Chapitre 3

Produit scalaire

Sommaire

Dans tout ce chapitre, E désigne un espace vectoriel réel.

3.1 Définitions

Définition 31. On appelle *produit scalaire* sur E toute application bilinéaire b symétrique telle que

- $\forall u \in E,\ b(u,u) \geqslant 0$ (b est *positive*),
- $\forall u \in E,\ b(u,u) = 0 \Rightarrow u = 0$ (b est *définie*).

Remarque. Dans la pratique, pour prouver qu'une application $b : E \times E \to \mathbb{R}$ est un produit scalaire, on vérifiera successivement que

- b est une application symétrique ($b(u,v) = b(v,u)$ pour tout $(u,b) \in E \times E$),
- b est linéaire en l'une des deux variables ($b(\lambda u + u', v) = \lambda b(u,v) + b(u',v)$),
- b est positive : $\forall u \in E, b(u,u) \geqslant 0$,
- b est définie : $\forall u \in E, b(u,u) = 0 \Rightarrow u = 0$.

Définition 32. Un espace vectoriel réel E muni d'un produit scalaire b est appelé *espace préhilbertien*. Si, de plus, E est de dimension finie, on dit que (E,b) est un *espace euclidien*.

Usuellement, pour noter un produit scalaire, on utilisera la notation $\langle u,v\rangle$ plutôt que $b(u,v)$. On trouve aussi les notations $u.v$, $(u|v)$ ou bien (en mécanique quantique particulièrement) $\langle u|v\rangle$. Dans ce cours, nous utiliserons uniquement la première notation : $\langle u,v\rangle$ (lire "produit scalaire de u avec v" ou, de manière abrégée "u scalaire v").

Il arrive fréquemment qu'un espace E soit muni d'un produit scalaire *canonique* (comme nous le verrons dans la section suivante). Dans ce cas, plutôt que noter $(E, \langle\cdot,\cdot\rangle)$, on omettra le produit scalaire et on parlera de l'espace préhilbertien (ou euclidien) E sans faire référence au produit scalaire.

3.2 Exemples de référence

1. L'exemple fondamental d'espace euclidien est l'espace $E = \mathbb{R}^n$ muni du produit scalaire
$$\langle u, v\rangle = u_1v_1 + \cdots + u_nv_n = \sum_{i=1}^{n} u_iv_i,$$
où on a noté $u = (u_1, \dots, u_n)$ et $v = (v_1, \dots, v_n)$ les coordonnées de u et v dans la base canonique de $\mathbb{R}^n$.

   Montrons que $\langle \cdot, \cdot\rangle$ est un produit scalaire. C'est tout d'abord une forme bilinéaire symétrique. Remarquons ensuite que, pour tout $u \in E$, $\langle u, u\rangle = u_1^2 + \cdots + u_n^2$ est une somme de termes positifs ou nuls. On a donc $\langle u, u\rangle \geqslant 0$. De plus, si $\langle u, u\rangle = 0$, cela signifie que chacun des termes de la somme est nul : $u_1^2 = u_2^2 = \cdots = u_n^2$, autrement dit $u = (0, \dots, 0) = 0_E$. Autrement dit $\langle \cdot, \cdot\rangle$ est défini positif. C'est donc un produit scalaire.

2. L'espace $E = M_{n,p}(\mathbb{R})$ où $n, p \in \mathbb{N}^*$ est muni du produit scalaire $\langle A, B\rangle = \mathrm{tr}({}^tA.B) = \sum_{i=1}^n \sum_{j=1}^p a_{ij}b_{ij}$. C'est un produit scalaire car, comme pour $\mathbb{R}^n$, on a que $\langle A, A\rangle = \sum_{i=1}^n \sum_{j=1}^p (a_{ij})^2$ est une somme de termes positifs ou nuls donc toujours positive et nulle si et seulement si tous les termes sont nuls.

   Remarquons que ce produit scalaire redonne lorsque $p = 1$ (resp. $n = 1$) le produit scalaire usuel sur $\mathbb{R}^n$ (resp. $\mathbb{R}^p$).

3. Soient $a, b \in \mathbb{R}$ avec $a < b$. Alors l'espace $E = C([a, b], \mathbb{R})$ des fonctions continues sur l'intervalle $[a, b]$ et à valeurs dans $\mathbb{R}$ est un espace préhilbertien pour le produit scalaire
$$\langle f, g\rangle = \int_a^b f(t)g(t)dt.$$
En effet, on voit que $\langle \cdot, \cdot\rangle$ est une forme bilinéaire symétrique. Par ailleurs, on a que, si $f = g$,
$$\langle f, f\rangle = \int_a^b f(t)^2dt \geqslant 0,$$
ce qui montre que $\langle \cdot, \cdot\rangle$ est positif. Reste à voir que, si $\langle f, f\rangle = 0$, on a $f = 0$. Ceci est une conséquence de la proposition suivante :

Proposition 33. *Soit g une fonction continue sur l'intervalle $[a, b]$ telle que $g \geqslant 0$ (c'est-à-dire, $\forall t \in [a, b], g(t) \geqslant 0$). Alors, si $\int_A^b g(t)dt = 0$, on a $g = 0$: $\forall t \in [a, b], g(t) = 0$.*

Démonstration. Puisque g est continue sur $[a, b]$, g admet une primitive G sur $[a, b]$, par exemple $G(x) = \int_a^x g(t)dt$. Si $x, y \in [a, b]$ avec $x < y$, on a
$$G(y) - G(x) = \int_x^y g(t)dt \geqslant 0,$$
donc G est croissante. Or, par hypothèse, $G(b) = G(a) = 0$. G est donc constante. Comme G est dérivable sur $]a, b[$ avec $G'(t) = g(t)$, on a $0 = G'(t) = g(t)$ pour tout $t \in]a, b[$. La fonction g étant continue, on a également $g(a) = g(b) = 0$ (pour voir que $g(a) = 0$, on peut dire, par exemple, que $a = \lim_{n\to\infty} a + 1/n$, donc $g(a) = g(\lim_{n\to\infty} a + 1/n) = \lim_{n\to\infty} g(a + 1/n) = 0$, idem pour b). □

3.3 Caractérisation matricielle

Soit b une forme bilinéaire sur E, un $\mathbb{R}$-espace vectoriel de dimension finie $n \geqslant 1$. Soit $\mathcal{B} = (e_1, \dots, e_n)$ une base de E. On rappelle que la matrice représentative de b dans la base $\mathcal{B}$ est $B = \mathrm{Mat}_{\mathcal{B}}(b) = (b(e_i, e_j))_{1\leqslant i,j\leqslant n}$. On a alors, si $u = u_1e_1 + \cdots + u_ne_n$ et $v = v_1e_1 + \cdots + v_ne_n$, $b(u, v) = {}^tUBV$ avec U (resp. V) la matrice colonne $U = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}$ (resp. $V = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$). On a alors que b est un produit scalaire si se seulement si les trois conditions suivantes sont satisfaites :

- B est *symétrique* : ${}^tB = B$,
- B est *positive* : $\forall X \in M_{1,n}(\mathbb{R})$, ${}^tX.B.X \geqslant 0$,

- B est *définie* : $\forall X \in M_{1,n}(\mathbb{R})$, ${}^tX.B.X = 0 \Rightarrow X = 0$.

On note

- $\mathcal{S}_n^+$ l'ensemble des matrices symétriques réelles positives et
- $\mathcal{S}_n^{++}$ l'ensemble des matrices symétriques réelles définies positives.

Par exemple, si a, b sont deux réels, on a

$$\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \in \mathcal{S}_n^+ \Leftrightarrow a \geqslant 0 \text{ et } b \geqslant 0; \qquad \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \in \mathcal{S}_n^{++} \Leftrightarrow a > 0 \text{ et } b > 0.$$

Donnons maintenant un critère pratique pour déterminer si une matrice symétrique est définie positive :

Théorème 34. *Soit $B \in M_n(\mathbb{R})$, $B = (b_{ij})_{1 \leqslant i,j \leqslant n}$ une matrice symétrique. Alors B est définie positive si et seulement si on a*

$$b_{11} > 0, \quad \begin{vmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{vmatrix} > 0, \quad \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{vmatrix} > 0, \dots \quad \begin{vmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{n1} & \cdots & b_{nn} \end{vmatrix} > 0.$$

La preuve de ce résultat s'inspire de techniques que nous développerons dans le prochain chapitre. Elle est d'un niveau plus élevé que les autres preuves de ce cours et peut donc être omise.

Soit b la forme bilinéaire symétrique associée à B : $b(X, Y) = {}^tXBY$. Nous allons montrer par récurrence sur k que b est définie positive sur $\mathrm{Vect}(e_1, \dots, e_k)$ avec $(e_1, \dots, e_n)$ la base canonique de $E = M_{n,1}(\mathbb{R})$. Le théorème 34 en découlera en prenant $k = n$.

Nous aurons besoin des deux définitions suivantes :

- b est dite *non-dégénérée* si le noyau de b

$$\mathrm{Ker}\ (b) := \{u \in E | \forall v \in E,\ b(u, v) = 0\}$$

est réduit à $\{0\}$.

- Si F est un sous-espace vectoriel de E, on appelle *orthogonal* de F le sous-espace vectoriel

$$F^\perp = \{u \in E | \forall v \in F,\ b(u, v) = 0\}.$$

Donnons maintenant deux lemmes qui nous seront utiles dans la preuve du théorème 34. Nous donnons une formulation volontairement plus générale que celle qui nous intéressera par la suite.

Lemme 35. *Soient b une forme bilinéaire sur un espace vectoriel E de dimension finie et $\mathcal{B} = (e_1, \dots, e_n)$ une base de E. Alors b est non-dégénérée si et seulement si $\det(B) \neq 0$ avec $B = \mathrm{Mat}_{\mathcal{B}}(b)$. On a alors que $u \mapsto b(u, \cdot)$ est un isomorphisme entre E et E^*.*

Démonstration. Si $\det(B) = 0$, il existe $U \in M_{n,1}(\mathbb{R})$, $U = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} \neq 0$ tel que $BU = 0$. En particulier, pour tout $V = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$, on a ${}^tVBU = 0$. Autrement dit, si $u = u_1e_1 + \dots + u_ne_n$ et $v = v_1e_1 + \cdots v_ne_n$ quelconque, on a $b(v, u) = 0 : 0 \neq u \in \mathrm{Ker}\ (b)$. La forme b est donc dégénérée.

Inversément, si $u \in \mathrm{Ker}\ (b)$, on a, en reprenant les notations précédentes, que pour tout $V \in M_{n,1}(\mathbb{R})$, ${}^tVBU = 0$. En particulier, si $V = e_i = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}$, on a (pour tout i) que la i-ème composante de BU est nulle, c'est-à-dire $BU = 0$ et donc $\det(B) \neq 0$.

Si b est non-dégénérée, l'application $u \mapsto b(u, \cdot)$ est injective. En effet, $b(u, \cdot) = 0 \Leftrightarrow \forall v \in E,\ b(u, v) = 0 \Leftrightarrow u \in \mathrm{Ker}\ (b)$. Comme E et E^* ont même dimension, cette application est un isomorphisme. $\square$

Lemme 36. *Soit b une forme bilinéaire sur un $\mathbb{R}$-espace vectoriel de dimension finie n. Alors, si F est un sous-espace vectoriel de E, $F^\perp$ est un sous-espace vectoriel de E de dimension* $\dim(F^\perp) = n - \dim(F)$.

Ce résultat sera montré à nouveau dans le cas (plus simple) où b est un produit scalaire (voir la proposition 51 et la section 4.3).

Démonstration. Soient $k = \dim(F)$ et $(e_1, \dots, e_k)$ une base de F. Comme $u \mapsto b(u, \cdot)$ est un isomorphisme, les formes linéaires $\varphi_i = b(e_i, \cdot)$ $(i = 1, \dots, k)$ forment une famille libre de E^* et $F^\perp = \bigcap_{i=1}^{k} \operatorname{Ker}(\varphi_i)$. Complétons $(\varphi_1, \dots, \varphi_k)$ en une base de E^* et notons $(f_1, \dots, f_n)$ la base antéduale.

Soit $G = \operatorname{Vect}(f_{k+1}, \dots, f_n)$. On a $\varphi_i(f_j) = 0$ si $i \in \{1, \dots, k\}$ et $j \in \{k+1, \dots, n\}$, donc les f_j, $j \geqslant k+1$ sont dans $F^\perp$: $G \subset F^\perp$. Inversément, si $u \notin G$, $u = u_1 f_1 + \dots + u_k f_k + u_{k+1} f_{k+1} + \dots + u_n f_n$, au moins un des $u_1, \dots, u_k$ est non nul, choisissons-en un : u_i. On a alors $\varphi_i(u) = u_i \neq 0$ (par définition de la base antéduale) et donc $u \notin F^\perp$.

Nous avons donc montré $F^\perp = G$, d'où $\dim(F^\perp) = \dim(G) = n - k$. □

Preuve du Théorème 34. Comme nous l'indiquions après l'énoncé du théorème, nous allons montrer que la forme bilinéaire b est définie positive sur chacun des $\operatorname{Vect}(e_1, \dots, e_k)$ par une récurrence sur k. Plus précisément, nous allons construire une base $\mathcal{B} = (f_1, \dots, f_n)$ de vecteurs tels que $b(f_i, f_j) = 0$ si $i \neq j$ et $b(f_i, f_i) > 0$, ce qui montrera que b, la forme linéaire associée à B est définie positive.

- Si $k = 1$, $F_1 = \operatorname{Vect}(e_1)$ est de dimension 1. Choisissons $f_1 \in F_1 \setminus \{0\}$ arbitrairement.
- Supposons que nous avons montré que b est définie positive sur F_k pour un certain $k \geqslant 1$, $k < n$. b est non-dégénérée sur $E = M_{n,1}(\mathbb{R})$ car, par hypothèse

$$\det(\operatorname{Mat}_{(e_1,\dots,e_n)}(b)) = \det(B) = \begin{vmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{n1} & \cdots & b_{nn} \end{vmatrix} > 0$$

(voir le lemme 35). En utilisant le lemme 2, nous en déduisons que $F_k^\perp$ est un sous-espace vectoriel de E de dimension $n - k$. Or, si $u \in F_k \cap F_k^\perp$, on a $b(u, u) = 0$ par définition de $F_k^\perp$. Comme b est définie positive sur F_k, cela impose $u = 0$. Nous avons donc montré que $F_k \cap F_k^\perp = \{0\}$. On a alors $\dim(F_k \oplus F_k^\perp) = \dim(F_k) + \dim(F_k^\perp) = k + (n-k) = n$, d'où

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

On a alors que $F_k^\perp \cap F_{k+1}$ est de dimension 1 ($\dim(F_{k+1}) + \dim(F_k^\perp) = n + 1 > n$ donc $\dim(F_k^\perp \cap F_{k+1}) > 0$ et $\dim(F_{k+1}) \geqslant \dim(F_k) + \dim(F_k^\perp \cap F_{k+1})$ i.e. $k + 1 \geqslant k + \dim(F_k^\perp \cap F_{k+1})$.

Soit $f_{k+1} \in F_k^\perp \cap F_{k+1}$, $f_{k+1} \neq 0$. Alors f_{k+1} est linéairement indépendant de $(f_1, \dots, f_k)$, la famille $(f_1, \dots, f_{k+1})$ est donc libre. De plus, comme $f_{k+1} \in F_k^\perp$ on a $b(f_{k+1}, f_i) = 0$ pour tout $i \leqslant k$. On a donc que la matrice de b (restreinte à F_{k+1}) dans la base $(f_1, \dots, f_{k+1})$ est diagonale :

$$\operatorname{Mat}_{(f_1,\dots,f_{k+1})}(b|_{F_{k+1}}) = \begin{pmatrix} c_1 & & 0 \\ & \ddots & \\ 0 & & c_{k+1} \end{pmatrix}.$$

Par hypothèse de récurrence, $c_i >$ pour tout $i \leqslant k$. Reste donc à voir $c_{k+1} > 0$. Soit P_{k+1} la matrice de passage de $(e_1, \dots, e_{k+1})$ à $(f_1, \dots, f_{k+1})$. Alors, en utilisant la formule (2.2.4), on a

$$\operatorname{Mat}_{(f_1,\dots,f_{k+1})}(b|_{F_{k+1}}) = {}^tP_{k+1} \operatorname{Mat}_{(e_1,\dots,e_{k+1})}(b|_{F_{k+1}}) P_{k+1}.$$

En prenant le déterminant :

$$c_1 \cdots c_k c_{k+1} = \det(P_{k+1})^2 \begin{vmatrix} b_{11} & \cdots & b_{1,k+1} \\ \vdots & \ddots & \vdots \\ b_{k+1,1} & \cdots & b_{k+1,k+1} \end{vmatrix} > 0.$$

Comme $c_i = b(f_i, f_i) > 0$ pour tout $i \leqslant k$, on en déduit $c_{k+1} = b(f_{k+1}, f_{k+1}) > 0$.

On a donc trouvé une base dans laquelle la matrice de b est diagonale avec uniquement des termes strictement positifs (sur la diagonale) : b est définie positive. □

3.4 Deux inégalités fondamentales

Définition 37. Soient $(E, \langle\cdot,\cdot\rangle)$ un espace préhilbertien réel et $u \in E$. Puisque $\langle u, u\rangle \geqslant 0$, on peut extraire sa racine carrée. On note donc $\|u\| = \sqrt{\langle u, u\rangle}$ la *norme* de u. Cette norme est dite *associée* au produit scalaire $\langle\cdot,\cdot\rangle$.

Dans le cas où $E = \mathbb{R}^2$ muni du produit scalaire usuel, on a $\|(x_1, x_2)\| = \sqrt{x_1^2 + x_2^2}$, c'est-à-dire que la norme représente la longueur du vecteur (x_1, x_2).

Remarquons que $\|u\| \geqslant 0$, $\|u\| = 0$ si et seulement si $u = 0$ et que $\|\lambda u\| = |\lambda|\ \|u\|$ si $\lambda \in \mathbb{R}$.

Définition 38. Un vecteur $u \in E$ est dit *unitaire* si $\|u\| = 1$.

On a les résultats fondamentaux suivants (Théorèmes 39 et 41) :

Théorème 39 (Inégalité de Cauchy-Schwarz)**.** *Soient $u, v \in E$, on a*

$$|\langle u, v\rangle| \leqslant \|u\|.\|v\|.$$

De plus, on a égalité ($|\langle u, v\rangle| = \|u\|.\|v\|$) si et seulement si la famille (u, v) est liée.

Démonstration. Soient u et v deux vecteurs de E. Si $u = 0$, on a $\langle u, v\rangle = 0$ et $\|u\|.\|v\| = 0$ donc l'inégalité de Cauchy-Schwarz est satisfaite dans ce cas. C'est même une égalité et la famille (u, v) est liée.

Nous pouvons donc supposer que $u \neq 0$. Considérons maintenant l'application p qui à tout $t \in \mathbb{R}$ associe

$$p(t) = \|tu + v\|^2 = \langle tu + v, tu + v\rangle.$$

Par bilinéarité du produit scalaire, on a

$$p(t) = t^2\langle u, u\rangle + 2t\langle u, v\rangle + \langle v, v\rangle = t^2\|u\|^2 + 2t\langle u, v\rangle + \|v\|^2.$$

Donc p est un trinôme en t. Comme $p(t) \geqslant 0$ pour tout $t \in \mathbb{R}$, le discriminant de p est négatif ou nul :

$$\Delta = 4\langle u, v\rangle^2 - 4\|u\|^2.\|v\|^2 \leqslant 0 \quad \text{soit} \quad \langle u, v\rangle^2 \leqslant \|u\|^2.\|v\|^2.$$

En prenant la racine carrée, nous obtenons l'inégalité de Cauchy-Schwarz. Examinons maintenant le cas d'égalité :

$$\begin{aligned} |\langle u, v\rangle| = \|u\|.\|v\| &\Leftrightarrow \Delta = 0 \\ &\Leftrightarrow p \text{ a une racine double } t_0 \in \mathbb{R} \\ &\Leftrightarrow \exists t_0 \in \mathbb{R},\ \|t_0 u + v\|^2 = 0 \\ &\Leftrightarrow \exists t_0 \in \mathbb{R},\ t_0 u + v = 0_E \qquad \Leftrightarrow u \text{ et } v \text{ sont liés.} \end{aligned}$$

□

Remarquons qu'on a toujours $\langle u, v\rangle \leqslant |\langle u, v\rangle|$. On en déduit le corollaire suivant :

Corollaire 40. *Soient $u, v \in E$, on a*

$$\langle u, v\rangle \leqslant \|u\|.\|v\|$$

avec égalité si et seulement si u et v sont positivement liés, c'est-à-dire si u et v sont tous les deux nuls ou si u (resp. v) est non nul et qu'il existe $\alpha \geqslant 0$ tel que $v = \alpha u$ (resp. $u = \alpha v$).

Démonstration. Nous avons déjà montré que l'inégalité est vraie, reste à voir le cas d'égalité. Excluons le cas $u = v = 0$ qui est évident. Ensuite, remarquons que, comme

$$\langle u, v\rangle \leqslant |\langle u, v\rangle| \leqslant \|u\|.\|v\|,$$

on doit avoir $|\langle u, v\rangle| = \|u\|.\|v\|$ (inégalité dans Cauchy-Schwarz) et $\langle u, v\rangle = |\langle u, v\rangle| \geqslant 0$. Supposons, par exemple, que $u \neq 0$. On a alors $v = \alpha u$ pour un certain $\alpha \in \mathbb{R}$ car la famille (u, v) est liée. Comme $0 \leqslant \langle u, v\rangle = \alpha\langle u, u\rangle = \alpha\|u\|^2$, on voit que $\alpha \geqslant 0$, ce qui conclut la preuve du corollaire. □

L'inégalité de Cauchy-Schwarz permet de montrer de nombreuses inégalités :

•3.4.1

•3.4.1: rg: Faire ici une liste d'exemples

Tournons-nous maintenant vers la seconde inégalité fondamentale :

Théorème 41 (Inégalité de Minkowski)**.** *Soient $u, v \in E$, on a*

$$\|u+v\| \leqslant \|u\| + \|v\|,$$

avec égalité si et seulement si u et v sont positivement liés.

Démonstration. On calcule

$$\|u+v\|^2 = \langle u+v, u+v\rangle = \langle u,u\rangle + 2\langle u,v\rangle + \langle v,v\rangle = \|u\|^2 + \|v\|^2 + 2\langle u,v\rangle.$$

En utilisant le corollaire 40, on a alors

$$\|u+v\|^2 \leqslant \|u\|^2 + \|v\|^2 + 2\|u\|.\|v\| = (\|u\| + \|v\|)^2.$$

En prenant la racine carrée de cette inégalité, on obtient l'inégalité de Minkowski.

Pour obtenir le cas d'égalité, remarquons que le seul moment où, dans les calculs, nous avons introduit une inégalité, est lorsque nous avons remplacé (ou plus exactement majoré) $\langle u,v\rangle$ par $\|u\|.\|v\|$. Nous aurons donc égalité dans l'inégalité de Minkowski si et seulement si nous avons $\langle u,v\rangle = \|u\|.\|v\|$, c'est-à-dire si u et v sont positivement liés. □

Remarquons que le concept de norme en mathématiques est plus étendu que celui que nous venons de voir :

Définition 42. Soit E un $\mathbb{R}$-espace vectoriel. Une application $\|\cdot\| : E \to \mathbb{R}$ est appelée *norme* sur E si elle satisfait les trois conditions suivantes :

- $\forall u \in E,\ \|u\| \geqslant 0,$
- $\forall u \in E,\ \|u\| = 0 \Leftrightarrow u = 0_E,$
- $\forall u \in E, \lambda \in \mathbb{R},\ \|\lambda u\| = |\lambda|\ \|u\|,$
- $\forall u, v \in E,\ \|u+v\| \leqslant \|u\| + \|v\|$ (inégalité triangulaire).

Remarquons une conséquence importante de l'inégalité triangulaire. On peut toujours écrire, pour tous $x, y \in E$, $\|x\| = \|x-y+y\| \leqslant \|x-y\| + \|y\|$, donc $\|x\| - \|y\| \leqslant \|x-y\|$. En inversant les rôles de x et y, on trouve de même $-\|x\| + \|y\| \leqslant \|x-y\|$, d'où

$$\big|\|x\| - \|y\|\big| \leqslant \|x-y\|. \tag{3.4.1}$$

Il existe des normes qui ne peuvent pas être obtenues à partir d'un produit scalaire, citons par exemple les deux normes suivantes sur $\mathbb{R}^n$ ($n \geqslant 2$) :

$$\|u\|_1 = \sum_{i=1}^n |x_i|, \qquad \|u\|_\infty = \max_{i \in \{1,\dots,n\}} |x_i|.$$

Pour montrer que ces normes ne peuvent être obtenues à partir d'un produit scalaire, remarquons le résultat suivant :

Proposition 43 (Identité du parallélogramme)**.** *Soit $(E, \langle\cdot,\cdot\rangle)$ un espace préhilbertien. Notons $\|\cdot\|$ la norme associée au produit scalaire ($\|u\|^2 = \langle u,u\rangle$). Alors on a l'identité suivante : (*identité du parallélogramme*)*

$$\|u+v\|^2 + \|u-v\|^2 = 2(\|u\|^2 + \|v\|^2) \tag{3.4.2}$$

Démonstration. La preuve est un simple calcul :

$$\begin{aligned}\|u+v\|^2 + \|u-v\|^2 &= \langle u+v, u+v\rangle + \langle u-v, u-v\rangle \\ &= (\langle u,u\rangle + \langle u,v\rangle + \langle v,u\rangle + \langle v,v\rangle) + (\langle u,u\rangle - \langle u,v\rangle - \langle v,u\rangle + \langle v,v\rangle) \\ &= 2\langle u,u\rangle + 2\langle v,v\rangle = 2(\|u\|^2 + \|v\|^2).\end{aligned}$$

□

Prenons par exemple les vecteurs $u = (1,1,0,\dots,0)$ et $v = (0,1,0,\dots,0)$ de $\mathbb{R}^n$. On a alors

$$\|u+v\|_\infty = \|(1,2,0,\dots,0)\|_\infty = 2, \quad \|u-v\|_\infty = \|(1,0,0,\dots,0)\|_\infty = 1.$$

Donc $\|u+v\|_\infty^2 + \|u-v\|_\infty^2 = 2^2 + 1^2 = 5$ alors que

$$2(\|u\|_\infty^2 + \|v\|_\infty^2) = 2(1^2 + 1^2) = 2 \neq 5,$$

ce qui montre que l'identité du parallélogramme n'est pas satisfaite pour la norme $\|\cdot\|_\infty$. De la même manière, si $u = (1,1,0,\dots,0)$ et $v = (1,-1,0,\dots,0)$ on a

$$\|u+v\|_1 = \|(2,0,0,\dots,0)\|_1 = 2, \quad \|u-v\|_1 = \|(0,2,0,\dots,0)\|_1 = 2.$$

Donc $\|u+v\|_1^2 + \|u-v\|_1^2 = 8$ alors que

$$2(\|u\|_1^2 + \|v\|_1^2) = 2(2^2 + 2^2) = 16 \neq 8.$$

La norme $\|\cdot\|_1$ ne satisfait donc pas non plus l'identité du parallélogramme. Ces deux normes ne sont sont donc associée à aucun produit scalaire sur $\mathbb{R}^n$.

On peut démontrer également que la réciproque est vraie : l'identité (3.4.2) est satisfaite si et seulement si la norme $\|\cdot\|$ est associée à un produit scalaire. C'est le théorème de Fréchet–von Neumann–Jordan :

Théorème 44. *Soit $\|\cdot\|$ une norme sur un espace euclidien E. Alors il existe un produit scalaire $\langle\cdot,\cdot\rangle$ telle que $\|\cdot\|$ dérive de ce produit scalaire ($\forall x \in E, \|x\|^2 = \langle x,x\rangle$) si et seulement si la norme $\|\cdot\|$ satisfait l'identité du parallélogramme :*

$$\forall x,y \in E, \|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2).$$

Démonstration. Nous avons déjà vu qu'une norme issue d'un produit scalaire satisfaisait l'identité du parallélogramme. Il nous reste donc à voir la réciproque. Pour cela, nous allons construire le candidat au produit scalaire : Posons

$$f(x,y) = \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2\right)$$

et remarquons que f est symétrique en x et y et que $f(x,0) = 0$. Montrons ensuite que

$$\forall x,y,z \in E, f(x,y+z) = f(x,y) + f(x,z).$$

On a

$$f(x,y+z) = \frac{1}{4}\left(\|x+y+z\|^2 - \|x-y-z\|^2\right).$$

Appliquons l'identité du parallélogramme à $\frac{x}{2}+y$ et $\frac{x}{2}+z$:

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

De même,

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

En soustrayant, on obtient alors

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

En prenant en particulier $z = 0$ (resp. $y = 0$), on trouve $f(x,y) = 2f(x/2,y)$ (resp. $f(x,z) = 2f(x/2,z)$), d'où

$$f(x,y+z) = f(x,y) + f(x,z).$$

Tout ce qu'il nous reste à faire est de voir que $f(x,\lambda y) = \lambda f(x,y)$ pour tout $\lambda \in \mathbb{R}$. Or, par une récurrence simple, on a que, pour tout $p \in \mathbb{Z}$, $f(x,py) = pf(x,y)$. Donc, pour tout $q \in \mathbb{N}^*$, $f(x,y) = f(x,qy/q) = qf(x,y/q)$ et $f(x,y/q) = 1/qf(x,y)$. En combinant ces deux faits, on obtient donc que $f(x,p/qy) = p/qf(x,y)$. Nous avons démontré que $f(x,\lambda y) = \lambda f(x,y)$ pour tout $\lambda \in \mathbb{Q}$.

Pour passer de $\lambda \in \mathbb{Q}$ à $\lambda \in \mathbb{R}$, il nous faut montrer que l'application $\lambda \mapsto f(x,\lambda y)$ est continue. Ceci dérive du fait que les applications $\lambda \mapsto \|x+\lambda y\|$ et $\lambda \mapsto \|x-\lambda y\|$ sont continues. Pour le voir, on utilise l'inégalité (3.4.1) :

$$|\|x+\lambda y\| - \|x+\lambda' y\|| \leqslant \|(x+\lambda y)-(x+\lambda' y)\| = \|(\lambda-\lambda')y = |\lambda-\lambda'|\|y\|$$

d'où on voit que $\lambda \mapsto \|x+\lambda y\|$ est Lipschitzienne. □

Disposer d'une norme sur un espace vectoriel E permet de définir une distance entre deux éléments de E par la formule suivante :

$$d(x,y) = \|x-y\|.$$

La notion de distance est une notion encore plus vaste que celle d'une norme. En effet, il est possible de définir une distance sur n'importe quel ensemble X. De manière axiomatique, une distance est une application $d : X \times X \to \mathbb{R}$ telle que

1. d est symétrique : $\forall x,y \in X,\ d(x,y) = d(y,x)$,
2. d est toujours positive : $\forall x,y \in X,\ d(x,y) \geqslant 0$,
3. d sépare les points : $\forall x,y \in X,\ d(x,y) = 0 \Leftrightarrow x = y$,
4. d satisfait l'inégalité triangulaire : $\forall x,y,z \in X,\ d(x,z) \leqslant d(x,y) + d(y,z)$.

On voit facilement que la distance $d(x,y) = \|x-y\|$ que nous avons définie plus haut satisfait ces quatre axiomes. Mais il est possible de définir d'autres distances, par exemple (sur n'importe quel ensemble X) :

$$d(x,y) = \begin{cases} 1 \text{ si } x \neq y, \\ 0 \text{ si } x = y. \end{cases}$$

L'étude des espaces métriques (i.e. des ensembles munis d'une distance) est un pan entier des mathématiques appelé topologie des espaces métriques.

3.5 Angle entre deux vecteurs

Définition 45. Soient u et v deux vecteurs non nuls de E. En utilisant l'inégalité de Cauchy-Schwarz, on a que $\dfrac{\langle u,v\rangle}{\|u\|\ \|v\|} \in [-1,1]$. Il existe donc un unique $\theta \in [0,\pi]$ tel que $\cos(\theta) = \dfrac{\langle u,v\rangle}{\|u\|\ \|v\|}$. θ est appelé l'*angle* entre les vecteurs u et v.

Exemple. Dans $\mathbb{R}^3$ muni du produit scalaire usuel, cnosiérons $x = (1,0,1)$ et $y = (1,1,0)$. Alors $\|x\| = \|y\| = \sqrt{2}$ et $\langle x,y\rangle = 1$, d'où

$$\cos(\theta) = \frac{\langle x,y\rangle}{\|x\|\|y\|} = \frac{1}{2}$$

et $\theta = \pi/3$ est l'angle géométrique entre x et y.

Proposition 46. *Soient u et v deux vecteurs non-nuls de E et θ l'angle entre ces deux vecteurs. Alors*

- $\theta = 0 \Leftrightarrow u$ *et* v *sont positivement liés* ($u = \lambda v$ *avec* $\lambda > 0$),
- $\theta = \pi \Leftrightarrow u$ *et* v *sont négativement liés* ($u = \lambda v$ *avec* $\lambda < 0$).

Démonstration. On utilise ici le cas d'égalité de l'inégalité de Cauchy-Schwarz. Si $\theta = 0$ ou π, on a $\dfrac{\langle u,v\rangle}{\|u\|\ \|v\|} = \pm 1$ donc $|\langle u,v\rangle| = \|u\|\ \|v\|$ d'où on tire que u et v sont colinéaires. Comme u et v sont non nuls, on a $u = \lambda v$ avec $\lambda \in \mathbb{R}^*$.

Si $\theta = 0$, $\lambda\|v\|^2 = \langle \lambda v, v\rangle = \langle u,v\rangle = \|u\|\ \|v\| > 0$ donc $\lambda > 0$.

De même si $\theta = \pi$, $\lambda\|v\|^2 = \langle \lambda v, v\rangle = \langle u,v\rangle = -\|u\|\ \|v\| < 0$ donc $\lambda < 0$. $\square$

Dans le cas particulier où $\theta = \pi/2$, c'est-à-dire si $\langle u,v\rangle = 0$, les vecteurs u et v sont dits *perpendiculaires* ou *orthogonaux*. Comme la condition $\langle u,v\rangle = 0$ fait sens même si u ou v est nul, on dira que u et v sont orthogonaux dès lors que $\langle u,v\rangle = 0$. On notera $u \perp v$ pour signifier que u et v sont orthogonaux.

3.6 Orthogonalité

Nous allons maintenant étudier plus en détail la notion d'orthogonalité. Le premier théorème (fondamental) est celui de Pythagore :

Théorème 47 (Théorème de Pythagore)**.** *Soient u et v deux vecteurs de E. Alors on a*

$$u \perp v \Leftrightarrow \|u+v\|^2 = \|u\|^2 + \|v\|^2 \Leftrightarrow \|u-v\|^2 = \|u\|^2 + \|v\|^2.$$

Démonstration. Pour la première équivalence, on écrit

$$\|u+v\|^2 = \langle u+v, u+v\rangle = \langle u,u\rangle + \langle u,v\rangle + \langle v,u\rangle + \langle v,v\rangle = \|u\|^2 + 2\langle u,v\rangle + \|v\|^2.$$

On voit donc que $u \perp v$ si et seulement si $\|u+v\|^2 = \|u\|^2 + \|v\|^2$. De même, on a

$$\|u-v\|^2 = \|u\|^2 - 2\langle u,v\rangle + \|v\|^2,$$

ce qui montre l'équivalence $u \perp v \Leftrightarrow \|u-v\|^2 = \|u\|^2 + \|v\|^2$. □

Remarquons également les deux équivalences suivantes :

Proposition 48. *Soient u et v deux vecteurs de E*

- $\|u\| = \|v\| \Leftrightarrow (u+v) \perp (u-v)$ *(identité du losange),*
- $\|u+v\| = \|u-v\| \Leftrightarrow u \perp v$ *(identité du rectangle).*

Démonstration. Pour démontrer l'identité du losange, il suffit de voir que $\|u\|^2 - \|v\|^2 = \langle u+v, u-v\rangle$. De même, pour l'identité du rectangle, $\|u+v\|^2 - \|u-v\|^2 = 4\langle u,v\rangle$. □

Remarquons que nous avons utilisé plusieurs identités permettant d'obtenir le produit scalaire $\langle u,v\rangle$ en fonction de certaines normes. Ces identités sont dites de polarisation. Regroupons-les ici :

Proposition 49 (Identités de polarisation)**.** *Soient u et v deux vecteurs de E, on a*

$$\langle u,v\rangle = \frac{1}{2}\left(\|u+v\|^2 - \|u\|^2 - \|v\|^2\right),$$
$$\langle u,v\rangle = \frac{1}{2}\left(\|u\|^2 + \|v\|^2 - \|u-v\|^2\right),$$
$$\langle u,v\rangle = \frac{1}{4}\left(\|u+v\|^2 - \|u-v\|^2\right).$$

Nous allons maintenant étendre la notion d'orthogonalité à des parties de E :

Définition 50. • Soient A et B deux partes de E. On dit que A et B sont *orthogonales* (noté $A \perp B$) si et seulement si tous les vecteurs de A sont orthogonaux aux vecteurs de B :

$$\forall u \in A, \forall v \in B, \langle u,v\rangle = 0.$$

• Soit A une partie de E, on appelle orthogonal de A l'ensemble

$$A^\perp = \{u \in E | \forall v \in A, \langle u,v\rangle = 0\}.$$

Proposition 51. *Soit A une partie de E. Alors on a les propriétés suivantes :*

1. $A \perp A^\perp$,
2. $A^\perp$ *est un sous-espace vectoriel de E,*
3. *Si* $A \subset B$, $B^\perp \subset A^\perp$,
4. $A^\perp = (\mathrm{Vect}(A))^\perp$,
5. $A \subset (A^\perp)^\perp$,
6. *si A est un sous-espace vectoriel de E, $A \cap A^\perp = \{0\}$. Si, de plus, E est de dimension finie, $E = A \oplus A^\perp$ et $(A^\perp)^\perp = A$.*

Démonstration. 1. Par définition, pour tout $u \in A^\perp$, on a $\forall v \in A, \langle u,v\rangle = 0$, ce qui montre que $A \perp A^\perp$.

2. Comme, pour tout vecteur $v \in E$, on a $\langle 0_E, v\rangle = 0$, on a $0_E \in A^\perp$. De plus, soient $u_1, u_2 \in A^\perp$ et $\lambda \in \mathbb{R}$, alors on a
$$\forall v \in A,\ \langle \lambda u_1 + u_2\rangle = \lambda\langle u_1, v\rangle + \langle u_2, v\rangle = 0,$$
car $u_1, u_2 \in A^\perp$, ce qui montre que $\lambda u_1 + u_2 \in A^\perp$. $A^\perp$ est donc un sous-espace vectoriel de E.

3. Soit $u \in B^\perp$, on a $\forall v \in B, \langle u, v\rangle = 0$. En particulier, $\forall v \in A, \langle u, v\rangle = 0$, c'est-à-dire $u \in A^\perp$. Ceci montre que $B^\perp \subset A^\perp$.

4. Comme $A \subset \mathrm{Vect}(A)$, on a $(\mathrm{Vect}(A))^\perp \subset A^\perp$. Reste donc à voir que $A^\perp \subset (\mathrm{Vect}(A))^\perp$. Or, si $v \in \mathrm{Vect}(A)$, il existe $v_1, \dots, v_k \in A$, $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ tels que $v = \lambda_1 v_1 + \cdots + \lambda_k v_k$. Donc, si $u \in A^\perp$, on a
$$\langle u, v\rangle = \langle u, \lambda_1 v_1 + \cdots + \lambda_k v_k\rangle = \lambda_1\langle u, v_1\rangle + \cdots + \lambda_k\langle u, v_k\rangle = 0$$
car, comme $v_1, \dots, v_k \in A$, on a $\langle u, v_1\rangle = \cdots = \langle u, v_k\rangle = 0$. On a donc, pour tout $u \in A^\perp$, $\forall v \in \mathrm{Vect}(A), \langle u, v\rangle = 0$, c'est-à-dire $u \in (\mathrm{Vect}(A))^\perp$.

5. Soit $u \in A$. Alors, par définition de $A^\perp$, on a, pour tout $v \in A^\perp$, $\langle u, v\rangle = 0$, c'est-à-dire $u \in (A^\perp)^\perp$.

6. Soit $u \in A \cap A^\perp$. Alors, comme $u \in A^\perp$, on a, pour tout $v \in A$, $\langle u, v\rangle = 0$. Or $u \in A$, donc, en particulier, $\langle u, u\rangle = 0$. Comme le produit scalaire est défini, on en déduit $u = 0_E$.

Pour démontrer les résultats lorsque E est de dimension finie, nous aurons besoin de plus d'outils. Voir la section 4.3. □

3.7 Exercices

Exercice 3.1. En utilisant l'inégalité de Cauchy-Schwarz (en précisant l'espace préhilbertien $(E, \langle ., .\rangle)$ dans lequel on travaille et les vecteurs concernés) établir les inégalités suivantes et étudier les cas d'égalité :

1. $\forall (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$, $|x_1 + x_2 + \ldots + x_n| \leqslant \sqrt{n}\left(x_1^2 + x_2^2 + \cdots + x_n^2\right)^{1/2}$;

2. $\forall (x_1, x_2, \dots, x_n) \in (\mathbb{R}_+^*)^n$, $(x_1 + x_2 + \ldots + x_n)\left(\frac{1}{x_1} + \frac{1}{x_2} + \cdots + \frac{1}{x_n}\right) \geqslant n^2$;

3. $\displaystyle\sum_{k=1}^{n} \sqrt{k} \leqslant n\sqrt{\frac{n+1}{2}}$;

4. $\forall M \in M_n(\mathbb{R})$, $\mathrm{tr}(M)^2 \leqslant n\,\mathrm{tr}({}^tM \cdot M)$;

5. Pour toute fonction f continue et strictement positive sur $[a, b]$, avec $a < b$, on a
$$\left(\int_a^b f(t)dt\right)\left(\int_a^b \frac{1}{f(t)}dt\right) \geqslant (b-a)^2.$$

Exercice 3.2. Soient $u = (1, 1, 1, 1)$ et $v = (1 - x, x - y, y - z, z)$ deux vecteurs de $\mathbb{R}^4$.

1. Montrer que $\langle u, v\rangle = 1$.

2. En utilisant l'inégalité de Cauchy-Schwarz, résoudre dans $\mathbb{R}^4$ l'équation suivante
$$(1-x)^2 + (x-y)^2(y-z)^2 + z^2 = \frac{1}{4}.$$

Exercice 3.3. Soient $E = \mathbb{R}^2$ et $a, b, c, d \in \mathbb{R}$. Notons φ la forme bilinéaire définie pour tous vecteurs $u = (x_1, x_2)$ et $v = (y_1, y_2)$ par $\varphi(u, v) = ax_1y_1 + bx_2y_2 + cx_2y_1 + dx_1y_2$. A quelles conditions sur a, b, c, d, a-t-on que φ est un produit scalaire sur E ?

Exercice 3.4. Soit $E = \mathbb{R}^2$.

1. Prouver que la forme bilinéaire b dont la matrice dans la base canonique $\mathcal{B} = (e_1, e_2)$ est $B = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$ est un produit scalaire sur E.

2. En déduire que $\forall (x, y) \in \mathbb{R}^2$, $|2x + 3y| \leqslant \sqrt{5}\sqrt{x^2 + 2xy + 2y^2}$.

Exercice 3.5. Soit $E = C^1([0, 1], \mathbb{R})$.

1. Prouver que l'application b de E^2 dans $\mathbb{R}$ définie par

$$b(f,g) = f(0)g(0) + \int_0^1 f'(t)g'(t)dt$$

est un produit scalaire sur E.

2. Définir la norme $\|\cdot\|$ issue de b.

3. Prouver que pour toute fonction $f \in E$, on a

$$\left|f(0) + \int_0^1 f'(t)dt\right| \leqslant \sqrt{2}\sqrt{f(0)^2 + \int_0^1 (f'(t))^2 dt}.$$