RÉSUMÉ DE COURS : PRODUIT SCALAIRE - ESPACES EUCLIDIENS

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

1 Notion de produit scalaire

Définition 1.1 *Une forme bilinéaire sur E est une application $\varphi : E \times E \longrightarrow \mathbb{R}$ telle que :*

1. $\forall (x_1, x_2, y) \in E^3$, $\forall (\lambda_1, \lambda_2) \in \mathbb{R}^2$, $\varphi(\lambda_1 x_1 + \lambda_2 x_2, y) = \lambda_1 \varphi(x_1, y) + \lambda_2 \varphi(x_2, y)$.
2. $\forall (x, y_1, y_2) \in E^3$, $\forall (\lambda_1, \lambda_2) \in \mathbb{R}^2$, $\varphi(x, \lambda_1 y_1 + \lambda_2 y_2) = \lambda_1 \varphi(x, y_1) + \lambda_2 \varphi(x, y_2)$.

En d'autres termes, une forme bilinéaire est une application qui est linéaire en chacun des deux arguments dont elle dépend (en l'occurence, des vecteurs de E). A noter que, pour toute forme bilinéaire f sur E, et pour tous $x, y \in E$, on a : $f(x, 0) = f(0, y) = 0$.

Définition 1.2 *Une forme bilinéaire φ sur E est dite :*

1. *symétrique si :* $\forall (x, y) \in E^2$, $\varphi(x, y) = \varphi(y, x)$.
2. *positive si :* $\forall x \in E$, $\varphi(x, x) \geq 0$.
3. *définie si :* $\forall x \in E$, $\varphi(x, x) = 0 \Longrightarrow x = 0$.

Définition 1.3 *Un produit scalaire sur E est une forme bilinéaire symétrique définie positive sur E.*

En général, si φ est un produit scalaire sur E, le réel $\varphi(x, y)$ est noté $(x|y)$ ou $\langle x, y \rangle$. Dans la pratique, pour déterminer qu'une application $\varphi : E \times E \longrightarrow \mathbb{R}$ est un produit scalaire, on doit vérifier quatre conditions, à savoir que φ est bilinéaire, puis symétrique, puis positive et enfin définie.

Définition 1.4 *Soit $\langle \ , \ \rangle$ un produit scalaire sur E. La norme euclidienne sur E associée à $\langle \ , \ \rangle$ est l'application qui, à tout $x \in E$, associe le réel $\|x\| = \sqrt{\langle x, x \rangle}$.*

A noter que la norme d'un vecteur $x \in E$ est par définition le réel $\|x\| = \sqrt{\langle x, x \rangle}$. A présent, on va présenter deux exemples fondamentaux de produits scalaires, à connaître par cœur.

Exemple 1.5 *Un premier exemple de produit scalaire est donné par* le produit scalaire canonique sur $\mathbb{R}^n$, *lequel est défini pour tout $x = (x_1, ..., x_n) \in \mathbb{R}^n$ et $y = (y_1, ..., y_n) \in \mathbb{R}^n$ par :*

$$\langle x, y \rangle = \sum_{k=1}^{n} x_k y_k = x_1 y_1 + ... + x_n y_n.$$

En particulier, la norme associée est définie pour tout $x = (x_1, ..., x_n) \in \mathbb{R}^n$ par : $\|x\| = \sqrt{x_1^2 + ... + x_n^2}$.

Exemple 1.6 *Un deuxième exemple de produit scalaire est donné par* le produit scalaire canonique sur $\mathcal{M}_{n,1}(\mathbb{R})$. *Plus précisément, si l'on identifie toute matrice de $\mathcal{M}_{1,1}(\mathbb{R})$ à un réel, alors le produit scalaire canonique sur $\mathcal{M}_{n,1}(\mathbb{R})$ est défini par :*

$$\forall X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathcal{M}_{n,1}(\mathbb{R}), \quad \forall Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in \mathcal{M}_{n,1}(\mathbb{R}), \quad \langle X, Y \rangle = {}^t X Y = \sum_{k=1}^{n} x_k y_k.$$

La norme associée est définie pour tout $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathcal{M}_{n,1}(\mathbb{R})$ *par :* $\|X\| = \sqrt{{}^t X X} = \sqrt{x_1^2 + ... + x_n^2}$.

A noter que, d'un point de vue géométrique, la norme d'un vecteur est la longueur de ce vecteur (mesurée à l'aide du produit scalaire). En particulier, si $E = \mathbb{R}^2$ ou $E = \mathbb{R}^3$, alors la norme d'un vecteur $x \in E$ (pour le produit scalaire canonique) est la longueur usuelle de x. De façon générale, le produit scalaire canonique sur $\mathbb{R}^2$ ou $\mathbb{R}^3$ n'est ni plus ni moins que le produit scalaire usuel utilisé en géométrie euclidienne, c'est-à-dire la géométrie apprise dans le secondaire. A noter que la distance euclidienne entre deux vecteurs x et y de E est définie comme le réel $d(x, y) = \|x - y\|$.

Théorème 1.7 *Soit $\langle \ , \ \rangle$ un produit scalaire sur E. Alors, pour tous $(x, y) \in E^2$, on a :*

$$\|x + y\|^2 = \|x\|^2 + 2\langle x, y \rangle + \|y\|^2.$$

A noter que cette égalité se retraduit sous la forme suivante (que l'on appelle *identité de polarisation*). Pour tous $(x, y) \in E^2$, on a :

$$\boxed{\langle x, y \rangle = \frac{1}{2}\left[\|x+y\|^2 - \|x\|^2 - \|y\|^2\right].}$$

En particulier, on voit qu'un produit scalaire est complètement déterminé par sa norme associée.

Théorème 1.8 (Inégalité de Cauchy-Schwarz) *Soit $\langle\ ,\ \rangle$ un produit scalaire sur E. Alors, pour tous $(x, y) \in E^2$, on a : $|\langle x, y \rangle| \leq \|x\|.\|y\|$, avec égalité si et seulement si x et y sont colinéaires.*

A noter que l'inégalité de Cauchy-Schwarz s'écrit de façon équivalente sous la forme $\langle x, y \rangle^2 \leq \|x\|^2.\|y\|^2$ pour tous $x, y \in E$. En particulier, pour le produit scalaire canonique sur $\mathbb{R}^n$, l'inégalité de Cauchy-Schwarz peut s'écrire sous la forme suivante :

$$\boxed{\forall((x_1, ..., x_n), (y_1, ..., y_n)) \in (\mathbb{R}^n)^2, \quad \left(\sum_{i=1}^n x_i y_i\right)^2 \leq \left(\sum_{i=1}^n x_i^2\right) \times \left(\sum_{i=1}^n y_i^2\right).}$$

D'un point de vue géométrique, on peut interpréter cette inégalité de la façon suivante. Si $E = \mathbb{R}^2$ ou $E = \mathbb{R}^3$, on sait que le produit scalaire canonique de deux vecteurs $x, y \in E$ est égal au produit des normes (c'est-à-dire des longueurs) des vecteurs x, y et du cosinus de l'angle entre ces deux vecteurs. Du coup, comme un cosinus est toujours compris entre -1 et 1, l'inégalité de Cauchy-Schwarz s'en déduit aussitôt. A noter que le cas d'égalité se produit si et seulement si le cosinus de cet angle est égal à ± 1, c'est-à-dire si l'angle entre x et y est un multiple entier de π, ou en d'autres termes si x et y sont colinéaires.

Théorème 1.9 *Soit $\langle\ ,\ \rangle$ un produit scalaire sur E, et soit $\|\ \|$ la norme associée. Alors :*

1. $\forall x \in E$, $\|x\| = 0 \iff x = 0$.
2. $\forall x \in E$, $\forall \lambda \in \mathbb{R}$, $\|\lambda.x\| = |\lambda|.\|x\|$.
3. $\forall (x, y) \in E^2$, $\|x + y\| \leq \|x\| + \|y\|$ *(Inégalité triangulaire).*

D'un point de vue géométrique, l'inégalité triangulaire signifie que, si l'on considère le triangle construit à l'aide des vecteurs $x, y, x + y$, alors la longueur du "côté" $x + y$ est inférieure ou égale à la somme des longueurs des "côtés" x et y. Dès lors, on voit que l'inégalité triangulaire devient une égalité si et seulement si les vecteurs x et y sont colinéaires et de même sens, c'est-à-dire s'il existe un réel $\lambda \geq 0$ tel que $x = \lambda y$.

2 Orthogonalité

A partir de maintenant, on désigne par E un espace vectoriel muni d'un produit scalaire noté $\langle\ ,\ \rangle$, et par $\|\ \|$ la norme euclidienne associée.

Définition 2.1 *Deux vecteurs x et y de E sont dits orthogonaux si $\langle x, y \rangle = 0$, et on note alors : $x \perp y$.*

A noter que *le seul vecteur orthogonal à tout vecteur de E est le vecteur nul.* En effet, si x est un vecteur orthogonal à tout vecteur de E, alors x est orthogonal à lui-même, ce qui entraine que $\|x\|^2 = \langle x, x \rangle = 0$, et donc $x = 0_E$ (vu que $\langle\ ,\ \rangle$ est un produit scalaire, et qu'il est en particulier défini). A noter aussi que, dans l'espace vectoriel $\mathbb{R}^n$ muni du produit scalaire canonique, les vecteurs de la base canonique sont deux à deux orthogonaux. D'un point de vue géométrique, la notion d'orthogonalité dans $\mathbb{R}^2$ et $\mathbb{R}^3$ (munis du produit scalaire canonique) correspond à celle que l'on connaît depuis le secondaire, à savoir que deux vecteurs du plan ou de l'espace sont orthogonaux si et seulement si l'angle entre ces vecteurs est un angle droit, c'est-à-dire de la forme $\frac{\pi}{2} + k\pi$, où $k \in \mathbb{Z}$. Plus généralement, on a la :

Définition 2.2 *Une famille $(e_1, ..., e_n)$ de E est dite orthogonale si :*

$$\forall (i, j) \in \{1, ..., n\}^2, \quad i \neq j \quad \Longrightarrow \quad \langle e_i, e_j \rangle = 0.$$

Théorème 2.3 (Pythagore)

1. *Pour tous $(x, y) \in E^2$, on a : $x \perp y \iff \|x + y\|^2 = \|x\|^2 + \|y\|^2$.*
2. *Plus généralement, si $(x_1, ..., x_n)$ est une famille orthogonale, alors : $\|\sum_{k=1}^n x_k\|^2 = \sum_{k=1}^n \|x_k\|^2$.*

A noter que la première assertion dans le théorème de Pythagore est bel et bien une équivalence, alors que la deuxième ne l'est pas (il s'agit juste une implication, la réciproque n'étant pas toujours vraie).

Théorème 2.4 *Toute famille orthogonale de E ne contenant pas le vecteur nul est libre.*

Définition 2.5 *Une famille* $(e_1, ..., e_n)$ *de* E *est dite orthonormée (ou orthonormale) si* $(e_1, ..., e_n)$ *est orthogonale et si de plus :* $\forall k \in \{1, ..., n\}$, $\|e_k\| = 1$.

Théorème 2.6 *Toute famille orthonormée de* E *est libre.*

A titre d'exemple, dans l'espace vectoriel $\mathbb{R}^n$ muni du produit scalaire canonique, la base canonique est orthonormale. A noter que, si $n \geq 2$, il existe bien d'autres bases orthonormées dans $\mathbb{R}^n$ que la seule base canonique. A noter aussi qu'un vecteur $u \in E$ est dit *unitaire* si $\|u\| = 1$. En d'autres termes, une famille $(e_1, ..., e_n)$ de E est orthonormale si elle est orthogonale et constituée de vecteurs unitaires. Dans ce qui suit, on va voir que l'on peut toujours "orthonormaliser" une famille libre de vecteurs, c'est-à-dire la transformer en une famille orthonormée. Plus précisément, on a le :

Théorème 2.7 (Orthonormalisation de Schmidt) *Soit* $(e_1, ..., e_n)$ *une famille libre de* E*. Posons :*

$$f_1 = \frac{e_1}{\|e_1\|} \quad \text{et} \quad \forall k \in \{2, ..., n\}, \;\; f_k = \frac{e_k - \sum_{i=1}^{k-1} \langle e_k, f_i \rangle f_i}{\|e_k - \sum_{i=1}^{k-1} \langle e_k, f_i \rangle f_i\|}.$$

Alors la famille $(f_1, ..., f_n)$ *est orthonormée et de plus :* $\forall k \in \{1, ..., n\}$, $\mathrm{Vect}(e_1, ..., e_k) = \mathrm{Vect}(f_1, ..., f_k)$.

A noter que, *si la famille* $(e_1, ..., e_n)$ *est déjà orthonormée, alors le procédé d'orthonormalisation de Schmidt ne modifie pas cette famille* (ce que l'on peut vérifier par récurrence). On retiendra surtout que ce procédé permet de transformer une famille libre de vecteurs donnée en une famille orthonormée qui engendre le même sous-espace vectoriel. A noter aussi que la notion d'orthogonalité se généralise aux sous-espaces vectoriels. Plus précisément :

Définition 2.8 *Deux sous-espaces vectoriels* F *et* G *de* E *sont dits orthogonaux (notation :* $F \perp G$*) si :* $\forall (u, v) \in F \times G$, $\langle u, v \rangle = 0$.

En d'autres termes, F et G sont orthogonaux si tout vecteur de F est orthogonal à tout vecteur de G. On dispose alors d'un critère pour déterminer si deux sous-espaces vectoriels sont orthogonaux, à savoir le :

Théorème 2.9 *Soient* $(e_1, ..., e_p)$ *et* $(f_1, ..., f_q)$ *deux familles de* E*, et posons* $F = \mathrm{Vect}(e_1, ..., e_p)$ *et* $G = \mathrm{Vect}(f_1, ..., f_q)$*. Alors* $F \perp G$ *si et seulement si :* $\forall (i, j) \in \{1, ..., p\} \times \{1, ..., q\}$, $\langle e_i, f_j \rangle = 0$.

Ce résultat permet notamment de vérifier facilement que deux sous-espaces vectoriels sont orthogonaux si l'on en a des familles génératrices.

3 Espaces euclidiens

3.1 Définitions et propriétés de base

Définition 3.1 *On appelle espace euclidien tout espace vectoriel réel* E *de dimension finie muni d'un produit scalaire.*

A partir de maintenant, on désigne par E *un espace euclidien, et par* $\langle \, , \, \rangle$ *le produit scalaire associé.*

Définition 3.2 *Une base orthonormée de* E *est une famille orthonormée et une base de* E*.*

A titre d'exemples, la base canonique de $\mathbb{R}^n$ (resp. $\mathcal{M}_{n,1}(\mathbb{R})$) est une base orthonormée de $\mathbb{R}^n$ (resp. $\mathcal{M}_{n,1}(\mathbb{R})$) pour le produit scalaire canonique. A partir des théorèmes de la base incomplète et d'orthonormalisation de Schmidt, on obtient les :

Théorème 3.3 *Tout espace euclidien* E *non réduit au vecteur nul admet une base orthonormée.*

Théorème 3.4 (de la base incomplète orthonormée) *Toute famille orthonormée de* E *peut être complétée en une base orthonormée de* E*.*

Théorème 3.5 *Toute famille orthonormée formée de* n *vecteurs de* E*, avec* $n = \dim(E)$*, est une base orthonormée de* E*.*

L'un des intérêts d'une base orthonormée $\mathcal{B}$ est que l'on peut facilement calculer les coordonnées d'un vecteur quelconque dans la base $\mathcal{B}$, c'est-à-dire sans avoir à résoudre un système linéaire. Plus précisément :

Théorème 3.6 *Soit $\mathcal{B} = (e_1, ..., e_n)$ une base orthonormée de E. Alors, pour tout $x \in E$, on a :*

$$x = \sum_{k=1}^{n} \langle x, e_k \rangle e_k \quad \textit{et} \quad \|x\|^2 = \sum_{k=1}^{n} \langle x, e_k \rangle^2.$$

En d'autres termes, le réel $\langle x, e_k \rangle$ est la k-ème coordonnée du vecteur x dans la base orthonormée $\mathcal{B}$. A noter que, si l'on identifie les éléments de $\mathcal{M}_{1,1}(\mathbb{R})$ avec les réels, alors on a le :

Théorème 3.7 (Expression matricielle du produit scalaire en base orthonormée) *Soit $\mathcal{B}$ une base orthonormée de E, et soient x, y deux vecteurs de E, de coordonnées respectives $(x_1, ..., x_n), (y_1, ..., y_n)$ dans la base $\mathcal{B}$. Alors :*

$$\langle x, y \rangle = {}^tXY \quad \textit{et} \quad \|x\|^2 = {}^tXX, \quad \textit{avec } X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \textit{ et } Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

En particulier, on obtient que, pour toute base orthonormée $\mathcal{B}$ de E et tous vecteurs x, y de E de coordonnées respectives $(x_1, ..., x_n), (y_1, ..., y_n)$ dans la base $\mathcal{B}$:

$$\boxed{\langle x, y \rangle = \sum_{k=1}^{n} x_k y_k \quad \text{et} \quad \|x\|^2 = \sum_{k=1}^{n} x_k^2.}$$

En d'autres termes, on voit que le calcul du produit scalaire de deux vecteurs x, y de E se ramène à la somme des produits des coordonnées de x, y dans une base orthonormée (comme pour le calcul du produit scalaire canonique). De même, on peut définir le produit scalaire à partir d'une matrice, comme suit :

Définition 3.8 *Soit $\mathcal{B} = (e_1, ..., e_n)$ une base quelconque de E. La matrice du produit scalaire $\langle \ , \ \rangle$ dans la base $\mathcal{B}$ est la matrice $A = (\langle e_i, e_j \rangle)_{1 \leq i,j \leq n}$.*

On dit aussi que $A = (\langle e_i, e_j \rangle)_{1 \leq i,j \leq n}$ est *la matrice de Gram associée à la famille* $(e_1, ..., e_n)$.

Théorème 3.9 (Expression matricielle du produit scalaire en base quelconque) *Soit $\mathcal{B}$ une base quelconque de E, soit A la matrice du produit scalaire dans la base $\mathcal{B}$ et soient x, y deux vecteurs de E, de coordonnées respectives $(x_1, ..., x_n), (y_1, ..., y_n)$ dans la base $\mathcal{B}$. Alors :*

$$\langle x, y \rangle = {}^tXAY \quad \textit{et} \quad \|x\|^2 = {}^tXAX, \quad \textit{avec } X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \textit{ et } Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

En particulier, on voit que, si la base $\mathcal{B}$ est orthonormée, alors $\langle e_i, e_j \rangle = 1$ si $i = j$ et $\langle e_i, e_j \rangle = 0$ si $i \neq j$, et donc $A = I_n$, ce qui nous redonne l'expression matricielle du produit scalaire en base orthonormée. A noter que la matrice d'un produit scalaire dans une base $(e_1, ..., e_n)$ quelconque est toujours symétrique réelle, puisque $\langle e_i, e_j \rangle = \langle e_j, e_i \rangle$ pour tous i, j. A noter aussi que la matrice d'un produit scalaire caractérise complètement ce dernier. Plus précisément, on montre (et nous admettrons) que, *si A, B sont deux matrices symétriques réelles de taille n telles que ${}^tXAY = {}^tXBY$ pour tous $X, Y \in \mathcal{M}_{n,1}(\mathbb{R})$, alors $A = B$.*

3.2 Matrices orthogonales

Définition 3.10 *On appelle matrice orthogonale toute matrice $P \in \mathcal{M}_n(\mathbb{R})$ telle que ${}^tPP = I_n$.*

A noter que, comme ${}^tPP = I_n$, une matrice orthogonale P est toujours inversible, et de plus ${}^tP = P^{-1}$. En particulier, si ${}^tPP = I_n$, alors on a la relation $P{}^tP = I_n$. A noter aussi qu'une matrice $P \in \mathcal{M}_n(\mathbb{R})$ est orthogonale si et seulement si ses vecteurs colonnes forment une famille orthonormée de $\mathcal{M}_{n,1}(\mathbb{R})$ (pour le produit scalaire canonique). Typiquement, les matrices orthogonales interviennent dans les changements de bases orthonormées. Plus précisément, on a le :

Théorème 3.11 *Toute matrice de passage entre deux bases orthonormées de E est orthogonale.*

Réciproquement, on montre (et nous admettrons) que, *si $\dim(E) = n$, alors toute matrice orthogonale $P \in \mathcal{M}_n(\mathbb{R})$ est la matrice de passage entre deux bases orthonormées de E.* En outre, on vérifie facilement par le calcul que *tout produit de matrices orthogonales est une matrice orthogonale.*

3.3 Orthogonal d'un sous-espace vectoriel

Définition 3.12 *Soit F un sous-espace vectoriel de E. Alors l'orthogonal $F^\perp$ de F (dans E) est l'ensemble défini par* $F^\perp = \{x \in E |\ \forall u \in F,\ \langle x, u\rangle = 0\}$.

Théorème 3.13 *Soit F un sous-espace vectoriel de E. Alors $F^\perp$ est un sous-espace vectoriel de E.*

En d'autres termes, l'orthogonal de F est l'ensemble des vecteurs de E orthogonaux à tous les vecteurs de F. A titre d'exemple, on vérifie facilement que $\{0\}^\perp = E$ et $E^\perp = \{0\}$. De plus, on voit que l'orthogonal d'une droite dans le plan est une droite du plan, et l'orthogonal d'un plan dans l'espace est aussi une droite. Dans ce cas, on parle aussi (en géométrie) de droite *normale*.

Théorème 3.14 *Soit F un sous-espace vectoriel de E. Alors $E = F \oplus F^\perp$.*

En d'autres termes, les sous-espaces vectoriels F et $F^\perp$ sont supplémentaires dans E. En particulier, on parle souvent de $F^\perp$ comme du *supplémentaire orthogonal de F* (dans E). Dans la pratique, pour trouver l'orthogonal d'un sous-espace vectoriel, on procède comme dans la démonstration du théorème 3.14. Plus précisément, on commence par déterminer une base orthonormée $(e_1, ..., e_p)$ de F (en partant d'une base de F à laquelle on applique le procédé d'orthonormalisation de Schmidt). Ensuite, on complète cette base de F en une base orthonormée $(e_1, ..., e_n)$ de E (en utilisant d'abord le théorème de la base incomplète, puis en modifiant la base obtenue à l'aide du procédé d'orthonormalisation de Schmidt). Enfin, on conclut en disant que la famille $(e_{p+1}, ..., e_n)$ est une base de l'orthogonal de F dans E.

Corollaire 3.15 *Soit F un sous-espace vectoriel de E. Alors :* $\dim F^\perp = \dim E - \dim F$.

Corollaire 3.16 *Soit F un sous-espace vectoriel de E. Alors :* $(F^\perp)^\perp = F$.