Espaces préhilbertiens et euclidiens

Si tu veux que je te dessine un mouton, indique moi d'abord le produit scalaire.

Le but de ce chapitre est de faire de la géométrie sur les espaces vectoriels pour généraliser la géométrie dans le plan et l'espace vue au lycée. On se limitera au corps de base $\mathbf{R}$.
Les illustrations du cours seront essentiellement en dimension finie ($\mathbf{R}^2$ ou $\mathbf{R}^3$), mais nous verrons également quelques belles applications sur des espaces fonctions (dimension infinie).
Pour faire de la géométrie, nous avons besoin de définir des angles et des longueurs : nous le ferons par l'intermédiaire du produit scalaire. Reste à choisir un produit scalaire adapté et trouver la bonne formalisation pour généraliser ces objets.
Bien entendu, nous trouverons des similitudes avec le chapitre sur le déterminant (lié à la surface).

Remarque : Dans ce chapitre et pour les exercices, il faut absolument faire des dessins pour visualiser : nous faisons de la géométrie !

1 Espaces préhilbertiens

A Produit scalaire

Définition 1.1 *(Produit scalaire)*

Soit E un $\mathbf{R}$-espace vectoriel.
On appelle **produit scalaire** sur E, toute application $f : E \times E \to \mathbf{R}$ qui est :

1. *bilinéaire* : $\forall y \in E,\ x \mapsto f(x,y)$ est linéaire,
   $\forall x \in E,\ y \mapsto f(x,y)$ est linéaire,
2. *symétrique* : $\forall (x,y) \in E^2,\ f(x,y) = f(y,x)$,
3. *définie* : $\forall x \in E,\ f(x,x) = 0 \iff x = 0_E$,
4. *positive* : $\forall x \in E,\ f(x,x) \geqslant 0$.

⚠ **vocabulaire.**

- Le terme « défini » a un sens bien particulier ici. Dire que f n'est pas définie, ne veut pas forcément dire qu'elle n'existe pas.
  C'est souvent ce point qui pose problème, quand on veut justifier qu'une application est un produit scalaire.
- Le caractère « positif » n'a rien à voir avec le *signe* de x : x est un vecteur est n'a donc pas de signe algébrique.

On remarquera que E est un $\mathbf{R}$-espace vectoriel. La notion de produit scalaire sur les $\mathbf{C}$-espaces vectoriels est différente et n'est pas au programme.

Méthode *(Astuce)*

Pour montrer qu'une application est un produit scalaire, on montre la symétrie avant la bilinéarité (car alors, il suffit de montrer la linéarité par rapport à une seule variable et conclure par symétrie).

Notation

Le produit scalaire de deux vecteurs $(x,y) \in E^2$ se note habituellement

$$< x, y > \qquad \text{ou} \qquad (x|y) \qquad \text{ou} \qquad x \cdot y.$$

Définition 1.2 *(Espaces préhilbertiens et euclidiens)*

Un $\mathbf{R}$-espace vectoriel muni d'un produit scalaire est un **espace préhilbertien**.
Un $\mathbf{R}$-espace vectoriel *de dimension finie* muni d'un produit scalaire est un **espace euclidien**.

⚠ Il existe plusieurs produits vectoriels possibles sur un espace vectoriel : parler d'espace préhilbertien ou d'espace euclidien, c'est faire le choix d'un produit scalaire.

Vocabulaire :
Le terme *pré*hilbertien laisse entendre qu'il existerait des espaces... hilbertiens. C'est vrai, mais on parle plutôt d'espaces de Hilbert. Ce sont des espaces préhilbertiens auxquels on ajoute une propriété de convergence pour certaines suites (les suites de Cauchy) ce qui permet de faire de l'analyse sur ces espaces, un peu comme on le fait sur $\mathbf{R}$ ou $\mathbf{C}$.
Cela sera étudié plus précisément en deuxième année. On peut noter qu'en dimension finie, les espaces euclidiens sont tous des espaces de Hilbert.
Le terme euclidien fait référence à la géométrie euclidienne classique qui sert de modèle pour ces espaces.

B Exemples à connaître

Exemple

Sur $E = \mathbf{R}^2$, le produit scalaire usuel $\big((x_1,\, y_1)\big|(x_2,\, y_2)\big) = x_1x_2 + y_1y_2$ est un produit scalaire.

Solution :
La forme est bien évidemment symétrique et bilinéaire.
$((x,\, y)|(x,\, y)) = x^2 + y^2 \geqslant 0$ avec égalité si, et seulement si $x = y = 0$ ce qui prouve le caractère défini positif de l'application.
$((x_1,\, y_1),\, (x_2,\, y_2)) \mapsto \big((x_1,\, y_1)\big|(x_2,\, y_2)\big) = x_1x_2 + y_1y_2$ définit donc bien un produit scalaire sur $\mathbf{R}^2$.

Exemple

Sur $E = \mathscr{C}^0([a, b], \mathbf{R})$, $(f,\, g) \mapsto \int_a^b fg$ est un produit scalaire.

Solution :
f et g étant supposées continues sur le segment, leur intégrale est bien définie et à valeurs réelle.
La fonction est évidemment symétrique par commutativité du produit dans $\mathbf{R}$.
Elle est bilinéaire par linéarité de l'intégrale :

$$\int_a^b (f_1 + \lambda f_2)g = \int_a^b (f_1g + \lambda f_2g) = \int_a^b f_1g + \lambda \int_a^b f_2g,$$

(et la symétrie pour avoir la linéarité par rapport à l'autre fonction).
On a vu enfin dans le cours d'intégration, la croissance de l'intégrale qui donne la positivité, et enfin qu'une fonction continue de signe constant est d'intégrale nulle si, et seulement si elle est identiquement nulle.

$$\int_a^b f^2 = 0 \Rightarrow f^2 = 0 \Rightarrow f = 0.$$

$(f,\, g) \mapsto \int_a^b fg$ définit donc bien un produit scalaire sur E.

Exemple

Sur $E = \mathscr{C}_{mx}([a, b], \mathbf{R})$, $(f,\, g) \mapsto \int_a^b fg$ n'est pas un produit scalaire.

Solution :
En effet, si on suppose seulement f continue par morceaux, alors on perd le caractère défini de l'intégrale.
Par exemple si $f(a) = 1$ et $f_{|]a,\, b]} = 0$ alors $\displaystyle\int_a^b f^2 = 0$ mais $f \neq 0$.

Exemple

Sur $E = \mathbf{R}_n[X]$, montrer que les trois applications forment chacune un produit scalaire :

1. $(P,\, Q) \mapsto \int_{-1}^1 PQ$.
2. $(\sum_{k=0}^n a_kX^k,\, \sum_{k=0}^n b_kX^k) \mapsto \sum_{k=0}^n a_kb_k$.
3. $(P,\, Q) \mapsto \sum_{k=0}^n P(k)Q(k)$.

Solution :

1. P et Q étant des polynômes, les fonctions polynomiales sont continues sur $[-1,\, 1]$ et on peut donc reprendre ce qu'on a vu avec le produit scalaire usuel sur $\mathscr{C}^0\left([-1,\, 1],\, \mathbf{R}\right)$.
Ceci donne la bilinéarité, la symétrie et la positivité.
Par contre, pour montrer le caractère défini, il faut rappeler que si $P_{|[-1,\, 1]} = 0$ alors P admet une infinité de racines, donc $P = 0$.
2.
   - La symétrie est immédiate par commutativité du produit dans $\mathbf{R}$: $a_kb_k = b_ka_k$.
   - La bilinéarité s'obtient avec celle de la somme.
   $$\sum_{k=0}^n \left(a_k + \lambda a'_k\right) b_k = \sum_{k=0}^n a_kb_k + \lambda a'_kb_k = \sum_{k=0}^n a_kb_k + \lambda \sum_{k=0}^n a'_kb_k,$$
   (et la symétrie pour avoir la linéarité par rapport à l'autre polynôme).
   - Si $P = Q = \sum\limits_{k=0}^n a_kX^k$ alors $(P|P) = \sum\limits_{k=0}^n a_k^2 \geqslant 0$.
   La somme est nulle si, et seulement si tous les termes sont nuls (car positifs).
   donc $(P|P) = 0 \Rightarrow P = 0$.
3. La symétrie et la bilinéarité se montrent comme au cas précédent.
$(P|P) = \sum\limits_{k=0}^n P(k)^2 \geqslant 0$.
La somme est nulle si, et seulement si tous les termes sont nuls (car positifs).
donc P admet $n + 1$ racines distinctes donc $P = 0$.

Exemple

Avec les matrices carrées, $(A, B) \in (\mathscr{M}_n(\mathbf{R}))^2$, le produit scalaire usuel est défini par $(A|B) = \mathrm{tr}({}^tAB) = \sum\limits_{i=1}^n \sum\limits_{j=1}^n a_{i,\, j}b_{i,\, j}$.

Exemple

Sur l'espace des variables aléatoires de $\Omega \to [\![1, n]\!]$, on peut définir le produit scalaire :

$$(X, Y) \mapsto \mathbf{E}(XY).$$

Solution :
On remarque d'abord que l'ensemble des variables aléatoires sur l'univers fini Ω forme bien un espace vectoriel.

- La fonction est bien symétrique car $XY = YX$.
- Par linéarité de l'espérance, elle est bilinéaire.
- $\mathbf{E}\left(X^2\right) \geqslant 0$ par positivité de l'espérance.
- Si $\mathbf{E}\left(X^2\right) = 0$, alors $\sum\limits_{x\in X(\Omega)} x^2 \mathbf{P}(X = x) = 0$ et comme il s'agit d'une somme de termes positifs, ils sont tous nuls.
  Donc $X(\Omega) = \{0\}$, donc la variable aléatoire est nulle.

Exemple *(La fonction de masse discrète)*
Sur $E = \mathbf{R}^n$, on définit $(w_1, w_2, \cdots, w_n) \in \mathbf{R}^n$ tel que $\forall k \in [\![1, n]\!], w_k > 0$.
$(x,\, y) \mapsto \sum_{k=1}^n a_k b_k w_k$ forme un produit scalaire sur E. w s'appelle la *fonction de masse*, elle sert à donner plus ou moins d'importance à différents coefficients du vecteur.

Exemple *(La fonction de masse continue)*
Sur $E = \mathscr{C}^0([a, b]\mathbf{R})$, on définit $w \in E$ à valeurs **strictement positives**.
$(f,\, g) \mapsto \int_a^b f(t)g(t)w(t)\,\mathrm{d}t$ forme un produit scalaire sur E.

C Inégalité de Cauchy-Schwarz

Théorème 1.3 *(Inégalité de Cauchy-Schwarz)*

Soit E un espace préhilbertien muni du produit scalaire $(.|.)$,

$$\forall (x, y) \in E^2, \qquad |(x|y)| \leqslant \sqrt{(x|x)}\sqrt{(y|y)}$$

avec égalité si et seulement si x et y sont colinéaires.

Preuve
Cette preuve a déjà été faite dans le cas particulier de l'intégrale.
On étudie $(\lambda x + y|\lambda x + y)$

$$\begin{aligned}(\lambda x + y|\lambda x + y) &= \lambda^2(x|x) + \lambda(x|y) + \lambda(y|y) + (y|y) \qquad \text{par linéarité}\\ &= \lambda^2(x|x) + 2\lambda(x|y) + (y|y) \qquad \text{par symétrie}\end{aligned}$$

Or par positivité du produit scalaire

$$\forall \lambda \in \mathbf{R}, (\lambda x + y|\lambda x + y) \geqslant 0.$$

Donc le polynôme réel $X^2(x|x) + 2X(x|y) + (y|y)$ admet au plus une racine.
⋆ Si $(x|x) = 0$, alors $x = 0$ et l'inégalité est évidemment vérifiée et c'est même une égalité. On voit dans ce cas que x et y sont liés.
⋆ Si $x \neq 0$, alors $(x|x) > 0$ et le discriminant est négatif ou nul : $(x|y)^2 - (x|x)(y|y) \leqslant 0$.
Ce que l'on peut aussi écrire sous la forme $|(x|y)| \leqslant \sqrt{(x|x)}\sqrt{(y|y)}$.
On a l'égalité lorsque le discriminant est nul, c'est-à-dire lorsqu'il existe $\lambda \in \mathbf{R}$ racine du polynôme qui vérifie donc $(\lambda x + y|\lambda x + y) = 0$.
D'après le caractère défini du produit scalaire, ceci n'est possible que si $\lambda x + y = 0$, c'est-à-dire si x et y sont liés. ■

Remarque : On voit que pour avoir l'inégalité de Cauchy-Schwarz, l'application n'a pas besoin d'être « définie ». Cette condition n'intervient que dans le cas d'égalité.

D Norme et distance

Définition 1.4 *(Norme)*

Soit E un $\mathbf{R}$-espace vectoriel,
on appelle **norme** sur E toute application $N : E \to \mathbf{R}_+$ qui est

1. *définie* (ou axiome de séparation) : $\forall x \in E,\, N(x) = 0 \iff x = 0$,
2. *absolument homogène* : $\forall x \in E,\, \forall \lambda \in \mathbf{R}, N(\lambda x) = |\lambda| N(x)$,
3. vérifie *l'inégalité triangulaire* : $\forall (x, y) \in E^2,\, N(x + y) \leqslant N(x) + N(y)$.

On appelle **espace vectoriel normé**, tout espace vectoriel muni d'une norme.

Exemple
Le modèle pour la définition d'une norme est la valeur absolue.

Exemple
Montrer que la positivité de la norme découle des autres axiomes.
Solution :

$$\begin{aligned}\text{Soit } x \in E,\ 0 &= N(0) && \text{(séparation)}\\ &= N(x - x)\\ &\leqslant N(x) + N(-x) && \text{(inégalité triangulaire)}\\ &\leqslant 2N(x) && \text{(absolue homogénéité)}\end{aligned}$$

Donc $N(x) \geqslant 0$.

Exemple
Sur $\mathbf{R}^n$, on définit les normes :

- $\|x\|_1 = \sum_{i=1}^n |x_i|$,
- $\|x\|_2 = \sqrt{\sum_{i=1}^n x_i^2}$,
- $\forall p > 1,\ \|x\|_p = \left(\sum_{i=1}^n |x_i|^p\right)^{\frac{1}{p}}$,
- $\|f\|_\infty = \sup_{i \in [\![1,n]\!]} |x_i|$.

Exemple

Sur $\mathscr{C}^0([a,b],\mathbf{R})$, on définit différentes normes sur le même modèle de que cas discret :

- $\|f\|_1 = \displaystyle\int_{[a,b]} |f|$,
- $\|f\|_2 = \sqrt{\displaystyle\int_{[a,b]} f^2}$,
- $\forall p > 1,\ \|f\|_p = \left(\displaystyle\int_{[a,b]} |f|^p\right)^{\frac{1}{p}}$ (voir exercice),
- $\|f\|_\infty = \sup_{[a,b]} |f|$.
  (voir l'exercice du TD d'intégration pour comprendre la notation « $\|\cdot\|_\infty$ »).

Solution :

- L'application $\|\cdot\|_1$ est bien définie sur $\mathscr{C}^0\,([a,\,b],\,\mathbf{R})$ car toute fonction continue sur un segment admet une primitive et si f est continue, alors par composition, $x \mapsto |f(x)|$ l'est aussi.
  - par positivité de l'intégrale $\|\cdot\|_1$ est positive, et vérifie l'axiome de séparation ($x \mapsto |f(x)|$ positive et continue).
  - $\|\lambda f\|_1 = \int_a^b |\lambda f| = |\lambda| \int_a^b |f| = |\lambda|\,\|f\|_1$.
  - la valeur absolue vérifie l'inégalité triangulaire, donc par croissance de l'intégrale

$$\|f+g\|_1 = \int_a^b |f+g| \leqslant \int_a^b |f| + |g| = \int_a^b |f| + \int_a^b |g| = \|f\|_1 + \|g\|_1\,.$$

- On verra plus loin que cette norme $\|\cdot\|_2$ « dérive » du produit scalaire.

Propriété 1.5 *(2ème inégalité triangulaire)*

Soit $(x,y) \in E^2$, et N une norme sur E,

$$N(x+y) \geqslant |N(x) - N(y)|.$$

Preuve

$N(x) = N(x+y-y) \leqslant N(x+y) + N(-y) = N(x+y) + N(y)$.
Donc $N(x) - N(y) \leqslant N(x+y)$.
Et par symétrie des rôles entre x et y, on a $N(y) - N(x) \leqslant N(x+y)$.
Donc $|N(x) - N(y)| \leqslant N(x+y)$. ∎

Définition 1.6 *(Distance)*

Soit E un $\mathbf{R}$-espace vectoriel,
on appelle **distance** sur E toute application $d : E^2 \to \mathbf{R}_+$ qui vérifie

1. (*axiome de séparation*) $\forall (x,y) \in E^2,\ d(x,y) = 0 \iff x = y$,
2. (*symétrie*) $\forall (x,y) \in E^2,\ d(x,y) = d(y,x)$,
3. (*inégalité triangulaire*) $\forall (x,y,z) \in E^3,\ d(x,z) \leqslant d(x,y) + d(y,z)$.

Remarque : La notion de distance n'est pas spécifique aux espaces vectoriels.

⚠ Les normes et les distances ne sont pas des applications linéaires.

Exemple

Le caractère positif découle-t-il aussi des autres axiomes, comme pour la norme, ou non ?

Solution :

Si $(x,\,y) \in E^2$ alors

$$0 = d(x,\,x) \leqslant d(x,\,y) + d(y,\,x) = 2d(x,\,y).$$

On retrouve donc bien la positivité d'après les autres axiomes comme pour la norme.

Propriété 1.7

Toute norme définit une distance par

$$d\ : (x,\,y) \mapsto N(y-x).$$

Preuve

Cette fonction vérifie évidemment les axiomes d'une distance :

- positivité d'après la positivité de la norme.
- séparation d'après celle de la norme :

$$d(x,\,y) = 0 \iff N(y-x) = 0 \iff y - x = 0 \iff y = x.$$

- symétrie d'après l'absolue homogénéité de la norme (utilisée avec -1).
- inégalité triangulaire directement d'après celle de la norme :

$$d(x,\,y) = N(y-x) = N(y-z+(z-x)) \leqslant N(y-z) + N(z-x) = d(y,\,z) + d(z,\,x).$$

∎

E Norme euclidienne

Théorème 1.8 *(Norme préhilbertienne ou euclidienne)*

Soit $(E, (.|.))$ un espace préhilbertien,

1. $x \mapsto \|x\| = \sqrt{(x|x)}$ définit une norme sur E. appelée **norme préhiblertienne ou euclidienne** associée à $(.|.)$ selon que l'espace est de dimension finie ou non.
2. $x \mapsto d(x,y) = \|x-y\|$ définit une distance sur E, appelée **distance préhilbertienne ou euclidienne** associée à $(.|.)$.

Exemple

La norme euclidienne usuelle sur $\mathbf{R}^2$ est celle que vous utilisez depuis le collège :

$$N : (x,y) \mapsto \sqrt{x^2+y^2}.$$

Elle dérive du produit scalaire usuel $((x_1,y_1)|(x_2,y_2)) \mapsto x_1x_2 + y_1y_2$.

Preuve

De par la positivité du produit scalaire, $x \mapsto \sqrt{(x|x)}$ existe bien (et positive).

- *séparation :* si $\sqrt{(x|x)} = 0$, alors $(x|x) = 0$ donc $x = 0$ (axiome de séparation du produit scalaire).
- *absolue homogénéité :* elle découle de la bilinéarité du produit scalaire. Si $\lambda \in \mathbf{R}$,
$$\|\lambda x\| = \sqrt{(\lambda x|\lambda x)} = \sqrt{\lambda^2(x|x)} = |\lambda|\sqrt{(x|x)} = |\lambda|\,\|x\|\,.$$
- *inégalité triangulaire :* elle découle de Cauchy-Schwarz.

$$\begin{aligned}
\|x+y\|^2 &= (x+y|x+y) \\
&= (x|x) + (y|y) + 2(x|y) && \text{bilinéarité} \\
&\leqslant \|x\|^2 + \|y\|^2 + 2\sqrt{(x|x)(y|y)} && \text{Cauchy-Schwarz} \\
&\leqslant \|x\|^2 + \|y\|^2 + 2\,\|x\|\,\|y\| \\
&\leqslant (\|x\| + \|y\|)^2\,.
\end{aligned}$$

Par positivité, on peut passer à la racine sans changer les inégalités, ce qui donne l'inégalité triangulaire.

■

Exemple

Pour les variables aléatoires à valeurs dans $[\![1,n]\!]$, on peut définir la norme associée au produit scalaire $(X,Y) \mapsto \mathbf{E}(XY)$. Elle s'exprime

$$\|X\| = \sqrt{\mathbf{E}\left(X^2\right)}.$$

On remarque que la distance de X à son espérance est donné par son écart type.

Remarque : On voit qu'il existe des normes ou des distances qui ne proviennent pas de produits scalaires : ce ne sont pas des normes (ou distances) euclidiennes.

Dans toute la suite du cours, les normes et distances désigneront des normes euclidiennes et des distances euclidiennes.

Propriété 1.9 *(Formules de polarisation)*

Soit E un espace préhilbertien muni de la norme euclidienne $\|.\|$.
$\forall (x,y) \in E^2$

$$\begin{aligned}
(x|y) &= \frac{1}{2}\left(\|x+y\|^2 - \|x\|^2 - \|y\|^2\right) \\
&= \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2\right).
\end{aligned}$$

Preuve

$$\|x+y\|^2 = (x+y|x+y) = (x|x) + 2(x|y) + (y|y) = \|x\|^2 + 2(x|y) + \|y\|^2\,.$$

Donc

$$(x|y) = \frac{1}{2}\left(\|x+y\|^2 - \|x\|^2 - \|y\|^2\right).$$

De même pour la deuxième égalité, on développe les normes. ■

Le théorème qui suit indique en particulier que pour un produit scalaire, il n'existe qu'une seule norme euclidienne associée.

Propriété 1.10 *(Caractérisation des normes euclidiennes)*

Soit N une norme sur E.
N est euclidienne si, et seulement si

$$\varphi\ : (x,\,y) \mapsto \frac{1}{2}\left(N(x+y)^2 - N(x)^2 - N(y)^2\right)$$

définit un produit scalaire sur E.

Preuve

Immédiat avec la formule de polarisation correspondante.
On peut remplacer la caractérisation avec l'autre formule de polarisation si besoin. ■

Exemple

Montrer que $N\ :\ (x_1,\,x_2) \mapsto \sqrt{x_1^2 + 2x_1x_2 + 3x_2^2}$ est une norme euclidienne sur $\mathbf{R}^2$.

Solution :

On pose $\varphi\ : (x,\,y) \mapsto \frac{1}{2}\left(N(x+y)^2 - N(x)^2 - N(y)^2\right)$.

Alors pour $(x,\,y)=((x_1,\,x_2),\,(y_1,\,y_2))$ on trouve

$$\begin{aligned}&\varphi(x,\,y)\\&=\frac{1}{2}\left((x_1+y_1)^2+2(x_1+y_1)(x_2+y_2)+3(x_2+y_2)^2-x_1^2-y_1^2-2x_1x_2\right.\\&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\left.-2y_1y_2-3x_2^2-3y_2^2\right)\\&=x_1y_1+x_1y_2+x_2y_1+3x_2y_2.\end{aligned}$$

On trouve clairement une forme bilinéaire symétrique définie positive.

Exemple

Montrer que $N_\infty:(x_1,\,\cdots,\,x_n)\mapsto\max_{i\in[\![1,\,n]\!]}|x_i|$ est une norme non euclidienne sur $\mathbf{R}^n$.

Solution :

Montrons que N est une norme.

- (positive) immédiat avec les valeurs absolues.
- (séparation)
  Soit $x=(x_i)_{i\in[\![1,\,n]\!]}$ tel que $N_\infty(x)=0$, alors $\max_{i\in[\![1,\,n]\!]}|x_i|=0$. Or comme tous les termes $|x_i|$ sont positifs, cela implique que $\forall i\in[\![1,\,n]\!]$, $x_i=0$.
  Donc $x=0$.
- (absolue homogénéité)
  Soit $x\in\mathbf{R}^n$, $\lambda\in\mathbf{R}$, $\max_{i\in[\![1,\,n]\!]}|\lambda x_i|=\max_{i\in[\![1,\,n]\!]}(|\lambda|\ |x_i|)=|\lambda|\ \max_{i\in[\![1,\,n]\!]}|x_i|$.
  Donc $N_\infty(\lambda x)=|\lambda|\ N_\infty(x)$.
- (inégalité triangulaire)
  Pour (x_i), (y_i) deux vecteurs de $\mathbf{R}^n$.
  Pour tout $i\in[\![1,\,n]\!]$, $|x_i+y_i|\leqslant|x_i|+|y_i|$, donc, par passage au maximum,
  $$N_\infty(x+y)\leqslant N_\infty(x)+N_\infty(y).$$

Si on utilise la formule de polarisation, on trouve

$$\varphi(x,\,y)=\frac{1}{2}\left(N_\infty(x+y)^2-N_\infty(x)-N_\infty(y)\right).$$

Montrons que φ n'est pas bilinéaire.
Notons $(e_i)_{i\in[\![1,\,n]\!]}$ la base canonique.

$$\varphi(e_1+e_2,\,e_2)=\frac{1}{2}\left(N_\infty(e_1+2e_2)^2-N_\infty(e_1+e_2)^2-N_\infty(e_2)^2\right)=\frac{1}{2}(4-4-1)=-\frac{1}{2}.$$

Et $\varphi(e_1,\,e_2)=\frac{1}{2}\left(N_\infty(e_1+e_2)^2-N_\infty(e_1)^2-N_\infty(e_2)^2\right)=\frac{1}{2}(4-1-1)=1$,
$\varphi(e_2,\,e_2)=N(e_2)^2=1$.
La formule de polarisation n'est donc pas vérifiée.

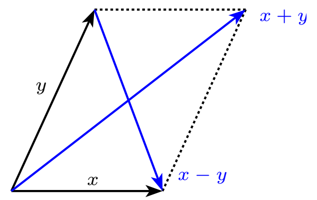
Exemple

Montrer que le caractère symétrique défini positif de φ est automatiquement vérifié si N est une norme, et qu'il suffit de vérifier que φ est bilinéaire pour montrer que c'est un produit scalaire.

Propriété 1.11 *(Identité du parallélogramme)*

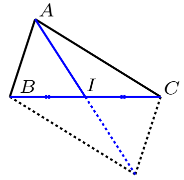
$\forall(x,y)\in E^2$,

$$\|x+y\|^2+\|x-y\|^2=2\left(\|x\|^2+\|y\|^2\right).$$



L'identité du parallélogramme n'est pas explicitement au programme de MPSI, mais comme elle est au programme du brevet des collèges, il serait bienvenu de la connaître.
Remarque : Reformulée différemment, elle donne le *lemme de la médiane.*

$$AB^2+AC^2=\frac{1}{2}BC^2+2AI^2.$$



Preuve

Trivial en développant les normes. ■

Exemple

Pour montrer que N_∞ n'est pas une norme euclidienne, on peut aussi montrer qu'elle ne vérifie pas l'identité du parallélogramme, ce qui est en général plus simple que de passer par la formule de polarisation.

Remarque : Le théorème de Fréchet - Von Neumann - Jordan affirme qu'une norme dérive d'un produit scalaire si, et seulement si elle vérifie l'identité du parallélogramme (voir exercice).

Théorème 1.12 *(Inégalité de Cauchy-Schwarz)*

Soit E un espace préhilbertien muni du produit scalaire $(.|.)$,

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

avec égalité si, et seulement si x et y sont colinéaires.
Autre formulation : Si x,y non nuls, alors

$$\left|\left(\frac{x}{\|x\|}\middle|\frac{y}{\|y\|}\right)\right|\leqslant 1.$$

Nous avons déjà démontré ce résultat plus haut. Il ne s'agit que d'une reformulation. Cette inégalité est très importante et se retrouve dans de nombreux contextes mathématiques. Quand vous avez une inégalité à montrer, pensez-y. Nous verrons plusieurs exemples en exercice.
C'est cette égalité qui permet de définir un angle (non orienté) entre deux vecteurs de $\mathbf{R}^n$ par $\widehat{x,y}=\theta=\operatorname{Arccos}\left(\left(\frac{x}{\|x\|}\middle|\frac{y}{\|y\|}\right)\right)$.

Nous avons donc vu que nous pouvons avoir l'enchaînement suivant

$$\text{Produit scalaire} \Rightarrow \text{Norme} \Rightarrow \text{Distance.}$$

Ainsi, un espace muni d'un produit scalaire possède naturellement une norme, et donc une distance.
Par contre, aucune de ces étapes ne peut se faire automatiquement dans l'autre sens et un espace vectoriel normé n'est pas nécessairement préhilbertien : nous avons vu qu'en général une norme ne dérive pas d'un produit scalaire. De même, on peut imaginer des distances qui ne proviennent d'aucune norme comme dans l'exemple qui suit.

Exemple

Soit E un espace vectoriel non réduit à 0.
On définit la distance d sur E par

$$\forall (x,\, y) \in E^2,\, d(x,\, x) = 0 \text{ et } d(x,\, y) = 1 \text{ si } x \neq y.$$

On vérifie aisément que d est bien un distance sur E.
Elle est à valeurs dans $\mathbf{R}_+$, symétrique, vérifie l'axiome de séparation et une simple disjonction des cas, montre qu'elle vérifie l'inégalité triangulaire.
Cependant, si d provenait d'une norme N alors $N(x) = d(x,\, 0)$. Pour $x \neq 0$, on a alors $N(2x) = d(2x,\, 0) = 1 = N(x)$, donc N n'est pas absolument homogène donc ce n'est pas une norme.

2 Orthogonalité

A Familles orthogonales

Définition 2.1 *(Vecteurs orthogonaux)*

Deux vecteurs x, y sont dits **orthogonaux**, si $(x|y) = 0$.
Deux vecteurs x, y sont dits **orthonormaux** si $(x|x) = (y|y) = 1$ et $(x|y) = 0$.

Théorème 2.2

Toute famille de vecteur **tous non nuls** et orthogonaux deux à deux est libre.
En particulier, toute famille orthonormale est libre.

⚠ Une famille de vecteurs orthogonaux n'est pas nécessairement libre car elle peut contenir le vecteur nul.

Preuve

Si $(x_1, x_2, \cdots, x_p)$ forme une famille de vecteurs non nuls orthogonaux deux à deux, alors si on pose $(\lambda_1, \lambda_2, \cdots, \lambda_p)$ tels que $\sum_{i=1}^{p} \lambda_i x_i = 0$,
$\forall k \in [\![1, p]\!]$,

$$0 = \Big(\sum_{i=1}^{p} \lambda_i x_i \Big| x_k\Big) = \lambda_k \left\|x_k\right\|^2.$$

Or $x_k \neq 0$, donc $\lambda_k = 0$.
Donc la famille est libre. ■

Théorème 2.3 *(Pythagore)*

Soient $(x_1, x_2, \cdots, x_n) \in E^n$, une famille de vecteurs orthogonaux deux à deux,

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

Preuve

On passe au produit scalaire :

$$\left\|\sum_{k=1}^{n} x_k\right\|^2 = \Big(\sum_{k=1}^{n} x_k \Big| \sum_{k=1}^{n} x_k\Big) = \sum_{k=1}^{n}\sum_{i=1}^{n} (x_k|x_i) = \sum_{k=1}^{n}\sum_{i=1}^{n} \delta_{k,i}(x_k|x_i) = \sum_{k=1}^{n} (x_k|x_k)$$
$$= \sum_{k=1}^{n} \left\|x_k\right\|^2.$$

■

Finalement, il est plus facile de démontrer un résultat en dimension n que de faire la preuve avec la géométrie *classique* en dimension 2.
C'est un exemple de la puissance des théories mathématiques que l'apprend maintenant. Au prix d'un effort conceptuel plus grand au départ de nombreux résultats deviennent triviaux.
Mais il faut reconnaître que cela ne rend pas caduques pour autant toutes les preuves traditionnelles avec leur élégance et qui procèdent d'une autre démarche (en fait, ces nouvelles théories ont été bâties de manière à intégrer ces résultats au cœur de leur axiomes ce qui n'a de sens que parce qu'on a montré leur caractère fondamental par ces moyens classiques).

Exemple

Soit Ω un ensemble fini et E l'ensemble des variables aléatoires réelles centrées sur Ω.

1. Vérifier que E est un espace vectoriel.
2. Montrer que $(X,\, Y) \mapsto (X|Y) = \mathbf{E}\,(XY)$ définit un produit scalaire sur E.
3. Comment exprimer simplement que deux variables aléatoires sont orthogonales ?
4. Énoncer le théorème de Pythagore dans cet espace préhilbertien.

Solution :

1. C'est le noyau de $X \mapsto \mathbf{E}\,(X)$ qui est une appliation linéaire sur l'espace vectoriel des variables aléatoires réelles.

2. On a montré que c'est un produit scalaire sur l'ensemble des variables aléatoires réelles.
   Restreint à E, il garde évidemment les mêmes propriétés.
3. On remarque que pour X, Y deux variables de E, $(X|Y) = \mathbf{E}(XY) = \mathbf{E}(XY) - \mathbf{E}(X)\mathbf{E}(Y)$ car $\mathbf{E}(X) = \mathbf{E}(Y) = 0$.
   Donc $(X|Y) = \mathbf{Cov}(X, Y)$.
   Ainsi deux variables sont orthogonales si, et seulement si elles sont décorrellées.
4. Dans le cas des variables centrées, la norme est donc l'écart type.
   Ainsi, pour des variables aléatoires sont décorrellées deux à deux, on obtient l'égalité de Bienaymé :
$$\mathbf{V}(X_1 + X_2 + \cdots + X_n) = \sum_{k=1}^{n} \mathbf{V}(X_i).$$

B Orthogonal d'un partie

Définition 2.4 *(Orthogonal d'une partie)*

Soit $(E, (\cdot|\cdot))$ un espace préhilbertien et X une partie de E (pas nécessairement un espace vectoriel).
L'orthogonal de X est défini par
$$X^{\perp} = \{x \in E,\ \forall y \in X,\ (x, y) = 0\}.$$

Exemple

1. $\{0\}^{\perp} = E$.
2. $E^{\perp} = \{0\}$.

Propriété 2.5 *(Propriété de l'orthogonal)*

1. $X^{\perp}$ est un sous-espace vectoriel de E.
2. Si $X \subset Y$, alors $Y^{\perp} \subset X^{\perp}$.
3. $X \subset (X^{\perp})^{\perp}$.

Preuve

1. Si $(x, x') \in \left(X^{\perp}\right)^2$, et $\lambda \in \mathbf{R}$, alors pour tout $y \in X$, $(x+\lambda x'|y) = (x|y)+\lambda(x'|y) = 0$.
   Donc $(x + \lambda x') \in X^{\perp}$.
   De plus $0 \in X^{\perp}$ (immédiat), donc $X^{\perp}$ est bien un sous-espace vectoriel de E.
2. Soit $y \in Y^{\perp}$, alors pour tout $x \in X$, $x \in Y$, donc $(y|x) = 0$, donc $y \in X^{\perp}$.
   Ainsi $Y^{\perp} \subset X^{\perp}$.
3. Si $x \in X$, alors pour tout $y \in X^{\perp}$, $(y|x) = 0$, donc $(x|y) = 0$, donc $x \in (X^{\perp})^{\perp}$.
   En général l'inclusion est stricte. En effet, si X n'est pas un espace vectoriel, comme $(X^{\perp})^{\perp}$ en est un, on n'a pas l'égalité (mais même pour X un espace vectoriel, on n'a pas l'égalité en général en dimension infinie).

■

Exemple

Montrer que $X^{\perp} = (\mathrm{Vect}(X))^{\perp}$.
Solution :
$X \subset \mathrm{Vect}(X)$ donc $(\mathrm{Vect}(X))^{\perp} \subset X^{\perp}$.
Montrons l'inclusion réciproque.
Soit $x \in X^{\perp}$, considérons $y \in \mathrm{Vect}(X)$, y peut s'écrire comme combinaison des x_i, avec $x_i \in X$.
Ainsi $(x|x_i) = 0$ et par linéarité du produit scalaire, on a bien $(x|y) = 0$, donc $x \in (\mathrm{Vect}(X))^{\perp}$ ce qui prouve l'inclusion réciproque.

Exemple

Montrer que $A^{\perp} = \bigcap_{a\in A} \{a\}^{\perp}$.
Solution :
$$\begin{aligned} x \in A^{\perp} &\iff \forall a \in A,\ (x|a) = 0 \\ &\iff \forall a \in A,\ x \in \{a\}^{\perp} \\ &\iff x \in \bigcap_{a\in A} \{a\}^{\perp}. \end{aligned}$$

Propriété 2.6 *(Orthogonal d'un sous-espace vectoriel)*

Si F est un sous-espace vectoriel de E, alors
$$F + F^{\perp} = F \oplus F^{\perp}.$$

Preuve

Si $x \in F \cap F^{\perp}$, alors $x \in F^{\perp}$, donc $\forall y \in F$, $(x|y) = 0$, en particulier pour $y = x \in F$, on a $(x|x) = 0$ et par le caractère défini du produit scalaire, on a donc $x = 0$.
Ainsi $F \cap F^{\perp} = \{0\}$. ■

Un sous-espace et son orthogonal ne sont pas nécessairement supplémentaires dans un espace préhilbertien (un exemple est proposé plus loin dans ce cours). Par contre, nous verrons qu'en dimension finie, ils le sont.

Exemple

Soit $u = (1, 2, 0)$ et $v = (1, 1, 1)$. Donner une équation de l'espace orthogonal de $\mathrm{Vect}(u, v)$.

Solution :
Cela revient à trouver $w = (x, y, z)$ tel que $(u|w) = (v|w) = 0$, donc à résoudre ${}^t uw = {}^t vw = 0$.

Si on met ensemble les deux équations cela donne le système :

$$\begin{pmatrix} 1 & 2 & 0 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Pour résoudre, on voit que l'on a une matrice de rang 2 dans un espace de dimension 3.
Il y aura donc un paramètre et deux inconnues principales.
Le plus simple est de prendre x comme paramètre.

Alors le vecteur solution est $x \begin{pmatrix} 1 \\ -\frac{1}{2} \\ -\frac{1}{2} \end{pmatrix}$.

Donc $\mathrm{Vect}\,(u,v)^{\perp} = \mathrm{Vect}\,(2,-1,-1)$.

C Théorème d'orthonormalisation de Gram-Schmidt

Définition 2.7 *(Base et famille orthonormales)*

Une famille $(e_1, e_2, ..., e_n)$ est dite **orthonormale** (ou orthonormée) si ses vecteurs sont orthogonaux deux à deux et de norme 1,
c'est-à-dire si

$$\forall (i,j) \in [\![1,n]\!]^2, (e_i|e_j) = \delta_{i,j}.$$

Si la famille est une base, on parle de **base orthonormale**, on utilise parfois l'abréviation « B.O.N. », ou plus simplement « BON ».

Propriété 2.8

Toute famille orthonormale à n éléments dans un espace euclidien de dimension n est une base orthonormale.

Preuve

C'est une famille libre car elle est orthonormale et elle contient autant d'éléments que la dimension de l'espace : c'est donc une base. ∎

Exemple

La base canonique de $\mathbf{R}^n$ est orthonormale avec le produit scalaire usuel sur les coordonnées.
Par contre, si on change de produit scalaire et que l'on prend par exemple $(x|y) \mapsto 2x_1x_2 + y_1y_2$ la base canonique n'est plus orthonormale.

Le théorème qui suit est intéressant non seulement pour ce qu'il énonce, mais surtout pour sa preuve qu'il vous faudra maîtriser. Elle donne une méthode algorithmique complète pour construire une base orthonormale à partir de n'importe quelle base. Dans la pratique, on vous demandera facilement d'appliquer ce théorème. Cela peut constituer un exercice algorithmique à coder sous Python.

Théorème 2.9 *(Théorème d'orthonormalisation de Gram-Schmidt)*

Soit $E, (.|.)$ un espace vectoriel préhilbertien.
Si $(x_1, x_2, \cdots, x_n)$ est une famille libre de E,
alors il existe une famille $(e_1, e_2, \cdots, e_n)$ **orthonormale**, telle que

$$\forall i \in [\![1,n]\!],\ \mathrm{Vect}\,(e_1, e_2, \cdots, e_i) = \mathrm{Vect}\,(x_1, x_2, \cdots, x_i).$$

On peut imposer de plus que $\forall i \in [\![1,n]\!],\ (e_i|x_i) > 0,$ dans ce cas, cette famille est unique.

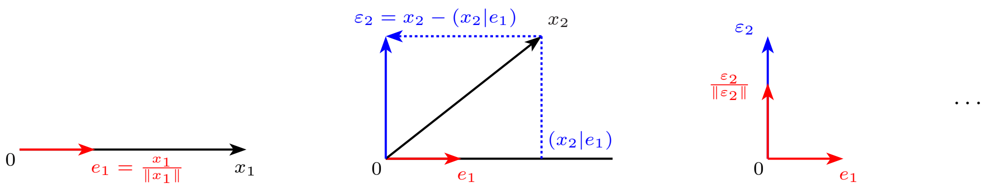
Preuve

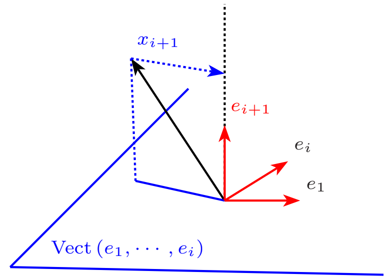
Existence : On procède par récurrence sur i :

$$\mathcal{P}(i) : \text{« On a construit } (e_1, e_2, \cdots, e_i) \text{ qui vérifie les hypothèses. »}$$

(*Initialisation*) pour $i = 1$:
On sait que la famille est une base, donc $\|x_1\| \neq 0$ et on peut poser $e_1 = \frac{x_1}{\|x_1\|}$.
Ainsi $\|e_1\| = 1$, $(e_1|x_1) > 0$ et $\mathrm{Vect}\,(e_1) = \mathrm{Vect}\,(x_1)$.
(*Hérédité*) on suppose le résultat vrai au rang $i \leqslant n-1$,
Pour définir e_{i+1}, on part de x_{i+1} auquel on veut déjà enlever toutes les composantes suivants les $(e_k)_{k \leqslant i}$ pour que la famille soit orthogonale (pour simplifier, on ne s'intéresse au caractère orthonormé qu'à la fin).

On pose donc $\varepsilon_{i+1} = x_{i+1} - \sum\limits_{k=1}^{i} (x_{i+1}|e_k)e_k$.





Vérifions que les propriétés sont vérifiées :

- *Égalité des espaces engendrés*
  $\mathrm{Vect}\,(e_1, e_2, \cdots, e_i, \varepsilon_{i+1}) \subset \mathrm{Vect}\,(x_1, x_2, \cdots, x_{i+1})$ (évident) et l'autre inclusion est aussi vérifiée car $x_{i+1} \in \mathrm{Vect}\,(e_1, \cdots, e_i, \varepsilon_{i+1})$.
- *famille orthogonale*

Pour tout $j \leqslant i$, on a

$$\begin{aligned}(\varepsilon_{i+1}|e_j) &= \left(x_{i+1} - \sum_{k=1}^{i}(x_{i+1}|e_k)e_k \;\middle|\; e_j\right)\\ &= (x_{i+1}|e_j) - \sum_{k=1}^{i}(x_{i+1}|e_k)\,(e_k|e_j) \qquad \text{(linéarité)}\\ &= (x_{i+1}|e_j) - \sum_{k=1}^{i}(x_{i+1}|e_k)\delta_{k,j} \qquad \text{(famille orthonormale)}\\ &= (x_{i+1}|e_j) - (x_{i+1}|e_j) = 0.\end{aligned}$$

- *Positivité du produit scalaire*
Ici, il faut penser à remplacer x_{i+1} par son expression en fonction des $(e_k)_{k\leqslant i}$ et ε_{i+1} et non pas le contraire.

$$x_{i+1} = \varepsilon_{i+1} + \sum_{k=1}^{i}(x_{i+1}|e_k)e_k.$$

$$\begin{aligned}(x_{i+1}|\varepsilon_{i+1}) &= \left(\varepsilon_{i+1} + \sum_{k=1}^{i}(x_{i+1}|e_k)e_k \;\middle|\; \varepsilon_{i+1}\right)\\ &= \|\varepsilon_{i+1}\|^2 + \sum_{k=1}^{i}(x_{i+1}|e_k)(e_k|\varepsilon_{i+1}) \qquad \text{(linéarité)}\\ &= \|\varepsilon_{i+1}\|^2 \qquad \text{(famille orthogonale)}\\ &> 0 \qquad \text{non nul comme vu juste après.}\end{aligned}$$

- *Norme*
$x_{i+1} \notin \mathrm{Vect}\,(e_1, e_2, \cdots, e_i)$ par hypothèse de récurrence, donc $\varepsilon_{i+1} \neq 0$.
Donc $\|\varepsilon_{i+1}\| \neq 0$ et on peut poser

$$e_{i+1} = \frac{\varepsilon_{i+1}}{\|\varepsilon_{i+1}\|}.$$

Alors $\|e_{i+1}\| = 1$ et il est évident que les autres propriétés restent vérifiées.

D'où le résultat vrai jusqu'au rang n par principe de récurrence.
Unicité :
On suppose l'existence d'une autre base $(e'_1, e'_2, \cdots, e'_n)$ qui vérifie les mêmes propriétés, et on montre par récurrence qu'elle est identique à la base $(e_1, e_2, \cdots, e_n)$ déjà construite.
(Initialisation) pour $i = 1$, $\mathrm{Vect}\,(x_1) = \mathrm{Vect}\,(e_1) = \mathrm{Vect}\,(e'_1)$, donc il existe $\lambda \in \mathbf{R}$ tel que $e'_1 = \lambda e_1$. Or $\|e_1\| = \|e'_1\|$, donc $\lambda = \pm 1$.
De plus, $(x_1|e'_1) = \lambda(x_1|e_1)$, et comme les deux produits scalaires sont positifs, $\lambda > 0$, donc $\lambda = 1$. Donc $e'_1 = e_1$.
(Hérédité) on suppose le résultat vrai jusqu'au rang i.
Alors $e'_{i+1} \in \mathrm{Vect}\,(x_1, x_2, \cdots, x_{i+1}) = \mathrm{Vect}\,(e_1, e_2, \cdots, e_{i+1})$.
Donc il existe $(\lambda_1, \lambda_2, \cdots, \lambda_{i+1}) \in \mathbf{R}^{i+1}$ tels que

$$e_{i+1'} = \sum_{k=1}^{i+1}\lambda_k e_k = \lambda_{i+1}e_{i+1} + \sum_{k=1}^{i}\lambda_k e'_k.$$

Par hypothèse de récurrence.
Or la famille des (e'_i) est orthonormale, donc pour tout $j \leqslant i$

$$\begin{aligned}0 &= (e'_{i+1}|e'_j)\\ &= \left(\lambda_{i+1}e_{i+1} + \sum_{k=1}^{i}\lambda_k e'_k \;\middle|\; e'_j\right)\\ &= \lambda_{i+1}(e_{i+1}|e'_j) + \sum_{k=1}^{i}\lambda_k(e'_k|e'_j) && \text{par linéarité}\\ &= \lambda_{i+1}(e_{i+1}|e'_j) + \sum_{k=1}^{i}\lambda_k\delta_{k,j} && \text{car la famille est orthonormée}\\ &= \lambda_{i+1}(e_{i+1}|e'_j) + \lambda_j\\ &= \lambda_{i+1}(e_{i+1}|e_j) + \lambda_j && \text{par hypothèse de récurrence } e_j = e'_j\\ &= \lambda_j && \text{car la famille est orthonormée.}\end{aligned}$$

Donc $e'_{i+1} = \lambda_{i+1}e_{i+1}$ et par des arguments similaires à ceux de l'initialisation, $e'_{i+1} = e_{i+1}$.
D'où l'unicité par principe de récurrence. ■

Remarque : Parfois nous parlons simplement de l'algorithme de Schmidt.
Ce procédé est également valable pour une famille libre en dimension quelconque, sans parler nécessairement de base.

Corollaire 2.10 *(Existence d'une base orthonormée)*

Tout espace vectoriel euclidien admet une base orthonormée.

Méthode *(Algorithme de Gram-Schmidt « à la main »)*

Dans la pratique pour les calculs à la main, on préfère en général garder la normalisation pour la fin (cela évite les quotients).
On construit donc une famille orthogonale $(\varepsilon_i)_{1\leqslant i\leqslant n}$ par récurrence :

- On pose $\varepsilon_1 = x_1$
- On pose $\varepsilon_{i+1} = x_{i+1} + \sum_{k=1}^{i}\lambda_k\varepsilon_k$.
- On trouve les λ_k
  - soit en en résolvant un système avec $(\varepsilon_{i+1}|\varepsilon_i) = 0$, puis $(\varepsilon_{i+1}|\varepsilon_{i-1}) = 0, \ldots, (\varepsilon_{i+1}|\varepsilon_1) = 0$,
  - ou directement $\lambda_k = -\frac{1}{\|\varepsilon_k\|}(x_{i+1}|\frac{\varepsilon_k}{\|\varepsilon_k\|})$.

On normalise ensuite la base $(\varepsilon_1, \cdots, \varepsilon_n)$ à la fin.

Exemple

Sur $E = \mathbf{R}_2[X]$, on définit le produit scalaire $(P|Q) = \int_{-1}^{1} PQ$.
Orthonormaliser la base canonique de $\mathbf{R}_2[X]$ par le procédé de Schmidt.
Solution :
La base canonique est $x_1 = 1, x_2 = X, x_3 = X^2$.

On pose donc $\varepsilon_1 = 1$, puis $\varepsilon_2 = X + \lambda\varepsilon_1 = X + \lambda$
On veut $0 = (\varepsilon_1|\varepsilon_2) = \int_{-1}^{1}(x+\lambda)\,\mathrm{d}x = 2\lambda$, donc $\varepsilon_2 = X$.
On pose enfin $\varepsilon_3 = X^2 + \lambda_2\varepsilon_2 + \lambda_1\varepsilon_1 = X^2 + \lambda_2 X + \lambda_1$.
Or $0 = (\varepsilon_1|\varepsilon_3) = \int_{-1}^{1}(x^2 + \lambda_2 x + \lambda_1)\,\mathrm{d}x = \frac{2}{3} + 2\lambda_1$, donc $\lambda_1 = -\frac{1}{3}$.
De plus $0 = (\varepsilon_2|\varepsilon_3) = \int_{-1}^{1} x^3 + \lambda_2 x^2 + \lambda_1 x\,\mathrm{d}x = \frac{2\lambda_2}{3}$, donc $\lambda_2 = 0$.
Ainsi $\varepsilon_3 = X^2 - \frac{1}{3}$.
Enfin on normalise : $\|\varepsilon_1\| = \sqrt{2}$, $\|\varepsilon_2\| = \sqrt{\frac{2}{3}}$ et $\|\varepsilon_3\| = \sqrt{\frac{8}{45}}$.
Donc la base orthonormée correspondante est :

$$\left(\frac{1}{\sqrt{2}}, \sqrt{\frac{3}{2}}X, \sqrt{\frac{45}{8}}\left(X^2 - \frac{1}{3}\right)\right).$$

Remarque : Pour construire ε_{i+1}, on impose qu'il soit orthogonal à tous les précédents : $(\varepsilon_j)_{j\in[\![1,\,i]\!]}$.
On peut voir que c'est équivalent au fait qu'il soit orthogonal à tous les $(x_j)_{j\in[\![1,\,i]\!]}$.
En effet, s'il est orthogonal à une famille de vecteurs, alors il est orthogonal à ses combinaisons linéaires, donc à son espace vectoriel engendré.
Par égalité entre Vect $(x_1, \cdots, x_i)$ et Vect $(\varepsilon_1, \cdots, \varepsilon_i)$ on a donc l'équivalence voulue.

Théorème 2.11 *(Théorème de la base orthonormée incomplète)*

Toute famille orthonormée d'un espace euclidien peut être complétée en une base orthonormée.

Preuve

Il suffit de compléter la famille (qui est libre) en une base, et d'orthonormaliser cette famille.
Les premiers vecteurs sont alors inchangés. ■

D Écriture matricielle en dimension finie

Théorème 2.12 *(Coordonnées dans une base orthonormée)*

Soit $\mathscr{E} = (e_1, e_2, \cdots, e_n)$ une base *orthonormée* (BON) de $(E, (.|.))$.
Si $x \in E$, alors

$$x = \sum_{i=1}^{n}(x|e_i)e_i$$

$$\mathrm{mat}_{\mathscr{E}}(x) = \begin{pmatrix}(x|e_1)\\(x|e_2)\\ \vdots \\(x|e_n)\end{pmatrix}.$$

Preuve

Puisque $\mathscr{E} = (e_1, e_2, \cdots, e_n)$ est une base, on peut écrire $x = \sum_{i=1}^{n} x_i e_i$.
Or $\forall j \in [\![1, n]\!]$, $(x|e_j) = \sum_{i=1}^{n} x_i(e_i|e_j) = x_j$. ■

Corollaire 2.13 *(Produit scalaire et norme dans une base orthonormée)*

Soit E un espace euclidien et $\mathscr{E} = (e_1, e_2, \cdots, e_n)$ une base **orthonormée** de E.
Si $x = \sum_{i=1}^{n} x_i e_i$ et $y = \sum_{i=1}^{n} y_i e_i$, alors

$$(x|y) = {}^tXY.$$

Preuve

$$(x|y) = \sum_{i=1}^{n}\sum_{j=1}^{n} x_i y_j (e_i|e_j) = \sum_{i=1}^{n} x_i y_i.$$

■

Exemple

La base canonique de $\mathbf{R}^n$ est orthonormée pour le produit scalaire usuel.
Par exemple pour $u = (x_1, y_1, z_1)$ et $v = (x_2, y_2, z_2)$, alors

$$(u|v) = {}^tuv = (x_1, y_1, z_1)\begin{pmatrix}x_1\\y_2\\z_2\end{pmatrix} = x_1x_2 + y_1y_2 + z_1z_2.$$

Théorème 2.14 *(Produit scalaire dans une base quelconque)*

Soit E un $\mathbf{R}$-espace vectoriel de dimension finie n, muni du produit scalaire $(.|.)$.
Soit $\mathscr{C} = (e_1, e_2, \cdots, e_n)$ une base **quelconque** de E.
Si $x = \sum_{i=1}^{n} x_i e_i$ et $y = \sum_{i=1}^{n} y_i e_i$, alors

$$(x|y) = \sum_{i=1}^{n}\sum_{j=1}^{n} x_i y_j (e_i|e_j).$$

Si on note $A = \big((e_i|e_j)\big)_{i,j} \in \mathscr{M}_n(\mathbf{R})$, alors

$$(x|y) = {}^tXAY.$$

E Complément hors programme

Définition 2.15 *(HP – Matrice du produit scalaire)*

Avec les notations du théorème précédent,
Si on note φ le produit scalaire $(.|.)$, alors on pose $A = \mathrm{mat}_{\mathscr{C}}(\varphi)$.

⚠ φ est une forme bilinéaire, il s'agit donc d'une nouvelle définition totalement différente de celle des matrices des applications linéaires.
Il ne faut pas confondre les deux.

Propriété 2.16 *(HP – Réduction de la matrice d'un produit scalaire)*

Soit φ un produit scalaire sur E euclidien.
Si $\mathscr{B}$ est une base orthonormée pour φ, alors

$$\mathrm{mat}_{\mathscr{B}}(\varphi) = I_n.$$

Preuve

Trivial, il suffit d'écrire les produits scalaires. ■

⚠ La réduction matricielle des formes quadratiques (dont les matrices de produits scalaires sont un cas particulier) n'est pas la même que celle des endomorphismes. La façon d'interpréter la matrice étant différente, l'opération de changement de base aussi. Si $A = \mathrm{mat}(\varphi, \mathscr{B})$ et $A' = \mathrm{mat}(\varphi, \mathscr{B}')$ sont les matrices d'un même produit scalaire dans deux bases différentes. On note $P = \mathrm{Pa}_{\mathscr{B}}^{\mathscr{B}'}$, ce qui donne :

$$A' = {}^tPAP.$$

L'inverse de la matrice P a été remplacé par la transposée pour avoir ${}^t(PX)A(PY) = X({}^tPAP)Y$.
D'après l'algorithme d'orthonormalisation de Gram-Schmidt, nous avons vu qu'il est toujours possible de construire une base en laquelle une matrice d'un produit scalaire s'écrit I_n.
Ainsi, toute matrice d'un produit scalaire peut s'écrire sous la forme

$$A = \left({}^tP\right)^{-1} P^{-1} = {}^t(P^{-1})P^{-1}.$$

On peut poser $Q = P^{-1}$ et on obtient alors que toute matrice de produit scalaire s'écrit $A = {}^tQQ$, avec Q une matrice de changement de base.
On peut y voir une caractérisation des matrices de produit scalaire.
On verra plus loin que dans le cadre d'une matrice de passage d'une base orthonormale vers une autre base orthonormale, vérifie ${}^tP = P^{-1}$.
En effet, entre les deux bases, le produit scalaire s'écrit de la même façon sous la forme de la matrice identité, ce qui donne donc bien $I = {}^tPIP$.

3 Projection et symétrie orthogonales

A Supplémentaire orthogonal

Propriété 3.1

Soit E un espace vectoriel préhilbertien.
Si F est un sous-espace vectoriel de **dimension finie**, alors

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

$F^\perp$ est **le supplémentaire orthogonal** de F dans E.

Preuve

F étant de dimension finie, on note $(e_1, \cdots, e_k)$ une base orthonormale de F.

Soit $x \in E$, $x_F = \sum_{i=1}^{k} (x|e_i)e_i \in F$., montrons que $x - x_F \in F^\perp$.

Soit $k \in [\![1, n]\!]$, $(x - x_F|e_j) = (x|e_j) - (x|e_j) = 0$, donc $x - x_F \in (\{e_1, \cdots, e_k\})^\perp = F^\perp$ car pour toute partie $A \subset E$; $A^\perp = (\mathrm{Vect}\,(A))^\perp$.
Ainsi $E = F \oplus F^\perp$. ■

⚠ Comme cela a été dit plus haut, ce résultat tombe en défaut si F n'est pas supposé de dimension finie.

Exemple

Soit $E = \mathscr{C}^0\left([-1, 1], \mathbf{R}\right)$ muni du produit scalaire défini par

$$(f|g) = \int_{-1}^{1} f(t)g(t)\ \mathrm{d}t.$$

On pose

$$F = \{f \in E, \forall t \in [-1, 0]\,,\ f(t) = 0\} \text{ et } G = \{g \in E, \forall t \in [0, 1]\,, g(t) = 0\}\,.$$

1. Montrer que $F^\perp = G$.

2. Les sous-espaces vectoriels F et G sont-ils supplémentaires ?

Solution :

1. Il est déjà clair que $G \subset F^\perp$.
   Réciproquement, si $\varphi \in F^\perp$, alors on peut poser f l'application de F telle que

$$f(t) = \begin{cases} 0 & \text{si } t < 0 \\ \varphi(t)\,t & \text{sinon} \end{cases}$$

   f est clairement continue et nulle sur $[-1, 0]$
   Comme $\varphi \in F^\perp$, on a $(\varphi|f) = 0$, donc $0 = \int_{-1}^{1} \varphi(t)f(t)\,\mathrm{d}t = \int_0^1 \varphi^2(t)t\,\mathrm{d}t$.
   Or,sur $[0, 1]$, $t \mapsto \varphi^2(t)t$ est positive, donc l'intégrale nulle donne

$$\forall t \in [0, 1],\ \varphi^2(t)t = 0$$

donc φ nulle sur $]0, 1]$ et par continuité, elle est aussi nulle en 0, ce qui donne $F^\perp \subset G$ et par double inclusion :

$$F^\perp = G.$$

2. Soit $f \in F$ et $g \in G$, alors pour tout $(\lambda, \mu) \in \mathbf{R}^2$ $(\lambda f + \mu g)(0) = 0$.
Donc si on considère, l'application constante égale à 1 (continue sur $[-1, 1]$, elle n'appartient pas à $F \oplus G$.
Les sous-espaces sont en somme directe, mais non supplémentaires dans E.

Propriété 3.2

Soit F un sous-espace vectoriel de E **euclidien**,

1. $E = F \oplus F^\perp$,
2. $\dim F^\perp = \dim E - \dim F$.
3. $(F^\perp)^\perp = F$,

Preuve

1. Théorème précédent.
2. dimension d'un supplémentaire.
3. $F \subset \left(F^\perp\right)^\perp$ et d'après le point précédent, $\dim \left(F^\perp\right)^\perp = \dim E - \dim F^\perp = \dim F$.

■

B Projecteurs et symétries

Définition 3.3 *(Projecteur orthogonal)*

Soient E un espace préhiblertien et F un sous-espace vectoriel de dimension **finie**.
On appelle projecteur orthogonal de E sur F **le** projecteur sur F parallèlement à $F^\perp$.
De plus, si $(e_1, \cdots, e_k)$ est une base orthonormée de F, alors

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

Preuve

Immédiat à partir de la définition précédente. ■

Définition 3.4 *(Symétrie orthogonale)*

Soient E un espace préhiblertien et F un sous-espace vectoriel de dimension **finie**.
On appelle symétrie orthogonale de E par rapport à F **la** symétrie par rapport à F parallèlement à $F^\perp$.
De plus, si $(e_1, \cdots, e_k)$ est une base orthonormée de F, alors

$$\forall x \in E, \quad s(x) = 2\sum_{i=1}^{k} (x|e_i)e_i x.$$

Bien noter l'unicité du projecteur ou de la symétrie (immédiat).

Remarque : Par unicité de l'orthogonal, il suffit de donner la base du projecteur ou de la symétrie pour avoir la direction. Ce n'est pas le cas lorsque le projecteur ou la symétrie n'est plus supposé orthogonal.

Propriété 3.5

s est une symétrie orthogonale de E si, et seulement si $p = \frac{1}{2}(s + \mathrm{Id})$ est une projection orthogonale de E.

Preuve

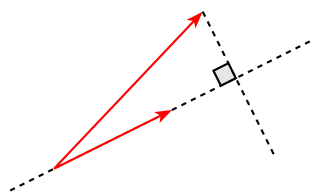
Si s est une symétrie orthogonale par rapport à F et parallèlement à $F^\perp$,
alors $\forall x \in E$, $x = x_F + x_{F^\perp}$ sur la somme directe.
On a donc $p(x) = \frac{1}{2}(x_F - x_{F^\perp} + x_F + x_{F^\perp}) = x_F$ ce qui montre bien que p est le projecteur orthogonal sur F.
Réciproquement, si p est un projecteur orthogonal sur F parallèlement à $F^\perp$, on montre de la même façon que $s = 2p - \mathrm{Id}$ est une symétrie orthogonale par rapport à F parallèlement à $F^\perp$. ■

Théorème 3.6 *(Expression de la projection orthogonale sur une droite)*

Soit E un espace vectoriel préhilbertien et $u \in E$, non nul.
Si p est un projecteur orthogonal de E sur $\mathrm{Vect}(u)$, alors

$$\forall x \in E, \quad p(x) = \frac{(x|u)}{\|u\|^2} u.$$

Preuve



Soit $x \in E$, on écrit $x = \lambda u + v$ avec $v \in (\mathrm{Vect}(u))^\perp$.

$p(x) = \lambda u$.
Or $(x|u) = (\lambda u + v|u) = \lambda \|u\|^2$, on trouve donc bien $\lambda = \frac{(x|u)}{\|u\|^2}$. ■

Théorème 3.7 *(Expression de la projection orthogonale sur un hyperplan)*

Soit E un espace vectoriel préhilbertien et $u \in E$, non nul.
Si p est un projecteur orthogonal de E sur $(\mathrm{Vect}\,(u))^\perp$, alors

$$\forall x \in E, \quad p(x) = x - \frac{(x|u)}{\|u\|^2} u.$$

Preuve

Comme pour la projection sur $\mathrm{Vect}\,(u)$, on trouve pour tout $x \in E$ $x = \frac{(x|u)}{\|u\|^2}u + v$ et $v \in (\mathrm{Vect}\,(u))^\perp$.
$p(x) = v = x - \frac{(x|u)}{\|u\|^2} u$. ■

Propriété 3.8 *(Réduction en dimension finie)*

Tout projecteur orthogonal de rang r dans E euclidien de dimension n peut s'écrire dans une certaine base orthonormée :

$$\left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right).$$

Toute symétrie orthogonale de rang r dans E euclidien de dimension n peut s'écrire dans une certaine base orthonormée :

$$\left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & -I_{n-r} \end{array}\right).$$

Preuve

Pour le projecteur :
Il suffit de choisir $(e_1, \cdots, e_r)$ une base orthonormée de $\mathrm{Im}\,p$
et de la compléter en $(e_1, \cdots, e_r, e_{r+1}, \cdots, e_n)$ une base de E.
Alors $\mathrm{Vect}\,(e_{r+1}, ..., e_n) = (\mathrm{Im}\,p)^\perp = \ker p$. D'où la forme voulue pour la matrice.
Pour la symétrie :
On choisit de la même façon les bases orthonormées correspondant au projecteur orthogonal associé.
Alors $\forall x \in \mathrm{Im}\,p, s(x) = x$ et $\forall x \in \ker p, s(x) = -x$. ■

Exemple *(Caractérisation des projecteurs et symétries orthogonaux)*

Soit E un espace vectoriel euclidien et $p \in \mathscr{L}(E)$.

1. Montrer que p est un projecteur orthogonal de E si et seulement si
   (a) $p \circ p = p$,
   (b) $\forall (x, y) \in E^2, (p(x)|y) = (x|p(y))$.

   Dans ce cas, p est le projecteur orthogonal sur $\mathrm{Im}\,p$ parallèlement à $\ker p = (\mathrm{Im}\,p)^\perp$.
   Un endomorphisme qui vérifie la deuxième propriété est dit *auto-adjoint.*

2. On montre de même que s est une symétrie orthogonale de E si et seulement si
   (a) $s \circ s = \mathrm{Id}_E$
   (b) $\forall x, y \in E^2, (s(x)|y) = (x|s(y))$

   Dans ce cas, s est la symétrie orthogonale par rapport $\ker(s - \mathrm{Id})$ parallèlement à $\ker(s + \mathrm{Id}) = (\ker(s - \mathrm{Id}))^\perp$.

3. En déduire que :
   P est la matrice d'une projection orthogonale dans une base orthonormée si, et seulement si $P^2 = P$ et ${}^tP = P$.
   S est la matrice d'une symétrie orthogonale dans une base orthonormée si, et seulement si $S^2 = I_n$ et ${}^tS = S$.

Solution :

1. *sens direct :* on a déjà vu que $p \circ p = p$
   $x = x_F + x_{F^\perp}$ et $y = y_F + y_{F^\perp}$, alors $(p(x)|y) = (x_F|y_F + y_{F^\perp}) = (x_F|y_F) + (x_F|y_{F^\perp}) = (x_F|y_F)$.
   Par symétrie, $(x|p(y)) = (x_F|y_F) = (p(x)|y)$.
   réciproque : $p \circ p = p$ donc p est un projecteur sur $F = \mathrm{Im}\,p$ parallèlement à $G = \ker p$ et en particulier ces deux espaces sont supplémentaires dans E. Montrons que $G = F^\perp$.
   Si $x \in F$, $y \in G$, alors $x = p(x)$ et $p(y) = 0$.
   Donc $(x|y) = (p(x)|y) = (x|p(y)) = 0$, donc x, y sont orthogonaux,
   Donc $G \subset F^\perp$ et par supplémentarité des espaces, on a égalité des dimensions.
   Donc $G = F^\perp$.
2. Utiliser le fait que $p = \frac{1}{2}(s + \mathrm{Id})$ est une projection orthogonale et réutiliser la question précédente par exemple.
   Sinon recommencer la même démonstration que pour la projection.
3. *(sens réciproque)* Si $P^2 = P$ alors P est la matrice d'une projection.
   Pour tout (x, y), le produit scalaire $(p(x)|y)$ dans une base orthonormée s'écrit matriciellement ${}^t(PX)Y = {}^tX{}^tPY = {}^tXPY$ ce qui correspond au produit scalaire $(x|p(y))$.

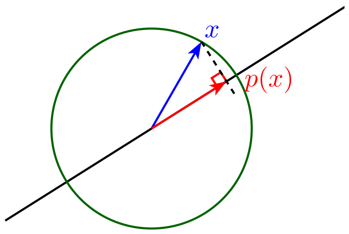
   (sens direct) $P^2 = P$ car la matrice représente une projection.
   Pour $x = e_i$ et $y = e_j$ deux vecteurs de la base orthonormée, on a $(p(e_i)|e_j) = (e_i|p(e_j))$ ce qui donne matriciellement ${}^t(PE_i)E_j = \left(\sum_{k=1}^{n} p_{k,\,i}{}^tE_k\right) E_j = p_{j,\,i}$.
   et ${}^tE_iPE_j = {}^tE_i \left(\sum_{k=1}^{n} p_{k,\,j}{}^tE_k\right) = p_{i,\,j}$.
   Donc par égalité des deux produits scalaires, on a bien $p_{i,\,j} = P_{j,\,i}$ donc $P = {}^tP$.

C Projecteurs et distances

Théorème 3.9 *(Inégalité de Bessel)*

Soient E préhilbertien et F un sous-espace vectoriel de dimension finie de E et p la projection orthogonale sur F.

$$\|p(x)\| \leqslant \|x\|.$$



Explications

La projection « raccourcit » les longueurs : elle sélectionne uniquement certaines coordonnées de p dans une base orthonormée ce qui a pour effet de rapprocher l'image de l'origine.
Si on s'intéresse un peu à l'analyse, on dit alors que

$$\forall (x,y) \in E^2, d(p(x),p(y)) \leqslant d(x,y).$$

La projection est 1-lipschitzienne et c'est donc une application continue !

Preuve

Soit $x \in E$,

$$\|x\|^2 = \|x-p(x)+p(x)\|^2.$$

Or $x-p(x) \in \operatorname{Ker} p$ et $p(x) \in F = \operatorname{Im} p$.
Donc $x-p(x)$ et $p(x)-y$ sont orthogonaux.
D'après le théorème de Pythagore,

$$\|x\|^2 = \|x-p(x)+p(x)\|^2 = \|x-p(x)\|^2 + \|p(x)\|^2 \geqslant \|p(x)\|^2.$$

D'où l'inégalité par passage à la racine carrée (croissante sur $\mathbf{R}_+$). ■

Exemple

On peut même montrer qu'une projection est 1-lipschitzienne si, et seulement si elle est orthogonale (voir feuille de TD).

Corollaire 3.10

Pour $(e_1, e_2, \cdots, e_p)$ une famille orthonormée de E, pour tout $x \in E$

$$\sum_{i=1}^{p} (x_k|e_k)^2 \leqslant \|x\|^2.$$

Preuve

On applique le théorème précédent au projecteur orthogonal sur $\operatorname{Vect}(e_1, \cdots, e_p)$. ■

Définition 3.11 *(Distance d'un point à un sous espace vectoriel)*

Soient E préhilbertien et F un sous-espace vectoriel de E de dimension finie.
On appelle distance de x à F et on note $d(x,F)$ le réel

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

Le théorème qui suit nous assure que cette distance existe et montre qu'elle correspond à un minimum.

Théorème 3.12

Soient E préhilbertien et F un sous-espace vectoriel de E de dimension finie.
Si on note p la projection orthogonale sur F, alors

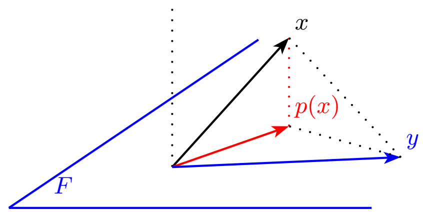
$$d(x,F) = \|x-p(x)\|.$$

$p(x)$ est l'**unique** vecteur de E qui réalise ce minimum :

$$\min_{y\in E} \|x-y\| = \|x-y_0\|.$$

La projection orthogonale donne la distance minimale : c'est la meilleure approximation au sens de la distance euclidienne.

Preuve



Soit $y \in F$,

$$\|x-y\| = \|x-p(x)+p(x)-y\|.$$

Or $x-p(x) \in \operatorname{Ker} p$ et $p(x)-y \in F = \operatorname{Im} p$.
Donc $x-p(x)$ et $p(x)-y$ sont orthogonaux, et d'après le théorème de Pythagore,

$$\|x-y\|^2 = \|x-p(x)+p(x)-y\|^2 = \|x-p(x)\|^2 + \|p(x)-y\|^2.$$

$\|p(x)-y\| \geqslant 0$, donc $\forall y \in E, \quad \|x-y\| \geqslant \|x-p(x)\|.$
Ainsi, la borne inférieure existe et est atteinte pour $y = p(x)$.
Donc $d(x,F)$ est bien définie et $d(x,F) = \|x-p(x)\|.$ ■

Exemple

Soit $E = \mathscr{C}^\infty([-1,1],\mathbf{R})$ muni du produit scalaire $(f|g) = \int_{-1}^{1} fg$.

Soit $f : x \mapsto e^x$ un vecteur de E, et $F = \mathbf{R}_2[X]$.
Alors la meilleur approximation de f par un polynôme de degré inférieur ou égal à 2 par rapport à la norme euclidienne de E, est donné par le polynôme P tel que $d(f, F) = d(f, P)$: le polynôme le plus proche de f au sens de la norme.
P s'obtient concrètement par $P = p(f)$ avec p le projecteur orthogonal sur $\mathbf{R}_2[X]$.
Une base orthonormée de $\mathbf{R}_2[X]$ s'obtient grâce au procédé d'orthonormalisation de Schmidt.
On a déjà trouvé $\left(e_1 = \frac{1}{\sqrt{2}}, e_2 = \sqrt{\frac{3}{2}}X, e_3 = \sqrt{\frac{45}{8}}\left(X^2 - \frac{1}{3}\right)\right)$.
Alors $P = (f|e_1)e_1 + (f|e_2)e_2 + (f|e_3)e_3$ que l'on calcule facilement avec l'expression du produit scalaire.

Remarques :

- Si on veut donner une importance particulière au voisinage de 0 par exemple dans notre approximation, on peut choisir un autre produit scalaire plus adapté. En l'occurrence, on introduirait une fonction de masse qui aurait des valeurs plus importantes au voisinage de 0 qu'aux bords de l'intervalle.
  Ainsi, lors de l'intégration, les valeurs proches de zéro, auront « plus de poids ».
- On pourrait voir que si on augmente le degré du polynôme, le projeté se rapprocherait infiniment de f, ce qui traduit que $\mathbf{R}[X]$ est dense dans $\mathbf{C}^\infty([-1, 1], \mathbf{R})$ pour la norme en question. Mais cela dépasse le cadre de ce cours.

Corollaire 3.13 *(Distance à un hyperplan vectoriel)*

Soit H un hyperplan de E euclidien. Alors il existe $a \in E$ tel que $H^\perp = \mathrm{Vect}\,(a)$.
Soit $x \in E$,
$$d(x, H) = \frac{|(x|a)|}{\|a\|}.$$

Méthode

Pour trouver la distance à un sous-espace vectoriel F de E.
On peut noter $\mathscr{B} = (e_1, \cdots, e_p)$ une base ON de F que l'on complète en une base ON de E $(e_1, \cdots, e_n)$.

- Si $p \leqslant n - p$ (la dimension de F est plus petite que celle de son orthogonale), alors on peut calculer la distance avec :
$$d(x, F) = \|x - p_F(x)\| = \left\| x - \sum_{i=1}^{p} (x|e_i)e_i \right\|.$$
- Sinon, si $p > n - p$, alors, on peut préférer l'expression :
$$d(x, F) = \|x - p_F(x)\| = \|x_{F^\perp}\| = \left\| \sum_{i=p+1}^{n} (x|e_i)e_i \right\| = \sqrt{\sum_{i=p+1}^{n} (x|e_i)^2}.$$

Exemple *(Un grand classique)*

Pour $a, b \in \mathbf{R}^2$, on pose $F(a, b) = \int_0^1 \left(t^2 - (at + b)\right)^2 \,\mathrm{d}t$.
Montrer que $\inf_{(a,b)\in\mathbf{R}^2} F(a, b)$ existe et est un minimum. Donner les valeurs de a, b pour le(s)quelle(s) il est atteint.

Solution :

Dans $E = \mathbf{R}_2[X]$, muni du produit scalaire $(P|Q) = \int_0^1 PQ$, la question revient à déterminer la distance minimale de X^2 au sous-espace vectoriel $\mathbf{R}_1[X]$.
Cette distance est atteinte en un unique point défini par le projeté de X^2 sur $\mathbf{R}_1[X]$.
Pour trouver $p(X^2) = \alpha X + \beta$ le projeté de X^2 sur $\mathbf{R}_1[X]$, on a plusieurs solutions :

- Soit on dispose d'une base orthonormée de $\mathbf{R}_1[X]$ et on calcule le projeté grâce à cette base.
- Soit on prend une base quelconque de $\mathbf{R}_1[X]$ et on cherche l'unique vecteur tel que $P - (\alpha X + \beta)$ soit orthogonal à tous les vecteurs de cette base.

Avec la deuxième méthode :
$$0 = \left(X^2 - (\alpha X + \beta) \middle| 1\right) = \int_0^1 t^2 - \alpha t - \beta \,\mathrm{d}t$$
$$= \frac{1}{3} - \frac{\alpha}{2} - \beta.$$
$$0 = \left(X^2 - (\alpha X + \beta) \middle| X\right) = \int_0^1 t^3 - \alpha t^2 - \beta t \,\mathrm{d}t$$
$$= \frac{1}{4} - \frac{\alpha}{3} - \frac{\beta}{2}.$$

La résolution donne $\alpha = 1$ et $\beta = -\frac{1}{6}$.
Donc la distance est atteinte pour un unique polynôme $X - \frac{1}{6}$.
La distance vaut
$$\min_{(a,b)} F(a, b) = F(\alpha, \beta) = \int_0^1 \left(t^2 - t - \frac{1}{6}\right)^2 \mathrm{d}t = \frac{1}{180}.$$

Exemple *(CCINP)*

On définit dans $\mathscr{M}_2(\mathbf{R}) \times \mathscr{M}_2(\mathbf{R})$ l'application $\varphi\ : (A,\, A') \mapsto \mathrm{tr}\,({}^tAA')$.
On rappelle que φ définit un produit scalaire sur $\mathscr{M}_2(\mathbf{R})$.
On note $\mathscr{F} = \left\{ \begin{pmatrix} a & b \\ -b & a \end{pmatrix},\ (a,\, b) \in \mathbf{R}^2 \right\}$.

1. Montrer que $\mathscr{F}$ est un sous-espace vectoriel de $\mathscr{M}_2(\mathbf{R})$.
2. Déterminer une base de $\mathscr{F}^\perp$.
3. Déterminer la projection de $J = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ sur $\mathscr{F}^\perp$.
4. Calculer la distance de J à $\mathscr{F}$.

Solution :

Ne pas hésiter à vérifier que l'on sait bien prouver que φ est un produit scalaire sur $\mathscr{M}_2(\mathbf{R})$.

1. On remarque que $\mathscr{F} = \text{Vect}\,(I_2,\, K)$ avec $K = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ ce qui prouve que c'est un sous-espace vectoriel de $\mathscr{M}_2(\mathbf{R})$.
2. I_2 et K forment une famille génératrice (aussi libre, mais sans importance ici) de $\mathscr{F}$, donc $M \in \mathscr{F}^\perp \iff \varphi(M,\, I_2) = \varphi(M,\, K) = 0$.
   On trouve alors, pour $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ le système d'équations
   $$\begin{cases} a + d = 0 \\ b - c = 0 \end{cases}.$$
   Ainsi M est de la forme (nécessaire et suffisant) $\begin{pmatrix} a & b \\ b & -a \end{pmatrix}$.
   Donc $\mathscr{F} = \text{Vect}\,(A,\, B)$ avec $A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ et $B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.
   De plus, la famille $(A,\, B)$ est libre, c'est donc une base de $\mathscr{F}^\perp$.
3. On remarque que $J = I + B$ avec $I \in \mathscr{F}$ et $B \in \mathscr{F}^\perp$ donc par unicité de la décomposition, le projeté orthogonal de J sur $\mathscr{F}^\perp$ est $B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.
4. Si on note $d(J,\, \mathscr{F})$ la distance euclidienne de J à $\mathscr{F}$ alors
   $$d(J,\, \mathscr{F}) = \|J - p_F(J)\| = \|B\| = \sqrt{2}.$$

Exemple

En dimension infinie, un hyperplan n'admet pas toujours de droite supplémentaire orthogonale.
On peut montrer que si H est un hyperplan, alors $H^\perp$ est soit une droite supplémentaire, soit $\{0\}$.
Reprenons l'exemple déjà fructueux vu plus haut avec $E = \mathscr{C}^0\,([-1,\, 1],\, \mathbf{R})$ muni du produit scalaire $(f|g) = \int_{-1}^1 fg$.
Soit $H = \ker\,(f \mapsto f(0))$ qui est bien un hyperplan comme noyau d'une forme linéaire non nulle.
Soit $g \in H^\perp$, alors $\forall h \in H,\ (h|g) = 0$, donc en particulier pour $h = t^2 g$ qui appartient bien à H.
On trouve alors $\int_{-1}^1 t^2 g^2 = 0$, donc g nulle sur $[-1,\, 0[\cup]0,\, 1]$ et par continuité $g = 0$.
Ainsi $H^\perp = \{0\}$, la réciproque est immédiate (ou on dit que $H^\perp$ est un espace vectoriel), donc $H^\perp = \{0\}$.