Chapitre 3

Espaces préhilbertiens réels

(version mise à jour le 20 juin 2020)

Sommaire

3.1 Produit scalaire et norme associée

3.1.1 Espace préhilbertien réel, espace euclidien

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.

Définition 3.1.1 (produit scalaire sur un $\mathbb{R}$-espace vectoriel)
Un *produit scalaire* est une application notée $\begin{array}{c} E \times E \to \mathbb{R} \\ (x,y) \mapsto (x \mid y) \end{array}$ et possédant les propriétés suivantes :

– caractère *bilinéaire* : $\begin{cases} (\alpha x + \beta x' \mid y) = \alpha\,(x \mid y) + \beta\,(x' \mid y) \\ (x \mid \alpha y + \beta y') = \alpha\,(x \mid y) + \beta\,(x \mid y') \end{cases}$
– caractère *symétrique* : $(x \mid y) = (y \mid x)$.
– caractère *défini positif* : $(x \mid x) \geqslant 0$, et $(x \mid x) = 0 \Leftrightarrow x = \overrightarrow{0}$.

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 $<x,y>$.

Définition 3.1.2 (espace préhilbertien réel, espace euclidien)
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*.

3.1.2 Exemples de référence

Proposition 3.1.1 (produit scalaire canonique sur $\mathbb{R}^n$)
Le produit scalaire canonique *sur* $\mathbb{R}^n$ *est défini par* $(x \mid y) = \displaystyle\sum_{i=1}^{n} x_i\, y_i$, *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 $(x \mid y) = X^\top Y$.

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

Pour le caractère « défini » on a $(x \mid x) = \displaystyle\sum_{i=1}^{n} x_i^2 = 0 \Rightarrow (\forall\, i \in [\![1,n]\!],\ x_i = 0)$ car les x_i sont des *réels*. □

Proposition 3.1.2 (produit scalaire canonique sur $\mathcal{M}_{n,p}(\mathbb{R})$)
Le produit scalaire canonique *sur* $\mathcal{M}_{n,p}(\mathbb{R})$ *est défini par* $(A \mid B) = \operatorname{tr}(A^\top B)$.

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

Commençons par une première remarque (rassurante) sur les notations :
– les matrices $A^\top B$ et $B^\top A$, carrées d'ordre p, sont transposées l'une de l'autre, donc ont même trace ;
– les matrices $BA^\top$ et $AB^\top$, carrées d'ordre n, sont transposées l'une de l'autre, donc ont même trace ;
– de plus $A^\top B$ et $BA^\top$ ont la même trace (propriété classique de la trace : voir prop.1.2.14) ;
– finalement, on peut écrire, sans distinction $(A \mid B) = \operatorname{tr}(A^\top B) = \operatorname{tr}(B^\top A) = \operatorname{tr}(A\,B^\top) = \operatorname{tr}(B\,A^\top)$.

On note ensuite que : $\operatorname{tr}(A^\top B) = \sum_{j=1}^{n}[A^\top B]_{j,j} = \sum_{j=1}^{n}\sum_{i=1}^{n}[A^\top]_{j,i}[B]_{i,j} = \sum_{i,j=1}^{n}[A]_{i,j}[B]_{i,j} = \sum_{i,j=1}^{n} a_{i,j}\, b_{i,j}$.
Cela revient à utiliser le produit scalaire canonique sur $\mathbb{R}^{np}$, si on convient d'identifier une matrice de $\mathcal{M}_{n,p}(\mathbb{R})$ au np-uplet de ses coefficients dans la base canonique. $\square$

Proposition 3.1.3 (un produit scalaire sur $\mathcal{C}^0([a,b],\mathbb{R})$)
Soit $\mathcal{C}^0([a,b],\mathbb{R})$ l'espace vectoriel des applications continues de $[a,b]$ dans $\mathbb{R}$, avec $a<b$.
En posant $(f \mid g) = \int_a^b f(t)\,g(t)\,\mathrm{d}t$, on définit un produit scalaire sur $\mathcal{C}^0([a,b],\mathbb{R})$.

Les propriétés de bilinéarité, symétrie, et positivité sont évidentes. Pour le caractère « défini » on a :
$(f \mid f) = \int_a^b f^2(t)\,\mathrm{d}t = 0 \Rightarrow (\forall\, t\in[a,b],\ f(t)=0)$ (car f^2 est continue, positive, d'intégrale nulle). $\square$

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

Là encore $(f \mid f)=0 \Rightarrow (\forall\, t\in[a,b], f^2(t)\omega(t)=0) \Rightarrow (\forall\, t\in[a,b], f(t)=0)$: la nullité de f est en effet acquise en tout point où ω ne s'annule pas, et pour les autres (dont on sait qu'ils sont isolés), on procède par passage à la limite en utilisant les hypothèses de continuité. $\square$

Proposition 3.1.4 (inégalité de Cauchy-Schwarz, et cas d'égalité)
Soit E un espace préhilbertien réel.
Pour tous vecteurs x,y de E, on a l'inégalité dite « de Cauchy-Schwarz » : $(x \mid y)^2 \leqslant (x \mid x)\,(y \mid y)$
Il y a égalité dans ce résultat si et seulement si x et y sont liés.

- On suppose que x et y sont liés, et par exemple qu'il existe λ dans $\mathbb{R}$ tel que $y=\lambda x$.
  Alors $(x \mid y)^2 = (x \mid \lambda x)^2 = \lambda^2\,(x \mid x)^2 = (x \mid x)\,(\lambda x \mid \lambda x) = (x \mid x)\,(y \mid y)$.
- Réciproquement, on suppose que x et y sont libres, et on pose : $\forall\,\lambda\in\mathbb{R},\ \varphi(\lambda) = (\lambda x - y \mid \lambda x - y)$.
  Par défini-positivité du produit scalaire, l'application φ est à valeurs dans $\mathbb{R}^{+*}$ (en effet $\lambda x - y \neq 0$).
  On développe par bilinéarité et symétrie, et on trouve : $\forall\,\lambda\in\mathbb{R},\ \varphi(\lambda) = (x \mid x)\,\lambda^2 - 2\,(x \mid y)\,\lambda + (y \mid y)$.
  Ainsi φ est un trinôme en λ qui reste strictement positif sur $\mathbb{R}$.
  Son discriminant Δ est donc strictement négatif. Mais $\Delta = 4((x \mid y)^2 - (x \mid x)\,(y \mid y))$.
  Donc écrire $\Delta<0$, c'est écrire l'inégalité de Cauchy-Schwarz (qui ici est stricte). $\square$

Exercice 3.1 ($\rightsquigarrow$ **corrigé**)
Soit A,B deux matrices symétriques réelles d'ordre n. Montrer que $\big(\operatorname{tr}(AB)\big)^2 \leqslant \operatorname{tr}(A^2)\operatorname{tr}(B^2)$.
Qu'obtient-on par exemple si $B=I_n$? Cas d'égalité ?

3.1.3 Norme associée à un produit scalaire

Définition 3.1.3 (norme associée à un produit scalaire)
Soit E un espace préhilbertien réel.
Pour tout vecteur x de E, on appelle *norme de* x la quantité $\|x\| = \sqrt{(x \mid x)}$.
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 : $\forall\,(x,y) \in E^2,\ |(x \mid y)| \leqslant \|x\|\ \|y\|$.

Proposition 3.1.5 (un développement usuel)
Soit E un espace préhilbertien réel.
Pour tous x, y de E, et tous réels α, β on a : $\|\alpha x + \beta y\|^2 = \alpha^2 \|x\|^2 + 2\,\alpha\beta\,(x \mid y) + \beta^2 \|y\|^2$.

C'est évident par bilinéarité et symétrie :

$$\begin{aligned}\|\alpha x + \beta y\|^2 &= (\alpha x + \beta y \mid \alpha x + \beta y) = \alpha^2\,(x \mid x) + \alpha\beta\,(x \mid y) + \beta\alpha\,(y \mid x) + \beta^2\,(y \mid y)\\ &= \alpha^2 \|x\|^2 + 2\,\alpha\beta\,(x \mid y) + \beta^2 \|y\|^2 .\end{aligned}$$

□

En particulier : $\|x+y\|^2 = \|x\|^2 + 2\,(x \mid y) + \|y\|^2$ et $\|x-y\|^2 = \|x\|^2 - 2\,(x \mid y) + \|y\|^2$.
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*.

Exercice 3.2 (⇝ **corrigé**)
Soit E préhilbertien réel. Montrer que : $\forall\,(x,y) \in E^2,\ 2 + \|x+y\|^2 \leqslant 2(1+\|x\|^2)(1+\|y\|^2)$.

Proposition 3.1.6 (identité de polarisation)
Soit E un espace préhilbertien réel.
Pour tous vecteurs x, y de E, on a : $(x \mid y) = \dfrac{1}{2}\left(\|x+y\|^2 - \|x\|^2 - \|y\|^2\right)$.

C'est en effet une simple réécriture de l'égalité : $\|x+y\|^2 = \|x\|^2 + 2\,(x \mid y) + \|y\|^2$. □

Proposition 3.1.7 (propriétés de la norme associée à un produit scalaire)
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 l'existence de λ dans $\mathbb{R}^+$ tel que $y = \lambda x$ ou $x = \lambda y$.

Le seul résultat qui mérite une démonstration est l'inégalité triangulaire.
- Si x est nul, alors x et y sont positivement liés ($x = 0y$) et on a l'égalité $\|x+y\| = \|x\| + \|y\|$.
- Si x est non nul et si x, y sont liés. Soit λ dans $\mathbb{R}$ tel que $y = \lambda x$.
  Alors $\|x+y\| = |1+\lambda|\,\|x\|$ et $\|x\| + \|y\| = (1+|\lambda|)\,\|x\|$.

Avec ces notations : $\|x+y\| \leqslant \|x\|+\|y\| \Leftrightarrow |1+\lambda| \leqslant 1+|\lambda| \Leftrightarrow \lambda \leqslant |\lambda|$ (par élévation au carré).

On a donc $\|x+y\| \leqslant \|x\|+\|y\|$, qui n'est une égalité que si $\lambda \geqslant 0$, c'est-à-dire x, y positivement liés.

– Il reste à examiner le cas où x et y sont libres :

$$\begin{aligned}\|x+y\|^2 &= \|x\|^2 + 2\,(x \mid y) + \|y\|^2 \leqslant \|x\|^2 + 2\,|(x \mid y)| + \|y\|^2 \\ &< \|x\|^2 + 2\,\|x\|\,\|y\| + \|y\|^2 \ \text{(Cauchy-Schwarz strict car } x, y \text{ sont libres)}\end{aligned}$$

On est ainsi arrivé à $\|x+y\|^2 < (\|x\|+\|y\|)^2$, c'est-à-dire à $\|x+y\| < \|x\|+\|y\|$. □

Remarque : pour tous x, y de E, on a l'encadrement : $\big|\, \|x\| - \|y\| \,\big| \leqslant \|x \pm y\| \leqslant \|x\| + \|y\|$.

On écrit $\|x\| = \|y + (x-y)\| \leqslant \|y\| + \|x-y\|$ donc $\|x\| - \|y\| \leqslant \|x-y\|$.
Par symétrie, on trouve $\|y\| - \|x\| \leqslant \|x-y\|$, et il en découle $|\,\|x\| - \|y\|\,| \leqslant \|x-y\|$.
Ainsi $|\,\|x\| - \|y\|\,| \leqslant \|x-y\| \leqslant \|x\| + \|y\|$ (et on ne change rien en remplaçant y par $-y$). □

Définition 3.1.4 (vecteurs unitaires)
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 $\pm\dfrac{x}{\|x\|}$ 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 $\|x\| = \Big(\displaystyle\sum_{i=1}^n x_i^2\Big)^{1/2}$

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

– Dans $E = \mathcal{C}^0([a,b], \mathbb{R})$ muni de $(f \mid g) = \displaystyle\int_a^b f(t)\,g(t)\,\mathrm{d}t$, pour tout f de E on a $\|f\| = \sqrt{\displaystyle\int_a^b f^2(t)\,\mathrm{d}t}$

L'inégalité de Cauchy-Schwarz s'écrit alors : $\Big(\displaystyle\int_a^b f(t)\,g(t)\,\mathrm{d}t\Big)^2 \leqslant \int_a^b f^2(t)\,\mathrm{d}t\ \int_a^b g^2(t)\,\mathrm{d}t$

Exercice 3.3 (⇝ **corrigé**)
Soit E un espace préhilbertien réel, et soit a un vecteur unitaire.
Pour quelles valeurs du réel λ l'application $\varphi\colon (x,y) \mapsto (x \mid y) + \lambda\,(x \mid a)\,(y \mid a)$ est-elle un produit scalaire ?

3.1.4 Produit scalaire usuel du plan ou de l'espace

Dans le plan orienté $\mathbb{R}^2$

On se place dans le plan $\mathbb{R}^2$ muni de son produit scalaire et son orientation habituelles (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 $(u \mid v) = \cos(\widehat{u,v})$.
Plus généralement, si u et v sont non nuls, on a $(u \mid v) = \|u\|\,\|v\| \cos(\widehat{u,v})$.

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 : $\cos(\theta) = (v \mid u)$.
Dans le cas général (u, v non nuls), on applique ce qui précède à $u' = \dfrac{u}{\|u\|}$ et $v' = \dfrac{v}{\|v\|}$ (unitaires).

Ainsi : $(u \mid v) = (\|u\|\, u' \mid \|v\|\, v') = \|u\|\, \|v\|\, (u' \mid v') = \|u\|\, \|v\| \cos(\widehat{u', v'}) = \|u\|\, \|v\| \cos(\widehat{u, v})$. $\square$

- 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 s'écrivent alors $ax + by = h$, avec h dans $\mathbb{R}$.
- 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 appelée « équation normale » de la droite affine $\mathcal{D}$.
  Dans cette équation (h, θ_0) est un couple de coordonnées polaires de la projection de O sur $\mathcal{D}$.

Dans l'espace orienté $\mathbb{R}^3$

On se place dans l'espace $\mathbb{R}^3$ muni de son produit scalaire et son orientation habituelles (pour laquelle 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 $\mathrm{Vect}(u, v)$.
  Alors $|(u \mid v)| = \|u\|\, \|v\| \cos \theta$.
- 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 $\mathcal{P}$.

3.2 Orthogonalité

3.2.1 Vecteurs orthogonaux, sous-espaces orthogonaux

Définition 3.2.1 (vecteurs orthogonaux)
Soit E un espace préhilbertien réel.
Deux vecteurs x et y de E sont dits *orthogonaux* si $(x \mid y) = 0$.

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.

Exercice 3.4 (⇝ **corrigé**)
Soit E un espace préhilbertien réel. Soit u et v deux vecteurs de E.
On suppose que, pour tout réel λ, on a : $\|u + \lambda v\| \geq \|u\|$. Montrer que u et v sont orthogonaux.

Exercice 3.5 (⇝ **corrigé**)
Soit E préhilbertien réel, et $f : E \to E$ avec : $\forall\, (x, y),\ (f(x) \mid y) = (x \mid f(y))$. Montrer que f est linéaire.

Définition 3.2.2 (sous-espaces vectoriels orthogonaux)
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 $x = \overrightarrow{0}$. □

Proposition 3.2.1 (orthogonal d'un sous-espace vectoriel)
Soit F un sous-espace vectoriel d'un espace préhilbertien réel E.
On note $F^\perp = \{y \in E,\ \forall x \in F,\ (x \mid y) = 0\}$.
En d'autres termes, $F^\perp$ 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.*

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

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, au sens de 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 $F = \mathrm{Vect}\{x_i,\ i \in I\}$, alors : $y \in F^\perp \Leftrightarrow \forall i \in I,\ (y \mid x_i) = 0$.

Si y est dans $F^\perp$ (s'il est orthogonal à tous les vecteurs de F), 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. □

Exercice 3.6
On munit $\mathbb{R}[X]$ du produit scalaire $(A \mid B) = \displaystyle\int_{-1}^{1} A(t)B(t)\,\mathrm{d}t$.

On définit $\varphi \colon A \mapsto \displaystyle\int_{-1}^{1} |t|\,A(t)\,\mathrm{d}t$, et on note 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 $\mathbb{R}[X] = H \oplus H^\perp$?

3.2.2 Familles orthogonales

Définition 3.2.3 (familles orthogonales ou orthonormales)
Soit E un espace préhilbertien réel.
On dit qu'une famille $(x_i)_{i\in I}$ de vecteurs de E est *orthogonale* si les x_i sont orthogonaux deux à deux.
Si de plus ils sont unitaires, alors la famille est dite *orthonormale* (ou orthonormée).

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

Exemples de référence

– La base canonique de $\mathbb{R}^n$ est orthonormale pour le produit scalaire canonique.
– De même, 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 $(f \mid g) = \displaystyle\int_0^{2\pi} f(t)\,g(t)\,\mathrm{d}t$.
La famille des $f_n : t \mapsto \cos(nt)$, avec n dans $\mathbb{N}$, est orthogonale pour ce produit scalaire.

Proposition 3.2.2 (une famille orthogonale de vecteurs non nuls est libre)
Soit E un espace préhilbertien réel.
Si une famille $(x_i)_{i\in I}$ est orthogonale et formée de vecteurs non nuls, alors c'est une famille libre.

On considère l'égalité $\displaystyle\sum_{i\in I} \lambda_i x_i = \overrightarrow{0}$, et on la multiplie scalairement par un vecteur x_j, avec j dans J.
Ainsi $0 = \Big(\displaystyle\sum_{i\in I} \lambda_i x_i \mid x_j\Big) = \sum_{i\in I} \lambda_i \underbrace{(x_i \mid x_j)}_{=0 \text{ si } i\neq j} = \lambda_j \left\|x_j\right\|^2$, donc $\lambda_j = 0$ (pour tout j de I) : les x_j sont libres. □

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.

Proposition 3.2.3 (théorème de Pythagore)
Soit E un espace préhilbertien réel.
Si la famille $(x_i)_{1\leqslant i\leqslant p}$ est orthogonale, alors $\Big\|\displaystyle\sum_{i=1}^{p} x_i\Big\|^2 = \sum_{i=1}^{p} \left\|x_i\right\|^2$ *(relation de Pythagore).*

On développe par bilinéarité : $\Big\|\displaystyle\sum_{i=1}^{p} x_i\Big\|^2 = \Big(\sum_{i=1}^{p} x_i \mid \sum_{j=1}^{p} x_j\Big) = \sum_{i=1}^{n} \left\|x_i\right\|^2 + \sum_{i\neq j} \underbrace{(x_i \mid x_j)}_{=0 \text{ car } i\neq j} = \sum_{i=1}^{n} \left\|x_i\right\|^2$. □

Attention, la réciproque n'est vraie que si $p = 2$. Ainsi : $\left\|x_1 + x_2\right\|^2 = \left\|x_1\right\|^2 + \left\|x_2\right\|^2 \ \Leftrightarrow\ (x_1 \mid x_2) = 0$.

Exercice 3.7 ($\rightsquigarrow$ **corrigé**)
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 $U_n(X) = (X^2 - X)^n$.

a) Montrer que $(P_n)_{n\geqslant 0}$ est une famille orthogonale (utiliser des intégrations par parties répétées).

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

c) Former une base orthonormale de $\mathbb{R}_4[X]$ (faire le rapprochement avec l'exercice 3.8).

3.3 Bases orthonormales d'un espace euclidien

3.3.1 Algorithme d'orthonormalisation de Gram-Schmidt

Définition 3.3.1 (principe de l'algorithme)
Soit E un espace préhilbertien réel. Soit $(e_i)_{1\leqslant i\leqslant p}$ une famille libre de E.
Pour tout i de $\{1,\ldots,p\}$, 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 $\varepsilon_1 = \dfrac{1}{\|e_1\|}\,e_1$

Proposition 3.3.1 (preuve de l'algorithme)
Avec les notations précédentes, l'algorithme de Gram-Schmidt se termine.
Pour tout i de $\{1,\ldots,p\}$, on a $\mathrm{Vect}\big(\{\varepsilon_1,\ldots,\varepsilon_i\}\big) = \mathrm{Vect}\big(\{e_1,\ldots,e_i\}\big)$, *et* $(\varepsilon_i \mid e_i) > 0$.
La famille $(\varepsilon_i)_{1\leqslant i\leqslant p}$ ainsi construite est une base orthonormale de $\mathrm{Vect}\big(\{e_1,\ldots,e_p\}\big)$.

Pour simplifier les notations, on pose $F_i = \mathrm{Vect}(\{e_1,\ldots,e_i\})$ 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 $\varepsilon_1,\ldots,\varepsilon_i$ est orthonormale, elle vérifie $\mathrm{Vect}(\{\varepsilon_1,\ldots,\varepsilon_i\}) = F_i$, et on a $(\varepsilon_i \mid e_i) > 0$ ».

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 les indications de la définition 3.3.1, on pose donc $u_i = e_i - v_i$, en notant $v_i = \displaystyle\sum_{j=1}^{i-1} (\varepsilon_j \mid e_i)\,\varepsilon_j$.
Le vecteur v_i est dans $\mathrm{Vect}(\{\varepsilon_1,\ldots,\varepsilon_{i-1}\}) = F_{i-1}$ 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.
Pour tout k de $[\![1,i[\![$, on a $(\varepsilon_k \mid v_i) = \displaystyle\sum_{j=1}^{i-1} (\varepsilon_j \mid e_i)\,\underbrace{(\varepsilon_k \mid \varepsilon_j)}_{\delta_{j,k}} = (\varepsilon_k \mid e_i)$ donc $(\varepsilon_k \mid u_i) = 0$, donc $(\varepsilon_k \mid \varepsilon_i) = 0$.
Ainsi la famille $\varepsilon_1,\ldots,\varepsilon_i$ est orthonormale.

Par construction, u_i donc ε_i est dans F_i. Puisque $\mathrm{Vect}(\{\varepsilon_1,\ldots,\varepsilon_{i-1}\}) = F_{i-1}$ (hypothèse de récurrence), on en déduit $\mathrm{Vect}(\{\varepsilon_1,\ldots,\varepsilon_i\}) \subset F_i$ (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 $\mathrm{Vect}(\{e_1,\ldots,e_p\})$. □

Illustration du procédé

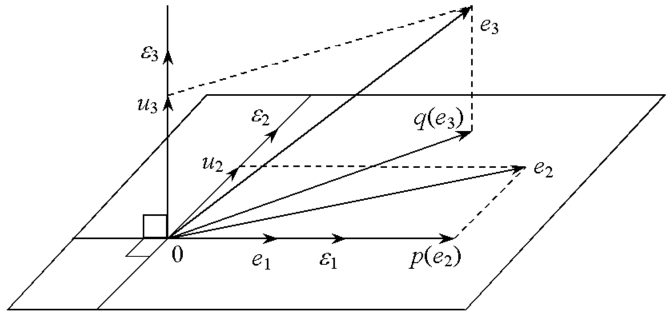
On va illustrer le passage d'une famille libre $e = (e_1, e_2, e_3)$ à une famille orthonormale $\varepsilon = (\varepsilon_1, \varepsilon_2, \varepsilon_3)$.

On a conservé les notations de la proposition 3.3.1 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 une famille $(e_i)_{1\leqslant i\leqslant p}$ est orthonormale, elle est invariante par l'algorithme de Gram-Schmidt.

Par récurrence. Avec les notations de la définition 3.3.1, on a bien sûr $\varepsilon_1 = e_1$.

Enfin si $\varepsilon_j = e_j$ pour $1 \leqslant j < i$, alors : $u_i = e_i - \displaystyle\sum_{j=1}^{i-1} (\varepsilon_j \mid e_i)\,\varepsilon_j = e_i - \sum_{j=1}^{i-1} \underbrace{(e_j \mid e_i)}_{=0} e_j = e_i$ donc $\varepsilon_i = e_i$. □

Exercice 3.8 (⇝ **corrigé**)

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

Former une base orthonormée de $\mathbb{R}_4[X]$ en appliquant le procédé de Gram-Schmidt à $1, X, X^2, X^3, X^4$.

Faire le rapprochement avec l'exercice 3.7.

3.3.2 Bases orthonormales d'un espace euclidien

Proposition 3.3.2 (existence de bases orthonormales dans un espace euclidien)
Soit E un espace euclidien (c'est-à-dire un espace préhilbertien réel de dimension finie).
Alors, dans l'espace vectoriel E, il existe des bases orthonormales.

Il suffit d'appliquer l'algorithme de Gram-Schmidt à une base quelconque $\mathcal{B}$ de E.

Proposition 3.3.3 (théorème de la base orthonormale incomplète)
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.

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. □

Proposition 3.3.4 (expression des coordonnées dans une base orthonormale)
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 : $x = \displaystyle\sum_{i=1}^{n} (e_i \mid x)\, e_i$

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

En d'autres termes, les *applications coordonnées* dans la base $\mathcal{B}$ sont les applications $e_i^* \colon x \mapsto (e_i \mid x)$.
Tout vecteur x de E est ainsi entièrement déterminé par ses produits scalaires sur les vecteurs de $\mathcal{B}$.

Proposition 3.3.5 (expression du produit scalaire et de la norme dans une base orthonormale)
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 $x = \displaystyle\sum_{i=1}^{n} x_i\, e_i$ *et* $y = \displaystyle\sum_{i=1}^{n} y_i\, e_i$*, on a :* $(x \mid y) = \displaystyle\sum_{i=1}^{n} x_i\, y_i$ et $\|x\|^2 = \displaystyle\sum_{i=1}^{n} x_i^2$.

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

Exercice 3.9 ($\rightsquigarrow$ **corrigé**)
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, $\|f(u)\| = \lambda\, \|u\|$.

▷ 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 : $(x \mid y) = X^\top Y$;
- dans le cas général d'une base $\mathcal{B}$ *quelconque*, alors on a seulement $(x \mid y) = \displaystyle\sum_{i=1}^{n} \sum_{j=1}^{n} x_i\, y_j\, (e_i \mid e_j)$;
  on peut écrire $(x \mid y) = X^\top G\, Y$, où $g_{i,j} = (e_i \mid e_j)$ pour $1 \leqslant i, j \leqslant n$.

Posons $x = \displaystyle\sum_{i=1}^{n} x_i\, e_i$ et $y = \displaystyle\sum_{j=1}^{n} y_j\, e_j$.
On a $(x \mid y) = \displaystyle\sum_{i=1}^{n} \sum_{j=1}^{n} x_i\, y_j\, (e_i \mid e_j) = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i\, y_j\, g_{i,j} = \sum_{i=1}^{n} x_i \Bigl(\sum_{j=1}^{n} g_{i,j} y_j \Bigr) = \sum_{i=1}^{n} x_i [GY]_i = X^\top G\, Y$
Si la famille $(e_i)_{1 \leqslant i \leqslant n}$ est orthonormale, on trouve bien sûr $G = I_n$ donc $(x \mid y) = X^\top Y$. □

Exercice 3.10
Soit A dans $\mathcal{M}_n(\mathbb{R})$. Montrer que les matrices $A^\top A$ et A ont le même noyau, donc le même rang.
Les matrices $A\,A^\top$ et A ont-elles le même rang ? le même noyau ?

Exercice 3.11 ($\rightsquigarrow$ **corrigé**)
Soit M dans $\mathcal{M}_n(\mathbb{R})$, telle que $M^2 = 0$. Montrer que $\operatorname{Ker}(M + M^\top) = \operatorname{Ker}(M) \cap \operatorname{Ker}(M^\top)$.

Exercice 3.12 ($\rightsquigarrow$ **corrigé**)
Soit M dans $\mathcal{M}_n(\mathbb{R})$, telle que $MM^\top = M^\top M$.
a) Montrer que M et $M^\top$ ont le même noyau, puis que $\operatorname{Sp}_{\mathbb{R}}(M) = \operatorname{Sp}_{\mathbb{R}}(M^\top)$.
b) Montrer que pour tout λ de $\operatorname{Sp}_{\mathbb{R}}(M)$, les sous-espaces propres de M et de $M^\top$ pour λ sont identiques.
c) Montrer que les sous-espaces propres de M dans $\mathcal{M}_{n,1}(\mathbb{R})$ sont orthogonaux deux à deux.

▷ Matrice d'une famille de vecteurs dans une base orthonormale

Soit $\mathcal{B} = (e_i)_{1\leqslant i\leqslant n}$ une base orthonormale de E, et soit $v = (v_j)_{1\leqslant j\leqslant p}$ 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 est un élément de $\mathcal{M}_{n,p}(\mathbb{K})$).
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 3.3.1, 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 $\operatorname{Vect}(\{e_1, \ldots, e_i\})$.
Il en découle que la matrice de passage $P_{\mathcal{B},\mathcal{B}'}$ de la base $\mathcal{B}$ à la base $\mathcal{B}'$ est triangulaire supérieure.
Il en est bien sûr de même de 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 proposition 3.3.1).
Les coefficients diagonaux de la matrice triangulaire $P_{\mathcal{B},\mathcal{B}'} = P^{-1}_{\mathcal{B}',\mathcal{B}}$ 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 $\mathcal{B}$. □

3.4 Projection orthogonale sur un sous-espace de dim finie

3.4.1 Supplémentaire orthogonal

Proposition 3.4.1 (supplémentaire orthogonal d'un sous-espace de dimension finie)
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.

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

Ainsi $\mathrm{F} = \mathrm{Vect}\{e_1, \ldots, e_p\}$, donc : $z \in F^{\perp} \Leftrightarrow \forall j \in [\![1, p]\!],\ (z \mid e_j) = 0$ (cf dernière remarque dans 3.2.1).

Soit x dans E. Il s'agit de prouver qu'il existe un couple unique (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 : $z \in \mathrm{F}^{\perp} \Leftrightarrow \forall j \in [\![1, p]\!],\ (z \mid e_j) = 0 \Leftrightarrow \forall j \in [\![1, p]\!],\ (x \mid e_j) = (y \mid e_j) = y_j$.

Ces conditions déterminent un unique y de F, donc un unique (y, z) de $F \times F^{\perp}$ tel que $x = y + z$.

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

Exercice 3.13
Soit F et G deux sous-espaces d'un espace euclidien.

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

Exercice 3.14 (⇝ **corrigé**)
(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 $(f \mid g) = \int_{-1}^{1} f(t) g(t)\, \mathrm{d}t$.

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 $E = F \oplus G$.

Proposition 3.4.2 (dimension du supplémentaire orthogonal en dimension finie)
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.

On sait que $E = F \oplus F^{\perp}$ (sous la seule hypothèse $\dim(F) < \infty$ d'ailleurs, cf 3.4.1).
Ainsi $\dim(F^{\perp}) = n - \dim(F)$. De même : $\dim((F^{\perp})^{\perp}) = n - \dim(F^{\perp}) = \dim(F)$.
Or on sait que $F \subset (F^{\perp})^{\perp}$ (voir remarque à la fin de la section 3.2.1) : il en résulte $F = (F^{\perp})^{\perp}$.
Ainsi la décomposition $E = F \oplus F^{\perp}$ peut aussi se lire $E = (F^{\perp})^{\perp} \oplus F^{\perp}$. □

Exercice 3.15

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

▷ Supplémentaire orthogonal et bases orthonormées

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

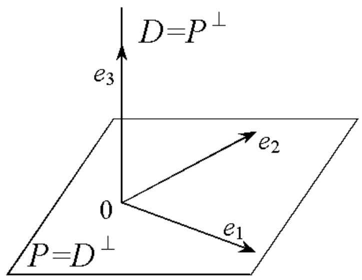
Si $\mathcal{B}$ est une base orthonormale de F et si $\mathcal{B}'$ est une base orthonormale de $F^\perp$, alors $\mathcal{B} \cup \mathcal{B}'$ (obtenue par juxtaposition) est une base orthonormale de E.

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 $E = F \oplus F^\perp$ (cf 1.1.8). □

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.



3.4.2 Projection orthogonale

Définition 3.4.1 (projection orthogonale sur un sous-espace de dimension finie)
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.

Proposition 3.4.3 (expression de la projection orthogonale, à l'aide d'une famille orthonormale)
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 : $p_F(x) = \displaystyle\sum_{i=1}^{p} (e_i \mid x)\, e_i$.

Cela résulte immédiatement de $E = F \oplus F^\perp$ et de la démonstration de la proposition 3.4.1. □

Exercice 3.16 (⇝ **corrigé**)
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 : $\forall\, x \in E,\ \|x\|^2 = \displaystyle\sum_{k=1}^{n} (e_k \mid x)^2 \quad (\star)$.
Montrer que les vecteurs $(e_k)_{1 \leqslant k \leqslant n}$ forment une base orthonormée de E.

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

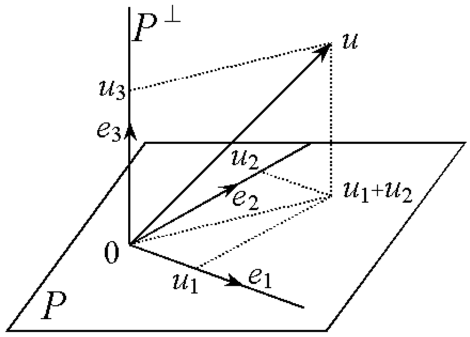
Bien sûr $\mathrm{Id} - p$ est la projection de E sur $F^\perp$ parallèlement à F. Mais on sait que $F = (F^\perp)^\perp$.
Ainsi, au sens de la proposition 3.4.3, $\mathrm{Id} - p$ est la projection orthogonale de E sur $F^\perp$. □

▷ Projection orthogonale sur une droite ou un hyperplan

– Soit $D = \mathbb{R}a$ la droite vectorielle engendrée par un vecteur a unitaire de E (préhilbertien réel).
  La projection orthogonale sur D s'écrit : $p_D(x) = (a \mid x)\, a$ (cas particulier de la prop. 3.4.3).
  Si on suppose seulement que a est non nul, alors $p_D(x) = \left(\dfrac{a}{\|a\|} \,\Big|\, x \right) \dfrac{a}{\|a\|} = \dfrac{(a \mid x)}{\|a\|^2}\, a$.
– On suppose ici que E est euclidien (donc de dimension finie).
  Soit a un vecteur non nul de E, et soit $H = (\mathbb{R}a)^\perp$ l'hyperplan orthogonal à la droite $D = \mathbb{R}a$.
  La projection orthogonale sur H est définie par $p_H(x) = x - \dfrac{(a \mid x)}{\|a\|^2}\, a$.
  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 3.3.1.

Soit E un espace préhilbertien réel. Soit $(e_i)_{1\leqslant i\leqslant p}$ une famille libre de E.
Le procédé de Gram-Schmidt 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 $v_i = \displaystyle\sum_{j=1}^{i-1} (\varepsilon_j \mid e_i)\, \varepsilon_j$.
– 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.

Proposition 3.4.4 (caractérisation des projections orthogonales par l'inégalité de Bessel)
Soit p une projection vectorielle de l'espace euclidien E.
L'application p est une projection orthogonale si et seulement si : $\forall x \in E,\ \|p(x)\| \leqslant \|x\|$.
Dans le cas d'une projection orthogonale, l'égalité $\|p(x)\| \leqslant \|x\|$ s'appelle inégalité de Bessel.

– 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 (cf proposition 3.2.3).
  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.
Pour tout réel λ, on a $p(\lambda y + z) = \lambda y$ et $\|p(\lambda y + z)\| \leqslant \|\lambda y + z\|$ donc $\|\lambda y\|^2 \leqslant \|\lambda y + z\|^2$.
On développe et il reste $2\lambda\,(y \mid z) + \|z\|^2 \geqslant 0$ pour tout réel λ, ce qui implique $(y \mid z) = 0$. □

Définition 3.4.2 (complément : symétrie orthogonale par rapport à un sous-espace de dimension finie)
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.

3.4.3 Distance à un sous-espace de dimension finie

Proposition 3.4.5 (distance d'un vecteur à un sous-espace de dimension finie)
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 ».

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.

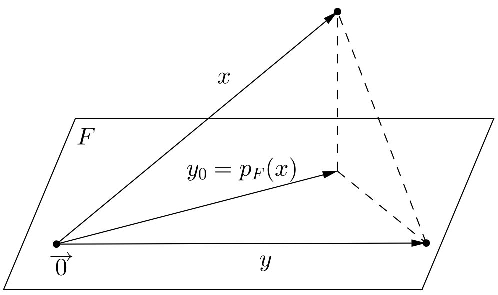
En effet, pour tout vecteur y de F, on a :

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

(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 $d(x,F)^2 = \|x - y_0\|^2 = \|x\|^2 - \|y_0\|^2$. □



Exercice 3.17 (⇝ **corrigé**)
Calculer le minimum de $I_{a,b} = \displaystyle\int_0^1 (t\ln(t) - at - b)^2\,\mathrm{d}t$ et dire pour quels a, b ce minimum est atteint.

▷ 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 :

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

3.5 Formes linéaires sur un espace euclidien

3.5.1 Formes linéaires et équations d'hyperplans

Proposition 3.5.1 (représentation des formes linéaires sur un espace euclidien)
Soit E un espace vectoriel euclidien.
Pour toute forme linéaire f sur E, il existe un unique vecteur a de E tel que : $\forall x \in E,\ f(x) = (a \mid x)$.

Pour tout vecteur a de E, l'application $f_a \colon x \mapsto (a \mid x)$ 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 : $f_a = 0 \Rightarrow f_a(a) = \|a\|^2 = 0 \Rightarrow a = \overrightarrow{0}$.
Du fait que $\dim(E) = \dim(\mathcal{L}(E,\mathbb{R}))$, il en résulte 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 : $\forall x \in E,\ f(x) = (a \mid x)$. □

Exercice 3.18 (⇝ **corrigé**)
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 $\deg(A) = n$.

Exercice 3.19 (⇝ **corrigé**)
(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, $\varphi(P) = (A \mid P)$.

Définition 3.5.1 (vecteur normal à un hyperplan)
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.

Proposition 3.5.2 (équation d'un hyperplan dans une base orthonormale)
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 $a = \displaystyle\sum_{i=1}^{n} a_i e_i$ un vecteur non nul de E. On a $(i) \Leftrightarrow (ii)$:

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

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

– *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 : $x \in H \Leftrightarrow x \in D^\perp \Leftrightarrow (a \mid x) = 0 \Leftrightarrow \sum_{i=1}^{n} a_i x_i = 0$ (voir la proposition 3.3.5).

– *ii)* $\Rightarrow$ *i)* : réciproquement, on suppose qu'une équation de H est $\sum_{i=1}^{n} a_i x_i = 0$ (cf proposition 1.4.4).
Puisque $\mathcal{B}$ est orthonormale, cette équation s'écrit $(a \mid x) = 0$, avec $a = \sum_{i=1}^{n} a_i e_i$ et $x = \sum_{i=1}^{n} x_i e_i$.
Tout x de H est donc orthogonal au vecteur a : en d'autres termes, a est normal à l'hyperplan H. □

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

Proposition 3.5.3 (distance d'un vecteur à un hyperplan vectoriel)
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 : $d(x, H) = |(a \mid x)|$.

On se reportera ici à la note « Projection orthogonale sur une droite ou un hyperplan » dans 3.4.2.
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 le vecteur normal a n'est pas unitaire, la distance de x à H s'écrit : $d(x, H) = \dfrac{1}{\|a\|}\, |(a \mid x)|$

Proposition 3.5.4 (distance d'un vecteur à une droite vectorielle)
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 : $d(x, D) = \|x - (a \mid x)\, a\|$.

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

En effet, c'est juste un cas particulier de la proposition 3.5.3, 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 : $d(u, P) = \dfrac{|\alpha x_0 + \beta y_0 + \gamma z_0|}{\sqrt{\alpha^2 + \beta^2 + \gamma^2}}$.

En effet, c'est juste un cas particulier de la proposition 3.5.3, 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 $d(u, D) = \sqrt{(u \mid e_1)^2 + (u \mid e_2)^2}$.

C'est un cas particulier de la note « Utilisation du supplémentaire orthogonal » dans la sous-section 3.4.3.

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

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 : $\|u \wedge \omega\|^2 = \alpha^2 + \beta^2 = d(u, D)^2$. □

Exercice 3.20 ($\rightsquigarrow$ **corrigé**)

Soit E un espace euclidien de dimension $n \geqslant 1$, muni d'une base orthonormale $\mathcal{B} = (e_k)_{1 \leqslant k \leqslant n}$.

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 $\Delta(\mathcal{U}) = \det(G(\mathcal{U}))$.

a) Avec les notations précédentes, montrer que $G(\mathcal{U}, \mathcal{V}) = U^\top V$, et que $\operatorname{rg}(G(\mathcal{U})) = \operatorname{rg}(\mathcal{U})$.

b) On suppose maintenant que $\mathcal{U}$ est une base d'un sous-espace F de E, avec $1 \leqslant \dim(F) = m \leqslant n$.
Pour tout x de E, on note $\mathcal{U}_x$ la famille obtenue en complétant la famille $\mathcal{U}$ par le vecteur x.
On définit comme précédemment la matrice $G(\mathcal{U}_x)$ (d'ordre $m+1$), et son déterminant $\Delta(\mathcal{U}_x)$.
Montrer que la distance de x à F s'obtient par la formule : $d(x, F)^2 = \dfrac{\Delta(\mathcal{U}_x)}{\Delta(\mathcal{U})}$.