Produits scalaires

1 Espaces préhilbertiens réels

1.1 Produit scalaire réel

Définition. Soit E un espace vectoriel réel. Un *produit scalaire sur* E est une forme bilinéaire symétrique définie positive sur E, c'est-à-dire une application

$$\varphi : \mathrm{E} \times \mathrm{E} \to \mathbb{R}$$

qui vérifie les propriétés suivantes
- pour tout x dans E, l'application $y \mapsto \varphi(x, y)$ est linéaire ;
- pour tout y dans E, l'application $x \mapsto \varphi(x, y)$ est linéaire ;
- $\forall (x, y) \in \mathrm{E}^2$, $\varphi(x, y) = \varphi(y, x)$;
- $\forall x \in \mathrm{E}$, $\varphi(x, x) \geqslant 0$;
- $\forall x \in \mathrm{E}$, $((\varphi(x, x) = 0) \Rightarrow x = 0_{\mathrm{E}})$.

Un espace préhilbertien réel est un espace vectoriel réel muni d'un produit scalaire.
Un espace euclidien est un espace préhilbertien réel de dimension finie.

On peut remarquer qu'après avoir vérifié le caractère symétrique (troisième point), chacun des deux premiers points (linéarité à droite ou à gauche) entraîne l'autre.

On peut aussi remarquer que le dernier point demande seulement de vérifier une implication. La réciproque découle de la bilinéarité, comme le montre le calcul suivant

$$\varphi(0_{\mathrm{E}}, 0_{\mathrm{E}}) = \varphi(0_{\mathrm{E}}, 0.0_{\mathrm{E}}) = 0 \times \varphi(0_{\mathrm{E}}, 0_{\mathrm{E}}) = 0.$$

1.2 Exemples de produits scalaires

Exemple 1. Sur $\mathbb{R}^n$, le *produit scalaire canonique* est défini comme suit. Étant donné deux vecteurs

$$x = (x_1, \ldots, x_n) \qquad \text{et} \qquad y = (y_1, \ldots, y_n),$$

leur produit scalaire vaut

$$(x|y) = \sum_{k=1}^{n} x_k y_k.$$

Exemple 2. Sur $\mathcal{M}_{n,p}(\mathbb{R})$, le *produit scalaire canonique* est défini comme suit. Étant donné deux matrices

$$\mathrm{A} = (a_{i,j})_{1 \leqslant i \leqslant n,\ 1 \leqslant j \leqslant p} \qquad \text{et} \qquad \mathrm{B} = (b_{i,j})_{1 \leqslant i \leqslant n,\ 1 \leqslant j \leqslant p}$$

de $\mathcal{M}_{n,p}(\mathbb{R})$, leur produit scalaire vaut

$$(\mathrm{A}|\mathrm{B}) = \sum_{i=1}^{n} \sum_{j=1}^{p} a_{i,j} b_{i,j}.$$

Propriété. La produit scalaire canonique de $\mathcal{M}_{n,p}(\mathbb{R})$ est donné par la formule

$$(\mathrm{A}|\mathrm{B}) = \mathrm{tr}(\mathrm{A}^{\mathrm{T}} \cdot \mathrm{B}).$$

Exemple 3. Notons E l'espace vectoriel réel $\mathcal{C}^0([a, b], \mathbb{R})$. Un produit scalaire classique sur E est défini par

$$\forall (f, g) \in \mathrm{E}^2, \qquad (f|g) = \int_a^b f(t)g(t)\ \mathrm{d}t.$$

Le caractère défini positif découle du fait qu'une fonction continue, positive, d'intégrale nulle sur un segment non trivial est identiquement nulle sur ce segment.

Exemple 4. Un produit scalaire sur $\mathbb{R}[X]$ est défini par

$$\forall (P, Q) \in (\mathbb{R}[X])^2, \qquad (P|Q) = \int_0^\pi P(\cos(t))Q(\cos(t))\ dt.$$

La seule vérification non immédiate est celle du caractère défini positif.

Prenons P tel que (P|P) soit nul. La fonction $t \mapsto (P(\cos(t)))^2$ est alors une fonction continue et positive sur $[0, 1]$, d'intégrale nulle, donc cette fonction est identiquement nulle. On en déduit que le polynôme P admet pour racines tous les éléments du segment $[-1, 1]$ car ce sont les valeurs prises par $\cos(t)$ quand t décrit le segment $[0, \pi]$. Le polynôme P admet donc une infinité de racines. C'est donc le polynôme nul.

Exemple 5. Un produit scalaire sur $\mathbb{R}_n[X]$ est défini par

$$\forall (P, Q) \in (\mathbb{R}_n[X])^2, \qquad (P|Q) = \sum_{k=0}^{n} P(k)Q(k).$$

Là encore, seul le caractère défini positif mérite une attention particulière.

Prenons P dans $\mathbb{R}_n[X]$ tel que (P|P) soit nul. On obtient alors

$$\sum_{k=0}^{n} P(k)^2 = 0.$$

Cette somme est nulle et tous ses termes sont positifs donc tous ses termes sont nuls. Le polynôme P admet donc pour racines les nombres $0, 1, \ldots, n$. C'est un polynôme de $\mathbb{R}_n[X]$ qui admet au moins $(n + 1)$ racines donc c'est le polynôme nul.

Exemple 6. Soit I un intervalle de $\mathbb{R}$ non trivial. Notons E(I) l'ensemble des fonctions continues sur I, à valeurs réelles, dont le carré est intégrable sur I. Cet ensemble est alors un sous-espace vectoriel de $\mathcal{C}^0(I, \mathbb{R})$ et l'application

$$(f, g) \mapsto \int_I f(t)g(t)\ dt$$

est un produit scalaire sur E(I). Les vérifications sont les mêmes que dans l'exemple 3.

Exemple 7. Les applications

$$(P, Q) \mapsto \int_0^{+\infty} P(t)Q(t)e^{-t}\ dt, \qquad (P, Q) \mapsto \int_{-\infty}^{+\infty} P(t)Q(t)e^{-t^2/2}\ dt, \qquad (P, Q) \mapsto \int_{-1}^{1} \frac{P(t)Q(t)}{\sqrt{1-t^2}}\ dt$$

sont des produits scalaires sur l'espace vectoriel $\mathbb{R}[X]$. Le caractère défini positif se justifie de la même manière que dans l'exemple 4. Il y a cependant un travail supplémentaire, à savoir démontrer que ces intégrales existent bien.

Exemple 8. L'application

$$(P, Q) \mapsto \sum_{n=0}^{+\infty} \frac{P(n)Q(n)}{2^n}$$

est un produit scalaire sur $\mathbb{R}[X]$. Là encore, il faut justifier l'existence de la somme écrite (la règle de d'Alembert fonctionne bien). Ensuite, les vérifications sont directes, à l'exception du caractère défini positif. Si (P|P) est nul, alors pour tout n dans $\mathbb{N}$, le nombre $P(n)^2/2^n$ est nul donc P possède une infinité de racines donc c'est le polynôme nul.

1.3 Inégalité de Cauchy-Schwarz

Théorème. Soit E un espace préhilbertien réel. Pour tout couple (x, y) de vecteurs de E, l'inégalité suivante est alors valable.

$$(x|y)^2 \leqslant (x|x)(y|y).$$

De plus, cet inégalité est une égalité si, et seulement si, les vecteurs x et y sont colinéaires.

Démonstration. Prenons un couple (x, y) de vecteurs de E. Remarquons d'abord que si x est nul, alors l'inégalité est vraie — c'est même une égalité dans ce cas. Supposons donc que x n'est pas nul. Considérons la fonction

$$\begin{array}{rccc} f : & \mathbb{R} & \to & \mathbb{R} \\ & t & \mapsto & (tx + y|tx + y). \end{array}$$

La bilinéarité et le caractère symétrique du produit scalaire permettent d'écrire

$$\forall t \in \mathbb{R}, \qquad f(t) = (x|x)t^2 + 2(x|y)t + (y|y).$$

Comme x n'est pas le vecteur nul, le nombre $(x|x)$ n'est pas nul. La fonction f est donc polynomiale de degré 2. Le caractère positif du produit scalaire donne maintenant

$$\forall t \in \mathbb{R}, \qquad f(t) \geqslant 0.$$

Ainsi, la fonction f est une fonction polynomiale de degré 2 qui ne change pas de signe. Son discriminant est donc forcément négatif. Or son discriminant vaut

$$4(x|y)^2 - 4(x|x)(y|y).$$

Le fait que ce discriminant soit négatif s'écrit

$$(x|y)^2 \leqslant (x|x)(y|y).$$

L'inégalité de Cauchy-Schwarz est démontrée.

Supposons maintenant que l'égalité est vraie. Si x est nul, alors les vecteurs x et y sont colinéaires. Supposons donc maintenant que x n'est pas le vecteur nul de E. Reprenons le cheminement du raisonnement précédent. Le fait que l'égalité

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

soit vraie signifie que le discriminant de f est nul. Par conséquent, ce polynôme possède une racine t_0, ce qui se réécrit

$$(t_0x + y|t_0x + y) = 0.$$

Le caractère défini positif du produit scalaire donne alors $t_0x + y = 0_{\mathrm{E}}$. Les vecteurs x et y sont donc colinéaires.

On a montré que le cas d'égalité implique que les vecteurs x et y soient colinéaires.

Réciproquement, supposons que x et y sont colinéaires. Sans perte de généralité, supposons que y est proportionnel à x et notons λ un facteur de proportionnalité. On obtient alors

$$(x|y)^2 = (x|\lambda x)^2 = \lambda^2(x|x)^2 \qquad \text{et} \qquad (x|x)(y|y) = (x|x)(\lambda x|\lambda x) = \lambda^2(x|x)^2 \qquad \text{donc} \qquad (x|y)^2 = (x|x)(y|y).$$

Le théorème est maintenant complètement démontré. □

Étudions maintenant l'écriture de cette inégalité dans le cadre des produits scalaires proposés précédemment.

Exemple 1. Prenons $(x_1, \ldots, x_n)$ et $(y_1, \ldots, y_n)$ dans $\mathbb{R}^n$. L'inégalité de Cauchy-Schwarz s'écrit

$$\left(\sum_{k=1}^{n} x_k y_k\right)^2 \leqslant \left(\sum_{k=1}^{n} (x_k)^2\right)\left(\sum_{k=1}^{n} (y_k)^2\right).$$

Remarque. Une autre manière de démontrer cette inégalité (et le cas d'égalité qui lui est attaché) est de prouver la *formule de Lagrange*

$$\left(\sum_{k=1}^{n} (x_k)^2\right)\left(\sum_{k=1}^{n} (y_k)^2\right) - \left(\sum_{k=1}^{n} x_k y_k\right)^2 = \sum_{1 \leqslant i < j \leqslant n} (x_i y_j - x_j y_i)^2.$$

Exemple 2. Prenons deux fonctions f et g dans $\mathcal{C}^0([a,b],\mathbb{R})$. L'inégalité de Cauchy-Schwarz donne

$$\left(\int_a^b f(t)g(t)\ \mathrm{d}t\right)^2 \leqslant \left(\int_a^b f(t)^2\ \mathrm{d}t\right)\left(\int_a^b g(t)^2\ \mathrm{d}t\right).$$

Remarque. Cette inégalité peut se démontrer à partir de la précédente à l'aide de sommes de Riemann.

Définition. Soit E un espace préhilbertien réel. Pour tout vecteur x de E, on note

$$||x|| = \sqrt{(x|x)}.$$

La fonction $||\ ||$ est *la norme euclidienne associée au produit scalaire* $(\ |\ )$. C'est effectivement une norme. La seule vérification non immédiate est l'inégalité triangulaire.

Propriété (inégalité de Minkowski). Soit E un espace préhilbertien. Pour tout couple (x, y) de vecteurs de E, l'inégalité

$$||x + y|| \leqslant ||x|| + ||y||$$

est valable.

Démonstration de l'inégalité de Minkowski. Prenons x et y quelconques dans E. En développant le produit scalaire par bilinéarité, on obtient

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

On applique ensuite l'inégalité de Cauchy-Schwarz

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

Comme les nombres $||x + y||$ et $||x|| + ||y||$ sont positifs, la croissance de la racine carrée donne

$$||x + y|| \leqslant ||x|| + ||y||. \qquad \heartsuit$$

1.4 Complément : identités de polarisation

Soient x et y deux vecteurs d'un espace préhilbertien E. En partant de la formule

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

on isole $(x|y)$. On obtient ainsi la première identité de polarisation

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

Une formule plus symétrique s'obtient en écrivant

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

En soustrayant cette formule à la première du paragraphe, on obtient la deuxième identité de polarisation

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

1.5 Vecteurs unitaires

Un *vecteur unitaire* est un vecteur de norme 1. Étant donné un vecteur x non nul, il existe exactement deux vecteurs unitaires proportionnels à x, à savoir

$$\frac{x}{||x||} \qquad \text{et} \qquad -\frac{x}{||x||}.$$

L'ensemble des vecteurs unitaires d'un espace préhilbertien s'appelle *la sphère unité* (la même notion existe évidemment plus généralement dans les espaces vectoriels normés).

2 Orthogonalité

2.1 Orthogonalité de deux vecteurs

Définition. Deux vecteurs x et y d'un espace préhilbertien E sont *orthogonaux* si, et seulement si, leur produit scalaire $(x|y)$ est nul.

Exemple 1. Dans $\mathbb{R}^3$ muni de son produit scalaire canonique, les vecteurs $(1, 2, 1)$ et $(1, -1, 1)$ sont orthogonaux.

Exemple 2. Dans $\mathcal{M}_{2,3}(\mathbb{R})$ muni de son produit scalaire canonique, les matrices

$$\begin{pmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \end{pmatrix} \qquad \text{et} \qquad \begin{pmatrix} 0 & 5 & 0 \\ 2 & 0 & 2 \end{pmatrix}$$

sont orthogonales entre elles.

Exemple 3. Dans $\mathcal{C}^0([0, 2\pi], \mathbb{R})$ muni du produit scalaire présenté au premier paragraphe, les fonctions sinus et cosinus sont orthogonales entre elles.

Formule de Pythagore. Soient $x_1, \ldots, x_p$ des vecteurs de E deux à deux orthogonaux.

$$||x_1 + \cdots + x_p||^2 = ||x_1||^2 + \cdots + ||x_p||^2.$$

2.2 Orthogonalité de deux sous-espaces vectoriels

Définition. Soient F et G deux sous-espaces vectoriels d'un espace préhilbertien E. Ces deux sous-espaces sont *orthogonaux* si, et seulement si, tout vecteur de F est orthogonal à tout vecteur de G. En formule, cette définition s'écrit

$$\mathrm{F} \perp \mathrm{G} \iff \forall (x, y) \in \mathrm{F} \times \mathrm{G}, \quad (x|y) = 0.$$

Propriété. Soient F et G deux sous-espaces vectoriels d'un espace préhilbertien E. On suppose que F et G sont orthogonaux. Ces deux espaces ont alors une intersection triviale.

Démonstration de la propriété. Soit $x \in \mathrm{F} \cap \mathrm{G}$. Le couple (x, x) est un élément de $\mathrm{F} \times \mathrm{G}$ donc le produit scalaire $(x|x)$ est nul, donc x est le vecteur nul. ♡

Corollaire. Si F et G sont deux sous-espaces vectoriels orthogonaux entre eux d'un espace euclidien E, alors ils vérifient l'inégalité

$$\dim(\mathrm{F}) + \dim(\mathrm{G}) \leqslant \dim(\mathrm{E}).$$

Question. Pourtant, il me semblait qu'il y avait des plans orthogonaux entre eux dans l'espace de dimension 3. N'y a-t-il pas quelque chose de contradictoire là-dedans ?

Réponse. Il y a en fait deux notions d'orthogonalité en concurrence ici. Celle utilisée en géométrie dans l'espace n'est pas compatible avec celle présentée dans ce chapitre. En fait, on dit plutôt que deux plans sont *perpendiculaires* dans ce cas.

Plus généralement, on peut définir la perpendicularité dans n'importe quel espace euclidien. Si F et G sont deux sous-espaces vectoriels d'un espace euclidien E, on dit qu'ils sont *perpendiculaires* si, et seulement si, l'orthogonal de $\mathrm{F} \cap \mathrm{G}$ dans F et l'orthogonal de $\mathrm{F} \cap \mathrm{G}$ dans G sont orthogonaux entre eux.

Cette notion ne figure toutefois pas dans le programme de PC.

Exemple 1. Dans $\mathbb{R}^3$ muni de son produit scalaire canonique, la droite engendrée par le vecteur $(1, 1, 0)$ et la droite engendrée par le vecteur $(0, 0, 1)$ sont orthogonales entre elles.

Exemple 2. Dans $\mathcal{M}_n(\mathbb{R})$ muni de son produit scalaire canonique, les sous-espaces vectoriels

$$\mathcal{S}_n(\mathbb{R}) = \{\mathrm{M} \in \mathcal{M}_n(\mathbb{R}) \ ; \ \mathrm{M}^{\mathrm{T}} = \mathrm{M}\} \qquad \text{et} \qquad \mathcal{A}_n(\mathbb{R}) = \{\mathrm{M} \in \mathcal{M}_n(\mathbb{R}) \ ; \ \mathrm{M}^{\mathrm{T}} = -\mathrm{M}\}$$

sont orthogonaux entre eux.

Démonstration. Prenons S dans $\mathcal{S}_n(\mathbb{R})$ et A dans $\mathcal{A}_n(\mathbb{R})$.

$$(\mathrm{S}|\mathrm{A}) = \mathrm{tr}(\mathrm{S}^{\mathrm{T}} \cdot \mathrm{A}) = \mathrm{tr}(\mathrm{SA}) \qquad \text{et} \qquad (\mathrm{S}|\mathrm{A}) = (\mathrm{A}|\mathrm{S}) = \mathrm{tr}(\mathrm{A}^{\mathrm{T}} \cdot \mathrm{S}) = \mathrm{tr}(-\mathrm{AS}) = -\mathrm{tr}(\mathrm{AS}) = -\mathrm{tr}(\mathrm{SA}).$$

Le nombre (S|A) est égal à son opposé donc il est nul.

Les espaces $\mathcal{S}_n(\mathbb{R})$ et $\mathcal{A}_n(\mathbb{R})$ sont donc orthogonaux entre eux. ♡

Exemple 3. Dans $\mathcal{C}^0([-1,1],\mathbb{R})$ muni du produit scalaire

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

les sous-espaces vectoriels formés des fonctions paires et des fonctions impaires sont orthogonaux entre eux.

2.3 Orthogonal d'une partie d'un espace préhilbertien

Définition. Soit A une partie d'un espace préhilbertien E. *L'orthogonal de* A est l'ensemble de tous les vecteurs de E qui sont orthogonaux à tous les éléments de A. En formule, cette définition s'écrit

$$\mathrm{A}^{\perp} = \{x \in \mathrm{E}\ ;\ \forall a \in \mathrm{A}, \quad (a|x) = 0\}.$$

Exemples de base.

$$\mathrm{E}^{\perp} = \{0_{\mathrm{E}}\} \qquad \text{et} \qquad \{0_{\mathrm{E}}\}^{\perp} = \mathrm{E}.$$

Propriété. Quelle que soit la partie A de E, l'ensemble $\mathrm{A}^{\perp}$ est un sous-espace vectoriel de E.

Propriété. Soient A et B deux parties de E vérifiant l'inclusion $\mathrm{A} \subset \mathrm{B}$.
Leurs orthogonaux vérifient alors l'inclusion $\mathrm{B}^{\perp} \subset \mathrm{A}^{\perp}$.

Propriété. Soit A une partie de E. L'orthogonal de A est égal à celui de Vect(A).

Remarque. Étant donné deux sous-espaces vectoriels F et G de E. Les trois propriétés suivantes sont alors équivalentes.

$$\mathrm{F} \perp \mathrm{G}, \qquad \mathrm{F} \subset \mathrm{G}^{\perp}, \qquad \mathrm{G} \subset \mathrm{F}^{\perp}.$$

Propriété. Si F et G sont orthogonaux entre eux et sont supplémentaires dans E, alors ils sont l'orthogonal l'un de l'autre.

Démonstration de la propriété. On connaît déjà les inclusions

$$\mathrm{F} \subset \mathrm{G}^{\perp} \qquad \text{et} \qquad \mathrm{G} \subset \mathrm{F}^{\perp}.$$

Prenons maintenant x dans $\mathrm{G}^{\perp}$. Comme F et G sont supplémentaires dans E, il existe un unique couple $(x_{\mathrm{F}}, x_{\mathrm{G}})$ de $\mathrm{F} \times \mathrm{G}$ vérifiant l'égalité $x = x_{\mathrm{F}} + x_{\mathrm{G}}$.
Les vecteurs x et x_{F} sont dans $\mathrm{G}^{\perp}$ donc x_{G} également. Ce vecteur est aussi dans G donc il est nul.
Il reste $x = x_{\mathrm{F}}$. Le vecteur x est donc dans F. L'inclusion $\mathrm{G}^{\perp} \subset \mathrm{F}$ est donc vraie.

On en déduit l'égalité $\mathrm{F} = \mathrm{G}^{\perp}$ par double inclusion.
L'égalité $\mathrm{G} = \mathrm{F}^{\perp}$ se démontre pareillement. ♡

Exemple. Dans $\mathcal{M}_n(\mathbb{R})$ muni de son produit scalaire canonique, les sous-espaces vectoriels $\mathcal{S}_n(\mathbb{R})$ et $\mathcal{A}_n(\mathbb{R})$ sont orthogonaux et supplémentaires donc chacun des deux est l'orthogonal de l'autre.

La réciproque est fausse. Il arrive que F et $\mathrm{F}^{\perp}$ ne soient pas supplémentaires dans E. Prenons l'exemple de l'espace vectoriel $\mathrm{E} = \mathcal{C}^0([0,1],\mathbb{R})$ muni du produit scalaire

$$(f,g) \mapsto \int_0^1 f(t)g(t)\ \mathrm{d}t.$$

Considérons le sous-espace vectoriel

$$\mathrm{F} = \{f \in \mathrm{E}\ ;\ f(0) = 0\}.$$

Montrons que $\mathrm{F}^\perp$ est trivial. Prenons g dans $\mathrm{F}^\perp$. La fonction $t \mapsto tg(t)$ est dans F donc elle est orthogonale à g.

$$\int_0^1 tg(t)^2 \, \mathrm{d}t = 0.$$

La fonction $t \mapsto tg(t)^2$ est continue et positive sur $[0,1]$, d'intégrale nulle, donc elle est identiquement nulle. On obtient donc

$$\forall t \in [0,1], \quad tg(t)^2 = 0, \qquad \text{puis} \qquad \forall t \in]0,1], \quad g(t) = 0.$$

Comme la fonction g est continue sur $[0,1]$, elle est en fait identiquement nulle sur $[0,1]$.

On a bien montré que $\mathrm{F}^\perp$ est le sous-espace vectoriel trivial $\{0_\mathrm{E}\}$. Comme F n'est pas égal à E (il manque par exemple la fonction constante égale à 1), les sous-espaces F et $\mathrm{F}^\perp$ ne sont pas supplémentaires dans E.

Cependant. On verra plus loin que si F est de dimension finie, alors F et $\mathrm{F}^\perp$ sont supplémentaires dans E.

2.4 Projection orthogonale, symétrie orthogonale

Définition. Une *projection orthogonale* est une projection dont les axes sont orthogonaux entre eux.

De même, une *symétrie orthogonale* est une symétrie dont les axes sont orthogonaux entre eux.

En particulier, si F est un sous-espace vectoriel de E tel que F et $\mathrm{F}^\perp$ soient supplémentaires, alors la projection sur F parallèlement à $\mathrm{F}^\perp$ est une projection orthogonale. On l'appelle *la projection orthogonale sur* F.

Exemple. Dans $\mathcal{M}_n(\mathbb{R})$ muni de son produit scalaire canonique, l'endomorphisme $\mathrm{M} \mapsto \mathrm{M}^\mathrm{T}$ est la symétrie orthogonale par rapport au sous-espace $\mathcal{S}_n(\mathbb{R})$.

Le projecteur orthogonal sur $\mathcal{S}_n(\mathbb{R})$ est l'endomorphisme $\mathrm{M} \mapsto \frac{1}{2}(\mathrm{M} + \mathrm{M}^\mathrm{T})$.

2.5 Orthogonal d'une droite vectorielle

Propriété. Soit D une droite de E. Son orthogonal $\mathrm{D}^\perp$ est alors un hyperplan de E. De plus, les espaces D et $\mathrm{D}^\perp$ sont supplémentaires dans E. Enfin, si a est un vecteur directeur de la droite D, alors le projecteur orthogonal sur D est donné par

$$x \mapsto \frac{(a|x)}{(a|a)} a.$$

Démonstration de la propriété. Soit x dans E. Un calcul direct donne

$$(x - \frac{(a|x)}{(a|a)} a | a) = (x|a) - \frac{(a|x)}{(a|a)} (a|a) = 0.$$

Le vecteur $x - \frac{(a|x)}{(a|a)} a$ est orthogonal à a donc il est dans $\mathrm{D}^\perp$. L'écriture

$$x = \frac{(a|x)}{(a|a)} a + \left(x - \frac{(a|x)}{(a|a)} a \right)$$

montre que x est dans $\mathrm{F} + \mathrm{F}^\perp$. C'est vrai pour tout x de E donc la somme $\mathrm{F} + \mathrm{F}^\perp$ vaut E. On sait déjà que F et $\mathrm{F}^\perp$ sont en somme directe donc ils sont supplémentaires dans E.

De plus, la décomposition obtenue ci-dessus montre que le projecteur orthogonal sur D est donné par la formule

$$x \mapsto \frac{(a|x)}{(a|a)} a. \quad \heartsuit$$

Exemple. Dans $\mathbb{R}^n$, notons h le vecteur $(1, \ldots, 1)$. L'orthogonal de la droite $\mathrm{Vect}(h)$ est alors l'ensemble

$$(\mathrm{Vect}(h))^\perp = \{(x_1, \ldots, x_n) \; ; \; x_1 + \cdots + x_n = 0.\}$$

Une base orthogonale de cet hyperplan est formée des vecteurs suivants

$$(1, -1, 0, \ldots, 0), \quad (1, 1, -2, 0, \ldots, 0), \quad (1, 1, 1, -3, 0, \ldots, 0), \quad , \ldots, \quad (1, \ldots, 1, -n+1).$$

2.6 Complément : orthogonal de l'image d'une matrice

Ce n'est pas un résultat du cours mais un exercice classique : pour toute matrice A de $\mathcal{M}_{n,p}(\mathbb{R})$, l'orthogonal de Im(A) dans $\mathcal{M}_{n,1}(\mathbb{R})$ est $\mathrm{Ker}(\mathrm{A}^{\mathrm{T}})$.

Démonstration. Notons $\mathrm{A}_1, \ldots, \mathrm{A}_p$ les colonnes de A. On sait que ces vecteurs colonnes engendrent Im(A) donc l'orthogonal de Im(A) est aussi l'orthogonal de l'ensemble $\{\mathrm{A}_1, \ldots, \mathrm{A}_p\}$. Soit U un vecteur de $\mathcal{M}_{n,1}(\mathbb{R})$. On voit que l'appartenance de U à $(\mathrm{Im}(\mathrm{A}))^\perp$ équivaut à

$$(\mathrm{A}_1)^{\mathrm{T}} \cdot \mathrm{U} = 0, \qquad \ldots, \qquad (\mathrm{A}_p)^{\mathrm{T}} \cdot \mathrm{U} = 0.$$

Maintenant, remarquons que $(\mathrm{A}_1)^{\mathrm{T}}, \ldots, (\mathrm{A}_p)^{\mathrm{T}}$ sont les lignes de la matrice A^{T}. On en déduit le calcul

$$\mathrm{A}^{\mathrm{T}} \cdot \mathrm{U} = \begin{pmatrix} (\mathrm{A}_1)^{\mathrm{T}} \cdot \mathrm{U} \\ \vdots \\ (\mathrm{A}_p)^{\mathrm{T}} \cdot \mathrm{U} \end{pmatrix}.$$

Ainsi, l'appartenance de U à $(\mathrm{Im}(\mathrm{A}))^\perp$ équivaut à l'égalité $\mathrm{A}^{\mathrm{T}} \cdot \mathrm{U} = 0$, qui elle-même équivaut à l'appartenance de U à $\mathrm{Ker}(\mathrm{A}^{\mathrm{T}})$. ♡

3 Bases orthogonales

3.1 Familles orthogonales

Définition. Une *famille orthogonale* d'un espace préhilbertien E est une famille $(x_i)_{i\in\mathrm{I}}$ dont les vecteurs sont deux à deux orthogonaux. En formule, cela s'écrit

$$\forall (i,j) \in \mathrm{I}^2, \qquad ((i \neq j) \Rightarrow (x_i|x_j) = 0).$$

Une *famille orthonormale* d'un espace préhilbertien est une famille orthogonale dont tous les vecteurs sont de norme 1. En formule, cela s'écrit

$$\forall (i,j) \in \mathrm{I}^2, \qquad (x_i|x_j) = \delta_{i,j}.$$

Propriété. Tout famille orthogonale de vecteurs **non nuls** est libre.
En particulier, toute famille orthonormale est libre.

Démonstration de la propriété. Contentons-nous du cas des familles finies. Soit $(x_1, \ldots, x_n)$ une famille orthogonale finie dont tous les vecteurs sont non nuls.

Soit $(\lambda_1, \ldots, \lambda_n) \in \mathbb{R}^n$ vérifiant l'égalité

$$\sum_{k=1}^{n} \lambda_k x_k = 0_{\mathrm{E}}.$$

Prenons j dans $[\![1,n]\!]$ et effectuons le produit scalaire du vecteur nul avec x_j. La bilinéarité du produit scalaire donne

$$\sum_{k=1}^{n} \lambda_k (x_j|x_k) = 0.$$

Dans cette somme, tous les termes pour $k \neq j$ sont nuls. Il reste

$$\lambda_j (x_j|x_j) = 0.$$

Comme x_j n'est pas le vecteur nul, le nombre $(x_j|x_j)$ est non nul. Il reste $\lambda_j = 0$.
C'est vrai pour tout indice j dans $[\![1,n]\!]$ donc la famille $(x_1, \ldots, x_n)$ est libre. ♡

Exemple. On prend $\mathrm{E} = \mathcal{C}^0([0,2\pi],\mathbb{R})$. Pour tout k dans $\mathbb{N}$, on introduit les fonctions

$$c_k : x \mapsto \cos(kx) \qquad \text{et} \qquad s_k : x \mapsto \sin(kx).$$

La concaténée des familles $(c_k)_{k\geqslant 0}$ et $(s_k)_{k\geqslant 1}$ est alors une famille orthogonale pour le produit scalaire

$$(f,g) \mapsto \int_0^{2\pi} f(t)g(t)\ \mathrm{d}t.$$

En particulier, cette famille est libre.

3.2 Bases orthogonales

Proposition. Soit F un sous-espace vectoriel de E qui possède une base orthogonale finie $\mathcal{D} = (d_1, \dots, d_p)$.
Alors, pour tout vecteur x de F, la décomposition de x dans cette base est

$$x = \sum_{k=1}^{p} \frac{(d_k|x)}{(d_k|d_k)} d_k.$$

Démonstration de la proposition. Soit x dans F. Introduisons sa décomposition dans la base $\mathcal{D}$

$$x = \sum_{k=1}^{p} x_k d_k.$$

Prenons j dans $[\![1, p]\!]$ et effectuons le produit scalaire de x avec d_j.

$$(d_j|x) = \sum_{k=1}^{p} x_k (d_j|d_k).$$

Dans cette somme, tous les termes d'indices $k \neq j$ sont nuls car la famille $\mathcal{D}$ est orthogonale. Il reste donc

$$(d_j|x) = x_j (d_j|d_j).$$

Le vecteur d_j n'est pas nul car il fait partie d'une base de F. On trouve donc $x_j = (d_j|x)/(d_j|d_j)$.
Cette relation est valable pour tout j dans $[\![1, p]\!]$. On en déduit la formule demandée. ♡

Proposition. Soit F un sous-espace vectoriel de E qui possède une base orthogonale finie $\mathcal{D} = (d_1, \dots, d_p)$.
Alors on peut affirmer que F et $F^\perp$ sont supplémentaires. De plus, la projection orthogonale sur F est l'application

$$p_F : x \mapsto \sum_{k=1}^{p} \frac{(d_k|x)}{(d_k|d_k)} d_k.$$

Démonstration de la proposition. Considérons l'endomorphisme p_F de E défini dans l'énoncé de la proposition.
Soit x dans E. Montrons que $x - p_F(x)$ est dans $F^\perp$. Pour cela, prenons y quelconque dans F et effectuons le produit scalaire

$$(x - p_F(x)|y) = (x|y) - \sum_{k=1}^{p} \frac{(d_k|x)}{(d_k|d_k)} (d_k|y).$$

Pour exprimer le produit scalaire $(x|y)$, faisons intervenir la décomposition de y donnée dans la proposition précédente

$$y = \sum_{k=1}^{p} \frac{(d_k|y)}{(d_k|d_k)} d_k.$$

On trouve donc

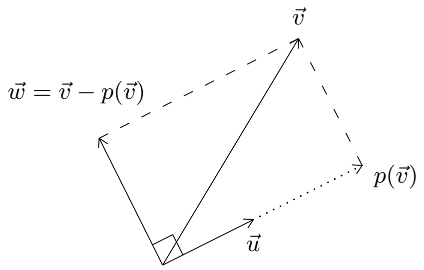
$$(x|y) = \sum_{k=1}^{p} \frac{(d_k|y)}{(d_k|d_k)} (x|d_k) \qquad \text{puis} \qquad (x - p_F(x)|y) = 0.$$

Ainsi, le vecteur $x - p_F(x)$ est un élément de $F^\perp$. Le vecteur $p_F(x)$ est un élément de F. Le vecteur x est donc la somme d'un vecteur de F et d'un vecteur de $F^\perp$.
C'est vrai pour tout vecteur x de E donc la somme $F + F^\perp$ vaut E tout entier. On sait déjà que cette somme est directe donc F et $F^\perp$ sont supplémentaires dans E.
On a montré au passage que l'unique décomposition d'un vecteur x dans la somme $F \oplus F^\perp$ est $p_F(x) + (x - p_F(x))$. Cela prouve que p_F est le projecteur sur F parallèlement à $F^\perp$. ♡

Remarque. Dans l'état actuel des connaissances, il est nécessaire de supposer que F possède une base orthogonale. L'existence d'une telle base va toutefois être démontrée, au paragraphe suivant, pour tout sous-espace vectoriel de dimension finie. Il sera donc ensuite possible d'énoncer à nouveau ces propriétés en raccourcissant les hypothèses.

3.3 Orthogonalisation de Gram-Schmidt

Commençons par une remarque géométrique simple. Dans le plan usuel, si on se donne deux vecteurs non colinéaires $\vec{u}$ et $\vec{v}$, il est possible de « redresser » la base $(\vec{u}, \vec{v})$ en une base orthogonale $(\vec{u}, \vec{w})$ du plan en ôtant à $\vec{v}$ son projeté orthogonal sur la droite engendrée par $\vec{u}$.



Ce projeté orthogonal, comme on l'a vu au paragraphe 2.6, est le vecteur $\frac{(u|v)}{(u|u)}u$.

Le procédé de Gram-Schmidt consiste à partir d'une famille libre quelconque et à « redresser » ses vecteurs les uns après les autres afin de rendre les angles droits.

Plus précisément, on part d'une famille finie $(u_0, \ldots, u_n)$ de vecteurs de E et on suppose que cette famille est libre. Pour tout k dans $[\![0, n]\!]$, on note

$$\mathrm{F}_k = \mathrm{Vect}(u_0, \ldots, u_k).$$

On construit par récurrence une famille orthogonale $(v_0, \ldots, v_n)$ vérifiant la propriété

$$\forall k \in [\![0, n]\!], \qquad \mathrm{Vect}(v_0, \ldots, v_k) = \mathrm{F}_k.$$

Pour cela, on prend $v_0 = u_0$ puis, si k est un indice dans $[\![1, n]\!]$ pour lequel les vecteurs $v_0, \ldots, v_{k-1}$ ont été construits, on pose

$$v_k = u_k - p_{\mathrm{F}_{k-1}}(u_k).$$

Notons que cette définition a bien un sens. En effet, à ce stade du raisonnement, la famille $(v_0, \ldots, v_{k-1})$ est une base orthogonale de F_{k-1} ; il est donc possible d'utiliser le raisonnement du paragraphe précédent. En particulier, le vecteur v_k est donné par

$$v_k = u_k - \sum_{j=0}^{k-1} \frac{(v_j|u_k)}{(v_j|v_j)} v_j.$$

Théorème. Tout sous-espace vectoriel de dimension finie d'un espace préhilbertien réel possède une base orthogonale.

Démonstration. Soit F un tel sous-espace vectoriel. Le cours d'algèbre linéaire nous enseigne que F possède au moins une base [1]. En appliquant le procédé de Gram-Schmidt, on construit alors une base orthogonale de F. ♡

Corollaire 1. Tout espace euclidien possède une base orthogonale.

Tout espace euclidien possède même une base orthonormale, obtenue en divisant chaque vecteur d'une base orthonormale par sa norme.

Corollaire 2. Soit F un sous-espace vectoriel de dimension finie d'un espace préhilbertien E. Les sous-espaces vectoriels F et $\mathrm{F}^\perp$ sont alors supplémentaires dans E.

Enfin, la projection orthogonale sur F peut être exprimée par la formule vue au paragraphe 3.2 relativement à n'importe quelle base orthogonale de F.

Corollaire 3. Soit F un sous-espace vectoriel d'un espace euclidien E. La formule suivante est alors valable

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

1. C'est un corollaire du *théorème de la base extraite*.

Théorème (orthonormalisation de Gram-Schmidt). Soit $(u_0, \ldots, u_n)$ une famille libre d'un espace préhilbertien E. Pour tout k dans $[\![0, n]\!]$, on note

$$\mathrm{F}_k = \mathrm{Vect}(u_0, \ldots, u_k).$$

Il existe alors une unique famille orthonormale $(e_0, \ldots, e_n)$ de E telle que
(1) pour tout $k \in [\![0, n]\!]$, la famille $(e_0, \ldots, e_k)$ soit une base orthonormale de F_k ;
(2) pour tout $k \in [\![0, n]\!]$, le produit scalaire $(e_k|u_k)$ soit strictement positif.

Démonstration du théorème. *Existence.* Considérons la famille $(v_0, \ldots, v_n)$ définie au début du paragraphe et posons

$$\forall k \in [\![0, n]\!], \quad e_k = \frac{v_k}{||v_k||}.$$

On obtient bien alors une famille orthonormale de E qui vérifie l'hypothèse (1).

De plus, le calcul donne

$$(e_0|u_0) = \frac{(u_0|u_0)}{||u_0||} = ||u_0|| > 0.$$

Soit $k \in [\![1, n]\!]$. Rappelons l'égalité $v_k = u_k - p_{\mathrm{F}_{k-1}}(u_k)$. Le vecteur $p_{\mathrm{F}_{k-1}}(u_k)$ est dans F_{k-1} et v_k est dans $\mathrm{F}_{k-1}^{\perp}$, de même que e_k donc

$$(e_k|u_k) = (e_k|v_k) = \frac{(v_k|v_k)}{||v_k||} = ||v_k|| > 0.$$

L'hypothèse (2) est également vérifiée.

Unicité. Soit $(f_0, \ldots, f_n)$ une famille orthonormale de E vérifiant les hypothèses (1) et (2).

Le vecteur f_0 est un vecteur unitaire de $\mathrm{Vect}(u_0)$ donc il vaut $u_0/||u_0||$ ou son opposé mais la condition $(f_0|u_0) > 0$ impose que f_0 soit égal à $u_0/||u_0||$ donc $f_0 = e_0$.

Soit $k \in [\![1, n]\!]$. La décomposition de f_k dans la base orthonormale $(e_0, \ldots, e_k)$ de F_k est

$$f_k = \sum_{j=0}^{k} (e_j|f_k) e_j.$$

Les vecteurs $e_0, \ldots, e_{k-1}$ sont dans F_{k-1} donc ce sont des combinaisons linéaires de $f_0, \ldots, f_{k-1}$. Ces vecteurs sont donc tous orthogonaux à f_k. Il reste donc

$$f_k = (e_k|f_k) e_k.$$

Les vecteurs f_k et e_k sont de norme 1 donc $(e_k|f_k)$ vaut 1 ou -1. En effectuant le produit scalaire avec u_k, il vient

$$(f_k|u_k) = (e_k|f_k)(e_k|u_k) \qquad \text{puis} \qquad (e_k|f_k) = \frac{(f_k|u_k)}{(e_k|u_k)} > 0$$

donc $(e_k|f_k) = 1$ donc $f_k = e_k$. L'unicité est démontrée. ♡

4 Distance d'un vecteur à un sous-espace vectoriel

4.1 Définition

Soit E un espace préhilbertien réel. Soit F un sous-espace vectoriel de E. Soit x un vecteur de E.
La distance de x à F est le nombre

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

Commençons par remarquer que l'ensemble considéré est une partie non vide de $\mathbb{R}$ qui est minorée par 0. Elle possède donc bien une borne inférieure.
Notons ensuite que cette borne inférieure n'est a priori pas un minimum.

4.2 Lien avec la projection orthogonale

Proposition. Soit F un sous-espace vectoriel d'un espace préhilbertien réel E. On suppose que F et $F^\perp$ sont supplémentaires dans E.

Alors, pour tout vecteur x de E, le nombre $d(x, \mathrm{F})$ est égal à la distance de x à son projeté orthogonal sur F.

$$d(x, \mathrm{F}) = ||x - p_\mathrm{F}(x)||.$$

Ainsi, la borne inférieure qui définit $d(x, \mathrm{F})$ est un minimum. De plus, le vecteur $p_\mathrm{F}(x)$ est le seul élément de F qui réalise ce minimum.

Démonstration de la proposition. Soit x dans E. Soit y dans F. Remarquons la décomposition

$$x - y = \underbrace{x - p_\mathrm{F}(x)}_{\in \mathrm{F}^\perp} + \underbrace{p_\mathrm{F}(x) - y}_{\in \mathrm{F}}.$$

La formule de Pythagore donne donc

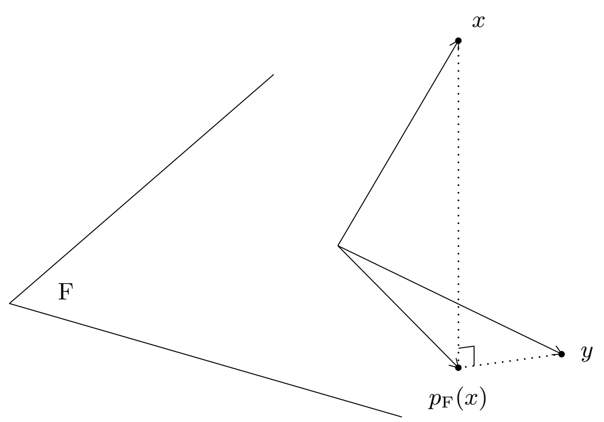
$$||x - y||^2 = ||x - p_\mathrm{F}(x)||^2 + ||p_\mathrm{F}(x) - y||^2.$$

On obtient donc l'inégalité

$$\forall y \in \mathrm{F}, \qquad ||x - y|| \geqslant ||x - p_\mathrm{F}(x)||.$$

Cette propriété montre que $||x - p_\mathrm{F}(x)||$ est le minimum des nombres de la forme $||x - y||$ quand y parcourt F. Ce nombre est donc la distance de x à F.

Par ailleurs, l'égalité $||x - y|| = d(x, \mathrm{F})$ équivaut à $||p_\mathrm{F}(x) - y|| = 0$, c'est-à-dire à $y = p_\mathrm{F}(x)$. ♡



Remarque. Les vecteurs $p_\mathrm{F}(x)$ et $x - p_\mathrm{F}(x)$ sont orthogonaux, ce qui donne, par application du théorème de Pythagore, la relation

$$||x - p_\mathrm{F}(x)||^2 = ||x||^2 - ||p_\mathrm{F}(x)||^2.$$

Dans les exercices pratiques, ça simplifie le calcul du minimum.

4.3 Détermination d'un projeté orthogonal

Problématique. On suppose qu'on connaît une base $(f_1, \ldots, f_n)$ quelconque de F et on cherche à exprimer le projeté orthogonal de x sur F.

Première méthode. On construit une base orthogonale $(e_1, \ldots, e_n)$ de F par la méthode de Gram-Schmidt puis on applique la formule du paragraphe 3.2 pour obtenir $p_\mathrm{F}(x)$.

L'avantage de cette méthode est que tout se fait algorithmiquement. L'inconvénient est que ça donne les coordonnées de $p_\mathrm{F}(x)$ dans la nouvelle base et non pas dans celle qui est donnée initialement.

Deuxième méthode. On introduit les coordonnées de $p_{\mathrm{F}}(x)$ dans la base $(f_1, \dots, f_n)$.

$$p_{\mathrm{F}}(x) = \sum_{k=1}^{n} y_k f_k.$$

On sait que le vecteur $x - p_{\mathrm{F}}(x)$ est dans $\mathrm{F}^{\perp}$ donc il est orthogonal à $f_1, \dots, f_n$, ce qui donne

$$\forall j \in [\![1, n]\!], \quad \sum_{k=1}^{n} y_k (f_k | f_j) = (x | f_j).$$

Il suffit alors de résoudre ce système pour obtenir les y_k.

L'avantage de cette méthode est qu'elle dispense de déterminer une nouvelle base. L'inconvénient est qu'elle requiert une inversion de matrice.

5 Représentations matricielles

5.1 Représentation des vecteurs

On considère un espace euclidien E dont on note le produit scalaire $(\ |\ )$.
On considère une base orthonormale $\mathcal{E} = (e_1, \dots, e_n)$ de cet espace euclidien.

Soit $x \in \mathrm{E}$. On sait que la décomposition du vecteur x dans la base orthonormale $\mathcal{E}$ est la suivante

$$x = \sum_{k=1}^{n} (e_k | x) e_k.$$

On sait aussi que pour tout couple (x, y) de vecteurs de E, le produit scalaire $(x|y)$ s'exprime comme suit

$$(x|y) = \sum_{k=1}^{n} (e_k | x)(e_k | y).$$

On peut remarquer que si l'on pose $\mathrm{X} = \mathrm{M}_{\mathcal{E}}(x)$ et $\mathrm{Y} = \mathrm{M}_{\mathcal{E}}(y)$, alors les relations suivantes sont valables

$$\mathrm{X} = \begin{pmatrix} (e_1|x) \\ \vdots \\ (e_n|x) \end{pmatrix}, \quad \mathrm{Y} = \begin{pmatrix} (e_1|y) \\ \vdots \\ (e_n|y) \end{pmatrix}, \qquad (x|y) = \mathrm{X}^{\mathrm{T}} \cdot \mathrm{Y}.$$

Techniquement, la notation $\mathrm{X}^{\mathrm{T}} \cdot \mathrm{Y}$ désigne un élément de $\mathcal{M}_1(\mathbb{R})$ et non un nombre. Il serait en fait plus rigoureux d'écrire $(x|y) = \mathrm{tr}(\mathrm{X}^{\mathrm{T}} \cdot \mathrm{Y})$ mais l'identification entre $\mathcal{M}_1(\mathbb{R})$ et $\mathbb{R}$ est couramment acceptée.

Remarquons que l'application $(\mathrm{X}, \mathrm{Y}) \mapsto \mathrm{X}^{\mathrm{T}} \cdot \mathrm{Y}$ est le produit scalaire usuel sur $\mathcal{M}_{n,1}(\mathbb{R})$. Il revient donc au même d'effectuer le produit scalaire des vecteurs x et y ou d'effectuer celui des vecteurs colonnes qui les représentent, *du moment que cette représentation se fait relativement à une même base orthonormale.*
On dit que l'application $x \mapsto \mathrm{M}_{\mathcal{E}}(x)$ est une *isométrie* de l'espace euclidien E sur l'espace euclidien $\mathcal{M}_{n,1}(\mathbb{R})$.
Remarquons alors que si l'on note $(\mathrm{E}_1, \dots, \mathrm{E}_n)$ la base canonique de $\mathcal{M}_{n,1}(\mathbb{R})$, alors celle-ci est une base orthonormale pour le produit scalaire canonique de $\mathcal{M}_{n,1}(\mathbb{R})$ et la décomposition d'un vecteur colonne U dans cette base est

$$\mathrm{U} = \sum_{k=1}^{n} (\mathrm{E}_k^{\mathrm{T}} \cdot \mathrm{U}) \mathrm{E}_k.$$

5.2 Représentation des endomorphismes

Soit f un endomorphisme de E. Notons $\mathrm{A} = (a_{i,j})_{1 \leqslant i,j \leqslant n}$ la matrice représentative de cet endomorphisme relativement à la base orthonormale $\mathcal{E}$.
Soit $(i, j) \in [\![1, n]\!]^2$ un couple d'indices. Le nombre $a_{i,j}$ est le i-ième coefficient de la j-ième colonne de A. C'est aussi le coefficient devant e_i dans la décomposition du vecteur $f(e_j)$ dans la base $\mathcal{E}$. On obtient donc les écritures suivantes pour ce coefficient

$$a_{i,j} = \mathrm{E}_i^{\mathrm{T}} \cdot \mathrm{A} \cdot \mathrm{E}_j = (e_i | f(e_j)).$$