux familles à deux vecteurs d'un espace préhilbertien réel.
- Attention, la réciproque n'est vraie que pour deux vecteurs.



Proposition 14.8. Théorème de Pythagore

Soit $(\mathrm{E}, \langle \mid \rangle)$ un espace préhilbertien réel.

a. *Théorème de Pythagore* : $\forall (u,v) \in \mathrm{E}^2,\ u \perp v \iff \|u+v\|^2 = \|u\|^2 + \|v\|^2$.

b. Si $(u_1, \dots, u_n)$ est orthogonale, alors $\left\| \sum_{i=1}^{n} u_i \right\|^2 = \sum_{i=1}^{n} \|u_i\|^2$.

- La réciproque du b. est fausse pour $n > 2$. Dans l'espace $\mathbb{R}^3$ muni du produit scalaire canonique les vecteurs suivants sont un contre-exemple $u_1 := (1,0,0)$, $u_2 := (2,1,0)$, $u_3 := (-1,1,0)$.

Pour les familles ne comportant pas le vecteur nul, l'orthogonalité est une condition suffisante de liberté. On en déduit évidemment que toute famille orthonormale est libre.

Proposition 14.9. Familles orthogonales sans vecteur nul

Soit $(\mathrm{E}, \langle \mid \rangle)$ un espace préhilbertien réel. Si $(u_1, \dots, u_n)$ est une famille orthogonale de E telle que $\forall i \in [\![1, n]\!],\ u_i \neq 0,$ alors $(u_1, \dots, u_n)$ est libre.

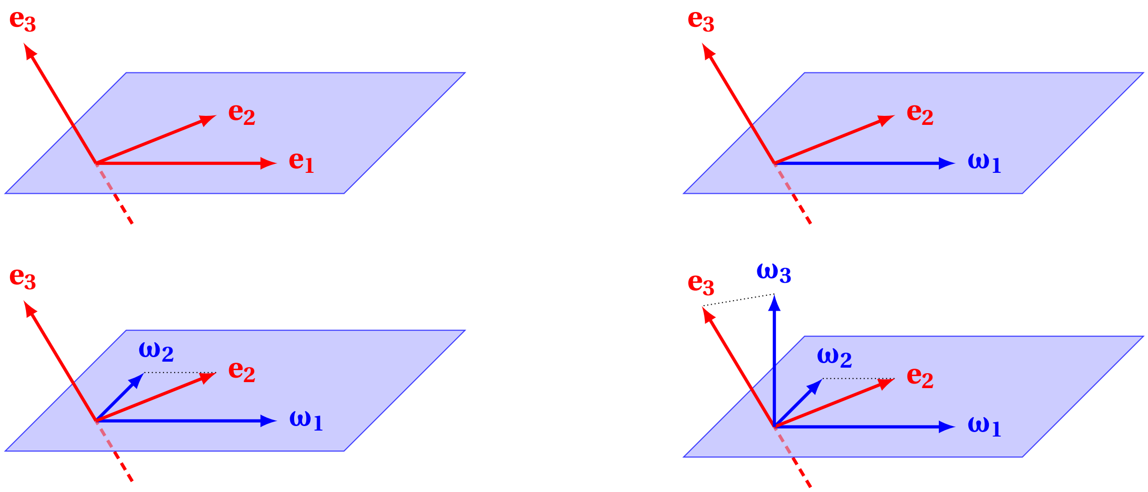
3.2. Bases orthonormées d'un espace euclidien

Nous commençons par un premier résultat fondamental.

Proposition 14.10. Existence de bases en dimension finie

Tout espace euclidien admet des bases orthonormées.

La preuve donnée de ce théorème est *constructive* et l'algorithme utilisé est connu sous le nom de *procédé d'orthonormalisation de Gram-Schmidt*. Voyons-le à l'œuvre dans l'espace :



- On « *redresse* » les vecteurs les uns après les autres.
- On choisit $\omega_1 := e_1$.
- On trouve λ_1 tel que $\omega_2 := e_2 + \lambda_1 \omega_1 \perp \omega_1$.
- On trouve μ_1 et μ_2 tels que $\omega_3 := e_3 + \mu_1 \omega_1 + \mu_2 \omega_2$ soit orthogonal à ω_1 et ω_2.
- On « *norme* » ensuite pour trouver (f_1, f_2, f_3) : $f_i := \dfrac{\omega_i}{\|\omega_i\|}$.
- On a $\mathrm{Vect}(e_1) = \mathrm{Vect}(f_1)$, $\mathrm{Vect}(e_1, e_2) = \mathrm{Vect}(f_1, f_2)$ et $\mathrm{Vect}(e_1, e_2, e_3) = \mathrm{Vect}(f_1, f_2, f_3)$.

Le résultat suivant a été démontré ci-dessus (on peut aussi montrer l'unicité de la famille $(f_1, \dots, f_n)$).

Proposition 14.11. Gram-Schmidt

Soit $(\mathrm{E}, \langle\cdot|\cdot\rangle)$ un espace euclidien et $\mathscr{B} = (e_1, \dots, e_n)$ une base de E. Il existe une base orthonormée $(f_1, \dots, f_n)$ de E telle que

$$\forall k \in [\![1, n]\!],\ \mathrm{Vect}(e_1, \dots, e_k) = \mathrm{Vect}(f_1, \dots, f_k) \quad \text{et} \quad \langle f_k | e_k \rangle > 0$$

Procédé d'orthonormalisation de Gram-Schmidt

On peut en construire une par le procédé de Gram-Schmidt décrit ci-dessous.

- On construit d'abord une famille orthogonale $(\omega_1, \dots, \omega_n)$ vérifiant les conditions indiquées ci-dessus. On procède par récurrence :
  - On pose $\omega_1 := e_1$.
  - Pour $k \in [\![2, n]\!]$, on pose $\omega_k := e_k + \lambda_1 \omega_1 + \cdots + \lambda_{k-1} \omega_{k-1}$. Le système d'équations $\langle \omega_k | \omega_i \rangle = 0$ pour $i < k$ admet une seule solution $(\lambda_1, \dots, \lambda_{k-1})$.
- On norme les vecteurs ω_i pour trouver les f_i.

On peut envisager une variante de cet algorithme qui consiste à normer les ω_i juste après leur calcul au lieu de le faire en une seule fois à la fin des calculs. On peut bien-sûr appliquer cet algorithme à toute famille libre de E (pas nécessairement une base de E). LE lecteur est renvoyé au test (⚡ **14.4**) afin d'approfondir.
‘

✘ On travaille dans l'espace $E = \mathbb{R}_2[X]$ muni du produit scalaire défini par $\langle P|Q\rangle = \int_0^1 P(t)Q(t)\mathrm{d}t$.
Orthonormaliser la base canonique $(1, X, X^2)$.

- ✔ On pose $Q_0 := 1$.
- ✔ On pose $Q_1 := X + \mu$. La condition $Q_1 \perp Q_0$ est équivalente à $\langle Q_1|Q_0\rangle = 0$, ie $\mu\langle 1|1\rangle = -\langle X|1\rangle$, ie $\mu = -1/2$.
- ✔ On pose $Q_2 := X^2 + aQ_1 + bQ_0$. On a

$$\begin{cases} \langle Q_2|Q_0\rangle = 0 \\ \langle Q_2|Q_1\rangle = 0 \end{cases} \iff \begin{cases} \langle X^2|Q_0\rangle + b\|Q_0\|^2 = 0 \\ \langle X^2|Q_1\rangle + a\|Q_1\|^2 = 0 \end{cases} \iff \begin{cases} \frac{1}{3} + b = 0 \\ \frac{1}{12} + \frac{a}{12} = 0 \end{cases} \iff Q_2 = X^2 - Q_1 - \frac{1}{3}Q_0$$

Par le théorème de Pythagore : $\|Q_2\|^2 = \|X^2\|^2 - \|Q_1\|^2 - \frac{1}{9}\|Q_0\|^2 = \frac{1}{5} - \frac{1}{12} - \frac{1}{9} = \frac{1}{36\times 5}$. On en déduit une bon de $\mathbb{R}_2[X]$:

$$\left(1, 2\sqrt{3}\left(X - \frac{1}{2}\right), 6\sqrt{5}\left(X^2 - X + \frac{1}{6}\right)\right)$$

✘ Le lecteur trouvera un autre exemple dans le test **(⚡ 14.5)**.

On voit à travers cet exemple que les calculs peuvent vite devenir « lourds » dès que le nombre de vecteurs dépasse deux.

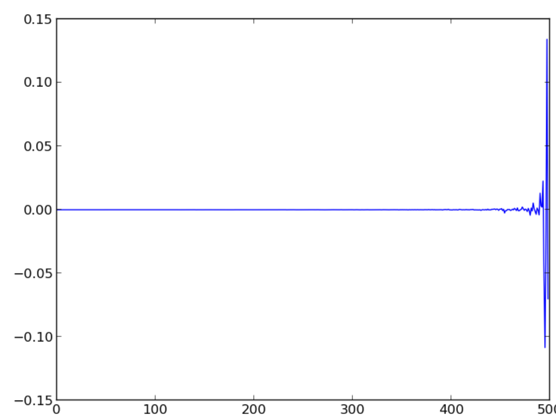
L'algorithme de Gram-Schmidt est numériquement instable

Tout comme l'algorithme du pivot de Gauss, cet algorithme pose des problème quand il est appliqué de manière approchée sur machine (à cause de la précision limitée des flottants). Il y a *divergence numérique* de l'algorithme dans certains cas :

Lors de l'application de l'algorithme sous Python pour 500 vecteurs de la forme

$$v_i = \begin{pmatrix} 1+\varepsilon_{i,1} \\ \vdots \\ 1+\varepsilon_{i,n} \end{pmatrix} \quad \text{où } |\varepsilon_{i,j}| \leqslant 10^{-5}$$

et formant une famille libre, on observe la divergence en fin d'algorithme. On a représenté ci-contre les produits scalaires $\langle f_k|f_{k+1}\rangle$ en fonction de k : on observe clairement « *la perte d'orthogonalité* » due aux erreurs d'arrondis successifs.



Proposition 14.12. Calculs dans une bon

Soit $(E, \langle\cdot|\cdot\rangle)$ un espace euclidien et $\mathscr{B} = (e_1, \dots, e_n)$ une base orthonormée de E. Pour tous vecteurs x et y, notons $X = [x_1, \dots, x_n]^T$ et $Y = [y_1, \dots, y_n]^T$ leurs coordonnées dans $\mathscr{B}$. On a

$$\langle x|y\rangle = \langle X|Y\rangle_2 = X^TY = \sum_{i=1}^n x_i y_i\ , \quad \|x\| = \|X\|_2 = \sqrt{\sum_{i=1}^n x_i^2} = \sqrt{X^TX}$$

où $\langle\cdot|\cdot\rangle_2$ et $\|\cdot\|_2$ désignent le produit scalaire et la norme canonique sur $\mathbb{R}^n$.

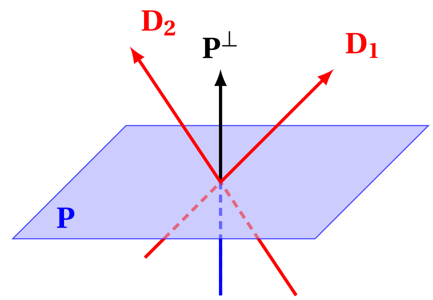
L'intérêt des bases orthonormées est de ramener *tous les calculs de produits scalaires (et donc de normes) à des calculs dans les espaces canoniques* $\mathbb{R}^n$. Si E est un espace euclidien de dimension n et de base orthonormée $\mathscr{B}$, l'application $\phi : \mathrm{E} \to \mathbb{R}^n$ définie par $x \mapsto \mathrm{mat}_{\mathscr{B}}(x)$ est *plus qu'un isomorphisme*, c'est *une isométrie*, ie une application linéaire qui conserve la norme, c'est-à-dire $\forall x \in \mathrm{E}$, $\|\phi(x)\| = \|x\|_2$.

3.3. Orthogonal d'une partie

Il s'agit dans cette section d'étendre la notion d'orthogonalité à des parties de E (et en particulier des sev) afin de décrire leur positions relatives.

Par exemple, dans l'espace $\mathbb{R}^3$ canonique, on sait qu'un plan vectoriel P admet une infinité de supplémentaires (toutes les droites qui ne sont pas contenues dans P). Parmi toutes ces droites, il en existe une seule, que nous noterons $\mathrm{P}^\perp$, orthogonale au plan P, ie vérifiant

$$\forall (x,y) \in \mathrm{D}\times\mathrm{P}\,,\ \ \langle x|y\rangle = 0$$



Définition 14.13. Orthogonal d'une partie

Soit $(\mathrm{E}, \langle\cdot|\cdot\rangle)$ un espace préhilbertien réel et $\mathrm{A} \subset \mathrm{E}$. On appelle orthogonal de A et l'on note $\mathrm{A}^\perp$ l'ensemble des vecteurs de E orthogonaux à tous les vecteurs de A. Ainsi

$$\mathrm{A}^\perp = \{u \in \mathrm{E}\,;\, \forall a \in \mathrm{A},\, u \perp a\}$$

- On considère l'espace $\mathbb{R}^3$ muni de sa structure euclidienne canonique. Soit $\mathrm{P}_1 := \mathrm{Vect}(u_1, u_2)$, où $u_1 = (1,1,1)$ et $u_2 = (1,1,0)$. Calculer $\mathrm{P}_1^\perp$. On a

$$\mathrm{P}_1^\perp = \left\{ v \in \mathbb{R}^3\,;\, \langle v|u_1\rangle = \langle v|u_2\rangle = 0 \right\} = \mathrm{Ker}\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix} \ \text{ d'où } \ \mathrm{P}_1^\perp = \mathrm{Vect}\, u_3 \ \text{ avec } \ u_3 := (1,-1,0)$$

- Dans le cas d'un produit scalaire non canonique, le calcul d'un orthogonal peut s'avérer astucieux. Par exemple, pour $\mathbb{R}[\mathrm{X}]$ muni du produit scalaire

$$\langle \mathrm{P}|\mathrm{Q}\rangle = \int_0^1 \mathrm{P}(t)\mathrm{Q}(t)\mathrm{d}t \ \ \text{ et } \ \ \mathrm{F} = \{\mathrm{P} \in \mathbb{R}[\mathrm{X}]\,;\, \mathrm{P}(0) = 0\}\ , \ \ \text{on prouve que } \mathrm{F}^\perp = \{0\}$$

  En effet, soit $\mathrm{Q} \in \mathrm{F}^\perp$. Comme $\mathrm{XQ} \in \mathrm{F}$, on a $\langle \mathrm{Q}|\mathrm{XQ}\rangle = \int_0^1 t\mathrm{Q}(t)^2 \mathrm{d}t = 0$. Comme $t \mapsto t\mathrm{Q}(t)^2$ est continue et positive sur $[0,1]$, on en déduit que $t\mathrm{Q}(t)^2 = 0$ pour tout $t \in [0,1]$. En particulier Q admet une infinité de racines (les éléments de $]0,1]$) donc est nul.

Proposition 14.14. Propriétés des orthogonaux

Soit $(\mathrm{E}, \langle\cdot|\cdot\rangle)$ un espace préhilbertien réel.

a. On a $\mathrm{E}^{\perp} = \{0\}$ et $\{0\}^{\perp} = \mathrm{E}$.

b. $\forall \mathrm{A} \subset \mathrm{E},\ \mathrm{A}^{\perp}$ est un sev de E.

c. $\forall \mathrm{A} \subset \mathrm{E}$, on a $\mathrm{A}^{\perp} = \mathrm{Vect}(\mathrm{A})^{\perp}$.

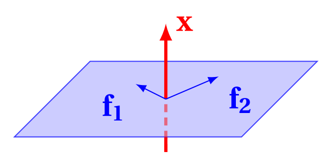
d. $\forall \mathrm{A} \subset \mathrm{E},\ \mathrm{A} \subset \mathrm{A}^{\perp\perp}$.

Attention au double orthogonal

L'inclusion du d. peut être stricte, même pour un sev de E : le double orthogonal $\mathrm{F}^{\perp\perp}$ d'un sev F n'est pas toujours égal à F. C'est le cas des sev de dimension finie comme nous le verrons plus bas. Reprenons le dernier exemple (cf. la seconde illustration du **??** ci-dessus) : on a vu que $\mathrm{F}^{\perp} = \{0\}$ et donc que $\mathrm{F} \subsetneq \mathrm{F}^{\perp\perp} = \mathbb{R}[\mathrm{X}]$.

Dans le cas de la dimension finie, appartenir à l'orthogonal se résume à un nombre fini d'équations : il suffit d'appliquer le c) de la proposition 3.3 à $\mathrm{A} := \{f_1, \ldots, f_p\}$ pour obtenir la caractérisation suivante :

Proposition 14.16. (Orthogonalité). *Pour* $(f_1, \ldots, f_p) \in \mathrm{E}^p$, $x \in Vect(f_1, \ldots, f_p)^{\perp} \iff \forall k \in [\![1, p]\!],\ \langle x | f_k \rangle = 0$.



Nous en venons au résultat généralisant l'exemple initial du plan P et de son orthogonal, la droite $\mathrm{P}^{\perp}$.

Proposition 14.16. Supplémentaire orthogonal

Soit $(\mathrm{E}, \langle | \rangle)$ un espace préhilbertien de dimension quelconque et F une sev de E de dimension finie.

a. $\mathrm{F} \oplus \mathrm{F}^{\perp} = \mathrm{E}$;

b. $\mathrm{F}^{\perp\perp} = \mathrm{F}$;

c. Si $\dim \mathrm{E} < +\infty$, alors

$$\dim \mathrm{F}^{\perp} = \dim \mathrm{E} - \dim \mathrm{F}$$

Parmi tous les supplémentaires de F, le sev $\mathrm{F}^{\perp}$ est appelé le supplémentaire orthogonal de F dans E.

Attention, si F est dimension infinie, alors F et $\mathrm{F}^{\perp}$ ne pas toujours supplémentaires

Reprenons le contre-exemple donné ci-dessus pour le double orthogonal (cf. le second exemple du **??** à la page **??**). On a $\mathrm{F}^{\perp} = \{0\}$ et donc $\mathrm{F} \oplus \mathrm{F}^{\perp} = \mathrm{F} \subsetneq \mathbb{R}[\mathrm{X}]$. En revanche, il est toujours vrai que F et $\mathrm{F}^{\perp}$ sont en somme directe car $\mathrm{F} \cap \mathrm{F}^{\perp} = \{0\}$.

Le lecteur est renvoyé aux exemples précédents pour de plus amples détails.

- Pour la structure euclidienne canonique de $\mathscr{M}_n(\mathbb{R})$, $\mathscr{S}_n(\mathbb{R})^{\perp} = \mathscr{A}_n(\mathbb{R})$. Voici deux preuves.
  - On a vu (cf. **??** page **??**) que $\mathscr{A}_n(\mathbb{R}) \subset \mathscr{S}_n(\mathbb{R})^{\perp}$. Cette inclusion est une égalité car

$$\dim \mathscr{S}_n(\mathbb{R})^{\perp} = n^2 - \dim \mathscr{S}_n(\mathbb{R}) = n^2 - \frac{n(n+1)}{2} = \frac{n(n-1)}{2} = \dim \mathscr{A}_n(\mathbb{R})$$

- ✔ Comme $\mathscr{S}_n(\mathbb{R})$ est engendré par les matrices de la forme $E_{i,j}+E_{j,i}$ pour $(i,j)\in[\![1,n]\!]^2$, on déduit de la proposition 14.16. que

$$X\in\mathscr{S}_n(\mathbb{R}) \iff \forall(i,j)\in[\![1,n]\!],\ \langle X,E_{i,j}\rangle = -\langle X,E_{j,i}\rangle \iff -X_{i,j} = X_{i,i}$$

car la base canonique de $\mathscr{M}_n(\mathbb{R})$ est orthonormée pour ce produit scalaire (cf. l'exemple **??** à la page **??**). On en déduit que $\mathscr{A}_n(\mathbb{R}) = \mathscr{S}_n(\mathbb{R})^\perp$.

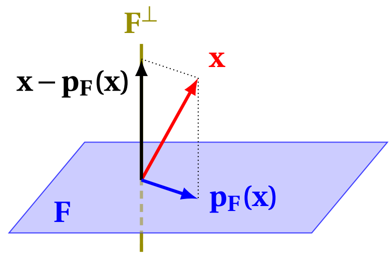
- ✘ Soit $P_2 := \{(x,y,z)\in\mathbb{R}^3\,;\, x-y+z=0\}$. Calculer P_2^T.On remarque que $P_2 = \mathrm{Vect}(u)^\perp$ où $u := (1,-1,1)$. Ainsi $P_2^\perp = \mathrm{Vect}(u)^{\perp\perp} = \mathrm{Vect}\,u$.

4. Projection orthogonale et approximation au sens des moindres carrés

Quand F est un sev de dimension finie, on distingue parmi toutes les projection sur F la projection parallélement à sont orthogonal.

Définition 14.17. Projection orthogonale

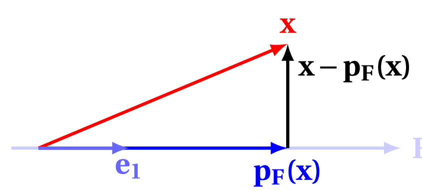
Soient $(E,\langle\cdot|\cdot\rangle)$ un espace préhilbertien réel et F un sev de E de dimension finie. On appelle projection orthogonale sur F la projection p_F sur F parallélement à $F^\perp$.



- ✘ Projection orthogonale sur une droite. Soit $F=\mathrm{Vect}(e_1)$ avec $e_1\neq 0$.
Il existe $\lambda_1\in\mathbb{R}$ tel que $p_F(x)=\lambda_1 e_1$. On obtient

$$0=\langle x-p_F(x)|e_1\rangle=\langle x-\lambda_1 e_1|e_1\rangle=\langle x|e_1\rangle-\lambda_1\|e_1\|^2$$

On a $p_F(x) = \dfrac{\langle x|e_1\rangle}{\|e_1\|^2}e_1 = \langle x|f_1\rangle f_1$ où $f_1 := \dfrac{e_1}{\|e_1\|}$.



Cette formule se généralise sans peine à des dimensions supérieures.

Proposition 14.18. Projections orthogonales

Avec les notations précédentes,

a. $\forall x\in E,\ \|p_F(x)\|\leqslant\|x\|$.

b. Si $\mathscr{B} = (f_1,,\dots,f_p)$ est une bon de F, alors $\forall x\in E,\ \ p_F(x) = \displaystyle\sum_{i=1}^{p}\langle x|f_i\rangle f_i$.

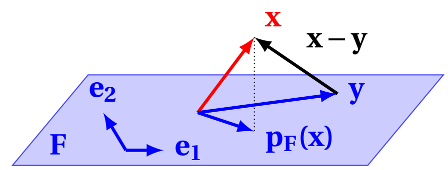
Calcul d'une projection orthogonale

Pour calculer le projeté orthogonal d'un vecteur x de E sur un sev F de E de dimension finie, on peut envisager deux pistes :

⇛ Appliquer le b. de la proposition 4 Pour cela, il faut disposer d'une bon de F (quitte à appliquer le procédé d'orthonormalisation de Gram-Schmidt) ;

⇛ On peut calculer $p_F(x)$ en ne connaissant qu'une base (quelconque) $(e_1,\dots,e_p)$ de F : soient $(\lambda_1,\dots,\lambda_p)\in\mathbb{R}^p$ et $y=\lambda_1e_1+\cdots+\lambda_pe_p$; $y=p_F(x)$ *si et seulement si* $x-y\in F^\perp=\mathrm{Vect}(e_1,\dots,e_p)^\perp$, ce qui équivalent à

$$\forall i\in[\![1,p]\!],\qquad \underbrace{\langle x-y|e_i\rangle}_{\langle x|e_i\rangle-\sum_{j=1}^{p}\lambda_j\langle e_i|e_j\rangle}=0$$



On obtient un système linéaire à p inconnues $\lambda_1,\dots,\lambda_p$ et p équations qui admet une unique solution.

✗ Calculons (par les deux méthodes) la projection orthogonale Q de X^2 sur $\mathbb{R}_1[X]$ pour le produit scalaire intégral sur $\mathbb{R}[X]$ défini à l'exemple **??** (cf. page **??**).

- ✔ Comme $\left(1,\sqrt{3}(2X-1)\right)$ est une bon de $\mathbb{R}_1[X]$, on a

$$Q=\langle X^2|1\rangle 1+\langle X^2|\sqrt{3}(2X-1)\rangle\sqrt{3}(2X-1)=\frac{1}{3}+\frac{3}{6}(2X-1)=X-\frac{1}{6}$$

- ✔ Il existe $(a,b)\in\mathbb{R}^2$ tel que $Q=a+bX$. On a

$$\begin{cases}\langle X^2-Q|1\rangle &= 0\\ \langle X^2-Q|X\rangle &= 0\end{cases}\iff\begin{cases}\frac{1}{3}-a-\frac{b}{2} &= 0\\ \frac{1}{4}-\frac{a}{2}-\frac{b}{3} &= 0\end{cases}\iff (a,b)=\left(-\frac{1}{6},1\right)\iff Q=X-\frac{1}{6}$$

✗ Soit $E=\mathbb{R}^4$ muni de sa structure euclidienne canonique et F le plan de E d'équations $x_1+2x_2+2x_3+x_1=0$, $2x_1+x_2+x_3+2x_4=0$. Calculons la matrice de p_F, projection orthogonale sur F, dans la base canonique de E.

- ✔ Le sev F a pour équations $x_1+x_4=x_2+x_3=0$ donc (u_1,u_2) en est une base, où $u_1=(1,0,0,-1)$ et $u_2:=(0,1,-1,0)$. Comme cette famille est orthogonale, il suffit de normer les vecteurs pour obtenir une bon (f_1,f_2) de F.
- ✔ En écrivant les vecteurs de $\mathbb{R}^4$ sous forme de matrice-colonnes :

$$p_F(x)=\langle x|f_1\rangle f_1+\langle x|f_2\rangle f_2=f_1(f_1^Tx)+f_2(f_2^Tx)=\left(f_1f_1^T+f_2f_2^T\right)x$$

par « associativité » du produit matriciel. On en déduit que la matrice recherchée vaut

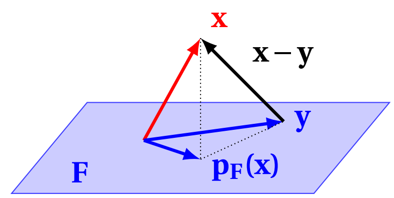
$$f_1f_1^T+f_2f_2^T=\frac{1}{2}\begin{pmatrix}1&0&0&-1\\0&1&-1&0\\0&-1&1&0\\-1&0&0&1\end{pmatrix}$$

✗ Le lecteur trouvera un autre exemple dans le test (⚡ **14.6**).

Recherchons à présent la meilleure approximation d'un vecteur x parmi les vecteurs de F. Il s'agit de trouver les vecteurs $y \in \mathrm{F}$ minimisant

$$\|x-y\|$$

Conformément à l'intuition géométrique, l'unique vecteur minimisant cette expression est $p_{\mathrm{F}}(x)$.



Proposition 14.19. Meilleure approximation au sens des moindres carrés

Soit $(\mathrm{E}, \langle\cdot|\cdot\rangle)$ un espace euclidien, F un sev de E et $x \in \mathrm{E}$.

⇛ Il existe un unique vecteur y de F tel que la distance $\|x-y\|$ soit minimale et il vaut $p_{\mathrm{F}}(x)$: on dit que c'est la meilleure approximation de x par les vecteurs de F.

⇛ On appelle distance de x à F, notée $d(x,\mathrm{F})$, le réel $d(x,\mathrm{F}) = \|x-p_{\mathrm{F}}(x)\| = \|p_{\mathrm{F}^\perp}(x)\| = \min\limits_{y\in\mathrm{F}}\|x-y\|$.

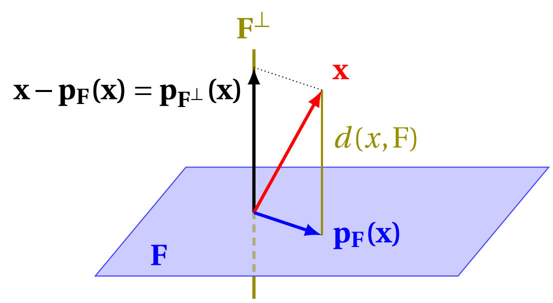
Calcul de $d(x,\mathrm{F})$

On reprend les notations précédentes.

⇛ Calculer la distance $d(x,\mathrm{F})$ revient à déterminer la projection orthogonale $p_{\mathrm{F}}(x)$ ou $p_{\mathrm{F}^\perp}(x)$.

⇛ On pourra calculer $d(x,\mathrm{F})$ en ayant recours au théorème de Pythagore :

$$d(x,\mathrm{F}) = \sqrt{\|x\|^2 - \|p_{\mathrm{F}}(x)\|^2} = \|p_{\mathrm{F}^\perp}(x)\|$$



La difficulté est parfois de reconnaître le bon cadre euclidien.

✗ Calculons $\mu := \inf\limits_{(a,b)\in\mathbb{R}^2} \int_0^1 \left(t^2+at+b\right)^2 \mathrm{d}t$.

✔ On se place dans l'espace euclidien $\mathrm{E} = \mathbb{R}_2[\mathrm{X}]$ pour le produit scalaire $\langle \mathrm{P}|\mathrm{Q}\rangle := \int_0^1 \mathrm{P}(t)\mathrm{Q}(t)\mathrm{d}t$. On reconnaît que $\mu = d\left(\mathrm{X}^2, \mathbb{R}_1[\mathrm{X}]\right)^2$.

✔ En notant p la projection orthogonale sur $\mathbb{R}_1[\mathrm{X}]$, on déduit des calculs du **??** (à la page **??**) que

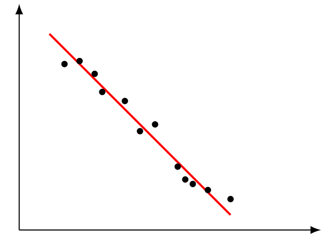
$$\mu = \left\|\mathrm{X}^2 - p\left(\mathrm{X}^2\right)\right\|^2 = \left\|\mathrm{X}^2\right\|^2 - \left\|p\left(\mathrm{X}^2\right)\right\|^2 = \left\|\mathrm{X}^2\right\|^2 - \left\|\mathrm{X}-\frac{1}{6}\right\|^2 = \frac{1}{5} - \frac{7}{36} = \frac{1}{180}$$

✗ Le lecteur poursuivra par le test (↯ **14.7**).

L'exemple suivant est à l'origine de la méthode de régression linéaire.

✗ On se place dans le cadre suivant : on dispose de n points $(x_1,y_1),\dots(x_n,y_n)$ du plan $\mathbb{R}^2$. En théorie, ces points sont alignés sur une droite Δ inconnue d'équation $y = ax+b$.

Cependant, pour des raisons d'erreurs de mesure, ces points ne sont pas connus de façon exacte. Pour déterminer Δ, il est donc illusoire de rechercher (a, b) solution du système linéaire à n équations $y_i = ax_i + b$ car ce dernier n'admettra probablement aucune solution.



On doit donc rechercher la droite Δ *la plus vraisemblable*. Matriciellement, ce système s'écrit

$$aX + bU = Y \quad \text{où} \quad U = \begin{pmatrix} a \\ b \end{pmatrix}, \; Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, \; X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad \text{et} \quad U = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$$

L'idée de déterminer $(a, b) \in \mathbb{R}^2$ tel que $\|Y - aX - bU\|$ soit minimale (pour la norme euclidienne usuelle de $\mathbb{R}^n$), à défaut d'être nulle. C'est ce qu'on appelle une pseudo-solution du système linéaire ci-dessus. On sait que ce minimum est atteint pour $aX + bU$ égal à la projection orthogonale de Y sur $F := \mathrm{Vect}(U, X)$. On suppose que $\mathrm{rg}(U, X) = 2$ (les points ne sont pas alignés sur une verticale, hypothèse raisonnable si les erreurs sur les points sont assez petites). On obtient :

$$\begin{cases} \langle Y - aX - bU | U\rangle = 0 \\ \langle Y - aX - bU | X\rangle = 0 \end{cases} \iff \begin{cases} a\langle X|U\rangle + b\langle U|U\rangle = \langle Y|U\rangle \\ a\langle X|X\rangle + b\langle U|X\rangle = \langle Y|X\rangle \end{cases} \iff \begin{pmatrix} a \\ b \end{pmatrix} = G^{-1}\begin{pmatrix} \langle Y|U\rangle \\ \langle Y|X\rangle \end{pmatrix}$$

où $G = \begin{bmatrix} \langle X|U\rangle & \langle U|U\rangle \\ \langle X|X\rangle & \langle U|X\rangle \end{bmatrix}$. On trouve $a = \dfrac{\langle X|U\rangle\langle Y|U\rangle - n\langle X|Y\rangle}{\langle X|U\rangle^2 - n\|X\|^2}$ et $b = \dfrac{\langle X|U\rangle\langle X|Y\rangle - \|X\|^2\langle Y|U\rangle}{\langle X|U\rangle^2 - n\|X\|^2}$.

Approcher la dimension infinie par la dimension finie

Ce résultat est également utilisé pour approcher, dans un espace E de dimension infinie, un vecteur x par une suite de vecteurs (f_n) correspondant à une suite croissante de sev de dimension finie (F_n). La réunion des F_n est dense dans E *si et seulement si*

$$\|x - f_n\| \xrightarrow[n\to\infty]{} 0$$

L'exemple le plus célèbre est certainement celui des séries de Fourier : ici E est l'ensemble des fonctions continues et 2π-périodique de $\mathbb{R}$ dans $\mathbb{R}$ muni du produit scalaire intégral

$$\langle f | g\rangle = \int_0^{2\pi} fg$$

et F_n est l'espace vectoriel engendré par les $x \mapsto \sin(kx)$ et $x \mapsto \cos(kx)$ pour $k \in [\![0, n]\!]$.

5. Énoncés des tests

14.1.

Soit $(\mathrm{E}, \langle\cdot|\cdot\rangle)$ un espace préhilbertien réel.

a. Développer, pour tous $x, y \in \mathrm{E}$, l'expression $\left\| \|y\|^2 x - \langle x|y\rangle y \right\|^2$.

b. En déduire une *nouvelle démonstration* de l'inégalité de Cauchy-Schwarz et de son cas d'égalité.

14.2.

On définit sur l'espace vectoriel réel E des fonctions de classe $\mathscr{C}^1$ de $[0,1]$ dans $\mathbb{R}$,

$$\langle f \mid g \rangle = f(1)g(1) + \int_0^1 f'(t)g'(t)\,\mathrm{d}t$$

a. Montrer que l'on définit ainsi un produit scalaire sur E.

b. Établir que $\forall f \in \mathrm{E}, \left(f(1) + \int_0^1 f'(t)\,\mathrm{d}t\right)^2 \leqslant 2\left(f^2(1) + \int_0^1 f'^2(t)\,\mathrm{d}t\right)$.

14.3.

Soit $\mathrm{M} \in \mathscr{M}_n(\mathbb{R})$. Montrer que les colonnes de M forment une bon de $\mathbb{R}^n$ (espace muni de sa structure canonique) *si et seulement si* $\mathrm{M}^\top \mathrm{M} = \mathrm{I}_n$.

14.4.

On reprend les notations de la méthode 3.2 de la page 8. Que dire de la matrice de passage des (e_i) aux (f_i) ?

14.5.

Sur l'espace vectoriel réel $\mathrm{E} = \mathbb{R}_2[\mathrm{X}]$, on définit $\langle \mathrm{P}|\mathrm{Q}\rangle = \mathrm{P}(-1)\mathrm{Q}(-1) + \mathrm{P}(0)\mathrm{Q}(0) + \mathrm{P}(1)\mathrm{Q}(1)$.

a. Montrer qu'il s'agit d'un produit scalaire.

b. Trouver une base orthonormée de E par le procédé d'orthonormalisation de Schmidt appliqué à la base canonique de E.

c. Trouver une autre base orthonormée de E en utilisant les polynômes interpolateurs de Lagrange.

14.6.

Donner la matrice dans la base canonique de la projection orthogonale p de $\mathbb{R}^4$ sur le sous-espace vectoriel F d'équations :

$$\begin{cases} x_1 + x_2 + x_3 + x_4 = 0 \\ x_1 - x_2 + x_3 - x_4 = 0 \end{cases}$$

14.7.

Soit $\mathrm{E} = \mathbb{R}_2[\mathrm{X}]$ muni de sa structure euclidienne canonique (ie la base canonique est orthonormée). On note F le sous-espace vectoriel de E des polynômes s'annulant en 1.

a. Déterminer une base de F.

b. Calculer $\delta = \inf_{\mathrm{P} \in \mathrm{F}} \|\mathrm{X} - \mathrm{P}\|$.

6. Solutions des tests

14.1.

a. Notons δ l'expression de l'énoncé. On a

$$\begin{aligned}\delta &= \|y\|^4\|x\|^2 - 2(x|y)^2\|y\|^2 + \langle x|y\rangle^2\|y\|^2\\ &= \|y\|^4\|x\|^2 - \langle x|y\rangle^2\|y\|^2\end{aligned}$$

b. Puisque $\delta \geqslant 0$, on obtient

$$\langle x|y\rangle^2\|y\|^2 \leqslant \|x\|^2\|y\|^4$$

c'est-á-dire

$$|\langle x|y\rangle|\|y\| \leqslant \|x\|\|y\|^2$$

Si $y=0$, on a banalement

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

puisque les deux membres sont nuls. Si $y \neq 0$, $\|y\| \neq 0$ et en divisant par cette quantité strictement positive dans l'inégalité obtenue ci-dessus, on aboutit à

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

Retrouvons le cas d'égalité dans l'inégalité de Cauchy-Schwarz : il y a égalité *si et seulement si* (x, y) est liée.

- Si (x, y) est liée, l'inégalité de Cauchy-Schwarz est en fait une égalité : la preuve en est facile, cf. le cours.
- D'après les calculs précédents, le cas d'égalité correspond à $\delta = 0$, ie

$$\|y\|^2x - \langle x|y\rangle y = 0$$

Si $y = 0$, la famille (x, y) est banalement liée. Sinon $\|y\| \neq 0$ et l'égalité précédente est une relation de liaison non banale entre x et y. Dans tous les cas la famille (x, y) est liée.

14.2.

a. Prouvons que $\langle\cdot|\cdot\rangle$ définit un produit scalaire sur E.

- L'application $(\cdot|\cdot)$ est bilinéaire par linéarité du produit sur $\mathbb{R}$, de la dérivation et de l'intégrale.
- La forme bilinéaire $\langle\cdot|\cdot\rangle$ est symétrique car le produit sur $\mathbb{R}$ est commutatif.
- Soit $f \in$ E. Par positivité de l'intégrale, on a

$$(f|f) = f(1)^2 + \int_0^1 f'(t)^2\mathrm{d}t \geqslant 0$$

La forme bilinéaire $\langle\,|\,\rangle$ est donc positive.

- Reprenons les notations précédentes. Puisque qu'une somme de réels positifs est nulle *si et seulement si* tous les termes sont nuls, l'égalité $\langle f|f\rangle = 0$ *implique*

$$f(1) = 0 \quad \text{et} \quad \int_0^1 f'^2(t)\mathrm{d}t = 0$$

La fonction f'^2 étant continue et positive, la deuxième condition est équivalente à $f' = 0$, la fonction f est donc constante et finalement nulle puisque $f(1) = 0$. La forme bilinéaire $\langle\cdot|\cdot\rangle$ est donc définie.

b. Appliquons l'inégalité de Cauchy-Schwarz aux fonctions de E définies par

$$f \in \mathrm{E},\ g : t \in [0,1] \mapsto t$$

Puisque

$$\langle f|g\rangle = f(1) + \int_0^1 f'(t)\mathrm{d}t$$

puis

$$\|f\|^2 = f^2(1) + \int_0^1 f^2(t)\mathrm{d}t$$

et finalement $\|g\|^2 = 1+1 = 2$ et l'inégalité s'écrit

$$\left(f(1) + \int_{[0,1]} f'\right)^2 \leqslant 2\left(f^2(1) + \int_{[0,1]} f'^2\right)$$

14.3.

En notant C_i les colonnes de M, on sait que

$$\forall (i,j) \in [\![1,n]\!]^2,\ \left(\mathrm{M}^\mathrm{T}\mathrm{M}\right)_{i,j} = \langle \mathrm{C}_i|\mathrm{C}_j\rangle$$

On en déduit l'équivalence demandée.

14.4.

Pour $i \in [\![0,]\!]$,

$$f_i \in \mathrm{Vect}(e_1, \dots, e_i)$$

Ainsi, la matrice de passage des (e_i) aux (f_i) est triangulaire supérieure.

14.5.

a. Prouvons que $\langle\cdot|\cdot\rangle$ définit un produit scalaire.

- $\langle\cdot|\cdot\rangle$ est bilinéaire par bilinéarité du produit sur $\mathbb{R}$ et par linéarité de l'évaluation.
- $\langle\cdot|\cdot\rangle$ est symétrique par commutativité du produit sur $\mathbb{R}$.
- Soit $P \in E$. On a $\langle P|P\rangle = P^2(-1) + P^2(0) + P^2(1) \geqslant 0$, la forme bilinéaire $\langle\cdot|\cdot\rangle$ est donc positive.
- Soit $P \in E$. Puisqu'une somme de réels positifs est nulle *si et seulement si* tous les termes sont nuls , on a

$$\langle P|P\rangle = P^2(-1) + P^2(0) + P^2(1) = 0$$

si et seulement si $0, 1, -1$ sont racines de P. Puisque P est de degré inférieur à deux , cette condition est équivalente à $P = 0$. La forme bilinéaire $\langle\cdot|\cdot\rangle$ est donc définie.

b. Orthonormalisons la base canonique de E par le procédé de Gram-Schmidt.

- *Première étape.* On pose

$$\Gamma_1 = \frac{1}{\|1\|} = \frac{1}{\sqrt{3}}$$

- *Deuxième étape.* Le vecteur défini par

$$n_1 = X - \langle X|\Gamma_1\rangle\Gamma_1 = X - 0\Gamma_1 = X$$

est orthogonal à Γ_1. Puisque $\|n_1\| = \sqrt{2}$, on complète (Γ_1) par

$$\Gamma_2 = \frac{n_1}{\|n_1\|} = \frac{X}{\sqrt{2}}$$

- *Troisième étape.* Le vecteur

$$\begin{aligned} n_3 &= X^2 - \langle X^2|\Gamma_1\rangle\Gamma_1 - \langle X^2|\Gamma_2\rangle\Gamma_2 \\ &= X^2 - \frac{2}{3}\Gamma_1 - 0\Gamma_2 \end{aligned}$$

appartient à $\mathrm{Vect}(\Gamma_1, \Gamma_2)^\perp$. Puisque $\|n_2\| = \sqrt{2/3}$, on complète (Γ_1, Γ_2) par

$$\Gamma_3 = \frac{n_2}{\|n_2\|} = \sqrt{3/2}(X^2 - 2/3)$$

La famille $(\Gamma_1, \Gamma_2, \Gamma_3)$ est une base orthonormée de E.

c. Posons

$$L_{-1} = \frac{X(X-1)}{2},\ L_1 = \frac{X(X-1)}{2},\ L_0 = 1 - X^2$$

Il s'agit des polynômes de Lagrange associés aux réels ± 1 et 0. Cette famille est clairement orthonormée pour le produit scalaire $\langle\cdot|\cdot\rangle$, elle est donc libre dans E qui est de dimension trois : il s'agit d'une base orthonormée de E.

Commentaire

Le procédé d'orthonormalisation de Gram-Schmidt est lourd en calculs au-delà de trois vecteurs... Il faut parfois avoir un peu de culture – voire du flair ! – mais surtout suivre les indications de l'énoncé pour trouver des bases orthonormées.

14.6.

Commençons par établir un plan de bataille : il nous faut calculer une base orthonormée de F afin de calculer la projection orthogonale p sur F. On commence donc par déterminer une base de F qu'il faudra ensuite orthonormaliser par le procédé de Schmidt.

- *Détermination d'une base de* F. Il est clair que le système d'équations définissant F est équivalent à

$$x_1 + x_3 = x_2 + x_4 = 0$$

Un vecteur X appartient donc à F *si et seulement si* il est de la forme

$$X = x_1(1, 0, -1, 0) + x_2(0, 1, 0, -1)$$

où $x_1, x_2 \in \mathbb{R}$. Posons

$$u = (1, 0, -1, 0)\ ,\ v = (0, 1, 0, -1)$$

La famille (u, v) est clairement libre et génératrice de F , il s'agit d'une base de ce sous-espace de $\mathbb{R}^4$.

⇛ *Détermination d'une base orthonormée de* F. La base (u, v) est clairement orthogonale. Puisque l'on a $\|u\| = \|v\| = \sqrt{2}$, la famille formée par

$$u' = (1/\sqrt{2}, 0, -1/\sqrt{2}, 0)\ ,\ \ v' = (0, 1/\sqrt{2}, 0, -1/\sqrt{2})$$

est une base orthonormée de F.

⇛ *Calcul de p*. Pour tout vecteur x de E, on a

$$p(x) = (x|u')u' + (x|v')v'$$

Ainsi, en notant $\mathscr{B} = (e_1, e_2, e_3, e_4)$ la base canonique de E, on a

$$p(e_1) = (1/2, 0, -1/2, 0),\ \ p(e_2) = (0, 1/2, 0, -1/2)$$

puis

$$p(e_3) = (-1/2, 0, 1/2, 0)\ \text{ et }\ p(e_4) = (0, -1/2, 0, 1/2)$$

Ainsi

$$\mathrm{mat}_{\mathscr{B}}(p) = \frac{1}{2}\begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \\ -1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{pmatrix}$$

14.7.

a. Pout tout $P \in E$, la formule de Taylor s'écrit

$$P = P(1) + P'(1)(X-1) + \frac{P''(1)}{2}(X-1)^2$$

Un polynôme P de F est donc combinaison linéaire des vecteurs

$$P_1 = X - 1,\ \ P_2 = (X-1)^2$$

Puisque ces deux polynômes appartiennent à F et ne sont pas proportionnels, la famille (P_1, P_2) est une base de F.

b. D'après le cours, la quantité $\|X - P\|^2$ est minimale lorsque $P = \pi_F(X)$, où π_F désigne la projection orthogonale sur F. Orthonormalisons la famille $(P_1 P_2)$ par le procédé de Gram-Schmidt. Posons

$$Q_1 = \frac{X-1}{\|P_1\|} = \frac{X-1}{\sqrt{2}}$$

Notons π_1 la projection orthogonale sur $\mathrm{Vect}(P_1)$. Posons

$$\begin{aligned} n &= (X-1)^2 - \pi_1((X-1)^2) \\ &= (X-1)^2 - \langle (X-1)^2 | Q_1 \rangle Q_1 \\ &= (X-1)^2 + \frac{3}{2}(X-1) \\ &= X^2 - X/2 - 1/2 \end{aligned}$$

Notons ensuite

$$Q_2 = \frac{n}{\|n\|} = \sqrt{2/3}\,n$$

On a

$$\begin{aligned} \pi_F(X) &= \langle X|Q_1\rangle Q_1 + \langle X|Q_2\rangle Q_2 \\ &= \frac{X-1}{2} - \frac{1}{3}(X^2 - X/2 - 1/2) \\ &= -X^2/3 + 2/3X - 1/3 \end{aligned}$$

Ainsi

$$X - \pi_F(X) = (X^2 + X + 1)/3$$

et

$$\delta = \|X - \pi_F(X)\| = \frac{1}{\sqrt{3}}$$

Commentaire.

On pouvait aussi remarquer que l'orthogonal de F est la droite vectorielle engendrée par $X^2 + X + 1$, ce qui est plus facile à voir lorsque l'on choisit $(X-1, X(X-1))$ comme base de F. Le calcul de la projection et de la distance de X à F s'en trouve simplifié.