

ESPACES PRÉHILBERTIENS

I. Produit scalaire et norme associée.

Sur ce sujet, presque tout a déjà été dit dans le chapitre “espaces vectoriels normés”, paragraphe **I.3**. Il y aura ici juste quelques petits rajouts.

Dans tout ce paragraphe, E est un espace préhilbertien réel dont le produit scalaire sera noté $(\cdot|\cdot)$, et on notera $\|\cdot\|$ la norme associée, définie par $\forall x \in E \quad \|x\| = \sqrt{(x|x)}$.

On rappelle l'**inégalité de Cauchy-Schwarz** $|(x|y)| \leq \|x\| \|y\|$, et le fait qu'il y a égalité si et seulement si les vecteurs x et y sont colinéaires.

Preuve. Voici une preuve différente de celle vue dans le cours sur les espaces vectoriels normés: Soient $x \in E, y \in E$. Si x ou y est nul, alors $|(x|y)| = \|x\| \|y\| = 0$. Sinon, considérons les vecteurs unitaires associés $u = \frac{x}{\|x\|}$ et $v = \frac{y}{\|y\|}$. On a alors

$$0 \leq \|u - v\|^2 = \|u\|^2 + \|v\|^2 - 2(u|v) = 2(1 - (u|v)) = 2\left(1 - \frac{(x|y)}{\|x\| \|y\|}\right),$$

on en déduit que $(x|y) \leq \|x\| \|y\|$. En faisant le même développement avec $\|u + v\|^2 \geq 0$, on obtient $-(x|y) \leq \|x\| \|y\|$, puis on conclut que $|(x|y)| \leq \|x\| \|y\|$.

Si x est nul ou $y = \lambda x$, alors $|(x|y)| = \|x\| \|y\|$ immédiatement. Réciproquement, si on a cette égalité, alors soit $x = 0$ ou $y = 0$, soit $u - v = 0$ ou $u + v = 0$ en reprenant les notations de la preuve ci-dessus, on en déduit dans tous les cas que x et y sont colinéaires.

On en déduit le cas d'égalité dans l'inégalité triangulaire: l'**égalité** $\|x + y\| = \|x\| + \|y\|$ a lieu si et seulement si les vecteurs x et y sont “positivement liés”, i.e.

$$\|x + y\| = \|x\| + \|y\| \iff (x = 0_E) \quad \text{ou} \quad (\exists \lambda \in \mathbb{R}_+ \quad y = \lambda x).$$

Preuve. Reprenons la démonstration de l'inégalité triangulaire: si $(x, y) \in E^2$,

$$\begin{aligned} \|x + y\|^2 &= (x + y|x + y) = \|x\|^2 + \|y\|^2 + 2(x|y) \\ &\leq \|x\|^2 + \|y\|^2 + 2|(x|y)| \quad \text{(1)} \\ &\leq \|x\|^2 + \|y\|^2 + 2\|x\| \|y\| \quad \text{(2)} \\ &= (\|x\| + \|y\|)^2. \end{aligned}$$

Il y a alors égalité entre les termes extrêmes $\|x + y\|^2$ et $(\|x\| + \|y\|)^2$ si et seulement s'il y a égalité en (1) et en (2). Or, (2) est l'inégalité de Cauchy-Schwarz dans laquelle l'égalité est réalisée si et seulement si x et y sont colinéaires. Et l'égalité (1) est une égalité si et seulement si $(x|y) \geq 0$ ce qui, dans le cas de vecteurs colinéaires, signifie qu'ils sont “de même sens” (au sens large, i.e. l'un des deux vecteurs pouvant être nul).

Remarque. Ce cas d'égalité dans l'inégalité triangulaire concerne uniquement les normes dérivant d'un produit scalaire, ce n'est pas vrai pour une norme quelconque.

Mentionnons maintenant quelques **identités remarquables** sur les normes associées à un produit scalaire. Des développements

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

et

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

on déduit (en ajoutant les deux) l'**identité du parallélogramme**:

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

et, en soustrayant les deux, l'**identité de polarisation**:

$$(x|y) = \frac{1}{4} (\|x + y\|^2 - \|x - y\|^2).$$

Cette dernière permet d'exprimer le produit scalaire de deux vecteurs à l'aide de normes, et aura des conséquences importantes (par exemple, un endomorphisme de E conserve la norme si et seulement s'il conserve le produit scalaire).

Remarque. L'identité du parallélogramme indique que, si $ABCD$ est un parallélogramme, alors la somme des longueurs de ses deux diagonales est égale à la somme des longueurs de ses quatre côtés:

$$AC^2 + BD^2 = AB^2 + BC^2 + CD^2 + DA^2 .$$

II. Orthogonalité.

1. Vecteurs orthogonaux.

Dans un espace préhilbertien E , on dit que deux vecteurs x et y sont **orthogonaux** lorsque leur produit scalaire $(x|y)$ est nul.

On peut caractériser l'orthogonalité de deux vecteurs par le

Théorème de Pythagore. Deux vecteurs x et y de E sont orthogonaux si et seulement si

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2 .$$

Preuve. Il suffit de développer $\|x + y\|^2 = \|x\|^2 + \|y\|^2 + 2(x|y)$, on a donc bien l'équivalence

$$(x|y) = 0 \iff \|x + y\|^2 = \|x\|^2 + \|y\|^2 .$$

Commentaire. On retrouve le théorème de Pythagore et sa réciproque de la géométrie de collège: un triangle ABC est rectangle en A si et seulement si $BC^2 = AB^2 + AC^2$, le carré de l'hypoténuse est égal à la somme des carrés des deux autres côtés.

Définition. Une famille $(x_i)_{i \in I}$ de vecteurs de E est dite **orthogonale** si les vecteurs x_i sont deux à deux orthogonaux, i.e.

$$\forall (i, j) \in I^2 \quad i \neq j \implies (x_i|x_j) = 0 .$$

Proposition. Toute famille (finie) orthogonale de vecteurs non nuls est libre.

Preuve. Soit $\mathcal{X} = (x_1, \dots, x_n)$ une famille orthogonale finie dans E préhilbertien, les vecteurs x_i étant tous non nuls. Supposons que l'on ait $\sum_{i=1}^n \lambda_i x_i = 0_E$, où $\lambda_1, \dots, \lambda_n$ sont des réels. Fixons un indice $j \in \llbracket 1, n \rrbracket$. En faisant le produit scalaire avec x_j de l'égalité ci-dessus, on obtient $(x_j | \sum_{i=1}^n \lambda_i x_i) = (x_j | 0_E) = 0$, soit $\sum_{i=1}^n \lambda_i (x_j | x_i) = 0$, soit $\lambda_j \|x_j\|^2 = 0$, donc $\lambda_j = 0$ puisque le vecteur x_j n'est pas nul.

Proposition. Soit (x_1, \dots, x_n) une famille orthogonale finie dans E préhilbertien. On a alors la **relation de Pythagore**:

$$\left\| \sum_{i=1}^n x_i \right\|^2 = \sum_{i=1}^n \|x_i\|^2 .$$

Preuve. Le développement par bilinéarité du produit scalaire et l'orthogonalité deux à deux des vecteurs donnent immédiatement

$$\left\| \sum_{i=1}^n x_i \right\|^2 = \left(\sum_{i=1}^n x_i \mid \sum_{j=1}^n x_j \right) = \sum_{i=1}^n \sum_{j=1}^n (x_i \mid x_j) = \sum_{i=1}^n \|x_i\|^2 .$$

Définition. On appelle **famille orthonormale** (ou **orthonormée**) de vecteurs toute famille orthogonale constituée de vecteurs unitaires. La famille $(x_i)_{i \in I}$ est donc orthonormale si et seulement si

$$\forall (i, j) \in I^2 \quad (x_i \mid x_j) = \delta_{i,j} .$$

Il résulte de ce qui précède que **toute famille orthonormale finie est libre**.

2. Sous-espaces orthogonaux.

Dans un espace préhilbertien E , on dit que deux sous-espaces vectoriels F et G sont **orthogonaux** lorsque tout vecteur de F est orthogonal à tout vecteur de G , i.e.

$$\forall (y, z) \in F \times G \quad (y \mid z) = 0 .$$

On note parfois $F \perp G$.

Il est immédiat que, si F et G sont orthogonaux, alors $F \cap G = \{0_E\}$, i.e. F et G sont en somme directe. *En effet, si $x \in F \cap G$, alors $(x \mid x) = 0$, donc $x = 0_E$ par le caractère défini du produit scalaire.*

Plus généralement,

Proposition. Si F_1, \dots, F_m sont des s.e.v. deux à deux orthogonaux, alors ils sont en somme directe.

Preuve. En effet, supposons que l'on ait $\sum_{i=1}^m x_i = 0_E$ avec $x_i \in F_i$ pour tout $i \in \llbracket 1, m \rrbracket$, il faut montrer que tous les vecteurs x_i sont nuls. Fixons alors un indice $j \in \llbracket 1, m \rrbracket$, on a

$$\left(x_j \mid \sum_{i=1}^m x_i \right) = (x_j \mid 0_E) = 0 ,$$

mais aussi, par linéarité à droite,

$$\left(x_j \mid \sum_{i=1}^m x_i \right) = \sum_{i=1}^m (x_j \mid x_i) = \|x_j\|^2 .$$

Donc $\|x_j\|^2 = 0$ puis $x_j = 0_E$ pour tout j , ce qu'il fallait prouver. On parle dans cette situation de **somme directe orthogonale**, et le sous-espace somme $S = \bigoplus_{i=1}^m F_i$ peut être noté

$$S = F_1 \overset{\perp}{\oplus} \dots \overset{\perp}{\oplus} F_m = \bigoplus_{1 \leq i \leq m} F_i .$$

3. Orthogonal d'une partie, d'un sous-espace.

Définition 1. Soit X une partie d'un espace préhilbertien E , on dit qu'un vecteur v de E est **orthogonal** à X s'il est orthogonal à tout vecteur de X , i.e. si $\forall x \in X \quad (v \mid x) = 0$.

Définition 2. Si X est une partie de E , on définit l'orthogonal de X , noté X^\perp , comme étant l'ensemble des vecteurs orthogonaux à X . Ainsi, X^\perp est l'ensemble des vecteurs qui sont orthogonaux à tout vecteur de X :

$$X^\perp = \{v \in E \mid \forall x \in X \quad (v|x) = 0\}.$$

Voici quelques propriétés élémentaires:

- X^\perp est un sous-espace vectoriel de E .
En effet, $0_E \in X^\perp$ et, si v et w sont dans X^\perp , si $\lambda \in \mathbb{R}$, alors pour tout $x \in X$, on a
 $(\lambda v + w|x) = \lambda(v|x) + (w|x) = 0$,
donc $\lambda v + w \in X^\perp$.
- $\{0_E\}^\perp = E$ et $E^\perp = \{0_E\}$.
- Si $X \subset Y$, alors $Y^\perp \subset X^\perp$.
- $X \subset (X^\perp)^\perp$.

Cette notion sera plus particulièrement utile lorsque la partie X est un sous-espace vectoriel de E . Notons à ce sujet que:

- Si x_1, \dots, x_n sont des vecteurs de E , alors

$$\{x_1, \dots, x_n\}^\perp = (\text{Vect}(x_1, \dots, x_n))^\perp.$$

Il est évident que $(\text{Vect}(x_1, \dots, x_n))^\perp \subset \{x_1, \dots, x_n\}^\perp$. De plus, la bilinéarité du produit scalaire entraîne que, si un vecteur v est orthogonal à chacun des x_i , alors il est orthogonal à toute combinaison linéaire des x_i , et cela fournit l'inclusion réciproque.

À retenir: pour qu'un vecteur soit orthogonal à un s.e.v. F de dimension finie, il suffit qu'il soit orthogonal à tous les vecteurs d'une famille génératrice de F .

Dans ce qui suit, F et G désignent des sous-espaces de E .

- Le sous-espace F^\perp est orthogonal à F (évident), et en conséquence $F \cap F^\perp = \{0_E\}$.
- $F \subset (F^\perp)^\perp$ (preuve facile)... mais l'inclusion inverse n'est pas toujours vraie, cf. exemple ci-dessous.
- Si F et G sont des s.e.v. tels que $F \subset G$, alors $G^\perp \subset F^\perp$.

Remarque 1. Deux sous-espaces F et G de E sont orthogonaux lorsque tout vecteur de l'un est orthogonal à tout vecteur de l'autre, ce qui signifie que chacun des deux sous-espaces est inclus dans l'orthogonal de l'autre (mais cela ne veut pas dire que $G = F^\perp$!). Donc

$$F \perp G \iff F \subset G^\perp \iff G \subset F^\perp.$$

Remarque 2. Si l'on a toujours $F \cap F^\perp = \{0_E\}$, autrement dit F et F^\perp sont en somme directe, il ne faut pas croire que F et F^\perp sont toujours supplémentaires, voici en effet un exemple où l'on a $F \oplus F^\perp \neq E$. Prenons $E = \mathbb{R}[X]$, le produit scalaire choisi étant $(P, Q) \mapsto (P|Q) = \int_0^1 P(t) Q(t) dt$, le lecteur vérifiera facilement que c'est bien un produit

scalaire. Soit le s.e.v. $H = X \cdot \mathbb{R}[X] = \{P \in E \mid P(0) = 0\}$, alors H est un hyperplan de E puisque $H = \text{Ker } \varphi$, où φ est la forme linéaire non nulle $P \mapsto \varphi(P) = P(0)$. Alors $H^\perp = \{0\}$: en effet, si $P \in H^\perp$, alors $(P|XP) = 0$, soit $\int_0^1 t P(t)^2 dt = 0$, ce qui entraîne que $\forall t \in [0, 1] \quad t P(t)^2 = 0$ (théorème de stricte positivité, la fonction $t \mapsto t P(t)^2$ étant continue et positive sur $[0, 1]$), puis que P est le polynôme nul puisqu'il a une infinité de racines. On voit déjà sur cet exemple que $(H^\perp)^\perp = \{0\}^\perp = E \neq H$, mais aussi que H^\perp n'est pas un supplémentaire de H dans E puisque $H \oplus H^\perp = H \oplus \{0\} = H \neq E$.

III. Un petit détour par les espaces euclidiens.

1. Existence de bases orthonormales.

Théorème. Dans tout espace euclidien E , il existe des bases orthonormales.

Preuve. Par récurrence sur la dimension $n = \dim(E)$.

- Initialisation: si $n = 1$, soit $x \in E$ un vecteur non nul, alors $\omega = \frac{x}{\|x\|}$ est unitaire, et (ω) est une base orthonormale de E .

- Hérité: soit $n \geq 2$, supposons la propriété vraie en dimension $n - 1$, soit E un espace euclidien de dimension n . De nouveau, soit $x \in E$ un vecteur non nul, posons $e_n = \frac{x}{\|x\|}$, soit $D = \text{Vect}(e_n)$ la droite vectorielle engendrée par le vecteur unitaire e_n . L'application $\varphi : E \rightarrow \mathbb{R}$, $y \mapsto (e_n|y)$ est une forme linéaire sur E , non nulle puisque $\varphi(e_n) = \|e_n\|^2 = 1$, son noyau $H = \text{Ker}(\varphi) = D^\perp$ est donc un hyperplan de E , donc $\dim(H) = n - 1$. Si l'on munit H de la restriction du produit scalaire de E , alors H est un espace euclidien de dimension $n - 1$. D'après l'hypothèse de récurrence, il existe dans H une base orthonormale (e_1, \dots, e_{n-1}) . Comme les vecteurs e_i , $1 \leq i \leq n - 1$, sont tous orthogonaux au vecteur e_n , il est immédiat que $(e_1, \dots, e_{n-1}, e_n)$ est une famille orthonormale dans E , donc libre, et c'est donc une base de E puisqu'elle est de cardinal n .

Commentaire. Nous (re)verrons un peu plus loin un procédé algorithmique, le procédé d'orthonormalisation de Gram-Schmidt, permettant la construction effective d'une base orthonormale de E en partant d'une base quelconque.

2. Supplémentaire orthogonal d'un s.e.v.

Proposition. Soit E un espace euclidien de dimension n , soit F un sous-espace vectoriel de E de dimension p . Alors F^\perp est un supplémentaire de F dans E et, en particulier, $\dim(F) = n - p$.

Preuve. D'après le théorème ci-dessus, comme F est un espace euclidien si on le munit de la restriction du produit scalaire de E , on sait qu'il existe dans F une base orthonormale (e_1, \dots, e_p) . Considérons l'application

$$\varphi : \begin{cases} E \rightarrow \mathbb{R}^p \\ x \mapsto ((e_1|x), \dots, (e_p|x)) \end{cases} .$$

Cette application φ est clairement linéaire. Elle est surjective, i.e. $\text{Im}(\varphi) = \mathbb{R}^p$ puisque, si l'on se donne un élément de l'espace d'arrivée $(\lambda_1, \dots, \lambda_p) \in \mathbb{R}^p$, il admet pour antécédent le vecteur $x = \lambda_1 e_1 + \dots + \lambda_p e_p$: on a en effet $(e_i|x) = \lambda_i$ pour tout $i \in [1, p]$. Enfin, il est

immédiat que $\text{Ker}(\varphi) = F^\perp$: en effet, un vecteur de E est orthogonal à F si et seulement s'il est orthogonal à tous les vecteurs d'une famille génératrice de F . En appliquant à φ le théorème du rang, on obtient que

$$\dim(F^\perp) = \dim(\text{Ker}(\varphi)) = \dim(E) - \dim(\text{Im}(\varphi)) = n - p.$$

Comme on sait que $F \cap F^\perp = \{0_E\}$, on déduit que $\dim(F \oplus F^\perp) = p + (n - p) = n = \dim(E)$, donc $F \oplus F^\perp = E$.

Commentaire. On dit, dans cette situation, que F^\perp est **le supplémentaire orthogonal** de F . Je rappelle que tout sous-espace vectoriel F de E (exceptés E et $\{0_E\}$) admet une infinité de supplémentaires, mais parmi tous ces supplémentaires, un seul est orthogonal à F d'où l'usage de l'article défini "le". Nous verrons dans le paragraphe suivant une situation plus générale dans laquelle l'orthogonal d'un sous-espace F est un supplémentaire de F .

Théorème de la base orthonormée incomplète. Si $\mathcal{X} = (x_1, \dots, x_p)$ est une famille orthonormale dans un espace euclidien E de dimension n , alors $p \leq n$, et on peut compléter \mathcal{X} en une base orthonormale $\mathcal{B} = (x_1, \dots, x_p, x_{p+1}, \dots, x_n)$ de E .

Preuve. La famille \mathcal{X} est orthonormale, donc libre, d'où $p \leq n$. Soit le s.e.v. $F = \text{Vect}(\mathcal{X})$ qui est de dimension p , la famille \mathcal{X} est alors une base orthonormale de F . Puis F^\perp est de dimension $n - p$, et il suffit de choisir dans F^\perp une base orthonormale (x_{p+1}, \dots, x_n) pour compléter \mathcal{X} en une b.o.n. de E .

3. Calculs en base orthonormale (BON).

Soit E un espace euclidien de dimension n , et $\mathcal{B} = (e_1, \dots, e_n)$ une base orthonormale de E .

Notons d'abord que, si $x = \sum_{i=1}^n x_i e_i$ est un vecteur de E , alors pour tout $j \in \llbracket 1, n \rrbracket$, on a

$$(e_j | x) = \left(e_j \mid \sum_{i=1}^n x_i e_i \right) = \sum_{i=1}^n x_i (e_i | e_j) = \sum_{i=1}^n x_i \delta_{i,j} = x_j.$$

On peut donc énoncer:

Proposition 1. Les coordonnées d'un vecteur x dans une base orthonormale \mathcal{B} sont les produits scalaires de ce vecteur x avec les vecteurs de la base \mathcal{B} :

$$\forall i \in \llbracket 1, n \rrbracket \quad x_i = (e_i | x).$$

On a donc, pour tout $x \in E$, la décomposition $x = \sum_{i=1}^n (e_i | x) e_i$.

Si $y = \sum_{i=1}^n y_i e_i$ est un autre vecteur de E , on a, par bilinéarité du produit scalaire,

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

On peut donc énoncer:

Proposition 2. Le produit scalaire de deux vecteurs x et y de E est la somme des produits deux à deux des coordonnées des deux vecteurs dans une base orthonormale \mathcal{B} :

$$(x|y) = \sum_{i=1}^n x_i y_i = X^\top Y,$$

en introduisant les matrices-colonnes $X = \text{Mat}_{\mathcal{B}}(x) = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, $Y = \text{Mat}_{\mathcal{B}}(y) = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$.

$$\text{Donc } \|x\| = \sqrt{\sum_{i=1}^n x_i^2} = (X^\top X)^{\frac{1}{2}}.$$

Soit maintenant u un endomorphisme de E , soit $A = (a_{i,j}) = \text{Mat}_{\mathcal{B}}(u)$ sa matrice relativement à une base orthonormale $\mathcal{B} = (e_1, \dots, e_n)$. On a donc $u(e_j) = \sum_{i=1}^n a_{i,j} e_i$ pour tout $j \in \llbracket 1, n \rrbracket$. Le coefficient $a_{i,j}$ est donc la i -ème coordonnée dans la base \mathcal{B} du vecteur $u(e_j)$, ce qui, d'après la **Proposition 1** ci-dessus, permet de l'exprimer comme un produit scalaire, à savoir

$$\forall (i, j) \in \llbracket 1, n \rrbracket^2 \quad a_{i,j} = (e_i | u(e_j)).$$

On en déduit par exemple que $\text{tr}(u) = \text{tr}(A) = \sum_{i=1}^n (e_i | u(e_i))$.

IV. Projection orthogonale sur un sous-espace de dimension finie.

1. Supplémentaire orthogonal.

Voici un résultat qui généralise la proposition du paragraphe **III.2.**:

Théorème. Soit V un sous-espace vectoriel de dimension finie d'un espace E préhilbertien. Alors V^\perp est un supplémentaire de V dans E .

Remarque. On dit parfois que V "admet un supplémentaire orthogonal".

Preuve. Par analyse-synthèse. Comme V , muni de la restriction du produit scalaire de E , est un espace euclidien, il existe dans V des bases orthonormales, choisissons-en une, $\mathcal{B} = (e_1, \dots, e_n)$.

Analyse. Soit $x \in E$, supposons que x admette une décomposition en $x = y + z$ avec $y \in V$ et $z \in V^\perp$. Comme $y \in V$, le vecteur y se décompose dans \mathcal{B} sous la forme $y = \sum_{i=1}^n (e_i | y) e_i$.

Mais, pour tout $i \in \llbracket 1, n \rrbracket$, comme $e_i \in V$ et $z \in V^\perp$, on a $(e_i | y) = (e_i | x) - (e_i | z) = (e_i | x)$.

Finalement, on a explicité y et z en fonction de x : (*)
$$\left\{ \begin{array}{l} y = \sum_{i=1}^n (e_i | x) e_i \\ z = x - \sum_{i=1}^n (e_i | x) e_i \end{array} \right., \text{ ce qui prouve}$$

l'unicité de la décomposition de ce vecteur x .

Synthèse. Soit $x \in E$, soient y et z définis par (*), il est alors clair que $x = y + z$ et que $y \in V$. Enfin, z est orthogonal à chaque vecteur e_j ($1 \leq j \leq n$) puisque

$$(e_j|z) = \left(e_j \left| x - \sum_{i=1}^n (e_i|x) e_i \right. \right) = (e_j|x) - \sum_{i=1}^n (e_i|x) (e_j|e_i) = (e_j|x) - \sum_{i=1}^n (e_i|x) \delta_{i,j} = 0 ,$$

le vecteur z est donc orthogonal à tous les vecteurs d'une famille génératrice de V , on a bien $z \in V^\perp$. La décomposition (*) convient.

Remarque. Dans cette situation, on a $(V^\perp)^\perp = V$. En effet, l'inclusion $V \subset (V^\perp)^\perp$ est toujours vraie. De plus, si $x \in (V^\perp)^\perp$, d'après le théorème ci-dessus, on peut le décomposer en $x = y + z$ avec $y \in V$ et $z \in V^\perp$. En effectuant le produit scalaire avec z , on a $(z|x) = (z|y) + \|z\|^2$, mais les produits scalaires $(z|x)$ et $(z|y)$ sont nuls tous les deux, il reste donc $\|z\|^2 = 0$, soit $z = 0_E$, donc $x = y$ et $x \in V$.

2. Projection orthogonale.

Définition. Soit V un sous-espace de dimension finie d'un espace préhilbertien E . Comme $E = V \oplus V^\perp$, on peut définir le projecteur sur V parallèlement à V^\perp , on l'appelle **projecteur orthogonal sur V** (ou **projection orthogonale sur V**), et on le note p_V . Pour tout $x \in E$, le vecteur $p_V(x)$ est le **projeté orthogonal** de x sur V .

De la preuve du théorème ci-dessus, il résulte que:

Si $\mathcal{B} = (e_1, \dots, e_n)$ est une base orthonormale de V , alors

$$\forall x \in E \quad p_V(x) = \sum_{i=1}^n (e_i|x) e_i .$$

Mentionnons aussi l'**inégalité de Bessel**:

Proposition. Si V est un sous-espace de dimension finie dans E préhilbertien, alors on a

$$\forall x \in E \quad \|p_V(x)\| \leq \|x\| ,$$

avec égalité si et seulement si $x \in V$.

Preuve. On écrit $x = p_V(x) + (x - p_V(x))$, c'est la décomposition du vecteur x selon la somme directe $E = V \oplus V^\perp$. Les vecteurs $p_V(x)$ et $x - p_V(x)$ sont orthogonaux l'un à l'autre, la relation de Pythagore donne alors

$$\|x\|^2 = \|p_V(x)\|^2 + \|x - p_V(x)\|^2 \geq \|p_V(x)\|^2 ,$$

avec égalité ssi $\|x - p_V(x)\| = 0$, i.e. ssi $x = p_V(x)$, i.e. ssi $x \in V$.

Remarque. On peut noter que le produit scalaire d'un vecteur avec son projeté orthogonal est toujours positif, et plus précisément que $(x | p_V(x)) = \|p_V(x)\|^2$. En effet, en décomposant le vecteur x comme ci-dessus et par bilinéarité du produit scalaire,

$$(x | p_V(x)) = (p_V(x) | p_V(x)) + (x - p_V(x) | p_V(x)) = \|p_V(x)\|^2 ,$$

le deuxième terme étant nul.

Autre forme de l'inégalité de Bessel. Soit (e_1, \dots, e_n) une famille orthonormale finie dans un espace préhilbertien E . On a alors

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

avec égalité si et seulement si $x \in \text{Vect}(e_1, \dots, e_n)$.

Preuve. Il suffit effectivement d'appliquer l'inégalité de Bessel prouvée ci-dessus en prenant $V = \text{Vect}(e_1, \dots, e_n)$. Comme $\mathcal{B} = (e_1, \dots, e_n)$ est alors une base orthonormale de V , on a

$$p_V(x) = \sum_{i=1}^n (e_i|x)e_i \quad \text{et} \quad \|p_V(x)\|^2 = \sum_{i=1}^n (e_i|x)^2, \quad \text{c'est donc bien la même inégalité.}$$

Pratique. Pour déterminer le projeté orthogonal d'un vecteur x de E sur un s.e.v. V de dimension finie, on peut:

- soit appliquer la formule $p_V(x) = \sum_{i=1}^n (e_i|x) e_i$ si l'on connaît une base **orthonormale** $\mathcal{B} = (e_1, \dots, e_n)$ de V ,
- soit obtenir les coordonnées de $p_V(x)$ dans une base quelconque de V en annulant des produits scalaires, cette méthode est détaillée ci-dessous.

Soit donc $(\varepsilon_1, \dots, \varepsilon_n)$ une base quelconque de V , soit $x \in E$, soit $y = p_V(x)$. Ce vecteur y est caractérisé par les conditions $\begin{cases} y \in V \\ x - y \in V^\perp \end{cases}$, soit $\begin{cases} \text{(1)} : y \in V \\ \text{(2)} : \forall i \in \llbracket 1, n \rrbracket \quad (\varepsilon_i|x - y) = 0 \end{cases}$.

La condition (1) nous amène à rechercher y sous la forme $y = \sum_{j=1}^n y_j \varepsilon_j$, avec les y_j réels.

La condition (2) se traduit alors par le système linéaire

$$\text{(S)} : \quad \forall i \in \llbracket 1, n \rrbracket \quad \sum_{j=1}^n (\varepsilon_i|\varepsilon_j) y_j = (\varepsilon_i|x)$$

de n équations à n inconnues. La matrice de ce système (S) est la **matrice de Gram** $G = ((\varepsilon_i|\varepsilon_j))_{1 \leq i, j \leq n} \in \mathcal{M}_n(\mathbb{R})$, soit

$$G = \begin{pmatrix} \|\varepsilon_1\|^2 & (\varepsilon_1|\varepsilon_2) & \cdots & (\varepsilon_1|\varepsilon_n) \\ (\varepsilon_2|\varepsilon_1) & \|\varepsilon_2\|^2 & \cdots & (\varepsilon_2|\varepsilon_n) \\ \vdots & \vdots & \ddots & \vdots \\ (\varepsilon_n|\varepsilon_1) & (\varepsilon_n|\varepsilon_2) & \cdots & \|\varepsilon_n\|^2 \end{pmatrix}.$$

Cette matrice symétrique est inversible. *En effet, supposons qu'une combinaison linéaire des colonnes C_1, \dots, C_n de cette matrice est nulle, soit $\sum_{j=1}^n \lambda_j C_j = 0$ (dans $\mathcal{M}_{n,1}(\mathbb{R}) \simeq \mathbb{R}^n$).*

Alors, en regardant coordonnée par coordonnée, on a $\forall i \in \llbracket 1, n \rrbracket \quad \sum_{j=1}^n \lambda_j (\varepsilon_i|\varepsilon_j) = 0$, soit $\forall i \in \llbracket 1, n \rrbracket \quad (\varepsilon_i | \sum_{j=1}^n \lambda_j \varepsilon_j) = 0$, le vecteur $\sum_{j=1}^n \lambda_j \varepsilon_j$, qui est dans V vu son écriture, est

aussi dans V^\perp , donc est nul, ce qui entraîne la nullité de tous les réels λ_j puisque la famille $(\varepsilon_1, \dots, \varepsilon_n)$ est libre. Les colonnes de la matrice G sont donc linéairement indépendantes, donc $\text{rg}(G) = n$, et G est inversible. Le système (S) est donc un système de Cramer, il admet une solution unique.

Exercice. On munit $E = \mathbb{R}[X]$ du produit scalaire $(P, Q) \mapsto (P|Q) = \int_0^1 P(t) Q(t) dt$. Déterminer le projeté orthogonal de $P = X^2$ sur le plan $V = \mathbb{R}_1[X] = \text{Vect}(1, X)$.

3. Distance à un sous-espace de dimension finie.

Dans ce paragraphe, V est toujours un sous-espace de dimension finie d'un espace préhilbertien réel E . Si x est un vecteur quelconque de E , son projeté orthogonal $p_V(x)$ est alors le vecteur de V "le plus proche" de x . Plus précisément, on a la

Proposition. Si $x \in E$, son projeté $p_V(x)$ est l'unique vecteur y_0 de V tel que

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

Preuve. Si $y \in V$, alors $x - y = (x - p_V(x)) + (p_V(x) - y)$, avec $x - p_V(x) \in V^\perp$ et $p_V(x) - y \in V$, ces deux vecteurs sont donc orthogonaux, d'où la relation de Pythagore:

$$\|x - y\|^2 = \|x - p_V(x)\|^2 + \|p_V(x) - y\|^2 \geq \|x - p_V(x)\|^2 ,$$

avec égalité si et seulement si $y = p_V(x)$.

Cette distance minimale du vecteur x fixé à un vecteur y décrivant le s.e.v. V est appelée **distance de x à V** , et notée $d(x, V)$. On a ainsi

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

Calcul pratique. Si $\mathcal{B} = (e_1, \dots, e_n)$ est une base orthonormale de V , on a

$$d(x, V)^2 = \|x - p_V(x)\|^2 = \|x\|^2 - \|p_V(x)\|^2 = \|x\|^2 - \sum_{i=1}^n (e_i|x)^2 ,$$

l'égalité du milieu résultant de Pythagore puisque les vecteurs $p_V(x)$ et $x - p_V(x)$ sont orthogonaux.

4. Une application: la régression linéaire (droite des moindres carrés).

Dans le plan euclidien rapporté à un repère orthonormal, on donne n points $M_1 \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, \dots, M_n \begin{pmatrix} x_n \\ y_n \end{pmatrix}$ et une droite affine \mathcal{D} d'équation cartésienne $y = ax + b$.

On projette verticalement les points M_i sur la droite \mathcal{D} en des points $H_i \begin{pmatrix} x_i \\ ax_i + b \end{pmatrix}$.

On suppose que les points M_i n'ont pas tous la même abscisse. L'objectif est de déterminer la droite \mathcal{D} de façon que la somme $\delta(a, b) = \sum_{i=1}^n M_i H_i^2$ soit minimale.

Pour formaliser cette recherche en termes de projection orthogonale sur un sous-espace d'un espace euclidien, on se place maintenant dans l'espace $E = \mathbb{R}^n$ muni du produit scalaire canonique noté $(\cdot | \cdot)$ et de la norme associée notée $\|\cdot\|$, et on introduit les vecteurs $\vec{x} = (x_1, \dots, x_n)$, $\vec{y} = (y_1, \dots, y_n)$ et $\vec{u} = (1, \dots, 1)$. On introduit enfin les nombres

$$X = \sum_{i=1}^n x_i \quad , \quad Y = \sum_{i=1}^n y_i \quad , \quad R = \sum_{i=1}^n x_i^2 \quad , \quad S = \sum_{i=1}^n x_i y_i \quad , \quad T = \sum_{i=1}^n y_i^2 .$$

On observe tout d'abord que $\delta(a, b) = \sum_{i=1}^n (y_i - (ax_i + b))^2 = \|\vec{y} - (a\vec{x} + b\vec{u})\|^2$.

Lorsque le couple (a, b) décrit \mathbb{R}^2 , le vecteur $a\vec{x} + b\vec{u}$ décrit le plan $P = \text{Vect}(\vec{x}, \vec{u})$ engendré par les vecteurs \vec{x} et \vec{u} . C'est bien un plan car, les x_i n'étant pas tous égaux par hypothèse, les vecteurs \vec{x} et \vec{u} de \mathbb{R}^n ne sont pas colinéaires. On sait que, lorsqu'un vecteur \vec{z} décrit le plan P , la distance $d(\vec{y}, \vec{z}) = \|\vec{y} - \vec{z}\|$ est minimale lorsque \vec{z} est le projeté orthogonal du vecteur \vec{y} sur le plan P . Or, un vecteur $\vec{z} = a\vec{x} + b\vec{u}$ de P est le projeté orthogonal de \vec{y} sur P si et seulement si $\vec{y} - \vec{z} \in P^\perp$, c'est-à-dire si et seulement si le vecteur $\vec{y} - \vec{z}$ est orthogonal aux deux vecteurs \vec{x} et \vec{u} qui engendrent le plan P , autrement dit si et seulement si les deux produits scalaires $(\vec{y} - (a\vec{x} + b\vec{u}) | \vec{x})$ et $(\vec{y} - (a\vec{x} + b\vec{u}) | \vec{u})$ sont nuls.

En développant les produits scalaires par linéarité à gauche, on obtient les deux équations permettant d'obtenir a et b pour que $\delta(a, b)$ soit minimal:

$$\begin{cases} \|\vec{x}\|^2 a + (\vec{u} | \vec{x}) b = (\vec{x} | \vec{y}) \\ (\vec{x} | \vec{u}) a + \|\vec{u}\|^2 b = (\vec{u} | \vec{y}) \end{cases} , \quad \text{soit} \quad \begin{cases} R a + X b = S \\ X a + n b = Y \end{cases} .$$

Le déterminant de ce système est non nul puisque, par Cauchy-Schwarz,

$$nR - X^2 = \|\vec{u}\|^2 \|\vec{x}\|^2 - (\vec{u} | \vec{x})^2 \geq 0$$

et l'inégalité est stricte puisque les vecteurs \vec{u} et \vec{x} sont non colinéaires. On a donc un système de Cramer, admettant une unique solution. La résolution fournit les formules

$$a = \frac{nS - XY}{nR - X^2} \quad \text{et} \quad b = \frac{RY - SX}{nR - X^2} .$$

V. Algorithme d'orthonormalisation de Gram-Schmidt.

Il s'agit ici, dans un espace préhilbertien, d'orthonormaliser une famille libre, c'est-à-dire de construire une famille orthonormale vérifiant certaines conditions. On va prouver l'unicité de cette construction, et donner pour cela un procédé algorithmique. Énonçons donc le

Théorème de Gram-Schmidt. Soit E un espace préhilbertien, soit $\mathcal{X} = (x_1, \dots, x_n)$ une famille libre finie de vecteurs de E , de cardinal n . Alors il existe une unique famille $\mathcal{E} = (e_1, \dots, e_n)$ de n vecteurs de E telle que:

(C1): la famille \mathcal{E} est orthonormale ;

(C2): pour tout $k \in \llbracket 1, n \rrbracket$, $\text{Vect}(e_1, \dots, e_k) = \text{Vect}(x_1, \dots, x_k)$;

(C3): pour tout $k \in \llbracket 1, n \rrbracket$, $(e_k | x_k) > 0$.

Définition. Cette famille \mathcal{E} est appelée l'orthonormalisée de la famille \mathcal{X} .

Preuve de l'existence. Pour tout $k \in \llbracket 1, n \rrbracket$, on posera $V_k = \text{Vect}(x_1, \dots, x_k)$. Comme la famille (x_1, \dots, x_k) est libre et de cardinal k , on a $\dim(V_k) = k$.

On construit les vecteurs e_k , avec $1 \leq k \leq n$, de la façon suivante:

- on pose d'abord $v_1 = x_1$ et $e_1 = \frac{v_1}{\|v_1\|} = \frac{x_1}{\|x_1\|}$;

- pour k de 2 à n , on note v_k la différence entre x_k et son projeté orthogonal sur V_{k-1} , soit $v_k = x_k - p_{V_{k-1}}(x_k)$, puis on "norme" ce vecteur v_k en introduisant $e_k = \frac{v_k}{\|v_k\|}$.

Tout d'abord ceci a bien un sens: en effet, le vecteur x_1 est non nul puisqu'il est extrait d'une famille libre ; ensuite, pour tout $k \in \llbracket 2, n \rrbracket$, le vecteur v_k est non nul car s'il était nul cela signifierait que $x_k \in V_{k-1}$ et la famille (x_1, \dots, x_k) serait liée ce qui est absurde.

Les vecteurs e_k sont tous unitaires par construction et, pour $k \in \llbracket 2, n \rrbracket$, on a $v_k \in V_{k-1}^\perp$ par construction. Il est par ailleurs immédiat que $v_k \in V_k$ pour tout $k \in \llbracket 1, n \rrbracket$. Donc, si $1 \leq i < j \leq n$, les vecteurs v_i et v_j sont orthogonaux car $v_i \in V_i$ alors que $v_j \in V_{j-1}^\perp \subset V_i^\perp$, cette inclusion résultant de l'inclusion $V_i \subset V_{j-1}$ car $i \leq j-1$. La famille (v_1, \dots, v_n) est donc orthogonale, et la famille $\mathcal{E} = (e_1, \dots, e_n)$ est orthonormale.

On a vu que, pour tout $k \in \llbracket 1, n \rrbracket$, on a $e_k \in V_k$, d'où immédiatement $\text{Vect}(e_1, \dots, e_k) \subset V_k$. La famille (e_1, \dots, e_k) étant libre car orthonormale, les sous-espaces $\text{Vect}(e_1, \dots, e_k)$ et $V_k = \text{Vect}(x_1, \dots, x_k)$ sont tous deux de dimension k , donc ils sont égaux.

Enfin, $(e_1 | x_1) = \|x_1\| > 0$ et, pour tout $k \in \llbracket 2, n \rrbracket$, en posant $w_k = p_{V_{k-1}}(x_k)$, on a $x_k = v_k + w_k$, les deux vecteurs étant orthogonaux, donc

$$(v_k | x_k) = (v_k | v_k + w_k) = \|v_k\|^2 > 0,$$

l'inégalité étant stricte puisque $x_k \notin V_{k-1}$, puis $(e_k | x_k) = \frac{(v_k | x_k)}{\|v_k\|} = \|v_k\| > 0$.

L'unicité de la famille \mathcal{E} sera prouvée ultérieurement.

À retenir. Procédé algorithmique. On pose d'abord $e_1 = \frac{x_1}{\|x_1\|}$ puis, pour k de 2 à n , on construit d'abord le vecteur v_k (différence entre x_k et son projeté orthogonal sur V_{k-1}) par $v_k = x_k - p_{V_{k-1}}(x_k) = x_k - \sum_{i=1}^{k-1} (e_i|x_k)e_i$, et on norme enfin ce vecteur: $e_k = \frac{v_k}{\|v_k\|}$.

Conséquence. Dans un espace euclidien de dimension n , on dispose d'un procédé permettant de construire une base orthonormale $\mathcal{E} = (e_1, \dots, e_n)$ en partant d'une base quelconque $\mathcal{X} = (x_1, \dots, x_n)$. Cette base "orthonormalisée" vérifiera en outre les propriétés

$$(\mathbf{C2}) : \quad \forall k \in \llbracket 1, n \rrbracket \quad \text{Vect}(e_1, \dots, e_k) = \text{Vect}(x_1, \dots, x_k)$$

et

$$(\mathbf{C3}) : \quad \forall k \in \llbracket 1, n \rrbracket \quad (e_k|x_k) > 0.$$

Si l'on note $R = P_{\mathcal{E}, \mathcal{X}} = (r_{i,j})$ la matrice de passage de la base orthonormalisée \mathcal{E} vers la base initiale \mathcal{X} , on a alors $x_j = \sum_{i=1}^n r_{i,j}e_i$ pour tout j . La condition **(C2)** montre que $x_j \in \text{Vect}(e_1, \dots, e_j)$ pour tout j , donc $r_{i,j}$ est nul dès que $i > j$, la matrice R est donc triangulaire supérieure. D'autre part, $r_{i,j}$ est la i -ième coordonnée du vecteur x_j dans la base orthonormale \mathcal{E} , donc $r_{i,j} = (e_i|x_j)$. La condition **(C3)** montre donc que les coefficients diagonaux $r_{i,i}$ de R sont strictement positifs. Notons que $\det(R) = \prod_{i=1}^n r_{i,i} > 0$, donc les bases \mathcal{X} et \mathcal{E} sont "de même sens", nous y reviendrons.

VI. Formes linéaires sur un espace euclidien.

1. Théorème de représentation de Riesz.

Théorème. Soit φ une forme linéaire sur un espace euclidien E . Alors il existe un unique vecteur a de E tel que

$$\forall x \in E \quad \varphi(x) = (a|x).$$

Preuve. Pour tout $a \in E$, soit $\varphi_a : E \rightarrow \mathbb{R}$ définie par $x \mapsto (a|x)$. De la linéarité à droite du produit scalaire, il résulte que φ_a est une forme linéaire sur E , autrement dit φ_a appartient à

$$E^* = \mathcal{L}(E, \mathbb{R}) \text{ espace dual de } E. \text{ Considérons maintenant l'application } \Phi : \begin{cases} E \rightarrow E^* \\ a \mapsto \varphi_a \end{cases}.$$

Il résulte maintenant de la linéarité à gauche du produit scalaire que Φ est une application linéaire. Elle est injective car, si $a \in \text{Ker}(\Phi)$, alors φ_a est la forme linéaire nulle, et en particulier $\varphi_a(a) = \|a\|^2 = 0$ donc $a = 0_E$. Comme Φ est linéaire et injective entre deux espaces vectoriels de même dimension finie, elle est donc bijective, c'est un isomorphisme, on a donc

$$\forall \varphi \in E^* \quad \exists ! a \in E \quad \varphi = \Phi(a) = \varphi_a,$$

et c'est ce que dit l'énoncé du théorème de Riesz.

2. Hyperplans d'un espace euclidien. Vecteur normal.

Si H est un hyperplan d'un espace euclidien E , alors $D = H^\perp$ est une droite vectorielle. Tout vecteur **non nul** a de D est appelé **vecteur normal** à l'hyperplan H . On a alors $H = D^\perp = (\text{Vect}(a))^\perp$, et $H = \text{Ker}(\varphi_a)$, où φ_a est la forme linéaire $x \mapsto (a|x)$ mentionnée dans le paragraphe précédent. Si $\mathcal{B} = (e_1, \dots, e_n)$ est une base orthonormale de E , si $a = \sum_{i=1}^n a_i e_i$ est un vecteur normal à H , alors une équation cartésienne de l'hyperplan H relativement à la base \mathcal{B} est: $a_1 x_1 + \dots + a_n x_n = 0$, qui traduit l'orthogonalité des vecteurs a et $x = \sum_{i=1}^n x_i e_i$.

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

Soit H un hyperplan de E euclidien, soit a un vecteur normal à H , soit $x \in E$. On a alors $d(x, H) = \|x - p_H(x)\|$. Or, le vecteur $x - p_H(x)$ est orthogonal (ou "normal") à l'hyperplan H , il est donc colinéaire à a : il existe λ réel tel que $x - p_H(x) = \lambda a$. Ce réel λ est déterminé par le fait que $x + \lambda a = p_H(x)$ appartient à H , donc $(a | x + \lambda a) = 0$, donc $\lambda = -\frac{(a|x)}{\|a\|^2}$.

On a donc $p_H(x) = x - \frac{(a|x)}{\|a\|^2} a$, et enfin

$$d(x, H) = \|x - p_H(x)\| = \frac{|(a|x)|}{\|a\|}, \quad \text{avec } a \text{ vecteur normal à } H.$$

Remarque. On peut retrouver l'expression de $p_H(x)$ en observant que $p_H(x) = x - p_D(x)$, et le projeté orthogonal de x sur D s'exprime par $p_D(x) = \left(\frac{a}{\|a\|} \mid x\right) \frac{a}{\|a\|} = \frac{(a|x)}{\|a\|^2} a$, puisque $\left(\frac{a}{\|a\|}\right)$ est une base orthonormale de la droite D . Donc

$$p_H(x) = x - \frac{(a|x)}{\|a\|^2} a.$$

Puis, par la relation de Pythagore, $d(x, D)^2 = \|x\|^2 - d(x, H)^2 = \|x\|^2 - \frac{(a|x)^2}{\|a\|^2}$.

Si l'hyperplan H a pour équation cartésienne $a_1 x_1 + \dots + a_n x_n = 0$ dans une base orthonormale $\mathcal{B} = (e_1, \dots, e_n)$, si $x = \sum_{i=1}^n x_i e_i$, alors $a = \sum_{i=1}^n a_i e_i$ est un vecteur normal à H , et on obtient

$$d(x, H) = \frac{|a_1 x_1 + \dots + a_n x_n|}{\sqrt{a_1^2 + \dots + a_n^2}}.$$