Espaces préhilbertiens réels PSI *

Produit scalaire et norme associée

Espace préhilbertien réel, espace euclidien

Ne pas oublier de glisser un morceau de l'exo 1138 ccp 21

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

De nombreuses définitions et propriétés sont des rappels du programme de MPSI/PCSI.

Def : Un *produit scalaire* est une application notée $\begin{array}{c} E\times E\to\mathbb{R} \\ (x,y)\mapsto < x,y > \end{array}$ et possédant les propriétés suivantes :

- caractère *bilinéaire* : $\begin{cases} < \alpha x+\beta x', y >= \alpha < x,y > +\beta < x',y > \\ < x, \alpha y+\beta y' >= \alpha < x,xy > +\beta < x,y' > \end{cases}$
- caractère *symétrique* : $< x,y >=< y,x >$.
- caractère *défini positif* : $< x,x >\geqslant 0$, et $< x,x >= 0 \Leftrightarrow x = \overrightarrow{0}$.

Un produit scalaire sur E est donc une "forme bilinéaire symétrique définie positive".

Pour désigner un produit scalaire, on utilise aussi les notations $x\cdot y$ ou $<x,y>$.

Def : Un $\mathbb{R}$-espace vectoriel E muni d'un produit scalaire est dit *préhilbertien réel*.

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

Exemple de référence fondamental

Le produit scalaire *canonique* sur $\mathbb{R}^n$ est défini par $< x,y >= \sum_{i=1}^{n} x_i\, y_i$, où $\begin{cases} x=(x_1,\ldots,x_n) \\ y=(y_1,\ldots,y_n) \end{cases}$

Notation matricielle : si on note X, Y les matrices-colonne associées à x et y,

alors $< x, y >= X^T.Y$.

Preuve : Les propriétés de bilinéarité, symétrie, et positivité sont évidentes.

Pour le caractère "défini" on a $< x, x >= \sum_{i=1}^{n} x_i^2 = 0 \Rightarrow (\forall i \in [\![1, n]\!],\ x_i = 0)$

car les x_i sont des *réels*.

Prop : Le produit scalaire *canonique* sur $\mathcal{M}_{n,p}(\mathbb{R})$ est défini par $< A, B >= tr\,(A^T.B)$.

En d'autres termes, pour $A = (a_{i,j})$ et $B = (b_{i,j})$, on a : $< A, B >= \sum_{i=1}^{n} \sum_{j=1}^{p} a_{ij}\, b_{ij}$.

Preuve : Tableau.

Prop : Soit $\mathcal{C}^0([a, b], \mathbb{R})$ l'e-v des applications continues de $[a, b]$ dans $\mathbb{R}$, avec $a < b$.

En posant $< f, g >= \int_a^b f(t)\, g(t)\, \mathrm{d}t$, on définit un produit scalaire sur $\mathcal{C}^0([a, b], \mathbb{R})$.

Preuve : Tableau.

On peut généraliser cette définition en utilisant une fonction ω dite "fonction poids".

Soit $t \mapsto \omega(t)$ une fonction continue positive sur $[a, b]$

(ne pouvant s'annuler qu'en des points isolés).

Alors $< f, g >= \int_a^b f(t)\, g(t)\, \omega(t)\, \mathrm{d}t$ définit un produit scalaire sur $\mathcal{C}^0([a, b], \mathbb{R})$.

Preuve : facile.

Prop : Soit E un espace préhilbertien réel.

Pour tous vecteurs x, y de E, on a l'inégalité dite "de Cauchy-Schwarz" :

$< x, y >^2 \leqslant < x, x >< y, y >$.

Il y a égalité dans ce résultat si et seulement si x et y sont liés.

Preuve : Facile.

Exo : Soit A, B deux matrices symétriques réelles d'ordre n.

Montrer que $\big(tr(AB)\big)^2 \leqslant tr(A^2).tr(B^2)$.

Qu'obtient-on par exemple si $B = I_n$? Cas d'égalité ?

Sol : CS , égalité si proportionnalité.

Norme associée à un produit scalaire

Def : Soit E un espace préhilbertien réel.

Pour tout vecteur x de E, on appelle *norme de* x la quantité $\|x\| = \sqrt{< x, x >}$.

L'application $x \mapsto \|x\|$, de E dans $\mathbb{R}^+$, est dite *norme associée* au produit scalaire de E.

Avec cette notation, l'inégalité de Cauchy-Schwarz devient :

$\forall (x, y) \in E^2,\ |< x, y >| \leqslant \|x\|\ \|y\|$.

Prop :Soit E un espace préhilbertien réel.

Pour tous x, y de E, et tous réels α, β on a :

$\|\alpha x + \beta y\|^2 = \alpha^2 \|x\|^2 + 2\,\alpha\beta < x, y > +\beta^2 \|y\|^2$.

Preuve : Clair.

En particulier :

$\|x + y\|^2 = \|x\|^2 + 2 < x, y > + \|y\|^2$ et $\|x - y\|^2 = \|x\|^2 - 2 < x, y > + \|y\|^2$.

Par addition, on en déduit : $\forall (x, y) \in E^2,\ \|x + y\|^2 + \|x - y\|^2 = 2\left(\|x\|^2 + \|y\|^2\right)$.

Cette égalité est connue sous le nom d'*identité du parallélogramme*.

Prop : Soit E un espace préhilbertien réel. Identité de polarisation.

Pour tous vecteurs x, y de E, on a :

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

Preuve : simple.

Prop : Soit E un espace préhilbertien réel.

– pour tout vecteur x de E, on a $\|x\| \geqslant 0$, et on a l'équivalence : $\|x\| = 0 \Leftrightarrow x = \overrightarrow{0}$

– pour tout vecteur x de E, et pour tout réel λ, on a : $\|\lambda x\| = |\lambda| \ \|x\|$

– pour tous vecteurs x, y de E, on a l'*inégalité triangulaire* : $\|x+y\| \leqslant \|x\| + \|y\|$

cette inégalité est une égalité si et seulement si x et y sont "positivement liés".

L'expression "positivement liés" signifie $\exists \lambda \in \mathbb{R}^+$ tel que $y = \lambda x$ ou $x = \lambda y$.

Preuve : CS en séparant le cas où l'un est nul .

On regarde le cas $x = \lambda y$.

Rq : CS strict car libre.

Remarque : pour tous x, y de E, on a l'encadrement :

$\big|\ \|x\| - \|y\|\ \big| \leqslant \|x \pm y\| \leqslant \|x\| + \|y\|$.

Def : Soit E un espace préhilbertien réel.

Un vecteur x de E est dit *unitaire* (ou encore *normé*) si $\|x\| = 1$.

Si x est non nul, les vecteurs $\pm\frac{x}{\|x\|}$ sont les seuls vecteurs unitaires de la droite $\mathbb{R}x$.

Retour sur deux exemples de référence

– Dans $\mathbb{R}^n$ avec son produit scalaire canonique, pour tout $x = (x_1, \ldots, x_n)$, on a

$\|x\| = \left(\sum_{i=1}^{n} x_i^2\right)^{1/2}$

L'inégalité de Cauchy-Schwarz s'écrit : $\left(\sum_{i=1}^{n} x_i\, y_i\right)^2 \leqslant \sum_{i=1}^{n} x_i^2 \sum_{i=1}^{n} y_i^2$

– Dans $E = \mathcal{C}^0([a,b],\mathbb{R})$ muni de $< f, g >= \int_a^b f(t)\, g(t)dt$,

pour tout f de E on a $\|f\| = \sqrt{\int_a^b f^2(t)dt}$.

L'inégalité de Cauchy-Schwarz s'écrit alors : $\left(\int_a^b f(t)\, g(t)dt\right)^2 \leqslant \int_a^b f^2(t)dt \int_a^b g^2(t)dt$.

Exo : Soit E un espace préhilbertien réel, et soit a un vecteur unitaire.

Pour quelles valeurs du réel λ l'application

φ: $(x,y) \mapsto < x,y > +\lambda < x,a >< y,a >$ est-elle un produit scalaire ?

Sol : CNS $\lambda > -1$. A clarifier ; Calcul prélable de $\varphi_a(a,a)$ qui entraine $\lambda > -1$.

Puis dans le bon cas , c'est positif par CS .

Puis la nullité donne en premier $< x,a >= 0$ on reporte $x = 0$.

Produit scalaire usuel du plan ou de l'espace

Dans le plan orienté $\mathbb{R}^2$

Dans $\mathbb{R}^2$ muni de son produit scalaire et son orientation habituelle

(pour laquelle la base canonique $e_1 = (1,0), e_2 = (0,1)$ est orthonormée directe).

– Soit u et v deux vecteurs unitaires, et soit $\theta = (\widehat{u,v})\ [2\pi]$. Alors $< u,v >= \cos(\widehat{u,v})$.

Plus généralement, si u et v sont non nuls, on a $< u,v >= \|u\|\,\|v\| \cos(\widehat{u,v})$.

Preuve :

On suppose que u et v sont unitaires. Soit $\theta = (\widehat{u,v})\ [2\pi]$.

Soit $\mathcal{B} = (u, w)$ la base orthonormale directe formée à partir du vecteur u.

On a : $v = \cos(\theta)u + \sin(\theta)w$ donc : $\cos(\theta) =< v, u >$.

Dans le cas général (u, v non nuls), on applique à $u' = \dfrac{u}{\|u\|}$ et $v' = \dfrac{v}{\|v\|}$ (unitaires).

Ainsi : $< u, v >=< \|u\| u', \|v\| v' >= \|u\| \|v\| < u', v' >=$

$\|u\| \|v\| \cos(\widehat{u', v'}) = \|u\| \|v\| \cos(\widehat{u, v})$.

– L'équation d'une droite vectorielle D est $ax + by = 0$,

où $\overrightarrow{n} = (a, b)$ est un "vecteur normal" à D.

Les équations des droites affines $\mathcal{D}$ de direction D sont $ax + by = h$, avec $h \in \mathbb{R}$.

– On peut imposer au vecteur $\overrightarrow{n}$ d'être unitaire, donc poser : $n = (\cos(\theta_0), \sin(\theta_0))$.

L'équation $x\cos(\theta_0) + y\sin(\theta_0) = h$ est alors "l'équation normale" de la dte affine $\mathcal{D}$.

Ici (h, θ_0) est un couple de coordonnées polaires de la projection de O sur $\mathcal{D}$.

Dans l'espace orienté $\mathbb{R}^3$.

Dans $\mathbb{R}^3$ muni de son ps et son orientation habituelle (la base canonique

$e_1 = (1, 0, 0), e_2 = (0, 1, 0), e_3 = (0, 0, 1)$ est orthonormée directe).

– Soit u et v deux vecteurs non nuls.

Si u, v sont liés, on pose $(\widehat{u, v}) = 0$ s'ils sont de même sens,

et $(\widehat{u, v}) = \pi$ s'ils sont de sens opposé.

Si (u, v) sont libres, on mesure $\theta = (\widehat{u, v})$ après avoir orienté le plan $\text{Vect}(u, v)$.

Alors $|< u, v >| = \|u\| \|v\| \cos\theta$.

– L'équation d'un plan vectoriel P est $ax + by + cz = 0$,

où $\overrightarrow{n} = (a, b, c)$ est un "vecteur normal" à P.

Les équations des plans affines $\mathcal{P}$ de direction P s'écrivent alors :

$ax + by + cz = h$, avec h dans $\mathbb{R}$.

– On peut imposer au vecteur $\overrightarrow{n}$ d'être unitaire.

L'équation $ax + by + cz = h$ est alors appelée "équation normale" du plan affine $\mathcal{P}$.

Dans ce cas, h est la mesure algébrique sur l'axe $(O, \overrightarrow{n})$ de la projection de O sur $\mathcal{P}$.

Vecteurs orthogonaux, sous-espaces orthogonaux

Def : Soit E un espace préhilbertien réel.

Deux vecteurs x et y de E sont dits *orthogonaux* si $< x, y >= 0$

Le seul vecteur à être orthogonal à lui-même est le vecteur nul.

A fortiori, le seul vecteur à être orthogonal à tous les vecteurs de E est le vecteur nul.

Exo : Soit E un espace préhilbertien réel. Soit u et v deux vecteurs de E.

On suppose que, pour tout réel λ, on a :

$\|u + \lambda v\| \geqslant \|u\|$. Montrer que u et v sont orthogonaux.

Sol : Par hypothèse, on a $\|u + \lambda v\|^2 \geqslant \|u\|^2$ pour tout λ c'est-à-dire :

$\forall \lambda \in \mathbb{R},\ 2\lambda < u, v > +\lambda^2 \|v\|^2 \geq 0.$

Nécessairement $< u, v >= 0$, sans quoi $\varphi(\lambda) = 2\lambda < u, v > +\lambda^2 \|v\|^2$

changerait de signe au voisinage de 0.

Exo : Soit E préhilbertien réel, et $f : E \to E$ avec : $\forall (x, y),\ < f(x), y >=< x, f(y) >$.

Montrer que f est linéaire.

Sol : On se donne trois vecteurs quelconque x, y, z, et deux scalaires quelconques α et β.

En utilisant plusieurs fois l'hypothèse sur f, on trouve :

$< f(\alpha x+\beta y),z >=< \alpha x+\beta y,f(z) >=$

$\alpha < x, f(z) > +\beta < y,f(z) >= \alpha < f(x),z > +\beta < f(y),z >=< \alpha f(x)+\beta f(y),z >$

Ainsi le vecteur $f(\alpha x + \beta y) - \big(\alpha f(x) + \beta f(y)\big)$

est orthogonal à tout z de E, donc il est nul.

En conclusion : $\forall (x, y) \in E^2,\ \forall (\alpha, \beta) \in \mathbb{R}^2,\ f(\alpha x + \beta y) = \alpha f(x) + \beta f(y)$.

L'application f est linéaire.

Def : Soit F et G deux sous-espaces vectoriels d'un espace préhilbertien réel E.

On dit que F et G sont orthogonaux si :

$\forall x \in F,\ \forall y \in G,\ (x \mid y) = 0.$

Remarque : si F et G sont orthogonaux, ils sont en somme directe.

En effet si x est dans $F \cap G$ alors $(x \mid x) = 0$ donc $x = \overrightarrow{0}$.

Prop : Soit F un sous-espace vectoriel d'un espace préhilbertien réel E.

On note $F^\perp = \{y \in E,\ \forall x \in F,\ (x \mid y) = 0\}$.

Ainsi, $F^\perp$ est l'ensemble des vecteurs de E qui sont orthogonaux à tous ceux de F.

L'ensemble $F^\perp$ est un sous-espace vectoriel de E. On l'appelle l'*orthogonal* de F.

Preuve : L'ensemble $F^\perp$ est non vide car il contient $\overrightarrow{0}$.

Ensuite, pour tous y, z de $F^\perp$, et pour tous réels λ, μ,

le vecteur $w = \lambda y + \mu z$ est dans $F^\perp$ car : $\forall x \in F,\ (w \mid x) = \lambda\,(y \mid x) + \mu\,(z \mid x) = 0.$

Autre idée : $F^\perp$ est l'intersection des noyaux des formes linéaires

$\varphi_x : y \mapsto (y \mid x)$ quand x parcourt F.

Rq1 : $F^\perp$ est donc un fermé .

Car c'est l'intersection (infinie) de fermés.

Les noyaux des φ_x sont fermés comme images réciproques de 0 par des applications continues (lip) par CS.

Remarques

– De manière évidente, on a $\{\overrightarrow{0}\}^\perp = E$, et $E^\perp = \{\overrightarrow{0}\}$.

– Dire que les sous-espaces F et G sont orthogonaux,

c'est dire que $F \subset G^\perp$, ou encore $G \subset F^\perp$.

L'espace $F^\perp$ est, pour l'inclusion, le plus grand sous-espace de E orthogonal à F.

– Si on a l'inclusion $F \subset G$, alors on a l'inclusion $G^\perp \subset F^\perp$.

Si $F \subset G$, et si x est dans $G^\perp$ (c'est-à-dire s'il est orthogonal à tous les vecteurs de G), alors il est a fortiori orthogonal à tous les vecteurs de F, donc x est dans $F^\perp$.

– Pour tout sous-espace de E, on a l'inclusion $F \subset (F^\perp)^\perp$

(F est inclus dans son *double orthogonal*).

On verra plus loin que si F est de dimension finie, il s'agit d'une égalité.

Soit x dans F. Pour tout y de $F^\perp$, on a $(x \mid y) = 0$.

Autrement dit, x est dans l'orthogonal de $F^\perp$.

– Si $F = \text{Vect}\{x_i,\ i \in I\}$, alors : $\ y \in F^\perp \Leftrightarrow \forall i \in I,\ (y \mid x_i) = 0$.

Si y est dans $F^\perp$, il est en particulier orthogonal aux x_i.

Réciproquement, si y est orthogonal aux x_i, il est orthogonal à leurs

combinaisons linéaires, c'est-à-dire à tous les vecteurs de F.

Exo : On munit $\mathbb{R}[X]$ du produit scalaire $(A \mid B) = \displaystyle\int_{-1}^{1} A(t)B(t)\,\mathrm{d}t$.

On définit $\varphi\colon A \mapsto \int_{-1}^{1} |t|\, A(t)\, \mathrm{d}t$, H l'hyperplan noyau de cette forme linéaire non nulle.

Pour tout A de $\mathbb{R}[X]$, montrer que $A - \varphi(A)$ est dans H. A-t'on l'égalité $\mathbb{R}[X] = H \oplus H^{\perp}$?

Sol :

– φ est bien une forme linéaire, non nulle car l'image de 1 est 1.

Son noyau est donc bien un hyperplan.

– Donc pour $Q \in H^{\perp}$, $\forall A \in \mathbb{R}[X]$, $\int_{-1}^{1} Q(t)((\varphi(A) - A)(t))dt = 0$.

Donc $\int_{-1}^{1} Q(t)A(t)dt = \varphi(A) \int_{-1}^{1} Q(t)dt$.

Donc $\int_{-1}^{1} \left(Q(t) - |t| \int_{-1}^{1} Q(t)dt \right) A(t)dt = 0$. En développant, rangeant...

Mais ceci est vérifié pour tous les polynômes A, donc

$t \mapsto \left(Q(t) - |t| \int_{-1}^{1} Q(t)dt \right) \in E^{\perp}$.

Donc cette fonction est identiquement nulle, $\left(\forall t \in [-1, 1], Q(t) = |t| \int_{-1}^{1} Q(t)dt \right)$.

Or un polynôme est dérivable, seul cas possible, $\int_{-1}^{1} Q(t)dt = 0$, donc $Q = \overrightarrow{0}$.

Ainsi $H^{\top} = \{\overrightarrow{0}\}$, $H \oplus H^{\top} = H \subsetneq E$.

On n'a donc pas l'égalité $\mathbb{R}[X] = H \oplus H^{\top}$ (la faute à la dimension infinie de H).

Familles orthogonales

Def : Soit E un espace préhilbertien réel.

On dit que $(x_i)_{i \in I}$ de vecteurs de E est *orthogonale* si les x_i sont $\perp$ 2 à 2.

Si de plus ils sont unitaires, alors la famille est dite *orthonormale* (ou orthonormée).

La famille $(x_i)_{i \in I}$ est orthonormale $\Leftrightarrow \forall (i, j) \in I^2,\ (x_i \mid x_j) = \delta_{ij}$ (Kronecker).

Exemples de référence

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

– La base canonique de $\mathcal{M}_n(\mathbb{R})$ est orthonormale pour le produit scalaire canonique.

– On se place dans $\mathcal{C}^0([0,2\pi],\mathbb{R})$ muni du produit scalaire $(f \mid g) = \displaystyle\int_0^{2\pi} f(t)\,g(t)\,\mathrm{d}t$.

Les $f_n : t \mapsto \cos(nt)$, avec $n \in \mathbb{N}$, est une famille orthogonale pour ce produit scalaire.

Prop : Soit E un espace préhilbertien réel.

Si une famille $(x_i)_{i\in I}$ est orthogonale de vecteurs tous non nuls, alors elle est libre.

Preuve : On considère l'égalité $\displaystyle\sum_{i\in I} \lambda_i x_i = \overrightarrow{0}$, et on la "multiplie" par un vecteur x_j.

Donc $0 = \Big(\displaystyle\sum_{i\in I} \lambda_i x_i \mid x_j\Big) = \sum_{i\in I} \lambda_i \underbrace{(x_i \mid x_j)}_{=0 \text{ si } i\neq j} = \lambda_j \|x_j\|^2$, $\lambda_j = 0$ ($\forall j \in I$) : (x_j) est libre.

Cas particulier évident : une famille orthonormale $(x_i)_{i\in I}$ est libre.

Si $\dim E = n \geqslant 1$, une famille orthonormale de n vecteurs est une base orthonormale.

Prop :Soit E un espace préhilbertien réel.
Si la famille $(x_i)_{1\leqslant i\leqslant p}$ est orthogonale, alors $\Big\|\displaystyle\sum_{i=1}^{p} x_i\Big\|^2 = \sum_{i=1}^{p} \|x_i\|^2$ *(relation de Pythagore).*

Preuve : On développe par bilinéarité :

$$\Big\|\sum_{i=1}^{p} x_i\Big\|^2 = \Big(\sum_{i=1}^{p} x_i \mid \sum_{j=1}^{p} x_j\Big) = \sum_{i=1}^{n} \|x_i\|^2 + \sum_{i\neq j} \underbrace{(x_i \mid x_j)}_{=0 \text{ car } i\neq j} = \sum_{i=1}^{n} \|x_i\|^2.$$

Attention, la réciproque n'est vraie que si $p = 2$.

Ainsi : $\|x_1 + x_2\|^2 = \|x_1\|^2 + \|x_2\|^2 \Leftrightarrow (x_1 \mid x_2) = 0$.

Exo : On munit $\mathbb{R}[X]$ du produit scalaire $(A \mid B) = \displaystyle\int_0^1 A(t)B(t)\,\mathrm{d}t$.
Pour tout n de $\mathbb{N}$, on pose $P_n = U_n^{(n)}$, avec $U_n(X) = (X^2 - X)^n$.

(1) Montrer que $(P_n)_{n\geqslant 0}$ est orthogonale (utiliser des IPP répétées).

(2) Calculer $J_{n,m} = \displaystyle\int_0^1 t^n(1-t)^m \,\mathrm{d}t$, avec $(n,m)^2$ dans $\mathbb{N}^2$. Puis $\|P_n\|$ pour tout n.

(3) Former une bon de $\mathbb{R}_4[X]$ (faire le rapprochement avec l'exercice qui vient).

Sol :

(1) On remarque que P_n est un polynôme de degré n, de coefficient dominant $\dfrac{(2n)!}{n!}$.

Les réels 0 et 1 sont racines de U_n de multiplicité n.

Il en résulte que, pour tout k de $\{0,\dots,n-1\}$, on a $U_n^{(k)}(0) = U_n^{(k)}(1) = 0$.

Après n intégrations par parties :

$$\forall Q \in \mathbb{R}[X],\ (P_n \mid Q) = \int_0^1 U_n^{(n)}(t)Q(t)\,\mathrm{d}t = (-1)^n \int_0^1 U_n(t)Q^{(n)}(t)\,\mathrm{d}t.$$

En particulier : $\forall n \in \mathbb{N},\ \forall Q \in \mathbb{R}_{n-1}[X],\ (P_n \mid Q) = 0$.

Les P_n, qui sont à degrés échelonnés, sont donc orthogonaux deux à deux.

(2) Posons $J_{n,m} = \displaystyle\int_0^1 t^n(1-t)^m \,\mathrm{d}t$, avec n et m dans $\mathbb{N}$.

$$\text{Si } m \geqslant 1 : J_{n,m} = \frac{1}{n+1}\Big[t^{n+1}(1-t)^m\Big]_0^1 + \frac{m}{n+1}\int_0^1 t^{n+1}(t-1)^{m-1}\,\mathrm{d}t = \frac{m}{n+1}J_{n+1,m-1}$$

$$\text{On obtient : } J_{n,m} = \frac{m}{n+1}\,\frac{m-1}{n+2}\cdots\frac{1}{n+m}J_{n+m,0} = \frac{m!\,n!}{(m+n)!}J_{n+m,0} = \frac{m!\,n!}{(m+n+1)!}$$

On réutilise l'intégration par parties mise en place dans la question précédente.

$$\|P_n\|^2 = (P_n \mid P_n) = (-1)^n\left(U_n \mid P_n^{(n)}\right) = (-1)^n(2n)!\int_0^1 U_n(t)\,\mathrm{d}t = (2n)!J_{n,n}.$$

Ainsi : $\|P_n\|^2 = (2n)!J_{n,n} = \dfrac{(n!)^2}{2n+1}$, donc $\|P_n\| = \dfrac{n!}{\sqrt{2n+1}}$.

(3) Une base orthonormale de $\mathbb{R}[X]$ est $L_n = \dfrac{P_n}{\|P_n\|} = \dfrac{\sqrt{2n+1}}{n!}\big((X^2-X)^n\big)^{(n)}$.

On trouve ainsi : $L_0(X) = 1$, $L_1(X) = \sqrt{3}(X^2-X)' = \sqrt{3}(2X-1)$.

Ensuite $L_2(X) = \dfrac{\sqrt{5}}{2}(X^4 - 2X^3 + X^2)'' = \sqrt{5}(6X^2 - 6X + 1)$.

Puis $L_3(X) = \dfrac{\sqrt{7}}{6}(X^6 - 3X^5 + 3X^4 - 1)''' = \sqrt{7}(20X^3 - 30X^2 + 12X - 1)$.

Finalement : $L_4(X) = 210X^4 - 420X^3 + 270X^2 - 60X + 3$.

Bases orthonormales d'un espace euclidien

Algorithme d'orthonormalisation de Gram-Schmidt

Le code Python Q-R.py annexe 32 ?

Def : Soit E un espace préhilbertien réel.

Soit $(e_i)_{1\leqslant i\leqslant p}$ une famille libre de E.

Pour tout i de $\{1,\dots,p\}$, on pose $\varepsilon_i = \dfrac{1}{\|u_i\|}\,u_i$, où $u_i = e_i - \displaystyle\sum_{j=1}^{i-1}(\varepsilon_j \mid e_i)\,\varepsilon_j$

La première étape de l'algorithme consiste bien sûr à poser $\varepsilon_1 = \dfrac{1}{\|e_1\|}\,e_1$

Prop : Avec les notations précédentes, l'algorithme de Gram-Schmidt se termine.

Pour tout i de $\{1,\dots,p\}$, on a $\mathrm{Vect}\big(\{\varepsilon_1,\dots,\varepsilon_i\}\big) = \mathrm{Vect}\big(\{e_1,\dots,e_i\}\big)$, et $(\varepsilon_i \mid e_i) > 0$.

La famille $(\varepsilon_i)_{1\leqslant i\leqslant p}$ ainsi construite est une base orthonormale de $\mathrm{Vect}\big(\{e_1,\dots,e_p\}\big)$.

Preuve : Pour simplifier , on pose $F_i = \mathrm{Vect}\big(\{e_1,\dots,e_i\}\big)$ pour tout i de $[\![1,p]\!]$.

Par récurrence finie sur i dans $[\![1,p]\!]$, on va montrer la propriété $\mathcal{P}(i)$ suivante :

$\mathcal{P}(i)$: « la famille $\varepsilon_1,\dots,\varepsilon_i$ est orthonormale, et $\mathrm{Vect}\big(\{\varepsilon_1,\dots,\varepsilon_i\}\big) = F_i$, et $(\varepsilon_i \mid e_i) > 0$ ».

On posant $\varepsilon_1 = \dfrac{1}{\|e_1\|}\,e_1$, on s'assure que la propriété $\mathcal{P}(1)$ est vraie.

On se donne i dans $[\![2,p]\!]$, et on suppose que $\mathcal{P}(i-1)$ est vraie.

Selon la définition , on pose $u_i = e_i - v_i$, avec $v_i = \displaystyle\sum_{j=1}^{i-1}(\varepsilon_j \mid e_i)\,\varepsilon_j$.

Le vecteur v_i est dans $\text{Vect}\big(\{\varepsilon_1, \dots, \varepsilon_{i-1}\}\big) = F_{i-1}$ et e_i n'y est pas, donc $u_i \neq \overrightarrow{0}$.

On peut donc poser $\varepsilon_i = \dfrac{1}{\|u_i\|}\, u_i$, qui est unitaire.

$\forall k \in [\![1, i[\![,\ (\varepsilon_k \mid v_i) = \displaystyle\sum_{j=1}^{i-1} (\varepsilon_j \mid e_i)\, \underbrace{(\varepsilon_k \mid \varepsilon_j)}_{\delta_{j,k}} = (\varepsilon_k \mid e_i),\ (\varepsilon_k \mid u_i) = 0,\ (\varepsilon_k \mid \varepsilon_i) = 0.$

Ainsi la famille $\varepsilon_1, \dots, \varepsilon_i$ est orthonormale.

Par construction, u_i donc ε_i est dans F_i. Puisque $\text{Vect}\big(\{\varepsilon_1, \dots, \varepsilon_{i-1}\}\big) = F_{i-1}$ (hypothèse de récurrence), on en déduit $\text{Vect}\big(\{\varepsilon_1, \dots, \varepsilon_i\}\big) \subset F_i$ (donc égalité car les dimensions sont les mêmes).

Enfin $(\varepsilon_i \mid e_i) = (\varepsilon_i \mid u_i + v_i) = (\varepsilon_i \mid u_i) + \underbrace{(\varepsilon_i \mid v_i)}_{=0} = (\varepsilon_i \mid \|u_i\|\, \varepsilon_i) = \|u_i\| > 0.$

On a ainsi prouvé la propriété au rang i, ce qui achève la récurrence finie.

L'algorithme s'arrête donc par la formation du vecteur ε_p, et finalement on a formé une famille $(\varepsilon_i)_{1 \leqslant i \leqslant p}$ qui constitue une base orthonormale de $\text{Vect}\big(\{e_1, \dots, e_p\}\big)$.

Illustration du procédé

Passage d'une famille libre $e = (e_1, e_2, e_3)$ à une famille orthonormale $\varepsilon = (\varepsilon_1, \varepsilon_2, \varepsilon_3)$.

On a conservé les notations de la proposition en ce qui concerne les vecteurs u_2 et u_3.

On a cependant noté $p(e_2) = (\varepsilon_1 \mid e_2)\, \varepsilon_1$, donc $u_2 = e_2 - p(e_2)$.

De même, on a noté $q(e_3) = (\varepsilon_1 \mid e_3)\, \varepsilon_1 + (\varepsilon_2 \mid e_3)\, \varepsilon_2$, donc $u_3 = e_3 - q(e_3)$.

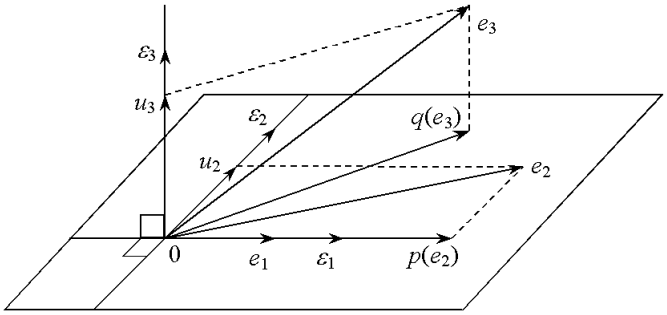
On voit bien, ce qui sera repris plus tard, que :

– le vecteur $p(e_2)$ est la « projection orthogonale » de e_2 sur la droite engendrée par ε_1

(donc par e_1)

– le vecteur $q(e_3)$ est la projection orthogonale de e_3 sur le plan engendré par $\varepsilon_1, \varepsilon_2$

(donc par e_1, e_2)



Remarque

Si $(e_i)_{1\leqslant i\leqslant p}$ est orthonormale, elle est invariante par Gram-Schmidt.

Par récurrence. Avec les notations de la définition , on a bien sûr $\varepsilon_1 = \mathrm{e}_1$.

Enfin si $\varepsilon_j = e_j$ pour $1 \leqslant j < i$, alors :

$$u_i = e_i - \sum_{j=1}^{i-1} (\varepsilon_j \mid e_i)\, \varepsilon_j = e_i - \sum_{j=1}^{i-1} \underbrace{(e_j \mid e_i)}_{=0} e_j = e_i \text{ donc } \varepsilon_i = e_i.$$

Exo : On munit $\mathbb{R}[X]$ du produit scalaire $(A \mid B) = \displaystyle\int_0^1 A(t)B(t)\,\mathrm{d}t$.

Former une bon de $\mathbb{R}_4[X]$ en appliquant le procédé de GS à $1, X, X^2, X^3, X^4$.

Faire le rapprochement avec l'exercice d'avant.

Sol : On note $e_k = X^k$, pour $0 \leqslant k \leqslant 4$, et on va former la base $(\varepsilon_k)_{0\leqslant k\leqslant 4}$.

Les résultats suivants ont été obtenus à l'aide d'un logiciel de calcul symbolique.

– Le polynôme e_0 est unitaire, donc $\varepsilon_0 = e_0 = 1$.

– On trouve $u_1 = e_1 - (\varepsilon_0 \mid e_1)\,\varepsilon_0 = X - (1 \mid X)\,1 = X - \dfrac{1}{2}$, $\|u_1\| = \dfrac{\sqrt{3}}{6}$,

donc $\varepsilon_1 = \dfrac{u_1}{\|u_1\|} = \sqrt{3}(2X - 1)$.

– On trouve $u_2 = e_2 - (\varepsilon_0 \mid e_2)\,\varepsilon_0 - (\varepsilon_1 \mid e_2)\,\varepsilon_1 = X^2 - X + \dfrac{1}{6}$, et $\|u_2\| = \dfrac{\sqrt{5}}{30}$,

donc $\varepsilon_2 = \sqrt{5}(6X^2 - 6X + 1)$.

– On trouve $u_3 = e_3 - (\varepsilon_0 \mid e_3)\,\varepsilon_0 - (\varepsilon_1 \mid e_3)\,\varepsilon_1 - (\varepsilon_2 \mid e_3)\,\varepsilon_2 = X^3 - \frac{3}{2}X^2 + \frac{3}{5}X - \frac{1}{20}$.

Ensuite $\|u_3\| = \frac{\sqrt{7}}{140}$, donc $\varepsilon_3 = \sqrt{7}(20X^3 - 30X^2 + 12X - 1)$.

– On trouve : $u_4 = e_4 - \sum_{j=0}^{3} (\varepsilon_j \mid e_4)\,\varepsilon_j = X^4 - 2X^3 + \frac{9}{7}X^2 - \frac{2}{7}X + \frac{1}{70}$.

Enfin $\|u_4\| = \frac{1}{210}$ donc $\varepsilon_4 = 210X^4 - 420X^3 + 270X^2 - 60X + 3$.

Bases orthonormales d'un espace euclidien

Prop : Soit E un espace euclidien (ie un espace préhilbertien réel de dimension finie).

Alors, dans l'espace vectoriel E, il existe des bases orthonormales.

Preuve : Appliquer l'algorithme de Gram-Schmidt à une base quelconque $\mathcal{B}$ de E.

Prop :Soit E un espace euclidien de dimension $n \geqslant 1$.

Soit $(e) = (e_1, e_2, \ldots, e_p)$ une famille orthonormale de E, avec $p < n$ (donc non génératrice).

Alors il est possible de compléter (e) en une base orthonormale de E.

Preuve : Avec le théorème de la base incomplète,

on prolonge (e) en une base $\mathcal{B} = (e_1, \ldots, e_p, e_{p+1}, \ldots, e_n)$ de E.

On applique l'algorithme d'orthonormalisation de Gram-Schmidt à la base $\mathcal{B}$.

On en déduit une base orthonormale $\mathcal{B}' = (e_1, \ldots, e_p, \varepsilon_{p+1}, \ldots, \varepsilon_n)$ de E.

On a bien sûr utilisé l'invariance de la famille orthonormale $(e_1, \ldots, e_p)$ par l'algorithme.

Prop : Soit E un espace euclidien de dimension $n \geqslant 1$,

et soit $\mathcal{B} = (e_i)_{1 \leqslant i \leqslant n}$ une base orthonormale de E.

Pour tout vecteur x de E, on a l'égalité : $x = \sum_{i=1}^{n} (e_i \mid x)\,e_i$.

Preuve : Posons $x = \sum_{i=1}^{n} x_i\, e_i$. Pour tout j de $[\![1, n]\!]$, on a : $(e_j \mid x) = \sum_{i=1}^{n} x_i \underbrace{(e_j \mid e_i)}_{=\delta_{i,j}} = x_j$.

Bref, les *applications coordonnées* dans la base $\mathcal{B}$ sont les $e_i^* : x \mapsto (e_i \mid x)$.

Tout vecteur x de E est entièrement déterminé par ses "ps" sur les vecteurs de $\mathcal{B}$.

Prop : Soit E un espace euclidien de dimension $n \geqslant 1$,

et soit $\mathcal{B} = (e_i)_{1 \leqslant i \leqslant n}$ une base orthonormale de E.

Pour tous vecteurs $x = \sum_{i=1}^{n} x_i\, e_i$ et $y = \sum_{i=1}^{n} y_i\, e_i$, on a :

$(x \mid y) = \sum_{i=1}^{n} x_i\, y_i$ et $\|x\|^2 = \sum_{i=1}^{n} x_i^2$.

Preuve : Posons $x = \sum_{i=1}^{n} x_i\, e_i$ et $y = \sum_{j=1}^{n} y_j\, e_j$.

Par bilinéarité : $(x \mid y) = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i y_j \underbrace{(e_i \mid e_j)}_{\delta_{i,j}} = \sum_{i=1}^{n} x_i y_i$.

Exo : Soit E un espace vectoriel euclidien.

Soit $f \in \mathcal{L}(E)$ tel que :

$\forall (u, v) \in E^2,\ (u \mid v) = 0 \Rightarrow (f(u) \mid f(v)) = 0$ (f « conserve l'orthogonalité »).

Montrer qu'il existe un réel $\lambda \geqslant 0$ tel que, pour tout vecteur u de E, $\|f(u)\| = \lambda \|u\|$.

Sol : Soit $(e_k)_{1 \leqslant k \leqslant n}$ une bon. Par hypothèse, les $f(e_k)$ sont orthogonaux deux à deux.

Nous allons montrer qu'ils ont même norme.

Pour cela, on se donne i et j distincts dans $[\![1, n]\!]$.

Les vecteurs $e_i + e_j$ et $e_i - e_j$ sont orthogonaux car $(e_i + e_j \mid e_i - e_j) = \|e_i\|^2 - \|e_j\|^2 = 0$.

On en déduit :

$0 = (f(e_i + e_j) \mid f(e_i - e_j)) = (f(e_i) + f(e_j) \mid f(e_i) - f(e_j)) = \|f(e_i)\|^2 - \|f(e_j)\|^2$.

Autrement dit, il existe $\lambda \geqslant 0$ tel que : $\forall i \in [\![1, n]\!],\ \|f(e_i)\| = \lambda$.

Pour $u = \sum_{i=1}^{n} x_i e_i$, on a

$f(u) = \sum_{i=1}^{n} x_i f(e_i)$ et $\|f(u)\|^2 = \sum_{i=1}^{n} x_i^2 \|f(e_i)\|^2 = \lambda^2 \sum_{i=1}^{n} x_i^2 = \lambda^2 \|u\|^2$ (Pythagore).

Expression matricielle du produit scalaire

Soit x et y deux vecteurs quelconques d'un espace euclidien E.

Soit X et Y les matrices-colonne des coordonnées de x, y dans une base $\mathcal{B}$ de E.

– si la base $\mathcal{B}$ est orthonormale, alors : $(x \mid y) = X^T.Y$;

– dans le cas général d'une base $\mathcal{B}$ *quelconque*, alors on a seulement $(x \mid y) = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i\, y_j\, (e_i \mid e_j)$;

on peut écrire $(x \mid y) = X^T G\, Y$, où $g_{i,j} = (e_i \mid e_j)$ pour $1 \leqslant i, j \leqslant n$.

Posons $x = \sum_{i=1}^{n} x_i\, e_i$ et $y = \sum_{j=1}^{n} y_j\, e_j$.

On a $(x \mid y) = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i\, y_j\, (e_i \mid e_j) = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i\, y_j\, g_{i,j} =$

$\sum_{i=1}^{n} x_i \Big(\sum_{j=1}^{n} g_{i,j} y_j \Big) = \sum_{i=1}^{n} x_i [GY]_i = X^T G\, Y$

Si la famille $(e_i)_{1 \leqslant i \leqslant n}$ est orthonormale, on a $G = I_n$ donc $(x \mid y) = X^T.Y$.

Exo : Soit A dans $\mathcal{M}_n(\mathbb{R})$.

Montrer que les matrices $A^T.A$ et A ont le même noyau, donc le même rang.

Les matrices $A.A^T$ et A ont-elles le même rang ? Le même noyau ?

Sol :

– Si X est dans le noyau de A (donc si $AX = \overrightarrow{0}$), alors $A^T AX = \overrightarrow{0}$,

donc X est dans le noyau de $A^T A$.

Réciproquement, soit X dans $\mathrm{Ker}(A^T A)$.

On a $A^TAX = \overrightarrow{0}$ donc $X^TA^TAX = 0$, donc $(AX)^TAX = 0$.

Mais cela s'écrit $\|AX\|^2 = 0$, donc conduit à $AX = \overrightarrow{0}$, c'est-à-dire $X \in \mathrm{Ker}(A)$.

Les deux matrices A^TA et A ont donc le même noyau, donc le même rang.

– La matrice $A.A^T$ a même noyau que A^T, même rang que A^T, même rang que A.

Mais en général, $\mathrm{Ker}(A.A^T)$ (c'est-à-dire $\mathrm{Ker}(A^T)$) est différent de $\mathrm{Ker}(A)$.

Exo : Soit M dans $\mathcal{M}_n(\mathbb{R})$, telle que $M^2 = 0$.

Montrer que $\mathrm{Ker}(M + M^T) = \mathrm{Ker}(M) \cap \mathrm{Ker}(M^T)$.

Sol : Soit X un vecteur du noyau de $\ker(M + M^T)$, donc tel que $MX = -M^TX$.

$\|MX\|^2 = (MX)^TMX = X^TM^TMX = -X^TM^TM^TX = -X^T(M^2)^TX = 0$,

donc $MX = 0$.

Il en découle $M^TX = 0$, donc $\mathrm{Ker}(M + M^T) \subset \mathrm{Ker}(M) \cap \mathrm{Ker}(M^T)$.

L'inclusion inverse est évidente.

Exo : Soit M dans $\mathcal{M}_n(\mathbb{R})$, telle que $MM^T = M^TM$.

(1) Montrer que M et M^T ont le même noyau, puis que $\mathrm{Sp}_{\mathbb{R}}(M) = \mathrm{Sp}_{\mathbb{R}}(M^T)$.

(2) Montrer que pour tout λ de $\mathrm{Sp}_{\mathbb{R}}(M)$,

les sous-espaces propres de M et de M^T pour λ sont identiques.

(3) Montrer que les sous-espaces propres de M dans $\mathcal{M}_{n,1}(\mathbb{R})$

sont orthogonaux deux à deux.

Sol :

(1) Pour tout vecteur-colonne X de $\mathcal{M}_{n,1}(\mathbb{R})$:

$$\|MX\|^2 = (MX \mid MX) = (MX)^TMX = X^TM^TMX = X^TMM^TX = \left\|M^TX\right\|^2.$$

Ainsi $\|MX\| = \left\|M^T X\right\|$, donc :

$X \in \mathrm{Ker}(M) \Leftrightarrow \|MX\| = 0 \Leftrightarrow \left\|M^T X\right\| = 0 \Leftrightarrow X \in \mathrm{Ker}\, M^T$.

Les deux matrices M et M^T ont donc le même noyau.

Pour λ dans $\mathbb{R}$, on applique ce qui précède à

$A = M - \lambda I_n$ et $B = A^T = M^T - \lambda I_n$ (elles commutent).

Pour tout réel λ, on a donc : $\mathrm{Ker}(M - \lambda I_n) = \mathrm{Ker}(M^T - \lambda I_n)$. Autrement dit, M et M^T ont les mêmes valeurs propres réelles, et les sous-espaces propres correspondants sont identiques.

(2) Soient λ et μ deux valeurs propres réelles distinctes de M (donc de M^T).

Soient X un vecteur propre de M (donc de M^T) pour la valeur propre λ et Y un vecteur propre de M (donc de M^T) pour la valeur propre μ.

Il faut montrer que X et Y sont orthogonaux.

Or $\begin{cases} (MX \mid Y) = (MX)^T Y = (\lambda X)^T Y = \lambda X^T Y \\ (MX \mid Y) = (MX)^T Y = X^T \left(M^T Y\right) = \mu X^T Y \end{cases}$

On a ainsi $\lambda\,(X \mid Y) = \mu\,(X \mid Y)$, avec $\lambda \neq \mu$.

Il en résulte l'orthogonalité des vecteurs X et Y.

On en déduit l'orthogonalité des sous-espaces propres E_λ et E_μ de M.

Matrice d'une famille de vecteurs dans une base orthonormale

$\mathcal{B} = (e_i)_{1 \leqslant i \leqslant n}$ une bon de E, et $v = (v_j)_{1 \leqslant j \leqslant p}$ une famille de p vecteurs de E.

Soit $A = (a_{i,j})$ la matrice de la famille v dans la base $\mathcal{B}$ (la matrice $A \in \mathcal{M}_{n,p}(\mathbb{K})$).

Alors, pour tout i de $[\![1, n]\!]$ et tout j de $[\![1, p]\!]$, on a : $a_{i,j} = (e_i \mid v_j)$.

Évident car $a_{i,j}$ est la coordonnée de v_j dans la base orthonormale $\mathcal{B} = (e_i)_{1 \leqslant i \leqslant n}$.

Cas particulier : si $\mathcal{B}'$ est la base orthonormale obtenue à partir d'une base quelconque $\mathcal{B}$ de E par l'algorithme de Gram-Schmidt, alors les matrices de passage $P_{\mathcal{B},\mathcal{B}'}$ (de $\mathcal{B}$ à $\mathcal{B}'$)

et $P_{\mathcal{B}',\mathcal{B}}$ (de $\mathcal{B}'$ à $\mathcal{B}$), inverses l'une de l'autre, sont triangulaires supérieures à coefficients diagonaux strictement positifs.

Reprenons les notations de la définition G-S, et notons $\mathcal{B}' = (\varepsilon_i)_{1\leqslant i\leqslant n}$

la base orthonormale obtenue à partir de la base quelconque $\mathcal{B} = (e_i)_{1\leqslant i\leqslant n}$.

Par construction $\varepsilon_i = \dfrac{1}{\|u_i\|}\, u_i$, où $u_i = e_i - \displaystyle\sum_{j=1}^{i-1} (\varepsilon_j \mid e_i)\, \varepsilon_j$.

En particulier, u_i (donc ε_i) est dans $\mathrm{Vect}\big(\{e_1, \ldots, e_i\}\big)$.

Il vient que la matrice de passage $P_{\mathcal{B},\mathcal{B}'}$ de $\mathcal{B}$ à $\mathcal{B}'$ est triangulaire supérieure.

Pareil pour la matrice de passage inverse $P_{\mathcal{B}',\mathcal{B}}$ (de la base $\mathcal{B}'$ à la base $\mathcal{B}$).

Mais, du fait que la base $\mathcal{B}'$ est orthonormale, les coefficients diagonaux de $P_{\mathcal{B}',\mathcal{B}}$ sont les produits scalaires $(e_j \mid \varepsilon_j)$ dont on sait qu'ils sont strictement positifs (cf propositionG-S).

Les coefficients diagonaux de la matrice triangulaire $P_{\mathcal{B},\mathcal{B}'} = P_{\mathcal{B}',\mathcal{B}}^{-1}$

sont donc aussi strictement positifs.

Matrice d'un endomorphisme dans une base orthonormale

Soit $\mathcal{B} = (e_i)_{1\leqslant i\leqslant n}$ une base orthonormale de E, et soit u un endomorphisme de E.

Soit $A = (a_{i,j})$ la matrice de u dans la base $\mathcal{B}$. Alors $a_{i,j} = (e_i \mid u(e_j))$ pour tous i, j.

Évident car $a_{i,j}$ est la coordonnée de $u(e_j)$ sur e_i dans la base orthonormale $\mathcal{B}$.

Projection orthogonale sur un sous-espace de dim finie

Supplémentaire orthogonal

Prop : Soit E un espace préhilbertien réel.

Soit F un sous-espace vectoriel de dimension finie de E.

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

Le sous-espace $F^\perp$ est appelé le *supplémentaire orthogonal* de F.

Le cadre du résultat précédent est : F sous-espace de dimension finie de E (sans hypothèse sur E lui-même). Évidemment, il n'y a pas d'hypothèse à faire sur F si E est lui-même de dimension finie.

Preuve : Posons $\dim(F) = p$. Soit $\mathcal{B} = (e_i)_{1 \leqslant i \leqslant p}$ une base orthonormale de F.

Ainsi $F = \text{Vect}\{e_1, \ldots, \text{e}_p\}$, donc : $z \in F^\perp \Leftrightarrow \forall j \in [\![1, p]\!],\ (z \mid e_j) = 0$

(cf dernière remarque dans prop formes linéaires avant...).

Soit x dans E. Il s'agit de prouver $\exists!(y, z)$ de $F \times F^\perp$ tel que $x = y + z$.

Pour cela, on pose $z = x - y$, où $y = \sum_{i=1}^{p} y_i e_i$ est un vecteur quelconque de F.

On a : $z \in F^\perp \Leftrightarrow \forall j \in [\![1, p]\!],\ (z \mid e_j) = 0 \Leftrightarrow \forall j \in [\![1, p]\!],\ (x \mid e_j) = (y \mid e_j) = y_j$.

Ceci détermine un unique y de F, donc un unique (y, z) de $F \times F^\perp$ tq $x = y + z$.

Retenons que la décomposition de x sur $E = F \oplus F^\perp$ est $x = y + z$ avec $y = \sum_{i=1}^{p} (e_i \mid x)\, e_i$.

Exo : Soit F et G deux sous-espaces d'un espace euclidien.

Montrer que $(F + G)^\perp = F^\perp \cap G^\perp$ et que $(F \cap G)^\perp = F^\perp + G^\perp$.

Sol :

– Si x est dans $F^\perp \cap G^\perp$, il est orthogonal à tous les vecteurs de F et de G donc à tous ceux de $F + G$ (par linéarité). Il est donc dans $(F + G)^\perp$, ce qui prouve l'inclusion $F^\perp \cap G^\perp \subset (F + G)^\perp$.

  Réciproquement, si x est orthogonal aux vecteurs de $F + G$, alors il est en particulier orthogonal à ceux de F et de G : il est donc dans $F^\perp \cap G^\perp$. On a donc bien l'égalité $(F + G)^\perp = F^\perp \cap G^\perp$.

– On applique ce qui précède à $F^\perp$ et $G^\perp$, et on trouve :

$(F^\perp + G^\perp)^\perp = (F^\perp)^\perp \cap (G^\perp)^\perp = F \cap G$.

Par passage à l'orthogonal, on en déduit l'égalité : $F^\perp + G^\perp = (F \cap G)^\perp$.

Exo : (sur l'importance de l'hypothèse de dimension finie)

On munit l'espace $E = \mathcal{C}([-1,1], \mathbb{R})$ du produit scalaire défini par $(f \mid g) = \displaystyle\int_{-1}^{1} f(t)g(t)\,\mathrm{d}t$.

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

Montrer que $F^\top = G$ et $G^\top = F$, mais qu'on n'a pas $E = F \oplus G$.

Sol : Il est clair que : $\forall (f,g) \in F \times G,\ (f \mid g) = 0$, donc $F \subset G^\perp$ et $G \subset F^\perp$.

Soit f un élément de $G^\perp$. Soit g définie sur $[-1,1]$ par : $\begin{cases} \forall t \in [-1,0],\ g(t) = tf(t) \\ \forall t \in [0,1],\ g(t) = 0 \end{cases}$.

La fonction g est continue sur $[-1,1]$, et nulle sur $[0,1]$: c'est donc un élément de G.

Ainsi $(f \mid g) = 0$, c'est-à-dire $\displaystyle\int_{-1}^{0} tf^2(t)\,\mathrm{d}t = 0$,

et donc (continuité, signe constant) : $\forall t \in [-1,0],\ f(t) = 0$.

La fonction f est donc dans F, et on a montré l'inclusion $G^\perp \subset F$ donc l'égalité $G^\perp = F$.

Idem pour l'égalité $F^\perp = G$.

Tous les éléments de $F \oplus G$ s'annulent en 0, donc $F \oplus G \subsetneq E$.

Prop : Soit E un espace euclidien, donc de dimension finie n.
Soit F un sous-espace vectoriel de E. Alors $\dim(F^\perp) = n - \dim(F)$.
De plus on a l'égalité $F = (F^\perp)^\perp$. Ainsi F est lui-même le supplémentaire orthogonal de $F^\perp$.

En dimension finie, on dira donc de F et $F^\perp$

qu'ils sont supplémentaires orthogonaux l'un de l'autre.

Preuve : On sait que $E = F \oplus F^\perp$ (sous l'hypothèse $\dim(F) < \infty$ d'ailleurs.

Ainsi $\dim(F^\perp) = n - \dim(F)$. De même : $\dim((F^\perp)^\perp) = n - \dim(F^\perp) = \dim(F)$.

Or on sait que $F \subset (F^\perp)^\perp$: il en résulte $F = (F^\perp)^\perp$.

Ainsi la décomposition $E = F \oplus F^\perp$ peut aussi se lire $E = (F^\perp)^\perp \oplus F^\perp$.

Exo : On munit $\mathcal{M}_n(\mathbb{R})$ de son produit scalaire canonique, défini par $(A \mid B) = tr(A^T B)$. Montrer que les espaces des matrices symétriques (resp. antisymétriques) sont supplémentaires orthogonaux.

Sol : On sait parfaitement que ces deux sous-espaces sont supplémentaires.

Il reste donc à vérifier que si A est symétrique, et B antisymétrique, alors $(A \mid B) = 0$.

Effectivement :

$(A \mid B) = tr(A^T B) = tr(AB) = tr(BA) = -tr(B^T A) = -(B \mid A) = -(A \mid B)$.

Supplémentaire orthogonal et bases orthonormées

Soit F un sous-espace vectoriel de l'espace euclidien E.

Si $\mathcal{B}$ est une base orthonormale de F et si $\mathcal{B}'$ est une base orthonormale de $F^\perp$,

alors $\mathcal{B} \cup \mathcal{B}'$ (obtenue par juxtaposition) est une base orthonormale de E.

Bien sûr $\mathcal{B} \cup \mathcal{B}'$ est orthonormale car les vecteurs de $\mathcal{B}$ (dans F)

sont orthogonaux à ceux de $\mathcal{B}'$ (dans $F^\perp$).

Enfin $\mathcal{B} \cup \mathcal{B}'$ est simplement une base de E adaptée à la somme directe $E = F \oplus F^\perp$.

Réciproquement, si on complète une base orthonormale $\mathcal{B}'(e_1, \ldots, e_p)$ de F en une base orthonormale $\mathcal{B}(e_1, \ldots, e_p, e_{p+1}, \ldots, e_n)$ de E,

alors $\mathcal{B}''(e_{p+1}, \ldots, e_n)$ est une base orthonormale de $F^\perp$.

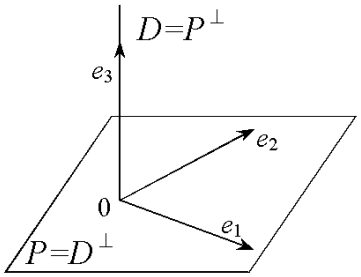
Les vecteurs de $\mathcal{B}''$ sont orthogonaux à ceux de $\mathcal{B}'$ donc ils sont dans $F^\perp$.

Pour finir, $\mathcal{B}''$ (libre car orthonormale) est de cardinal $n - p = \dim(F^\perp)$

donc est une base de $F^\perp$.

Plaçons nous dans un espace euclidien E de dimension 3.

Ici, le plan vectoriel P et la droite vectorielle D sont supplémentaires orthogonaux l'un de l'autre.



Si (e_1, e_2) est une base de P et si e_3 est une base de D, alors la famille (e_1, e_2, e_3) est une base orthonormale de E si et seulement si (e_1, e_2) est une base orthonormale de P et e_3 est unitaire.

Projection orthogonale

Def : Soit E un espace préhilbertien réel. Soit F un sous-espace vectoriel de dimension finie de E.

Avec ces hypothèses, on sait que $E = F \oplus F^\perp$.

La projection p_F sur F parallèlement à $F^\perp$ est appelée *projection orthogonale* sur F.

Prop : Soit F un sous-espace de dimension finie d'un espace préhilbertien réel E.

Soit $\mathcal{B} = (e_1, \ldots, e_p)$ une base orthonormale de F.

Pour tout x de E, on a : $p_F(x) = \sum_{i=1}^{p} (e_i \mid x)\, e_i$.

Preuve : Cela résulte de $E = F \oplus F^\perp$ et de la démonstration de la proposition supl orthog.

Exo : Soit E un espace préhilbertien sur $\mathbb{R}$.

On suppose qu'il existe n vecteurs unitaires $(e_k)_{1 \leqslant k \leqslant n}$ tels que :

$\forall x \in E,\ \|x\|^2 = \sum_{k=1}^{n} (e_k \mid x)^2 \quad (\star)$.

Montrer que les vecteurs $(e_k)_{1 \leqslant k \leqslant n}$ forment une base orthonormée de E.

Sol :

– En posant $x = e_j$ dans $(\star)$, on trouve : $1 = \sum_{k=1}^{n} (e_k \mid e_j)^2 = 1 + \sum_{k \neq j} (e_k \mid e_j)^2$, donc :

$\forall k \neq j,\ (e_k \mid e_j) = 0$.

La famille $(e_k)_{1\leqslant k\leqslant n}$ est orthonormale.

– Soit F le sev de $E =< (e_k)_{1\leqslant k\leqslant n} >$, et p la projection orthogonale de E sur F.

Pour tout u de E, $p(u) = \sum_{k=1}^{n} (e_k \mid u)\, e_k$ et donc $\|p(u)\|^2 = \sum_{k=1}^{n} (e_k \mid u)^2 = \|u\|^2$.

On sait d'autre part que $\|u\|^2 = \|p(u)\|^2 + \|u - p(u)\|^2$ (Pythagore).

Ainsi $u - p(u) = \overrightarrow{0}$. Autrement dit, pour tout u de E, $u = p(u) \in F$. Donc $F = E$.

La famille $(e_k)_{1\leqslant k\leqslant n}$ est donc une base orthonormée de E.

Remarque :

Si p est la projection orthogonale sur F, alors $\mathrm{Id} - p$ est la projection orthogonale sur $F^\perp$.

Bien sûr $\mathrm{Id} - p$ est la projection de E sur $F^\perp$ parallèlement à F.

Mais on sait que $F = (F^\perp)^\perp$.

Ainsi, $\mathrm{Id} - p$ est la projection orthogonale de E sur $F^\perp$.

Projection orthogonale sur une droite ou un hyperplan

– Soit $D = \mathbb{R}a$ la droite vectorielle engendrée par un vecteur a unitaire de E .

La projection orthogonale sur D s'écrit : $p_D(x) = (a \mid x)\, a$.

Si on suppose seulement que a est non nul, alors $p_D(x) = \left(\frac{a}{\|a\|} \,\middle|\, x\right) \frac{a}{\|a\|} = \frac{(a \mid x)}{\|a\|^2}\, a$.

– On suppose ici que E est euclidien (donc de dimension finie).

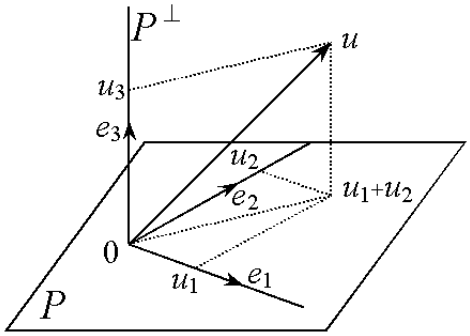
Soit $a \in E$ non nul, et $H = (\mathbb{R}a)^\perp$ l'hyperplan orthogonal à la droite $D = \mathbb{R}a$.

La projection orthogonale sur H est définie par $p_H(x) = x - \frac{(a \mid x)}{\|a\|^2}\, a$.

Si a est unitaire, cela se simplifie en $p_H(x) = x - (a \mid x)\, a$.

Illustration en dimension 3

On suppose ici $\dim E = 3$.
Soit e_1, e_2 une base orthonormale du plan P.
Soit e_3 un vecteur unitaire de la droite $P^\perp$.
On s'est donné un vecteur u de E.
Chaque $u_i = (e_i \mid u)\, e_i$ est la projection orthogonale de u sur la droite engendrée par e_i.
Le vecteur $u_1 + u_2 = (e_1 \mid u)\, e_1 + (e_2 \mid u)\, e_2$ est la projection orthogonale de u sur le plan P.



Retour à l'algorithme de Gram-Schmidt

On reprend ici les notations de la définition .

Soit E un espace préhilbertien réel.

Soit $(e_i)_{1 \leqslant i \leqslant p}$ une famille libre de E.

Le procédé de G-S transforme la famille $(e_i)_{1 \leqslant i \leqslant p}$ en une famille orthonormale $(\varepsilon_i)_{1 \leqslant i \leqslant p}$.

La formation du vecteur ε_i peut être interprétée de la manière suivante :

– Soit $F_{i-1} = \mathrm{Vect}(e_1, \ldots, e_{i-1}) = \mathrm{Vect}(\varepsilon_1, \ldots, \varepsilon_{i-1})$.

La projection orthogonale v_i de e_i sur F_{i-1} est donnée par $v_i = \sum_{j=1}^{i-1} (\varepsilon_j \mid e_i)\, \varepsilon_j$.

– On en déduit $u_i = e_i - v_i$, orthogonal à F_{i-1} et non nul.

Il suffit alors de normer le vecteur u_i pour obtenir le vecteur ε_i.

Prop : Soit p une projection vectorielle de l'espace euclidien E.

L'application p est une projection orthogonale si et seulement si :

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

Pour'une projection orthogonale, l'égalité $\|p(x)\| \leqslant \|x\|$ s'appelle inégalité de Bessel.

Preuve :
– On suppose que p est la projection orthogonale p_F sur un sous-espace F de E.

Pour tout x de E, écrivons $x = y + z$, avec $y = p_F(x)$, et z dans $F^\perp$.

On a $(y \mid z) = 0$ donc $\|x\|^2 = \|y\|^2 + \|z\|^2$ par Pythagore .

Il en découle $\|y\| \leqslant \|x\|$, avec égalité si et seulement si $z = \overrightarrow{0}$ c'est-à-dire $x \in F$.

– Réciproquement on suppose $E = F \oplus G$.

Soit p la projection vectorielle sur F, parallèlement à G.

On suppose que $\|p(x)\| \leqslant \|x\|$ pour tout x de E. On veut montrer que $G = F^\perp$.

Puisque $\dim(G) = \dim(F^\perp)$, il suffit de vérifier $G \subset F^\perp$,

c'est-à-dire : $\forall (y, z) \in F \times G,\ (y \mid z) = 0$.

On se donne donc y dans F et z dans G.

$\forall \lambda$, $p(\lambda y + z) = \lambda y$ et $\|p(\lambda y + z)\| \leqslant \|\lambda y + z\|$, $\|\lambda y\|^2 \leqslant \|\lambda y + z\|^2$.

On développe, $2\lambda\,(y \mid z) + \|z\|^2 \geqslant 0$ $\forall \lambda$, ce qui implique $(y \mid z) = 0$.

Def : Soit E un espace préhilbertien réel.

Soit F un sous-espace vectoriel de dimension finie de E.

La symétrie s_F par rapport F, parallèlement à $F^\perp$, est appelée *symétrie orthogonale* par rapport à F.

Distance à un sous-espace de dimension finie

Prop : Soit E un espace préhilbertien réel.

Soit F un sous-espace vectoriel de dimension finie de E.

Soit x un vecteur de E, et soit $y_0 = p_F(x)$ la projection orthogonale de x sur F.

Alors $\|x - y_0\| = \min\limits_{y \in F} \|x - y\|$, et y_0 est le seul vecteur de F à posséder cette propriété.

Ce minimum est appelé *distance de x à F*.

Il est noté $d(x, F)$, et il vérifie : $d(x, F)^2 = \|x\|^2 - \|y_0\|^2$.

On exprime la spécificité de $y_0 = p_F(x)$ par :

« y_0 est la meilleure approximation quadratique de x ».

Preuve :

On voit ici la projection orthogonale $y_0 = p_F(x)$ d'un vecteur x de E sur le sous-espace F.

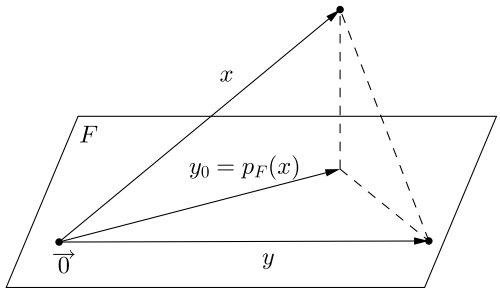
Parmi tous les vecteurs y de F, le projeté orthogonal de x est celui qui est

« le plus proche » de x.

En effet, pour tout vecteur y de F, on a :

$$\begin{aligned}\|x-y\|^2 &= \Big\| \underbrace{x-y_0}_{\in F^\perp} + \underbrace{y_0-y}_{\in F} \Big\|^2 \\ &= \|x-y_0\|^2 + \|y_0-y\|^2 \geqslant \|x-y_0\|^2\end{aligned}$$

(avec égalité si et seulement si $y = y_0$).



Enfin : $x = \underbrace{x-y_0}_{\in F^\perp} + \underbrace{y_0}_{\in F}$ donc $\|x\|^2 = \|x-y_0\|^2 + \|y_0\|^2$.

On en déduit $d(x,F)^2 = \|x-y_0\|^2 = \|x\|^2 - \|y_0\|^2$.

Exo : Calculer le minimum de $I_{a,b} = \displaystyle\int_0^1 (t\ln(t) - at - b)^2 dt$

et dire pour quels a, b ce minimum est atteint.

Sol : On définit un produit scalaire sur $E = \mathcal{C}([0,1],\mathbb{R})$ en posant : $(f \mid g) = \displaystyle\int_0^1 f(t)g(t)\,\mathrm{d}t$.

Soit F le sous-espace de E formé des polynômes de degré inférieur ou égal à 1.

Notons φ l'application de E définie par $\varphi(t) = t\ln(t)$, avec $\varphi(0) = 0$.

Ça équivaut donc à la recherche de P dans F qui minimise la quantité $\|\varphi - P\|$.

On sait que ce problème a une solution unique : la projection orthogonale P de φ sur F.

Cette projection est l'unique vecteur :

$P(t) = a + bt$ de F tel que $(P - \varphi \mid 1) = 0$ et $(P - \varphi \mid t) = 0$.

Il s'agit donc de résoudre le système (S) : $(P \mid 1) = (\varphi \mid 1)$, $(P \mid t) = (\varphi \mid t)$

Par des intégrations par parties très simples, on trouve :

$$(\varphi \mid 1) = \int_0^1 t \ln(t)\,\mathrm{d}t = \left[\frac{t^2}{2}\ln(t)\right]_0^1 - \int_0^1 \frac{t}{2}\,\mathrm{d}t = -\frac{1}{4},$$

et

$$(\varphi \mid t) = \int_0^1 t^2 \ln(t)\,\mathrm{d}t = \left[\frac{t^3}{3}\ln(t)\right]_0^1 - \int_0^1 \frac{t^2}{3}\,\mathrm{d}t = -\frac{1}{9}.$$

Enfin, en posant $P(t) = a + bt$, on trouve : $(P \mid 1) = a + \dfrac{b}{2}$ et $(P \mid t) = \dfrac{a}{2} + \dfrac{b}{3}$.

(S) est donc équivalent à : $a + \dfrac{b}{2} = -\dfrac{1}{4}$ et $\dfrac{a}{2} + \dfrac{b}{3} = -\dfrac{1}{9}$, et : $a = -\dfrac{1}{3}, b = \dfrac{1}{6}$.

Par « Pythagore », le minimum recherché $\|\varphi - P\|$ est tel que $\|\varphi - P\|^2 = \|\varphi\|^2 - \|P\|^2$.

$$\|\varphi\|^2 = \int_0^1 t^2(\ln(t))^2 dt = \left[\frac{t^3}{3}(\ln(t))^2\right]_0^1 - \frac{2}{3}\int_0^1 t^2\ln(t)dt = -\frac{2}{3}\int_0^1 t^2\ln(t)dt = \frac{2}{27}.$$

De même, $\|P\|^2 = \dfrac{1}{36}\displaystyle\int_0^1 (-2+t)^2 dt = \frac{7}{108}$.

Le minimum m des $I(a,b)$ vérifie donc $m^2 = \dfrac{2}{27} - \dfrac{7}{108} = \dfrac{1}{108}$. Ainsi $m = \sqrt{\dfrac{1}{108}} = \dfrac{\sqrt{3}}{18}$.

Utilisation du supplémentaire orthogonal

Soit F et G deux sous-espaces d'un espace euclidien E,

supplémentaires orthogonaux l'un de l'autre.

En d'autres termes, on suppose que $G = F^\perp$ (ou encore : $F = G^\perp$).

On note parfois $E = F \overset{\perp}{\oplus} G$.

Pour tout vecteur x de E, on a $x = p_F(x) + p_G(x)$ (somme orthogonale).

On a alors $d(x, F) = \|x - p_F(x)\| = \|p_G(x)\|$, et $d(x, G) = \|x - p_G(x)\| = \|p_F(x)\|$.

En particulier soit $\mathcal{B}(e_1, \ldots, e_p, e_{p+1}, \ldots, e_n)$ une base orthonormale de E, obtenue par concaténation d'une base orthonormale $\mathcal{B}_F(e_1, \ldots, e_p)$ de F et d'une base orthonormale

$\mathcal{B}_G(e_{p+1}, \ldots, e_n)$ de G.

Avec ces notations, et pour tout vecteur x de E, on a :

$$d(x,F)^2 = \|p_G(x)\|^2 = \sum_{i=p+1}^{n} (e_i \mid x)^2 \qquad \text{et} \qquad d(x,G)^2 = \|p_F(x)\|^2 = \sum_{i=1}^{p} (e_i \mid x)^2$$

Formes linéaires sur un espace euclidien

Formes linéaires et équations d'hyperplans

Prop : Soit E un espace vectoriel euclidien.

Pour toute forme linéaire f sur E,

il existe un unique vecteur a de E tel que : $\forall x \in E,\ f(x) = (a \mid x)$.

Preuve :

Pour tout vecteur a de E, l'application $f_a : x \mapsto (a \mid x)$ est une forme linéaire sur E.

Le noyau de la forme linéaire f_a est E si $a = \overrightarrow{0}$, et c'est l'hyperplan $(\mathbb{R}a)^\perp$ sinon.

L'application $a \mapsto f_a$ est linéaire de E dans $\mathcal{L}(E, \mathbb{R})$.

En effet, pour tous a, b de E, tous λ, μ de $\mathbb{R}$, et tout x de E, on a :

$f_{\lambda a+\mu b}(x) = (\lambda a+\mu b \mid x) = \lambda\,(a \mid x)+\mu\,(b \mid x) = \lambda f_a(x)+\mu f_b(x) = (\lambda f_a+\mu f_b)(x)$,

donc $f_{\lambda a+\mu b} = \lambda f_a+\mu f_b$.

Cette application est injective.

En effet, pour tout a de E : $f_a = 0 \Rightarrow f_a(a) = \|a\|^2 = 0 \Rightarrow a = \overrightarrow{0}$.

$\dim(E) = \dim(\mathcal{L}(E, \mathbb{R}))$, il vient que $a \mapsto f_a$ est un isomorphisme de E sur $\mathcal{L}(E, \mathbb{R})$.

Ainsi pour tout f de $\mathcal{L}(E, \mathbb{R})$, il existe un unique a de E tel que : $\forall x \in E,\ f(x) = (a \mid x)$.

Exo : On munit l'espace vectoriel $E = \mathbb{R}_n[X]$ du produit scalaire $(P \mid Q) = \displaystyle\int_0^1 P(t)Q(t)\,\mathrm{d}t$.

Soit φ la forme linéaire sur E définie par $\varphi(P) = P(0)$.

Montrer qu'il existe un unique A de E tel que : $\forall P \in E,\ \varphi(P) = (A \mid P)$.

Montrer que $\deg(A) = n$.

Sol : L'existence et l'unicité de A découle de la proposition ....

Par l'absurde, on suppose $\deg(A) < n$.

Le polynôme $P = XA$ est donc encore dans E.

On en déduit $0 = \varphi(P) = (A \mid P) = \displaystyle\int_0^1 t\,A^2(t)\,\mathrm{d}t$.

La fonction $t \mapsto tA^2(t)$ est continue, de signe constant, d'intégrale nulle sur $[0,1]$,

donc est nulle sur $[0,1]$.

Ainsi A est identiquement nul sur $]0,1]$, donc c'est le polynôme nul (absurde).

Exo : (importance de l'hypothèse de dimension finie).

On munit l'espace vectoriel $E = \mathbb{R}[X]$ du produit scalaire $(P \mid Q) = \displaystyle\int_0^1 P(t)Q(t)\,\mathrm{d}t$.

Soit φ la forme linéaire sur E définie par $\varphi(P) = P(0)$.

Montrer qu'il n'existe pas de vecteur A de E tel que, pour tout P de E, $\varphi(P) = (A \mid P)$.

Sol : Par l'absurde.

Avec $P = XA$, on alors $(XA \mid A) = \varphi(XA) = 0$, c'est-à-dire $\displaystyle\int_0^1 tA^2(t)dt = 0$.

$t \mapsto tA^2(t)$ est continue, de signe constant, d'intégrale nulle sur $[0,1]$,

donc est nulle sur $[0,1]$.

Ainsi A donc nul sur $]0,1]$, donc $A = 0$, donc $\varphi = 0$ (absurde car par exemple $\varphi(1) = 1$).

Def : Soit H un hyperplan d'un espace euclidien E.

Son supplémentaire orthogonal $H^\perp$ est une droite D, dite *normale* à l'hyperplan H.

On dit qu'un vecteur directeur a de D est un vecteur normal à l'hyperplan H.

Prop : Soit E un espace euclidien muni d'une base orthonormale $\mathcal{B} = (e_1, \ldots, e_n)$.

Soit H un hyperplan de E, et soit $a = \sum_{i=1}^{n} a_i e_i$ un vecteur non nul de E. On a $(i) \Leftrightarrow (ii)$:

(1) le vecteur a est normal à l'hyperplan H.

(2) une équation de H dans $\mathcal{B}$ est $(a \mid x) = 0$, c'est-à-dire $\sum_{i=1}^{n} a_i x_i = 0$.

Preuve :

– $i) \Rightarrow ii)$: on suppose que le vecteur a, non nul, est normal à l'hyperplan H.

Dans ces conditions il dirige la normale D à H et on a $E = H \oplus D$

(somme directe orthogonale).

Posons $a = \sum_{i=1}^{n} a_i e_i$, et soit $x = \sum_{i=1}^{n} x_i e_i$ un vecteur quelconque de E.

On a alors : $x \in H \Leftrightarrow x \in D^\perp \Leftrightarrow (a \mid x) = 0 \Leftrightarrow \sum_{i=1}^{n} a_i x_i = 0$ ...

– $ii) \Rightarrow i)$: réciproquement, on suppose qu'une équation de H est

$\sum_{i=1}^{n} a_i x_i = 0$ voir formes linéaires.

Puisque $\mathcal{B}$ est orthonormale, cette équation s'écrit $(a \mid x) = 0$, avec $a = \sum_{i=1}^{n} a_i e_i$ et $x = \sum_{i=1}^{n} x_i e_i$.

$\forall x \in H$ est $\perp$ au vecteur a : en d'autres termes, a est normal à l'hyperplan H.

Distance d'un vecteur à un hyperplan, à une droite

Prop : Soit H un hyperplan d'un espace euclidien E. Soit a un vecteur unitaire, normal à H.

Pour tout vecteur x de E, la distance de x à l'hyperplan H est : $d(x, H) = |(a \mid x)|$.

Preuve : On se reportera ici à la note

« Projection orthogonale sur une droite ou un hyperplan » .

Soit $D = \mathbb{R}a$. Par hypothèse, on a $E = D \oplus H$ (somme directe orthogonale).

Pour tout vecteur x de E, on a $x = p_D(x) + p_H(x)$, avec $p_D(x) = (a \mid x)\, a$.

On a alors $d(x, H) = \|x - p_H(x)\| = \|p_D(x)\| = |(a \mid x)|$.

Si a n'est pas unitaire, la distance de x à H s'écrit : $d(x, H) = \dfrac{1}{\|a\|} |(a \mid x)|$

Prop : Soit D une droite d'un espace euclidien E. Soit a un vecteur directeur *unitaire* de D.

Pour tout vecteur x de E, la distance de x à D est : $d(x, D) = \|x - (a \mid x)\, a\|$.

Preuve : Posons $E = D \oplus H$ (somme directe orthogonale),

où H est l'hyperplan orthogonal à la droite $D = \mathbb{R}a$.

Pour tout x de E, on a $x = p_D(x) + p_H(x)$, avec $p_D(x) = (a \mid x)\, a$.

On a alors $d(x, D) = \|x - p_D(x)\| = \|x - (a \mid x)\, a\|$.

Si le vecteur normal a n'est pas unitaire, alors $d(x, D) = \left\| x - \dfrac{(a \mid x)}{\|a\|^2} a \right\|$.

Si $\mathcal{B}_H = (e_1, \ldots, e_{n-1})$ est une base orthonormale de $H = D^\perp$, alors $d(x, D) = \sqrt{\sum_{i=1}^{n-1} (e_i \mid x)^2}$.

Distance à une droite dans $\mathbb{R}^2$

On se place dans l'espace euclidien $\mathbb{R}^2$, avec son produit scalaire canonique.

Soit D une droite de $\mathbb{R}^2$ (un hyperplan de $\mathbb{R}^2$) d'équation $\alpha x + \beta y = 0$.

Soit $u(x_0, y_0)$ un vecteur quelconque. Alors la distance de u à D est :

$$d(u, D) = \frac{|\alpha x_0 + \beta y_0|}{\sqrt{\alpha^2 + \beta^2}}.$$

Preuve : C'est un cas particulier de la prop dist-hyperplan, avec ici $H = D$ et $a = (\alpha, \beta)$.

Distance à un plan dans $\mathbb{R}^3$

On se place dans l'espace euclidien $\mathbb{R}^3$, avec son produit scalaire canonique.

Soit P un plan de $\mathbb{R}^3$ (un hyperplan), d'équation $\alpha x + \beta y + \gamma z = 0$.

Soit $u(x_0, y_0, z_0)$ un vecteur quelconque. La distance de u à P est :

$$d(u, P) = \frac{|\alpha x_0 + \beta y_0 + \gamma z_0|}{\sqrt{\alpha^2 + \beta^2 + \gamma^2}}.$$

C'est un cas particulier de la prop-dist-hyperplan, avec ici $H = P$ et $a = (\alpha, \beta, \gamma)$.

Distance à une droite dans $\mathbb{R}^3$

Soit D une droite de $\mathbb{R}^3$ muni de son produit scalaire et de son orientation canoniques.

Si (e_1, e_2) est une base orthonormale de $D^\perp$, alors $d(u, D) = \sqrt{(u \mid e_1)^2 + (u \mid e_2)^2}$.

Preuve : C'est un cas particulier de la note :Utilisation du supplémentaire orthogonal.

Ici la famille (e_1, e_2, ω) est un base orthonormale de E.

La projection orthogonale de u sur D s'écrit $p_D(u) = (\omega \mid u)\, u$.

Quand à la projection orthogonale de u sur le plan $H = D^\perp$,

elle s'écrit $p_H(u) = (e_1 \mid u)\, e_1 + (e_2 \mid u)\, e_2$.

On a alors $d(u, D)^2 = \|u - p_D(u)\|^2 = \|p_H(u)\|^2 = (e_1 \mid u)^2 + (e_2 \mid u)^2$.

Si ω est unitaire, une expression de la distance de u à la droite $D = \mathbb{R}\omega$ est :

$d(u, D) = \|u \wedge \omega\|$.

Preuve : On forme une base orthonormale directe (e_1, e_2, ω).

Si $u = \alpha e_1 + \beta e_2 + \gamma\omega$, on sait que $d(u, D)^2 = \alpha^2 + \beta^2$.

On trouve $u \wedge w = (\alpha e_1 + \beta e_2 + \gamma\omega) \wedge w = -\alpha e_2 + \beta e_3$. Ainsi : $\|u \wedge \omega\|^2 = \alpha^2 + \beta^2 = d(u, D)^2$.

Exo : Soit E un eve de dim $n \geqslant 1$, muni d'une bon $\mathcal{B} = (e_k)_{1 \leqslant k \leqslant n}$.

Pour toutes familles $\mathcal{U} = (u_i)_{1 \leqslant i \leqslant p}$, et $\mathcal{V} = (v_j)_{1 \leqslant j \leqslant q}$ de vecteurs de E :

– on note U la matrice de la famille $\mathcal{U}$ dans la base $\mathcal{B}$.

– $G(\mathcal{U}, \mathcal{V})$ la matrice, élément de $\mathcal{M}_{p,q}(\mathbb{R})$, dont le terme général est $g_{i,j} = (u_i \mid v_j)$.

– on note $G(\mathcal{U})$ plutôt que $G(\mathcal{U},\mathcal{U})$, et on pose $\Delta(\mathcal{U}) = \det(G(\mathcal{U}))$.

(1) Avec les notations précédentes, montrer que $G(\mathcal{U}, \mathcal{V}) = U^T V$,

et que $rg(G(\mathcal{U})) = rg(\mathcal{U})$.

(2) On suppose maintenant que $\mathcal{U}$ est une base d'un sous-espace F de E,

avec $1 \leqslant \dim(F) = m \leqslant n$.

$\forall x \in E$, on note $\mathcal{U}_x$ la famille obtenue en complétant $\mathcal{U}$ par le vecteur x.

On définit la matrice $G(\mathcal{U}_x)$ (d'ordre $m + 1$), et de déterminant $\Delta(\mathcal{U}_x)$.

Montrer que la distance de x à F s'obtient par la formule : $d(x, F)^2 = \dfrac{\Delta(\mathcal{U}_x)}{\Delta(\mathcal{U})}$.

Sol :

(1) Soit $u_i = \displaystyle\sum_{k=1}^{n} a_{k,i} e_k$ et $v_j = \sum_{k=1}^{n} b_{k,j} e_k$

les écritures de u_i et v_j dans la base orthonormale $\mathcal{B}$.

On a : $\big[U^T V\big]_{i,j} = \displaystyle\sum_{k=1}^{n} \big[U^T\big]_{i,k} \big[V\big]_{k,j} = \sum_{k=1}^{n} \big[U\big]_{k,i} \big[V\big]_{k,j} = \sum_{k=1}^{n} a_{k,i} b_{k,j} = (u_i \mid v_j)$.

On a donc l'égalité $G(\mathcal{U}, \mathcal{V}) = U^T V$.

Le résultat $rg(U^T U) = rg(U) = rg(\mathcal{U})$ est quant à lui classique.

(2) On note X la colonne des coordonnées de x dans la base orthonormale $\mathcal{B}$.

La matrice associée à la famille $\mathcal{U}_x$ s'écrit donc $U_x = (U \mid X)$.

On en déduit l'écriture par blocs : $G(\mathcal{U}_x) = U_x^T\, U_x = \begin{pmatrix} U^T \\ X^T \end{pmatrix} (U\ X) = \begin{pmatrix} U^TU & U^TX \\ X^TU & \|X\|^2 \end{pmatrix}$.

On décompose x en $x = a + b$, avec a dans F et b dans $F^\perp$

(donc matriciellement $X = A + B$).

On a $U^TB = 0$ (ou encore $B^TU = 0$), donc $\Delta(\mathcal{U}_x) = \det(G(\mathcal{U}_x)) = \begin{vmatrix} G(\mathcal{U}) & U^TA \\ A^TU & \|A\|^2 + \|B\|^2 \end{vmatrix}$.

On linéarise la dernière colonne en écrivant $\begin{pmatrix} U^TA \\ \|A\|^2 + \|B\|^2 \end{pmatrix} = \begin{pmatrix} U^TA \\ \|A\|^2 \end{pmatrix} + \begin{pmatrix} 0 \\ \|B\|^2 \end{pmatrix}$.

On développe le déterminant : $\Delta(\mathcal{U}_x) = \begin{vmatrix} G(\mathcal{U}) & U^TA \\ A^TU & \|A\|^2 \end{vmatrix} + \begin{vmatrix} G(\mathcal{U}) & 0 \\ A^TU & \|B\|^2 \end{vmatrix}$.

$\begin{vmatrix} G(\mathcal{U}) & U^TA \\ A^TU & \|A\|^2 \end{vmatrix}$ représente $\Delta(\mathcal{U}_a)$, qui est nul car $\mathcal{U}_a$ est une famille liée.

Le deuxième déterminant est triangulaire par blocs, et vaut $\Delta(\mathcal{U})\,\|B\|^2$,

c'est-à-dire $\Delta(\mathcal{U})d(x,F)^2$.

On a ainsi obtenu l'égalité : $d(x,F)^2 = \dfrac{\Delta(\mathcal{U}_x)}{\Delta(\mathcal{U})}$.