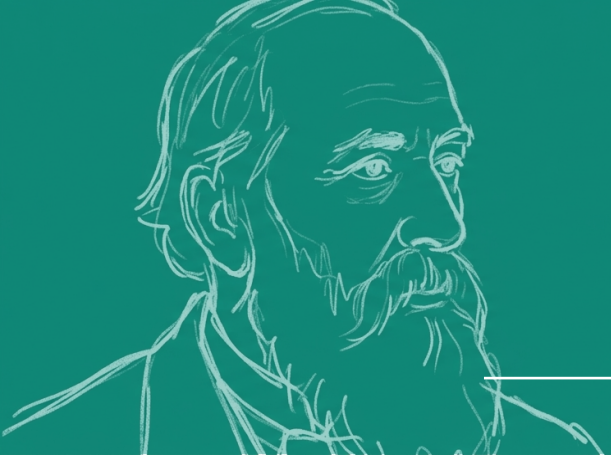


2

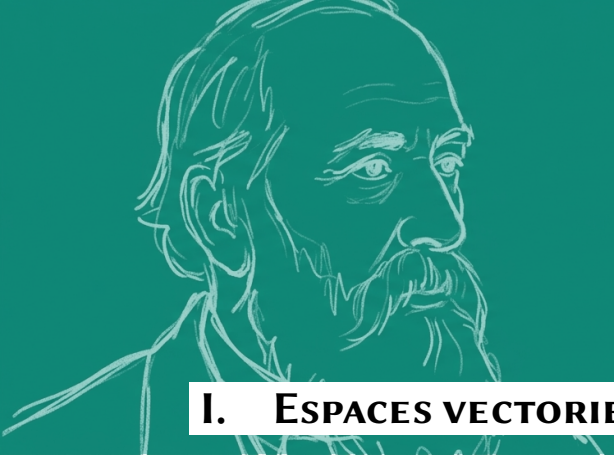
Révisions d'algèbre linéaire

James Joseph Sylvester (1814 – 1897)

L'algèbre linéaire de première année est le socle de presque tout le programme de seconde année. La réduction des endomorphismes, les espaces euclidiens et les polynômes annulateurs en sont le prolongement direct ; mais on la retrouve aussi, plus discrètement, dans l'étude des équations différentielles linéaires ou des fonctions de plusieurs variables. Il est donc indispensable d'y revenir avant d'aller plus loin.

Nous avons retenu trois thèmes. D'abord les espaces vectoriels, avec le vocabulaire des familles libres, des bases et de la dimension, sans lequel aucune des notions de cette année ne peut même être formulée. Ensuite les applications linéaires et leurs matrices, du théorème du rang aux formules de changement de base, ainsi que le calcul des déterminants. Enfin les produits scalaires, l'orthonormalisation et les projections orthogonales.

C'est un chapitre de révisions. Nous ne redémontrons donc pas tout : chaque section rappelle les résultats qu'il faut maîtriser, avant de les mettre en œuvre dans les exercices.



2

Révisions d'algèbre linéaire

I. Espaces vectoriels

1. Généralités

Dans tout ce chapitre, $\mathbb{K}$ désigne $\mathbb{R}$ ou $\mathbb{C}$.

Définition 2.1 (Espace vectoriel)

On appelle $\mathbb{K}$-*espace vectoriel* tout ensemble E muni :

- d'une loi de composition interne $+ : E \times E \longrightarrow E$, appelée *addition* ;
- d'une loi de composition externe $\cdot : \mathbb{K} \times E \longrightarrow E$, appelée *multiplication par un scalaire* ;

vérifiant les propriétés suivantes.

1. L'addition vérifie :
   - $\forall (x, y, z) \in E^3, \quad (x + y) + z = x + (y + z)$ (associativité)
   - $\forall (x, y) \in E^2, \quad x + y = y + x$ (commutativité)
   - il existe $0_E \in E$ tel que $\forall x \in E,\ x + 0_E = x$ (élément neutre)
   - $\forall x \in E,\ \exists x' \in E, \quad x + x' = 0_E$ (opposé)
2. La loi externe vérifie :
   - $\forall x \in E, \quad 1 \cdot x = x$
   - $\forall (\lambda, \mu) \in \mathbb{K}^2,\ \forall x \in E, \quad \lambda \cdot (\mu \cdot x) = (\lambda\mu) \cdot x$
   - $\forall (\lambda, \mu) \in \mathbb{K}^2,\ \forall x \in E, \quad (\lambda + \mu) \cdot x = \lambda \cdot x + \mu \cdot x$
   - $\forall \lambda \in \mathbb{K},\ \forall (x, y) \in E^2, \quad \lambda \cdot (x + y) = \lambda \cdot x + \lambda \cdot y$

Les éléments de E sont appelés *vecteurs*, ceux de $\mathbb{K}$ *scalaires*. L'opposé de x est unique ; on le note $-x$.

Remarque

On écrit λx plutôt que $\lambda \cdot x$ dès que le contexte est clair. De même, on note souvent 0 au lieu de 0_E lorsqu'il n'y a pas d'ambiguïté, mais il faut garder à l'esprit que le vecteur nul 0_E et le scalaire nul $0 \in \mathbb{K}$ ne vivent pas dans le même ensemble.

Remarque

Cette liste d'axiomes peut sembler intimidante. Rassurez-vous : il est *extrêmement* rare d'avoir à la vérifier. En pratique, tous les espaces que vous rencontrerez seront ou bien l'un des exemples de référence ci-dessous, ou bien un sous-espace vectoriel de l'un d'eux — et il est alors bien plus économique de montrer que c'est un sous-espace vectoriel (voir la proposition page 3).

Exemple 2.2

- Pour $n \in \mathbb{N}^*$, $\mathbb{K}^n$ est un $\mathbb{K}$-espace vectoriel pour les lois définies terme à terme ; on y retrouve $\mathbb{R}^n$ et $\mathbb{C}^n$, ainsi que $\mathbb{K}$ lui-même pour $n = 1$.
- Pour $(n, p) \in (\mathbb{N}^*)^2$, $\mathcal{M}_{n,p}(\mathbb{K})$ est un $\mathbb{K}$-espace vectoriel pour les lois définies terme à terme.
- $\mathbb{K}[X]$ est un $\mathbb{K}$-espace vectoriel.
- Pour tout ensemble non vide X, $\mathcal{F}(X, \mathbb{K})$ est un $\mathbb{K}$-espace vectoriel pour les lois définies point par point.
- Soit I un intervalle de $\mathbb{R}$. Pour $n \in \mathbb{N}$, les ensembles $\mathcal{C}^0(I, \mathbb{K})$, $\mathcal{D}(I, \mathbb{K})$, $\mathcal{C}^n(I, \mathbb{K})$ et $\mathcal{C}^\infty(I, \mathbb{K})$ sont des $\mathbb{K}$-espaces vectoriels ; ce n'est qu'une reformulation des théorèmes d'analyse sur les combinaisons linéaires de fonctions.
- $\mathbb{K}^{\mathbb{N}}$ est un $\mathbb{K}$-espace vectoriel pour les lois définies terme à terme ; ce n'est autre que $\mathcal{F}(\mathbb{N}, \mathbb{K})$.
- L'ensemble des suites de $\mathbb{K}^{\mathbb{N}}$ dont la série converge est un $\mathbb{K}$-espace vectoriel.
- Pour Ω fini, l'ensemble des variables aléatoires réelles définies sur Ω est un $\mathbb{R}$-espace vectoriel ; ce n'est autre que $\mathcal{F}(\Omega, \mathbb{R})$.

Remarque

Un même ensemble peut être muni de plusieurs structures d'espace vectoriel. Ainsi $\mathbb{C}$ est un $\mathbb{C}$-espace vectoriel, mais aussi un $\mathbb{R}$-espace vectoriel (en n'autorisant que la multiplication par des réels). Ces deux structures sont bien différentes : la première est de dimension 1, la seconde de dimension 2. Il faut donc toujours préciser si l'on travaille sur $\mathbb{R}$ ou sur $\mathbb{C}$.

2. Sous-espaces vectoriels

Définition 2.3 (Sous-espace vectoriel)

Soit E un $\mathbb{K}$-espace vectoriel et F une partie de E. On dit que F est un *sous-espace vectoriel* de E lorsque F, muni des lois de E restreintes à F, est lui-même un $\mathbb{K}$-espace vectoriel.

Proposition 2.4 (Caractérisation d'un sous-espace vectoriel)

Soit E un $\mathbb{K}$-espace vectoriel et F un ensemble. Alors F est un sous-espace vectoriel de E si et seulement si :

1. $F \subset E$;
2. $F \neq \varnothing$;
3. $\forall (x, y) \in F^2,\ \forall \lambda \in \mathbb{K},\quad \lambda x + y \in F$.

Remarque

En pratique, on ne vérifie jamais le deuxième point sous cette forme : on montre que $0_E \in F$. C'est en effet une condition nécessaire, puisqu'un sous-espace vectoriel de E contient toujours le vecteur nul, et elle est le plus souvent immédiate à vérifier.

Exemple 2.5

- $\{(x, y, z) \in \mathbb{R}^3 \mid x + y + z = 0\}$ est un sous-espace vectoriel de $\mathbb{R}^3$.
- Pour $n \in \mathbb{N}$, $\mathbb{K}_n[X]$ est un sous-espace vectoriel de $\mathbb{K}[X]$.
- L'ensemble des matrices symétriques et l'ensemble des matrices antisymétriques sont des sous-espaces vectoriels de $\mathcal{M}_n(\mathbb{K})$.
- Soit I un intervalle de $\mathbb{R}$. Les ensembles $\mathcal{C}^0(I, \mathbb{K})$, $\mathcal{D}(I, \mathbb{K})$, $\mathcal{C}^n(I, \mathbb{K})$ et $\mathcal{C}^\infty(I, \mathbb{K})$ sont des sous-espaces vectoriels de $\mathcal{F}(I, \mathbb{K})$.
- L'ensemble des suites convergentes et l'ensemble des suites de limite nulle sont des sous-espaces vectoriels de $\mathbb{K}^\mathbb{N}$.
- L'ensemble des suites dont la série converge et l'ensemble des suites dont la série converge absolument sont des sous-espaces vectoriels de $\mathbb{K}^\mathbb{N}$.
- Pour $(a, b) \in \mathbb{K}^2$, l'ensemble des solutions sur I de l'équation différentielle $y'' + ay' + by = 0$ est un sous-espace vectoriel de $\mathcal{C}^2(I, \mathbb{K})$.

Exemple 2.6 (Commutant d'une matrice)

Soit $n \in \mathbb{N}^*$ et $A \in \mathcal{M}_n(\mathbb{K})$. On appelle *commutant* de A l'ensemble des matrices qui commutent avec A :

$$\mathcal{C}(A) = \{M \in \mathcal{M}_n(\mathbb{K}) \mid AM = MA\}.$$

Montrons que $\mathcal{C}(A)$ est un sous-espace vectoriel de $\mathcal{M}_n(\mathbb{K})$ en vérifiant les trois points de la proposition page 3.

1. $\mathcal{C}(A) \subset \mathcal{M}_n(\mathbb{K})$: c'est immédiat, puisque $\mathcal{C}(A)$ est défini comme un ensemble de matrices de $\mathcal{M}_n(\mathbb{K})$.
2. $0 \in \mathcal{C}(A)$: la matrice nulle vérifie $A \times 0 = 0 = 0 \times A$.
3. Soit $(M, N) \in \mathcal{C}(A)^2$ et $\lambda \in \mathbb{K}$. Alors :

$$A(\lambda M + N) = \lambda AM + AN = \lambda MA + NA = (\lambda M + N)A,$$

où l'on a utilisé $AM = MA$ et $AN = NA$. Donc $\lambda M + N \in \mathcal{C}(A)$.

Ainsi, $\mathcal{C}(A)$ est un sous-espace vectoriel de $\mathcal{M}_n(\mathbb{K})$. Il n'est jamais réduit à $\{0\}$: il contient I_n, ainsi que A et toutes ses puissances.

3. Base et dimension

Dans toute cette sous-section, E désigne un $\mathbb{K}$-espace vectoriel.

Définition 2.7 (Famille libre)

Une famille $(v_1, \ldots, v_p)$ de vecteurs de E est *libre* (ou *linéairement indépendante*) si :

$$\sum_{i=1}^{p} \lambda_i v_i = 0_E \implies \lambda_1 = \cdots = \lambda_p = 0.$$

Exemple 2.8

- La famille $((1,2),(2,1))$ est libre dans $\mathbb{R}^2$.
- La famille $((1,0),(0,1),(1,1))$ n'est pas libre dans $\mathbb{R}^2$: on a en effet $(1,0)+(0,1)-(1,1)=(0,0)$, ce qui est une combinaison linéaire nulle à coefficients non tous nuls.
- Toute famille $(P_0,P_1,\dots,P_n)$ de polynômes non nuls *échelonnée en degré*, c'est-à-dire telle que $\deg P_0 < \deg P_1 < \cdots < \deg P_n$, est libre dans $\mathbb{K}[X]$.

Remarque

Une famille (v) réduite à un vecteur est libre si et seulement si $v \neq 0_E$. Une famille (u,v) de deux vecteurs est libre si et seulement si u et v ne sont pas colinéaires. Attention : cela ne se généralise pas. À partir de trois vecteurs, une famille peut être formée de vecteurs deux à deux non colinéaires sans être libre : c'est le cas de $((1,0),(0,1),(1,1))$ ci-dessus.

Exemple 2.9 (La famille (exp, sin, cos) est libre)

Montrons que $(\exp,\sin,\cos)$ est libre dans $\mathcal{F}(\mathbb{R},\mathbb{R})$. Soit $(a,b,c) \in \mathbb{R}^3$ tel que $a\exp + b\sin + c\cos = 0$, c'est-à-dire :

$$\forall t \in \mathbb{R}, \quad ae^t + b\sin(t) + c\cos(t) = 0.$$

Évaluons cette relation en trois points bien choisis :

- en $t=0$: $a+c=0$;
- en $t=\frac{\pi}{2}$: $ae^{\pi/2}+b=0$;
- en $t=\pi$: $ae^{\pi}-c=0$.

La première égalité donne $c=-a$. En reportant dans la troisième, il vient $ae^{\pi}+a=0$, soit $a(e^{\pi}+1)=0$. Comme $e^{\pi}+1\neq 0$, on obtient $a=0$, puis $c=-a=0$, et enfin $b=-ae^{\pi/2}=0$ grâce à la deuxième égalité.
Les trois coefficients sont nuls : la famille est libre.

Définition 2.10 (Famille génératrice)

Une famille $(v_1,\dots,v_p)$ de vecteurs de E est *génératrice* si tout vecteur de E est combinaison linéaire des v_i :

$$\forall x \in E,\ \exists(\lambda_1,\dots,\lambda_p)\in\mathbb{K}^p, \quad x=\sum_{i=1}^{p}\lambda_i v_i.$$

Exemple 2.11

- Pour $1\leq i\leq n$, soit e_i le n-uplet de $\mathbb{K}^n$ dont la i-ième composante vaut 1 et les autres 0. La famille $(e_1,\dots,e_n)$ est génératrice de $\mathbb{K}^n$, puisque $(x_1,\dots,x_n)=\sum_{i=1}^n x_i e_i$.
- La famille $(1,X,\dots,X^n)$ est génératrice de $\mathbb{K}_n[X]$.
- Pour $1\leq i\leq n$ et $1\leq j\leq p$, soit $E_{i,j}$ la matrice de $\mathcal{M}_{n,p}(\mathbb{K})$ dont le coefficient d'indice (i,j) vaut 1 et les autres 0. La famille $(E_{i,j})$ est génératrice de $\mathcal{M}_{n,p}(\mathbb{K})$.

— En notant $j = e^{2i\pi/3}$, la famille $(1, j, j^2)$ est génératrice de $\mathbb{C}$ vu comme $\mathbb{R}$-espace vectoriel, mais elle n'est pas libre : on a $1 + j + j^2 = 0$. Une famille génératrice n'est donc pas nécessairement libre.

Définition 2.12 (Base)

Une famille est une *base* de E si elle est à la fois libre et génératrice.

Définition 2.13 (Dimension)

On dit que E est de *dimension finie* s'il admet une famille génératrice finie. Dans ce cas, E admet des bases, et elles ont toutes le même cardinal : ce cardinal est appelé la *dimension* de E et se note $\dim E$.

Exemple 2.14 (Bases canoniques)

— $(e_1, \dots, e_n)$ est une base de $\mathbb{K}^n$, appelée *base canonique* ; ainsi $\dim \mathbb{K}^n = n$.

— $(E_{i,j})_{1 \leq i \leq n,\, 1 \leq j \leq p}$ est une base de $\mathcal{M}_{n,p}(\mathbb{K})$; ainsi $\dim \mathcal{M}_{n,p}(\mathbb{K}) = np$.

— $(1, X, \dots, X^n)$ est une base de $\mathbb{K}_n[X]$; ainsi $\dim \mathbb{K}_n[X] = n + 1$.

Proposition 2.15 (Caractérisations d'une base en dimension finie)

Soit E un espace vectoriel de dimension finie n et $(v_1, \dots, v_p)$ une famille de vecteurs de E.

— Si $(v_1, \dots, v_p)$ est libre, alors c'est une base de E si et seulement si $p = n$.

— Si $(v_1, \dots, v_p)$ est génératrice, alors c'est une base de E si et seulement si $p = n$.

Exemple 2.16 (Une base de $\mathbb{R}^2$)

Montrons que $((1, 1), (1, -1))$ est une base de $\mathbb{R}^2$.
Cette famille compte 2 vecteurs et $\dim \mathbb{R}^2 = 2$. D'après la proposition précédente, il suffit donc de montrer qu'elle est libre.
Soit $(a, b) \in \mathbb{R}^2$ tel que $a(1, 1) + b(1, -1) = (0, 0)$. En identifiant les composantes, cela s'écrit :

$$\begin{cases} a + b = 0, \\ a - b = 0. \end{cases}$$

En sommant ces deux équations, on obtient $2a = 0$, donc $a = 0$; la première équation donne alors $b = 0$. La famille est donc libre.
Étant libre et de cardinal $2 = \dim \mathbb{R}^2$, c'est une base de $\mathbb{R}^2$.

Proposition 2.17

Soit E un espace vectoriel de dimension finie n et $(v_1, \dots, v_p)$ une famille de vecteurs de E.

— Si $p > n$, alors la famille n'est pas libre.

— Si $p < n$, alors la famille n'est pas génératrice.

Exemple 2.18

Dans $\mathbb{R}^2$, qui est de dimension 2 :

- la famille $((1,2),(3,4),(5,6))$ n'est pas libre, car elle compte $3 > 2$ vecteurs ;
- la famille $((1,1))$ n'est pas génératrice, car elle compte $1 < 2$ vecteur.

Aucun calcul n'est nécessaire pour le justifier.

Remarque

Une famille n'a pas de dimension : elle a un *cardinal*, c'est-à-dire un nombre de vecteurs. La dimension est une propriété des *espaces vectoriels*. On ne dira donc jamais « ~~la dimension de la famille~~ », mais « le cardinal de la famille ».

Proposition 2.19 (Dimension d'un sous-espace vectoriel)

Soit E un espace vectoriel de dimension finie et F un sous-espace vectoriel de E. Alors F est de dimension finie et :

$$\dim F \leq \dim E.$$

De plus, $\dim F = \dim E$ si et seulement si $F = E$.

Définition 2.20 (Sous-espace engendré)

Soit A une partie de E. On appelle *sous-espace vectoriel engendré par* A, et on note $\mathrm{Vect}(A)$, l'intersection de tous les sous-espaces vectoriels de E contenant A :

$$\mathrm{Vect}(A) = \bigcap_{\substack{F \text{ sous-espace vectoriel de } E \\ A \subset F}} F.$$

Lorsque $A = \{v_1, \ldots, v_p\}$ est fini, on écrit simplement $\mathrm{Vect}(v_1, \ldots, v_p)$.

Proposition 2.21

Soit A une partie de E. Alors $\mathrm{Vect}(A)$ est un sous-espace vectoriel de E contenant A, et c'est le plus petit au sens de l'inclusion : si F est un sous-espace vectoriel de E tel que $A \subset F$, alors $\mathrm{Vect}(A) \subset F$.

Démonstration

Notons $\mathcal{S}$ l'ensemble des sous-espaces vectoriels de E contenant A. Comme E est un sous-espace vectoriel de lui-même et contient A, on a $E \in \mathcal{S}$: cet ensemble n'est pas vide, et $\mathrm{Vect}(A) = \bigcap_{F \in \mathcal{S}} F$ est bien défini.
Vérifions les trois points de la proposition page 3.

1. $\mathrm{Vect}(A) \subset E$: puisque $E \in \mathcal{S}$, l'intersection est en particulier incluse dans E.
2. $0_E \in \mathrm{Vect}(A)$: tout $F \in \mathcal{S}$ est un sous-espace vectoriel de E, donc contient 0_E. Le vecteur nul appartient donc à tous les F, donc à leur intersection.
3. Soient $(x, y) \in \mathrm{Vect}(A)^2$ et $\lambda \in \mathbb{K}$. Soit $F \in \mathcal{S}$. Comme $x \in F$, $y \in F$ et que F est un sous-espace vectoriel, on a $\lambda x + y \in F$. Ceci valant pour tout $F \in \mathcal{S}$, on obtient $\lambda x + y \in \mathrm{Vect}(A)$.

Donc $\mathrm{Vect}(A)$ est un sous-espace vectoriel de E.
De plus, tout $F \in \mathcal{S}$ contient A, donc leur intersection aussi : $A \subset \mathrm{Vect}(A)$.

Enfin, si F est un sous-espace vectoriel de E contenant A, alors $F \in \mathcal{S}$, et donc $\text{Vect}(A) \subset F$ par définition d'une intersection. □

Proposition 2.22

Soit $(v_1, \dots, v_p)$ une famille de vecteurs de E. Alors :

$$\text{Vect}(v_1, \dots, v_p) = \left\{ \sum_{i=1}^{p} \lambda_i v_i \mid (\lambda_1, \dots, \lambda_p) \in \mathbb{K}^p \right\}.$$

Démonstration

Notons $G = \left\{ \sum_{i=1}^{p} \lambda_i v_i \mid (\lambda_1, \dots, \lambda_p) \in \mathbb{K}^p \right\}$.
Montrons d'abord que G est un sous-espace vectoriel de E.

1. $G \subset E$: une combinaison linéaire de vecteurs de E est un vecteur de E.
2. $0_E \in G$: il suffit de prendre tous les λ_i nuls.
3. Soient $x = \sum_{i=1}^{p} \lambda_i v_i$ et $y = \sum_{i=1}^{p} \mu_i v_i$ deux éléments de G, et $\alpha \in \mathbb{K}$. Alors :

$$\alpha x + y = \sum_{i=1}^{p} (\alpha \lambda_i + \mu_i) v_i \in G.$$

Donc G est un sous-espace vectoriel de E.
Il contient chacun des v_j : il suffit de prendre $\lambda_j = 1$ et $\lambda_i = 0$ pour $i \neq j$. Ainsi $\{v_1, \dots, v_p\} \subset G$, et la minimalité de $\text{Vect}(v_1, \dots, v_p)$ donne $\text{Vect}(v_1, \dots, v_p) \subset G$.
Réciproquement, $\text{Vect}(v_1, \dots, v_p)$ est un sous-espace vectoriel qui contient $v_1, \dots, v_p$; il est donc stable par combinaison linéaire, et une récurrence immédiate montre qu'il contient toute somme $\sum_{i=1}^{p} \lambda_i v_i$. Ainsi $G \subset \text{Vect}(v_1, \dots, v_p)$.
Les deux inclusions donnent l'égalité. □

Remarque

Une famille $(v_1, \dots, v_p)$ de vecteurs de E est génératrice de E si et seulement si $\text{Vect}(v_1, \dots, v_p) = E$.

Remarque (Interprétation d'une famille libre)

Soit $(v_1, \dots, v_p)$ une famille *liée*, c'est-à-dire non libre. Il existe alors des scalaires $\lambda_1, \dots, \lambda_p$ *non tous nuls* tels que :

$$\sum_{i=1}^{p} \lambda_i v_i = 0_E.$$

Quitte à renuméroter les vecteurs, supposons $\lambda_p \neq 0$. On peut alors diviser par λ_p et isoler v_p :

$$v_p = -\frac{1}{\lambda_p} \sum_{i=1}^{p-1} \lambda_i v_i, \qquad \text{c'est-à-dire} \qquad v_p \in \text{Vect}(v_1, \dots, v_{p-1}).$$

Cela traduit l'intuition attachée à ces notions : une famille est liée lorsque l'un de ses vecteurs ne « sort » pas de l'espace engendré par les autres. De manière équivalente, une famille est libre lorsque chacun de ses vecteurs « sort » de l'espace engendré par les autres : il apporte une direction nouvelle.

4. Somme directe et supplémentaires

Définition 2.23 (Somme, somme directe)

Soient F et G deux sous-espaces vectoriels de E. On appelle *somme* de F et G l'ensemble :

$$F + G = \{x + y \mid x \in F,\ y \in G\}.$$

C'est un sous-espace vectoriel de E.
On dit que F et G sont *en somme directe* lorsque $F \cap G = \{0_E\}$; on note alors leur somme $F \oplus G$.

Remarque

Cette définition est propre au cas de *deux* sous-espaces vectoriels : la condition $F \cap G = \{0_E\}$ ne se généralise pas telle quelle à trois sous-espaces ou plus. Nous verrons plus tard dans l'année une définition de la somme directe valable pour un nombre quelconque de sous-espaces.

Exemple 2.24 (Matrices symétriques et antisymétriques)

Soit $n \in \mathbb{N}^*$. Notons $\mathcal{S}_n(\mathbb{K})$ l'ensemble des matrices symétriques et $\mathcal{A}_n(\mathbb{K})$ l'ensemble des matrices antisymétriques de $\mathcal{M}_n(\mathbb{K})$. Montrons que leur somme est directe.
Soit $M \in \mathcal{S}_n(\mathbb{K}) \cap \mathcal{A}_n(\mathbb{K})$. Alors $M^\top = M$ et $M^\top = -M$, d'où $M = -M$, puis $2M = 0$ et donc $M = 0$.
Ainsi $\mathcal{S}_n(\mathbb{K}) \cap \mathcal{A}_n(\mathbb{K}) = \{0\}$: la somme $\mathcal{S}_n(\mathbb{K}) \oplus \mathcal{A}_n(\mathbb{K})$ est directe.

Exemple 2.25 (Fonctions d'intégrale nulle)

Soit $E = \mathcal{C}^0([0,1], \mathbb{R})$. Posons :

$$F = \left\{ f \in E \;\middle|\; \int_0^1 f(t)\,\mathrm{d}t = 0 \right\} \qquad \text{et} \qquad G = \mathrm{Vect}(\mathbb{1}),$$

où $\mathbb{1}$ désigne la fonction constante égale à 1. Montrons que leur somme est directe.
Soit $f \in F \cap G$. Comme $f \in G$, il existe $\lambda \in \mathbb{R}$ tel que $f = \lambda \mathbb{1}$. Comme $f \in F$:

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

Donc $\lambda = 0$, puis $f = 0$. Ainsi $F \cap G = \{0\}$: la somme $F \oplus G$ est directe.

Proposition 2.26 (Formule de Grassmann)

Pour deux sous-espaces vectoriels F et G de dimension finie :

$$\dim(F + G) = \dim F + \dim G - \dim(F \cap G).$$

Démonstration

Notons $p = \dim F$, $q = \dim G$ et $r = \dim(F \cap G)$. Comme $F \cap G$ est un sous-espace vectoriel de F et de G, on a $r \leq p$ et $r \leq q$.

Soit $(u_1, \dots, u_r)$ une base de $F \cap G$. Cette famille est libre dans F ; d'après le théorème de la base incomplète, on peut la compléter en une base de F :

$$(u_1, \dots, u_r, f_1, \dots, f_{p-r}).$$

Elle est également libre dans G, et on peut de même la compléter en une base de G :

$$(u_1, \dots, u_r, g_1, \dots, g_{q-r}).$$

(Si $F \cap G = \{0_E\}$, alors $r = 0$ et la famille $(u_1, \dots, u_r)$ est la famille vide ; ce qui suit reste valable.)
Montrons que la famille :

$$\mathcal{B} = (u_1, \dots, u_r, f_1, \dots, f_{p-r}, g_1, \dots, g_{q-r})$$

est une base de $F + G$.
$\mathcal{B}$ est génératrice de $F + G$. Chacun de ses vecteurs appartient à F ou à G, donc à $F + G$. Réciproquement, soit $x \in F+G$: il existe $y \in F$ et $z \in G$ tels que $x = y+z$. Or y est combinaison linéaire de $(u_1, \dots, u_r, f_1, \dots, f_{p-r})$ et z est combinaison linéaire de $(u_1, \dots, u_r, g_1, \dots, g_{q-r})$. Donc $x = y + z$ est combinaison linéaire des vecteurs de $\mathcal{B}$.
$\mathcal{B}$ est libre. Soient des scalaires a_i, b_j, c_k tels que :

$$\sum_{i=1}^{r} a_i u_i + \sum_{j=1}^{p-r} b_j f_j + \sum_{k=1}^{q-r} c_k g_k = 0_E.$$

Posons $w = \sum_{k=1}^{q-r} c_k g_k$. D'une part $w \in G$, puisque les g_k sont dans G. D'autre part, la relation ci-dessus donne :

$$w = -\sum_{i=1}^{r} a_i u_i - \sum_{j=1}^{p-r} b_j f_j,$$

et les u_i et les f_j sont dans F, donc $w \in F$. Ainsi $w \in F \cap G$, et w se décompose dans la base $(u_1, \dots, u_r)$ de $F \cap G$: il existe des scalaires d_i tels que $w = \sum_{i=1}^{r} d_i u_i$. On en déduit :

$$\sum_{i=1}^{r} d_i u_i - \sum_{k=1}^{q-r} c_k g_k = 0_E.$$

Comme $(u_1, \dots, u_r, g_1, \dots, g_{q-r})$ est une base de G, elle est libre, donc tous les d_i et tous les c_k sont nuls. La relation initiale se réduit alors à :

$$\sum_{i=1}^{r} a_i u_i + \sum_{j=1}^{p-r} b_j f_j = 0_E,$$

et comme $(u_1, \dots, u_r, f_1, \dots, f_{p-r})$ est une base de F, elle est libre : tous les a_i et tous les b_j sont nuls. La famille $\mathcal{B}$ est donc libre.
Ainsi $\mathcal{B}$ est une base de $F + G$, et elle compte $r + (p - r) + (q - r)$ vecteurs. D'où :

$$\dim(F + G) = p + q - r = \dim F + \dim G - \dim(F \cap G).$$

□

Exemple 2.27

Dans $\mathbb{K}[X]$, posons $F = \mathrm{Vect}(1, X)$ et $G = \mathrm{Vect}(1, X^2)$. On a $\dim F = 2$ et $\dim G = 2$. Déterminons $F \cap G$. Un élément de F s'écrit $a + bX$ et un élément de G s'écrit $c + dX^2$, avec $(a, b, c, d) \in \mathbb{K}^4$. L'égalité $a + bX = c + dX^2$ impose, par identification des coefficients, $b = 0$, $d = 0$ et $a = c$. Donc $F \cap G = \mathrm{Vect}(1)$, qui est de dimension 1.

La formule de Grassmann donne alors :

$$\dim(F+G) = 2+2-1 = 3,$$

ce que l'on retrouve directement puisque $F+G = \mathrm{Vect}(1, X, X^2) = \mathbb{K}_2[X]$.

Définition 2.28 (Supplémentaires)

Soient F et G deux sous-espaces vectoriels de E. On dit que F et G sont *supplémentaires dans* E lorsque :

1. $F \cap G = \{0_E\}$, c'est-à-dire que la somme de F et G est directe ;
2. $F + G = E$, c'est-à-dire que cette somme est E tout entier.

On note alors $E = F \oplus G$. Cela équivaut à dire que tout $x \in E$ se décompose de manière unique en $x = x_F + x_G$ avec $x_F \in F$ et $x_G \in G$.

Remarque

Être supplémentaires n'a de sens que *relativement à un espace*. On ne dit jamais que F et G sont supplémentaires : on dit qu'ils sont supplémentaires *dans* E.

Exemple 2.29 (Fonctions d'intégrale nulle, suite)

Reprenons $E = \mathcal{C}^0([0,1], \mathbb{R})$, ainsi que F et $G = \mathrm{Vect}(\mathbb{1})$ définis plus haut. Nous avons déjà montré que la somme de F et G est directe ; montrons que $F + G = E$.
Soit $f \in E$. Posons $\lambda = \int_0^1 f(t)\,\mathrm{d}t$ et écrivons :

$$f = (f - \lambda\mathbb{1}) + \lambda\mathbb{1}.$$

D'une part $\lambda\mathbb{1} \in G$. D'autre part :

$$\int_0^1 (f(t) - \lambda)\,\mathrm{d}t = \int_0^1 f(t)\,\mathrm{d}t - \lambda = 0,$$

donc $f - \lambda\mathbb{1} \in F$. Ainsi $f \in F + G$, et $E = F + G$.
Finalement, F et G sont supplémentaires dans E.

Proposition 2.30 (Caractérisation en dimension finie)

Soit E un espace vectoriel de dimension finie et F, G deux sous-espaces vectoriels de E. Alors :

$$E = F \oplus G \iff \begin{cases} F \cap G = \{0_E\}, \\ \dim F + \dim G = \dim E. \end{cases}$$

Démonstration

Comme E est de dimension finie, F et G le sont aussi, et la formule de Grassmann s'applique :

$$\dim(F+G) = \dim F + \dim G - \dim(F \cap G).$$

Sens direct. Supposons $E = F \oplus G$. Par définition, $F \cap G = \{0_E\}$, ce qui est la première condition. De plus $\dim(F \cap G) = \dim\{0_E\} = 0$ et $F + G = E$, donc la formule de Grassmann donne :

$$\dim E = \dim(F + G) = \dim F + \dim G - 0 = \dim F + \dim G.$$

Sens réciproque. Supposons $F \cap G = \{0_E\}$ et $\dim F + \dim G = \dim E$. Alors $\dim(F \cap G) = 0$, et la formule de Grassmann donne :

$$\dim(F + G) = \dim F + \dim G - 0 = \dim E.$$

Or $F + G$ est un sous-espace vectoriel de E, et un sous-espace vectoriel de E ayant la même dimension que E est égal à E. Donc $F + G = E$. Avec $F \cap G = \{0_E\}$, on conclut que $E = F \oplus G$. □

Exemple 2.31 (Matrices symétriques et antisymétriques, suite)

Reprenons $\mathcal{S}_n(\mathbb{K})$ et $\mathcal{A}_n(\mathbb{K})$. Nous avons déjà montré que $\mathcal{S}_n(\mathbb{K}) \cap \mathcal{A}_n(\mathbb{K}) = \{0\}$. Par ailleurs :

$$\dim \mathcal{S}_n(\mathbb{K}) = \frac{n(n+1)}{2}, \qquad \dim \mathcal{A}_n(\mathbb{K}) = \frac{n(n-1)}{2},$$

de sorte que :

$$\dim \mathcal{S}_n(\mathbb{K}) + \dim \mathcal{A}_n(\mathbb{K}) = \frac{n(n+1)}{2} + \frac{n(n-1)}{2} = n^2 = \dim \mathcal{M}_n(\mathbb{K}).$$

D'après la proposition précédente, $\mathcal{S}_n(\mathbb{K})$ et $\mathcal{A}_n(\mathbb{K})$ sont supplémentaires dans $\mathcal{M}_n(\mathbb{K})$.

Méthode 2.32 (Montrer que deux sous-espaces sont supplémentaires)

Soient F et G deux sous-espaces vectoriels de E.

- **Si E est de dimension finie** et que l'on connaît $\dim F$, $\dim G$ et $\dim E$: montrer que $F \cap G = \{0_E\}$, puis que $\dim F + \dim G = \dim E$.
- **Sinon** (si E n'est pas de dimension finie, ou si les dimensions ne sont pas connues) : raisonner par *analyse-synthèse*. On se donne $x \in E$ et l'on cherche à l'écrire $x = x_F + x_G$ avec $x_F \in F$ et $x_G \in G$.
  - *Analyse* : on suppose une telle écriture donnée et on en déduit les seules valeurs possibles de x_F et x_G. Cela établit l'unicité.
  - *Synthèse* : on vérifie que les valeurs ainsi obtenues conviennent effectivement. Cela établit l'existence.

Exemple 2.33 (Fonctions paires et impaires)

Soit $E = \mathcal{F}(\mathbb{R}, \mathbb{R})$, P l'ensemble des fonctions paires et I l'ensemble des fonctions impaires. Comme E n'est pas de dimension finie, raisonnons par analyse-synthèse. *Analyse.* Soit $f \in E$ et supposons $f = p + i$ avec $p \in P$ et $i \in I$. Alors, pour tout $x \in \mathbb{R}$:

$$f(x) = p(x) + i(x) \qquad \text{et} \qquad f(-x) = p(-x) + i(-x) = p(x) - i(x).$$

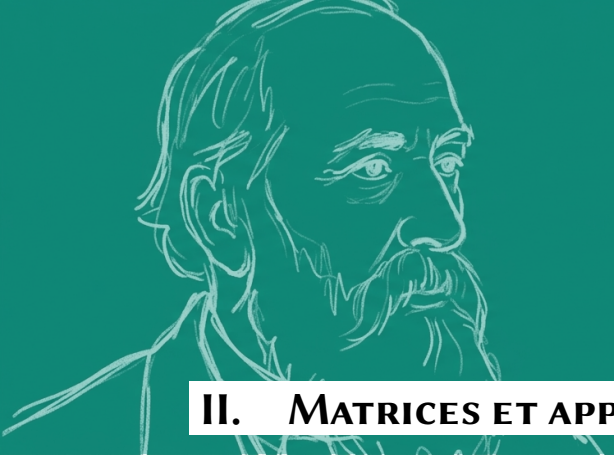
En sommant puis en soustrayant ces deux égalités, on obtient nécessairement :

$$p(x) = \frac{f(x) + f(-x)}{2} \qquad \text{et} \qquad i(x) = \frac{f(x) - f(-x)}{2}.$$

La décomposition, si elle existe, est donc unique.
Synthèse. Définissons p et i par les formules ci-dessus. Pour tout $x \in \mathbb{R}$:

$$p(-x) = \frac{f(-x) + f(x)}{2} = p(x), \qquad i(-x) = \frac{f(-x) - f(x)}{2} = -i(x),$$

donc $p \in P$ et $i \in I$. Enfin $p(x) + i(x) = f(x)$ pour tout $x \in \mathbb{R}$, c'est-à-dire $f = p + i$.
L'existence et l'unicité de la décomposition donnent $\mathcal{F}(\mathbb{R}, \mathbb{R}) = P \oplus I$.



2

Révisions d'algèbre linéaire

II. Matrices et applications linéaires

1. Généralités

Dans toute cette sous-section, E, F et G désignent des $\mathbb{K}$-espaces vectoriels.

Définition 2.34 (Application linéaire)

Une application $f : E \to F$ est *linéaire* lorsque :

$$\forall (x, y) \in E^2,\ \forall \lambda \in \mathbb{K}, \quad f(\lambda x + y) = \lambda f(x) + f(y).$$

On note $\mathcal{L}(E, F)$ l'ensemble des applications linéaires de E dans F.

Remarque

Cette condition unique remplace les deux conditions $f(x + y) = f(x) + f(y)$ et $f(\lambda x) = \lambda f(x)$: elles lui sont équivalentes. En prenant $\lambda = -1$ et $y = x$, on obtient en particulier $f(0_E) = 0_F$, ce qui fournit un test de non-linéarité très économique.

Définition 2.35 (Vocabulaire)

Soit $f : E \to F$ une application linéaire.

- Si $F = E$, on dit que f est un *endomorphisme* de E ; on note $\mathcal{L}(E) = \mathcal{L}(E, E)$.
- Si f est bijective, on dit que f est un *isomorphisme*, et que E et F sont *isomorphes*.
- Si f est un endomorphisme bijectif, on dit que f est un *automorphisme* de E.
- Si $F = \mathbb{K}$, on dit que f est une *forme linéaire* sur E.

Exemple 2.36

- L'application $f : \mathbb{R}^2 \to \mathbb{R}^2$ définie par $f(x, y) = (x + y,\ x - y)$ est linéaire.
- L'application $\mathrm{Id}_E : x \mapsto x$ est un automorphisme de E, et plus généralement $x \mapsto \lambda x$ pour $\lambda \in \mathbb{K}^*$ (*homothétie*).
- L'application $D : \mathbb{K}[X] \to \mathbb{K}[X]$, $P \mapsto P'$, est un endomorphisme de $\mathbb{K}[X]$.
- L'application $\Phi : C^0(\mathbb{R}, \mathbb{R}) \to C^1(\mathbb{R}, \mathbb{R})$ définie par

$$\Phi(f) : x \longmapsto \int_0^x f(t)\,\mathrm{d}t$$

  est linéaire : elle est bien définie par le théorème fondamental de l'analyse, et sa linéarité est celle de l'intégrale.
- $P \mapsto P(0)$ est une forme linéaire sur $\mathbb{K}[X]$, et $f \mapsto \int_0^1 f(t)\,\mathrm{d}t$ est une forme linéaire sur $C^0([0, 1], \mathbb{R})$.

Proposition 2.37 (Opérations sur les applications linéaires)

- $\mathcal{L}(E,F)$ est un sous-espace vectoriel de $\mathcal{F}(E,F)$; c'est en particulier un $\mathbb{K}$-espace vectoriel.
- Si $f \in \mathcal{L}(E,F)$ et $g \in \mathcal{L}(F,G)$, alors $g \circ f \in \mathcal{L}(E,G)$.
- Si $f : E \to F$ est un isomorphisme, alors f^{-1} est linéaire ; c'est donc un isomorphisme de F sur E.

Démonstration

Sous-espace vectoriel. L'application nulle est linéaire, donc $\mathcal{L}(E,F)$ est une partie non vide de $\mathcal{F}(E,F)$. Soient f et g dans $\mathcal{L}(E,F)$, $\mu \in \mathbb{K}$, puis $(x,y) \in E^2$ et $\lambda \in \mathbb{K}$:

$$(\mu f + g)(\lambda x + y) = \mu\big(\lambda f(x) + f(y)\big) + \lambda g(x) + g(y) = \lambda(\mu f + g)(x) + (\mu f + g)(y).$$

Composée. Soient $f \in \mathcal{L}(E,F)$, $g \in \mathcal{L}(F,G)$, $(x,y) \in E^2$ et $\lambda \in \mathbb{K}$:

$$(g \circ f)(\lambda x + y) = g\big(\lambda f(x) + f(y)\big) = \lambda(g \circ f)(x) + (g \circ f)(y).$$

Réciproque. Soit f un isomorphisme, et soient $(u,v) \in F^2$ et $\lambda \in \mathbb{K}$. Posons $x = f^{-1}(u)$ et $y = f^{-1}(v)$. Par linéarité de f, on a $f(\lambda x + y) = \lambda u + v$, d'où

$$f^{-1}(\lambda u + v) = \lambda x + y = \lambda f^{-1}(u) + f^{-1}(v).$$

□

Définition 2.38 (Noyau, image)

Soit $f : E \to F$ une application linéaire. On appelle :

- *noyau* de f l'ensemble $\operatorname{Ker} f = \{x \in E \mid f(x) = 0_F\}$;
- *image* de f l'ensemble $\operatorname{Im} f = \{f(x) \mid x \in E\}$.

Proposition 2.39

Soit $f : E \to F$ une application linéaire. Alors $\operatorname{Ker} f$ est un sous-espace vectoriel de E et $\operatorname{Im} f$ est un sous-espace vectoriel de F.

Démonstration

Vérifions les trois points de la caractérisation des sous-espaces vectoriels.
Pour $\operatorname{Ker} f$. Par construction $\operatorname{Ker} f \subset E$. Comme $f(0_E) = 0_F$, on a $0_E \in \operatorname{Ker} f$. Enfin, soient $(x,y) \in (\operatorname{Ker} f)^2$ et $\lambda \in \mathbb{K}$:

$$f(\lambda x + y) = \lambda f(x) + f(y) = \lambda \cdot 0_F + 0_F = 0_F,$$

donc $\lambda x + y \in \operatorname{Ker} f$.
Pour $\operatorname{Im} f$. Par construction $\operatorname{Im} f \subset F$, et $0_F = f(0_E) \in \operatorname{Im} f$. Soient enfin $(u,v) \in (\operatorname{Im} f)^2$ et $\lambda \in \mathbb{K}$: il existe x et y dans E tels que $u = f(x)$ et $v = f(y)$, et alors

$$\lambda u + v = \lambda f(x) + f(y) = f(\lambda x + y) \in \operatorname{Im} f.$$

□

Proposition 2.40 (Caractérisation de l'injectivité)

Soit $f : E \to F$ une application linéaire. Alors :

$$f \text{ est injective} \iff \operatorname{Ker} f = \{0_E\}.$$

Démonstration

Sens direct. Supposons f injective. L'inclusion $\{0_E\} \subset \operatorname{Ker} f$ est acquise puisque $f(0_E) = 0_F$. Réciproquement, soit $x \in \operatorname{Ker} f$: alors $f(x) = 0_F = f(0_E)$, et l'injectivité de f donne $x = 0_E$.
Sens réciproque. Supposons $\operatorname{Ker} f = \{0_E\}$ et soient $(x, y) \in E^2$ tels que $f(x) = f(y)$. Par linéarité :

$$f(x - y) = f(x) - f(y) = 0_F,$$

donc $x - y \in \operatorname{Ker} f = \{0_E\}$, c'est-à-dire $x = y$. Ainsi f est injective. □

Remarque

C'est un gain considérable : pour une application quelconque, l'injectivité se teste sur *tous* les couples (x, y) ; pour une application linéaire, il suffit de résoudre l'équation $f(x) = 0_F$.

Exemple 2.41

Reprenons $f : \mathbb{R}^2 \to \mathbb{R}^2$, $f(x, y) = (x + y,\ x - y)$. Soit $(x, y) \in \operatorname{Ker} f$:

$$\begin{cases} x + y = 0 \\ x - y = 0 \end{cases} \iff \begin{cases} x = 0 \\ y = 0. \end{cases}$$

Donc $\operatorname{Ker} f = \{(0, 0)\}$ et f est injective.
En revanche, $D : \mathbb{K}[X] \to \mathbb{K}[X]$, $P \mapsto P'$, ne l'est pas : son noyau est l'ensemble des polynômes constants, qui n'est pas réduit à $\{0\}$.

Proposition 2.42 (Caractérisation de la surjectivité)

Soit $f : E \to F$ une application linéaire et $(e_1, \dots, e_p)$ une famille génératrice de E. Alors :

$$\operatorname{Im} f = \operatorname{Vect}\left(f(e_1), \dots, f(e_p)\right).$$

En particulier, f est surjective si et seulement si la famille $\left(f(e_1), \dots, f(e_p)\right)$ est génératrice de F.

Démonstration

Inclusion directe. Soit $y \in \operatorname{Im} f$: il existe $x \in E$ tel que $y = f(x)$. Comme $(e_1, \dots, e_p)$ est génératrice, il existe des scalaires $\lambda_1, \dots, \lambda_p$ tels que $x = \sum_{i=1}^{p} \lambda_i e_i$. Par linéarité :

$$y = f\left(\sum_{i=1}^{p} \lambda_i e_i\right) = \sum_{i=1}^{p} \lambda_i f(e_i) \in \operatorname{Vect}\left(f(e_1), \dots, f(e_p)\right).$$

Inclusion réciproque. Chaque $f(e_i)$ appartient à $\operatorname{Im} f$, qui est un sous-espace vectoriel de F. Or $\operatorname{Vect}\left(f(e_1), \dots, f(e_p)\right)$ est le plus petit sous-espace vectoriel contenant ces vecteurs : il est donc contenu dans $\operatorname{Im} f$.
Les deux inclusions donnent l'égalité. La dernière assertion en découle : f est surjective si et seulement si $\operatorname{Im} f = F$, c'est-à-dire si et seulement si $\operatorname{Vect}\left(f(e_1), \dots, f(e_p)\right) = F$. □

Exemple 2.43

L'endomorphisme $D : \mathbb{K}[X] \to \mathbb{K}[X]$, $P \mapsto P'$, est surjectif : la famille $(X^k)_{k\in\mathbb{N}}$ est génératrice de $\mathbb{K}[X]$, et

$$D(1) = 0, \qquad D\left(\frac{X^{k+1}}{k+1}\right) = X^k \quad \text{pour tout } k \in \mathbb{N}.$$

L'image de D contient donc tous les X^k, et par conséquent $\operatorname{Im} D = \mathbb{K}[X]$.
Voilà donc un endomorphisme surjectif mais non injectif : nous verrons plus bas que cela est impossible en dimension finie.

2. Applications linéaires en dimension finie

Proposition 2.44 (Détermination par l'image d'une base)

Soit E un espace vectoriel de dimension finie, $\mathcal{B} = (e_1, \dots, e_p)$ une base de E et $(y_1, \dots, y_p)$ une famille quelconque de vecteurs de F. Il existe une unique application linéaire $f : E \to F$ telle que $f(e_j) = y_j$ pour tout $j \in [\![1, p]\!]$.

Définition 2.45 (Rang)

Soit $f : E \to F$ une application linéaire telle que $\operatorname{Im} f$ soit de dimension finie. On appelle *rang* de f l'entier $\operatorname{rg} f = \dim \operatorname{Im} f$.

Théorème 2.46 (Théorème du rang)

Soit E un espace vectoriel de dimension finie et $f : E \to F$ une application linéaire. Alors $\operatorname{Ker} f$ et $\operatorname{Im} f$ sont de dimension finie et :

$$\dim E = \dim \operatorname{Ker} f + \operatorname{rg} f.$$

Démonstration

Notons $n = \dim E$. Comme $\operatorname{Ker} f$ est un sous-espace vectoriel de E, il est de dimension finie ; posons $r = \dim \operatorname{Ker} f$, de sorte que $r \leq n$.
Soit $(e_1, \dots, e_r)$ une base de $\operatorname{Ker} f$. C'est une famille libre de E ; d'après le théorème de la base incomplète, on peut la compléter en une base de E :

$$\mathcal{B} = (e_1, \dots, e_r, e_{r+1}, \dots, e_n).$$

Montrons que $(f(e_{r+1}), \dots, f(e_n))$ est une base de $\operatorname{Im} f$.
Elle est génératrice de $\operatorname{Im} f$. Comme $\mathcal{B}$ est génératrice de E, la proposition page 16 donne $\operatorname{Im} f = \operatorname{Vect}(f(e_1), \dots, f(e_n))$. Or $f(e_i) = 0_F$ pour $i \leq r$, puisque ces vecteurs sont dans $\operatorname{Ker} f$. Il reste donc $\operatorname{Im} f = \operatorname{Vect}(f(e_{r+1}), \dots, f(e_n))$.
Elle est libre. Soient $\lambda_{r+1}, \dots, \lambda_n$ des scalaires tels que $\sum_{i=r+1}^{n} \lambda_i f(e_i) = 0_F$. Par linéarité :

$$f\left(\sum_{i=r+1}^{n} \lambda_i e_i\right) = 0_F,$$

donc $x = \sum_{i=r+1}^{n} \lambda_i e_i$ appartient à $\operatorname{Ker} f$. Il se décompose alors dans la base $(e_1, \dots, e_r)$ de $\operatorname{Ker} f$: il existe des scalaires $\mu_1, \dots, \mu_r$ tels que $x = \sum_{i=1}^{r} \mu_i e_i$. On en déduit :

$$\sum_{i=1}^{r} \mu_i e_i - \sum_{i=r+1}^{n} \lambda_i e_i = 0_E.$$

Comme $\mathcal{B}$ est libre, tous les μ_i et tous les λ_i sont nuls. La famille $(f(e_{r+1}), \dots, f(e_n))$ est donc libre.
C'est ainsi une base de $\operatorname{Im} f$, qui est par conséquent de dimension finie avec $\operatorname{rg} f = n - r$. D'où $\dim E = n = r + (n - r) = \dim \operatorname{Ker} f + \operatorname{rg} f$. □

Corollaire 2.47 (Noyau d'une forme linéaire non nulle)

Soit E un $\mathbb{K}$-espace vectoriel de dimension finie $n \geq 1$ et φ une forme linéaire *non nulle* sur E. Alors $\operatorname{Ker} \varphi$ est un *hyperplan* de E, c'est-à-dire un sous-espace vectoriel de dimension $n - 1$.

Démonstration

L'image $\operatorname{Im} \varphi$ est un sous-espace vectoriel de $\mathbb{K}$, qui est de dimension 1 : elle vaut donc $\{0\}$ ou $\mathbb{K}$. Comme φ n'est pas l'application nulle, il existe $x \in E$ tel que $\varphi(x) \neq 0$: ainsi $\operatorname{Im} \varphi = \mathbb{K}$ et $\operatorname{rg} \varphi = 1$.
Le théorème du rang donne alors $n = \dim \operatorname{Ker} \varphi + \operatorname{rg} \varphi = \dim \operatorname{Ker} \varphi + 1$, d'où $\dim \operatorname{Ker} \varphi = n - 1$.
□

Proposition 2.48 (Caractérisation de la bijectivité en dimension finie)

Soient E et F deux espaces vectoriels de dimension finie *de même dimension*, et $f : E \to F$ une application linéaire. Alors :

$$f \text{ injective} \iff f \text{ surjective} \iff f \text{ bijective.}$$

Démonstration

Posons $n = \dim E = \dim F$. Le théorème du rang s'écrit $n = \dim \operatorname{Ker} f + \operatorname{rg} f$. Or :

- f est injective si et seulement si $\operatorname{Ker} f = \{0_E\}$ (proposition page 16), c'est-à-dire si et seulement si $\dim \operatorname{Ker} f = 0$;
- f est surjective si et seulement si $\operatorname{Im} f = F$. Comme $\operatorname{Im} f$ est un sous-espace vectoriel de F, cela équivaut à $\operatorname{rg} f = \dim F = n$.

Le théorème du rang montre alors que ces deux conditions sont équivalentes : $\dim \operatorname{Ker} f = 0 \iff \operatorname{rg} f = n$. L'équivalence avec la bijectivité est immédiate. □

Remarque

L'hypothèse $\dim E = \dim F < +\infty$ est essentielle. Nous avons vu plus haut que $P \mapsto P'$ est surjectif mais non injectif sur $\mathbb{K}[X]$, qui n'est pas de dimension finie.

Exemple 2.49

Soit $n \in \mathbb{N}$ et $\varphi : \mathbb{K}_n[X] \to \mathbb{K}_n[X]$ définie par $\varphi(P) = P - P'$. Elle est linéaire, et c'est un endomorphisme de $\mathbb{K}_n[X]$ car $\deg(P - P') \leq n$.
Montrons qu'elle est injective. Soit $P \in \operatorname{Ker} \varphi$, c'est-à-dire $P = P'$. Si $P \neq 0$, en notant $d = \deg P$, on aurait $\deg P' = d - 1 \neq d$, ce qui est absurde. Donc $P = 0$ et $\operatorname{Ker} \varphi = \{0\}$.
Comme $\mathbb{K}_n[X]$ est de dimension finie et que φ est un endomorphisme, la proposition précédente assure que φ est bijective. En particulier, pour tout $Q \in \mathbb{K}_n[X]$, l'équation $P - P' = Q$ admet une unique solution dans $\mathbb{K}_n[X]$.

Proposition 2.50 (Isomorphisme et dimension)

Soient E et F deux espaces vectoriels et $f : E \to F$ un isomorphisme. Si E est de dimension finie, alors F aussi et :

$$\dim F = \dim E.$$

Démonstration

Soit $\mathcal{B} = (e_1, \dots, e_n)$ une base de E. Montrons que $(f(e_1), \dots, f(e_n))$ est une base de F.
Elle est génératrice. Comme f est surjective, la proposition page 16 donne $F = \operatorname{Im} f = \operatorname{Vect}(f(e_1), \dots, f(e_n))$.
Elle est libre. Soient $\lambda_1, \dots, \lambda_n$ des scalaires tels que $\sum_{i=1}^n \lambda_i f(e_i) = 0_F$. Par linéarité, $f\left(\sum_{i=1}^n \lambda_i e_i\right) = 0_F$, donc $\sum_{i=1}^n \lambda_i e_i \in \operatorname{Ker} f = \{0_E\}$ puisque f est injective. La liberté de $\mathcal{B}$ donne alors $\lambda_1 = \cdots = \lambda_n = 0$.
Ainsi F admet une base à n éléments : il est de dimension finie et $\dim F = n = \dim E$. □

Remarque

La réciproque est vraie : deux espaces vectoriels de même dimension finie sont toujours isomorphes. En particulier, tout $\mathbb{K}$-espace vectoriel de dimension n est isomorphe à $\mathbb{K}^n$.

Exemple 2.51 (Suites récurrentes linéaires d'ordre 2)

Soient $(a, b) \in \mathbb{K}^2$ et S l'ensemble des suites $u \in \mathbb{K}^{\mathbb{N}}$ vérifiant :

$$\forall n \in \mathbb{N}, \quad u_{n+2} = a u_{n+1} + b u_n.$$

C'est un sous-espace vectoriel de $\mathbb{K}^{\mathbb{N}}$: la suite nulle y appartient, et si $u, v \in S$ et $\lambda \in \mathbb{K}$, la suite $\lambda u + v$ vérifie la même relation. Montrons que $\dim S = 2$.
Considérons l'application :

$$\Psi : \begin{array}{ccc} S & \longrightarrow & \mathbb{K}^2 \\ u & \longmapsto & (u_0, u_1). \end{array}$$

Elle est linéaire, puisque les opérations sur les suites sont définies terme à terme.
Ψ *est injective.* Soit $u \in \operatorname{Ker} \Psi$, c'est-à-dire $u_0 = u_1 = 0$. Montrons par récurrence double que $u_n = 0$ pour tout $n \in \mathbb{N}$. Les deux premiers termes sont nuls ; et si $u_n = u_{n+1} = 0$ pour un certain n, alors $u_{n+2} = a u_{n+1} + b u_n = 0$. Ainsi $u = 0$ et $\operatorname{Ker} \Psi = \{0\}$.
Ψ *est surjective.* Soit $(\alpha, \beta) \in \mathbb{K}^2$. Définissons une suite u par $u_0 = \alpha$, $u_1 = \beta$ et la relation de récurrence ci-dessus : elle est bien définie, appartient à S et vérifie $\Psi(u) = (\alpha, \beta)$.
Donc Ψ est un isomorphisme de S sur $\mathbb{K}^2$, et la proposition précédente donne $\dim S = \dim \mathbb{K}^2 = 2$.
C'est ce résultat qui justifie la méthode de résolution vue en PCSI : il suffit d'exhiber *deux* solutions non colinéaires pour les avoir toutes.

3. Matrices et applications linéaires

Définition 2.52 (Matrice d'une application linéaire)

Soient E et F deux espaces vectoriels de dimension finie, $\mathcal{B} = (e_1, \dots, e_p)$ une base de E et $\mathcal{C} = (f_1, \dots, f_n)$ une base de F. Soit $u \in \mathcal{L}(E, F)$.
On appelle *matrice de u dans les bases $\mathcal{B}$ au départ et $\mathcal{C}$ à l'arrivée* la matrice $\mathrm{Mat}_{\mathcal{C}\leftarrow\mathcal{B}}(u) \in \mathcal{M}_{n,p}(\mathbb{K})$ dont la j-ième colonne est formée des coordonnées de $u(e_j)$ dans la base $\mathcal{C}$.

Remarque (Sur la notation)

La flèche se lit comme celle d'une application : $\mathrm{Mat}_{\mathcal{C}\leftarrow\mathcal{B}}(u)$ transforme des coordonnées dans $\mathcal{B}$ en coordonnées dans $\mathcal{C}$. L'intérêt est que les bases se simplifient alors deux à deux dans tous les produits, exactement comme dans une composée.
Attention : le programme n'impose aucune notation pour la matrice d'une application linéaire, et celle-ci n'est pas la plus répandue. Vous rencontrerez souvent $\mathrm{Mat}_{\mathcal{B},\mathcal{C}}(u)$, où c'est la base de *départ* qui est écrite en premier. Dans un devoir ou une épreuve, suivez la notation du sujet.

Remarque

Retenez la règle : *les images des vecteurs de la base de départ se lisent en colonnes.* Le nombre de colonnes est donc $\dim E$ et le nombre de lignes $\dim F$.

Remarque

Lorsque u est un endomorphisme de E, on choisit presque toujours la *même* base au départ et à l'arrivée ; on note alors simplement $\mathrm{Mat}_{\mathcal{B}}(u) = \mathrm{Mat}_{\mathcal{B}\leftarrow\mathcal{B}}(u)$, qui est une matrice carrée de taille $\dim E$. C'est le cas de loin le plus fréquent, et c'est le seul que nous considérerons dans la suite de cette sous-section.

Exemple 2.53

Reprenons $f : \mathbb{R}^2 \to \mathbb{R}^2$, $f(x, y) = (x + y,\ x - y)$, et notons $\mathcal{B} = (e_1, e_2)$ la base canonique de $\mathbb{R}^2$. On a :

$$f(e_1) = f(1, 0) = (1, 1) = e_1 + e_2, \qquad f(e_2) = f(0, 1) = (1, -1) = e_1 - e_2.$$

D'où :

$$\mathrm{Mat}_{\mathcal{B}}(f) = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

Exemple 2.54 (Opérateur de dérivation)

Soit $n \in \mathbb{N}^*$ et $D : \mathbb{K}_n[X] \to \mathbb{K}_n[X]$, $P \mapsto P'$. Dans la base canonique $\mathcal{B} = (1, X, \dots, X^n)$, on a $D(1) = 0$ et $D(X^k) = kX^{k-1}$ pour $1 \leq k \leq n$. D'où :

$$\mathrm{Mat}_{\mathcal{B}}(D) = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 2 & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & & & \ddots & n \\ 0 & \cdots & & \cdots & 0 \end{pmatrix}.$$

On lit sur cette matrice que D n'est ni injective (la première colonne est nulle) ni surjective (la dernière ligne est nulle) : le rang de D vaut n, et non $n+1$.

Proposition 2.55 (Compatibilité avec les opérations)

Soient E, F et G de dimension finie, munis respectivement des bases $\mathcal{B}$, $\mathcal{C}$ et $\mathcal{D}$.

1. Si X est la colonne des coordonnées de $x \in E$ dans $\mathcal{B}$, alors $\mathrm{Mat}_{\mathcal{C}\leftarrow\mathcal{B}}(u)X$ est la colonne des coordonnées de $u(x)$ dans $\mathcal{C}$.
2. Pour $u, v \in \mathcal{L}(E,F)$ et $\lambda \in \mathbb{K}$:

   $$\mathrm{Mat}_{\mathcal{C}\leftarrow\mathcal{B}}(\lambda u + v) = \lambda\,\mathrm{Mat}_{\mathcal{C}\leftarrow\mathcal{B}}(u) + \mathrm{Mat}_{\mathcal{C}\leftarrow\mathcal{B}}(v).$$

   Plus précisément, $u \mapsto \mathrm{Mat}_{\mathcal{C}\leftarrow\mathcal{B}}(u)$ est un isomorphisme de $\mathcal{L}(E,F)$ sur $\mathcal{M}_{n,p}(\mathbb{K})$.
3. Pour $u \in \mathcal{L}(E,F)$ et $v \in \mathcal{L}(F,G)$:

   $$\mathrm{Mat}_{\mathcal{D}\leftarrow\mathcal{B}}(v \circ u) = \mathrm{Mat}_{\mathcal{D}\leftarrow\mathcal{C}}(v)\ \mathrm{Mat}_{\mathcal{C}\leftarrow\mathcal{B}}(u).$$

4. Pour $u \in \mathcal{L}(E)$: u est un automorphisme si et seulement si $\mathrm{Mat}_{\mathcal{B}}(u)$ est inversible, et dans ce cas $\mathrm{Mat}_{\mathcal{B}}(u^{-1}) = \mathrm{Mat}_{\mathcal{B}}(u)^{-1}$.

Remarque

Le point 2 donne $\dim \mathcal{L}(E,F) = np = \dim E \times \dim F$. Le point 3 explique enfin la définition, si peu naturelle au premier abord, du produit matriciel : il est *fait* pour traduire la composition.

Exemple 2.56 (Un projecteur lu sur sa matrice)

Soit $\mathcal{B}$ la base canonique de $\mathbb{R}^2$ et p l'endomorphisme de $\mathbb{R}^2$ dont la matrice dans $\mathcal{B}$ est :

$$A = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \qquad \text{c'est-à-dire} \qquad p(x,y) = \left(\frac{x+y}{2}, \frac{x+y}{2}\right).$$

Un calcul immédiat donne :

$$A^2 = \frac{1}{4}\begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix} = A.$$

D'après le point 3 de la proposition, $\mathrm{Mat}_{\mathcal{B}}(p \circ p) = A^2 = A = \mathrm{Mat}_{\mathcal{B}}(p)$, et comme $u \mapsto \mathrm{Mat}_{\mathcal{B}}(u)$ est injective, $p \circ p = p$: p est un *projecteur*.

Ses éléments caractéristiques se lisent directement :

$$\operatorname{Im} p = \operatorname{Vect}((1,1)), \qquad \operatorname{Ker} p = \{(x,y) \in \mathbb{R}^2 \mid x+y=0\} = \operatorname{Vect}((1,-1)).$$

Il s'agit donc de la projection sur $\operatorname{Vect}((1,1))$ parallèlement à $\operatorname{Vect}((1,-1))$.

Définition 2.57 (Matrice de passage)

Soient $\mathcal{B} = (e_1, \dots, e_n)$ et $\mathcal{B}' = (e'_1, \dots, e'_n)$ deux bases de E. On appelle *matrice de passage* de $\mathcal{B}$ à $\mathcal{B}'$ la matrice $P \in \mathcal{M}_n(\mathbb{K})$ dont la j-ième colonne est formée des coordonnées de e'_j dans la base $\mathcal{B}$.

Remarque

Autrement dit, $P = \operatorname{Mat}_{\mathcal{B} \leftarrow \mathcal{B}'}(\mathrm{Id}_E)$: c'est bien une matrice d'application linéaire, mais avec deux bases différentes de part et d'autre. Attention à l'ordre : la *nouvelle* base est écrite en colonnes dans l'*ancienne*.

Proposition 2.58 (Formules de changement de base)

Avec les notations ci-dessus, P est inversible et P^{-1} est la matrice de passage de $\mathcal{B}'$ à $\mathcal{B}$. De plus :

1. si X et X' sont les colonnes des coordonnées d'un même vecteur $x \in E$ dans $\mathcal{B}$ et $\mathcal{B}'$, alors $X = PX'$;
2. pour tout endomorphisme u de E :

$$\operatorname{Mat}_{\mathcal{B}'}(u) = P^{-1} \operatorname{Mat}_{\mathcal{B}}(u)\, P.$$

Exemple 2.59 (Une symétrie)

Soit s la symétrie de $\mathbb{R}^2$ par rapport à $\operatorname{Vect}((1,1))$ parallèlement à $\operatorname{Vect}((1,-1))$, et $\mathcal{B}$ la base canonique. Déterminer directement $s(e_1)$ et $s(e_2)$ demanderait un peu de travail ; il est bien plus économique de commencer par une base adaptée.
Posons $\varepsilon_1 = (1,1)$ et $\varepsilon_2 = (1,-1)$. La famille $\mathcal{B}' = (\varepsilon_1, \varepsilon_2)$ est une base de $\mathbb{R}^2$, et par définition de s :

$$s(\varepsilon_1) = \varepsilon_1, \qquad s(\varepsilon_2) = -\varepsilon_2, \qquad \text{d'où} \qquad \operatorname{Mat}_{\mathcal{B}'}(s) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = D.$$

La matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$ est :

$$P = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

On vérifie que $P^2 = 2I_2$, ce qui donne $P^{-1} = \frac{1}{2}P$ sans aucun calcul d'inverse. La formule de changement de base s'écrit $D = P^{-1}\operatorname{Mat}_{\mathcal{B}}(s)P$, soit $\operatorname{Mat}_{\mathcal{B}}(s) = PDP^{-1}$:

$$\operatorname{Mat}_{\mathcal{B}}(s) = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 0 & 2 \\ 2 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Autrement dit $s(x,y) = (y,x)$: on reconnaît bien la symétrie par rapport à la droite d'équation $y = x$.

Définition 2.60 (Matrices semblables)

Deux matrices A et B de $\mathcal{M}_n(\mathbb{K})$ sont *semblables* lorsqu'il existe $P \in \mathrm{GL}_n(\mathbb{K})$ telle que :

$$B = P^{-1}AP.$$

Remarque

D'après la formule de changement de base, deux matrices sont semblables si et seulement si elles représentent un *même* endomorphisme dans deux bases éventuellement différentes. La similitude est donc la relation qui identifie ce qui, dans une matrice, relève de l'endomorphisme lui-même et non du choix arbitraire d'une base.

4. Déterminant

Théorème 2.61 (Existence et unicité du déterminant)

Soit E un $\mathbb{K}$-espace vectoriel de dimension n et $e = (e_1, \dots, e_n)$ une base de E. Il existe une unique application

$$\det_e : E^n \longrightarrow \mathbb{K}$$

qui soit linéaire par rapport à chacune de ses n variables, *alternée* — c'est-à-dire nulle dès que deux de ses arguments sont égaux — et telle que $\det_e(e_1, \dots, e_n) = 1$.
De plus, toute application $f : E^n \to \mathbb{K}$ linéaire par rapport à chaque variable et alternée est un multiple de $\det_e$.

Remarque

La démonstration de ce théorème est hors programme, de même que la théorie générale des formes multilinéaires. Retenez le principe : le déterminant n'est pas défini par une formule, mais *caractérisé* par trois propriétés. Toutes les règles de calcul que nous verrons en découlent.
Une conséquence immédiate du caractère alterné : échanger deux arguments change le déterminant en son opposé.

Proposition 2.62 (Changement de base)

Si e et e' sont deux bases de E, alors :

$$\det_{e'} = \det_{e'}(e) \times \det_e.$$

Proposition 2.63 (Caractérisation des bases)

Soit e une base de E et $(x_1, \dots, x_n)$ une famille de n vecteurs de E. Alors :

$$(x_1, \dots, x_n) \text{ est une base de } E \iff \det_e(x_1, \dots, x_n) \neq 0.$$

Définition 2.64 (Déterminant d'une matrice, d'un endomorphisme)

— Le *déterminant* d'une matrice $A \in \mathcal{M}_n(\mathbb{K})$ est le déterminant de la famille de ses colonnes dans la base canonique de $\mathcal{M}_{n,1}(\mathbb{K})$. On le note $\det(A)$, ou encore entre barres verticales.

— Le *déterminant* d'un endomorphisme u de E est $\det(u) = \det_e(u(e_1), \dots, u(e_n))$, où e est une base quelconque de E : la proposition de changement de base montre que le résultat ne dépend pas du choix de e.

Remarque

Avec ces définitions, $\det(u) = \det(\mathrm{Mat}_{\mathcal{B}}(u))$ pour toute base $\mathcal{B}$ de E. En particulier, *deux matrices semblables ont le même déterminant* : le déterminant est une propriété de l'endomorphisme, pas de la base choisie pour le représenter.

Déterminants d'ordre 2 et 3

Proposition 2.65

En dimension 2 et 3, le déterminant s'exprime en fonction des coordonnées par :

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc,$$

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = aei + bfg + cdh - ceg - afh - bdi.$$

Remarque

La formule d'ordre 3 se retient par la *règle de Sarrus* : on recopie les deux premières colonnes à droite du tableau, puis on somme les produits des trois diagonales descendantes et on retranche ceux des trois diagonales montantes.
Attention : cette règle est *fausse* à partir de l'ordre 4. Au-delà de 3, on passe par les opérations élémentaires et le développement par rapport à une ligne ou une colonne. Plus encore, il vaut mieux *éviter* d'y recourir. L'étudiant qui calcule tous ses déterminants d'ordre 3 par Sarrus ne pratique jamais les opérations élémentaires ni le développement ; le jour où il rencontre un ordre 4, ou un déterminant avec paramètre, il ne sait rien faire. Prenez l'habitude des méthodes générales dès l'ordre 3 : elles sont d'ailleurs le plus souvent plus rapides.

Remarque

Dans $\mathbb{R}^2$, $|\det_e(x_1, x_2)|$ s'interprète comme l'aire du parallélogramme construit sur x_1 et x_2 ; dans $\mathbb{R}^3$, $|\det_e(x_1, x_2, x_3)|$ est le volume du parallélépipède construit sur x_1, x_2 et x_3. Le signe indique l'orientation. Un déterminant nul traduit donc un aplatissement : les vecteurs sont coplanaires, ou colinéaires.

Exemple 2.66

— $\begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} = 1 \times 4 - 2 \times 3 = -2.$

— $\begin{vmatrix} 2 & -1 \\ 4 & -2 \end{vmatrix} = 2 \times (-2) - (-1) \times 4 = 0$: les deux colonnes sont colinéaires, elles ne

forment pas une base de $\mathbb{K}^2$.

— Par la règle de Sarrus :

$$\begin{vmatrix} 1 & 2 & 3 \\ 0 & 1 & 4 \\ 5 & 6 & 0 \end{vmatrix} = 0 + 40 + 0 - 15 - 24 - 0 = 1.$$

Exemple 2.67 (Reconnaître une base)

La famille $((1,0,5),\ (2,1,6),\ (3,4,0))$ de $\mathbb{R}^3$ est-elle une base ? Ces trois vecteurs sont exactement les colonnes de la matrice précédente, dont le déterminant vaut $1 \neq 0$. C'est donc une base de $\mathbb{R}^3$.

Propriétés de calcul

Proposition 2.68 (Propriétés algébriques)

Soient A et B dans $\mathcal{M}_n(\mathbb{K})$ et $\lambda \in \mathbb{K}$.

1. $\det(I_n) = 1$.
2. $\det(A^\top) = \det(A)$.
3. $\det(\lambda A) = \lambda^n \det(A)$.
4. $\det(AB) = \det(A)\det(B)$.
5. A est inversible si et seulement si $\det(A) \neq 0$, et dans ce cas :

$$\det\left(A^{-1}\right) = \frac{1}{\det(A)}.$$

Les énoncés correspondants valent pour les endomorphismes : $\det(v \circ u) = \det(v)\det(u)$, et u est un automorphisme si et seulement si $\det(u) \neq 0$.

Remarque

Attention : le déterminant n'est *pas* linéaire. En général $\det(A+B) \neq \det(A)+\det(B)$. Ainsi, pour $A = I_2$ et $B = -I_2$:

$$\det(A+B) = \det(0) = 0, \qquad \text{alors que} \qquad \det(A) + \det(B) = 1 + 1 = 2.$$

Le facteur λ^n du point 3 relève du même piège : multiplier la matrice entière par λ, c'est multiplier chacune de ses n colonnes par λ.

Proposition 2.69 (Effet des opérations élémentaires)

Soit $A \in \mathcal{M}_n(\mathbb{K})$. Les opérations sur les lignes (ou les colonnes) agissent ainsi sur le déterminant :

— échanger deux lignes, $L_i \leftrightarrow L_j$: le déterminant est changé en son opposé ;

— multiplier une ligne par λ, $L_i \leftarrow \lambda L_i$: le déterminant est multiplié par λ ;

— ajouter à une ligne une combinaison linéaire des *autres* lignes, par exemple $L_i \leftarrow L_i + \lambda L_j$ avec $j \neq i$: le déterminant est inchangé.

Remarque

C'est la troisième opération qui fait tout le travail : elle permet de faire apparaître des zéros sans rien changer au résultat. Prenez garde à la deuxième, en revanche : contrairement au pivot de Gauss appliqué à un système, diviser une ligne par 2 ici *divise le déterminant par* 2, et il faut en tenir la comptabilité.

Exemple 2.70

Si $A \in \mathcal{M}_3(\mathbb{R})$ vérifie $\det(A) = 5$, alors $\det(2A) = 2^3 \times 5 = 40$, et non 10.

Exemple 2.71

Soit $A \in \mathcal{M}_n(\mathbb{K})$ telle que $A^2 = I_n$. En passant au déterminant :

$$\det(A)^2 = \det\left(A^2\right) = \det(I_n) = 1,$$

donc $\det(A) = 1$ ou $\det(A) = -1$. C'est le cas par exemple de toute matrice de symétrie.

Exemple 2.72

Calculons, à l'aide des opérations élémentaires :

$$\Delta = \begin{vmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 4 & 9 \end{vmatrix}.$$

Les opérations $L_2 \leftarrow L_2 - L_1$ et $L_3 \leftarrow L_3 - L_1$ ne changent pas le déterminant :

$$\Delta = \begin{vmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 3 & 8 \end{vmatrix}.$$

Puis $L_3 \leftarrow L_3 - 3L_2$:

$$\Delta = \begin{vmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 2 \end{vmatrix} = 2,$$

la dernière égalité résultant de la proposition ci-dessous sur les matrices triangulaires.

Matrices triangulaires

Proposition 2.73 (Déterminant d'une matrice triangulaire)

Le déterminant d'une matrice triangulaire (supérieure ou inférieure) est le produit de ses coefficients diagonaux :

$$\begin{vmatrix} a_{1,1} & \star & \cdots & \star \\ 0 & a_{2,2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \star \\ 0 & \cdots & 0 & a_{n,n} \end{vmatrix} = \prod_{i=1}^{n} a_{i,i}.$$

C'est en particulier vrai pour une matrice diagonale.

Remarque

Combinée aux opérations élémentaires, cette proposition fournit la méthode de calcul la plus sûre : on triangule la matrice en n'utilisant que des opérations $L_i \leftarrow L_i + \lambda L_j$, qui ne changent rien, puis on multiplie les coefficients diagonaux.
Une conséquence utile : une matrice triangulaire est inversible si et seulement si tous ses coefficients diagonaux sont non nuls.

Exemple 2.74

- $\begin{vmatrix} 2 & 5 & -1 \\ 0 & 3 & 7 \\ 0 & 0 & 4 \end{vmatrix} = 2 \times 3 \times 4 = 24.$
- $\det(I_n) = 1$ et $\det(\lambda I_n) = \lambda^n$, ce qui est bien cohérent avec la propriété $\det(\lambda A) = \lambda^n \det(A)$.
- La matrice $\begin{pmatrix} 1 & 8 & 3 \\ 0 & 0 & 5 \\ 0 & 0 & 2 \end{pmatrix}$ a un déterminant nul : elle n'est pas inversible.

Développement par rapport à une ligne ou une colonne

Définition 2.75 (Mineur, cofacteur)

Soit $A = (a_{i,j}) \in \mathcal{M}_n(\mathbb{K})$ avec $n \geq 2$. Pour (i, j) fixé, on note $A_{i,j}$ la matrice de $\mathcal{M}_{n-1}(\mathbb{K})$ obtenue en supprimant la i-ième ligne et la j-ième colonne de A.
Le scalaire $\det(A_{i,j})$ est appelé *mineur* d'indice (i, j), et $(-1)^{i+j} \det(A_{i,j})$ le *cofacteur* d'indice (i, j).

Proposition 2.76 (Développement par rapport à une ligne ou une colonne)

Soit $A \in \mathcal{M}_n(\mathbb{K})$ avec $n \geq 2$. Pour toute ligne i fixée :

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{i,j} \det(A_{i,j}),$$

et pour toute colonne j fixée :

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{i,j} \det(A_{i,j}).$$

Remarque

La démonstration n'est pas exigible. Les signes $(-1)^{i+j}$ se retiennent en damier, en commençant par un + en haut à gauche :

$$\begin{pmatrix} + & - & + & \cdots \\ - & + & - & \cdots \\ + & - & + & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Le résultat ne dépend pas de la ligne ou de la colonne choisie : on a donc tout intérêt à retenir celle qui contient le plus de zéros.

Méthode 2.77 (Calculer un déterminant)

1. En taille 2 ou 3, appliquer directement les formules.
2. Sinon, chercher une ligne ou une colonne contenant beaucoup de zéros ; s'il n'y en a pas, en faire apparaître par des opérations $L_i \leftarrow L_i + \lambda L_j$ (ou l'analogue sur les colonnes), qui ne changent pas le déterminant.
3. Développer par rapport à cette ligne ou cette colonne, et recommencer sur les mineurs obtenus.

Exemple 2.78 (Développement suivant une colonne)

Reprenons le déterminant d'ordre 3 calculé plus haut par la règle de Sarrus, en développant cette fois par rapport à la première colonne, qui contient un zéro :

$$\begin{vmatrix} 1 & 2 & 3 \\ 0 & 1 & 4 \\ 5 & 6 & 0 \end{vmatrix} = 1 \times \begin{vmatrix} 1 & 4 \\ 6 & 0 \end{vmatrix} - 0 \times \begin{vmatrix} 2 & 3 \\ 6 & 0 \end{vmatrix} + 5 \times \begin{vmatrix} 2 & 3 \\ 1 & 4 \end{vmatrix}.$$

Soit :

$$= 1 \times (0 - 24) + 5 \times (8 - 3) = -24 + 25 = 1,$$

ce qui confirme le calcul précédent.

Exemple 2.79 (Un déterminant d'ordre 4)

Calculons :

$$\Delta = \begin{vmatrix} 1 & 0 & 2 & 0 \\ 3 & 1 & 0 & 4 \\ 0 & 0 & 5 & 0 \\ 2 & 0 & 1 & 3 \end{vmatrix}.$$

La troisième ligne ne comporte qu'un seul coefficient non nul : développons par rapport à elle. Le cofacteur en position $(3,3)$ a pour signe $(-1)^{3+3} = +1$, d'où :

$$\Delta = 5 \times \begin{vmatrix} 1 & 0 & 0 \\ 3 & 1 & 4 \\ 2 & 0 & 3 \end{vmatrix}.$$

Ce déterminant d'ordre 3 a une deuxième colonne presque nulle : développons par rapport à elle. Le seul coefficient non nul est en position $(2,2)$, de signe $(-1)^{2+2} = +1$:

$$\begin{vmatrix} 1 & 0 & 0 \\ 3 & 1 & 4 \\ 2 & 0 & 3 \end{vmatrix} = 1 \times \begin{vmatrix} 1 & 0 \\ 2 & 3 \end{vmatrix} = 3.$$

Finalement $\Delta = 5 \times 3 = 15$. Notez qu'un développement suivant la première ligne aurait demandé deux déterminants d'ordre 3 au lieu d'un : le choix de la ligne n'est pas neutre.

Exemple 2.80 (Un déterminant avec paramètre)

Soit $x \in \mathbb{K}$ et :

$$\Delta(x) = \begin{vmatrix} x & 1 & 1 \\ 1 & x & 1 \\ 1 & 1 & x \end{vmatrix}.$$

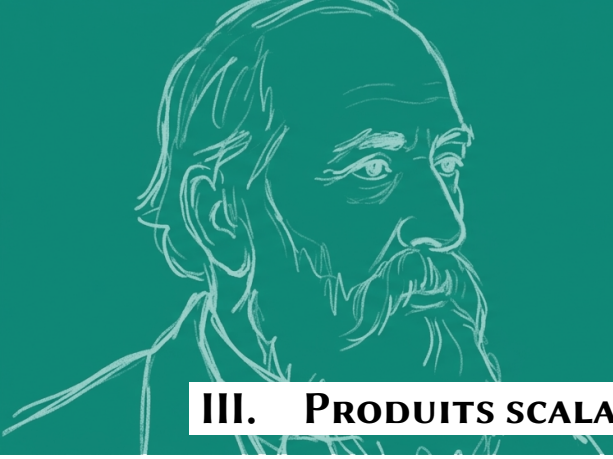
Chaque colonne a la même somme $x + 2$: l'opération $C_1 \leftarrow C_1 + C_2 + C_3$ laisse le déterminant inchangé et fait apparaître une colonne constante, que la linéarité par rapport à la première colonne permet de factoriser :

$$\Delta(x) = \begin{vmatrix} x+2 & 1 & 1 \\ x+2 & x & 1 \\ x+2 & 1 & x \end{vmatrix} = (x+2) \begin{vmatrix} 1 & 1 & 1 \\ 1 & x & 1 \\ 1 & 1 & x \end{vmatrix}.$$

Puis $L_2 \leftarrow L_2 - L_1$ et $L_3 \leftarrow L_3 - L_1$:

$$\Delta(x) = (x+2) \begin{vmatrix} 1 & 1 & 1 \\ 0 & x-1 & 0 \\ 0 & 0 & x-1 \end{vmatrix} = (x+2)(x-1)^2,$$

la matrice obtenue étant triangulaire.
On en déduit que la matrice est inversible si et seulement si $x \neq -2$ et $x \neq 1$.



2 Révisions d'algèbre linéaire

III. Produits scalaires

1. Généralités

Définition 2.81 (Produit scalaire)

Un *produit scalaire* sur un $\mathbb{R}$-espace vectoriel E est une application $\langle \cdot, \cdot \rangle : E \times E \to \mathbb{R}$ vérifiant :

- *Bilinéarité* : linéaire en chaque variable.
- *Symétrie* : $\forall (x, y) \in E^2,\ \langle x, y \rangle = \langle y, x \rangle$.
- *Positivité* : $\forall x \in E,\ \langle x, x \rangle \geq 0$.
- *Définie* : $\langle x, x \rangle = 0 \implies x = 0_E$.

Remarque

La bilinéarité et la symétrie se recoupent : une application symétrique et linéaire par rapport à sa première variable l'est automatiquement par rapport à la seconde. En pratique, on vérifie donc la symétrie d'abord, puis la linéarité à gauche seulement.

Exemple 2.82 (Produit scalaire canonique sur $\mathbb{R}^n$)

Pour $x = (x_1, \ldots, x_n)$ et $y = (y_1, \ldots, y_n)$ dans $\mathbb{R}^n$, on pose :

$$\langle x, y \rangle = \sum_{i=1}^{n} x_i y_i.$$

C'est un produit scalaire sur $\mathbb{R}^n$, appelé *produit scalaire canonique.*

Démonstration

Symétrie. Pour tous $x, y \in \mathbb{R}^n$, $\langle x, y \rangle = \sum_i x_i y_i = \sum_i y_i x_i = \langle y, x \rangle$, car le produit dans $\mathbb{R}$ est commutatif.
Linéarité à gauche. Soient $x, x', y \in \mathbb{R}^n$ et $\lambda \in \mathbb{R}$. La i-ième coordonnée de $\lambda x + x'$ est $\lambda x_i + x'_i$, donc :

$$\langle \lambda x + x', y \rangle = \sum_{i=1}^{n} (\lambda x_i + x'_i) y_i = \lambda \sum_{i=1}^{n} x_i y_i + \sum_{i=1}^{n} x'_i y_i = \lambda \langle x, y \rangle + \langle x', y \rangle .$$

Avec la symétrie, l'application est bien bilinéaire.
Positivité. Pour tout $x \in \mathbb{R}^n$, $\langle x, x \rangle = \sum_{i=1}^{n} x_i^2 \geq 0$ comme somme de carrés.
Caractère défini. Si $\langle x, x \rangle = 0$, alors $\sum_i x_i^2 = 0$. Une somme de réels positifs est nulle si et seulement si tous ses termes le sont, donc $x_i^2 = 0$, c'est-à-dire $x_i = 0$, pour tout i. Ainsi $x = 0$. □

Exemple 2.83 (Produit scalaire intégral sur $C^0([0,1],\mathbb{R})$)

Sur $E = C^0([0,1],\mathbb{R})$, on pose :

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

C'est un produit scalaire sur E.

Démonstration

L'intégrale est bien définie : si f et g sont continues sur le segment $[0,1]$, leur produit l'est aussi.
Symétrie. $\langle f, g\rangle = \int_0^1 fg = \int_0^1 gf = \langle g, f\rangle$.
Linéarité à gauche. Pour $f, f', g \in E$ et $\lambda \in \mathbb{R}$, la linéarité de l'intégrale donne :

$$\begin{aligned}\langle \lambda f + f', g\rangle &= \int_0^1 (\lambda f(t) + f'(t))\,g(t)\,\mathrm{d}t \\ &= \lambda \int_0^1 f(t)g(t)\,\mathrm{d}t + \int_0^1 f'(t)g(t)\,\mathrm{d}t = \lambda\,\langle f, g\rangle + \langle f', g\rangle\,.\end{aligned}$$

Positivité. Pour $f \in E$, la fonction f^2 est continue et positive sur $[0,1]$, donc $\langle f, f\rangle = \int_0^1 f(t)^2\,\mathrm{d}t \geq 0$ par croissance de l'intégrale.
Caractère défini. C'est le seul point délicat. Supposons $f \neq 0$: il existe $t_0 \in [0,1]$ tel que $f(t_0) \neq 0$. La fonction f^2 est continue en t_0 et $f(t_0)^2 > 0$; en appliquant la définition de la continuité avec $\varepsilon = \frac{1}{2}f(t_0)^2 > 0$, on obtient un $\eta > 0$ tel que :

$$\forall t \in [0,1]\,, \quad |t - t_0| \leq \eta \implies f(t)^2 \geq \tfrac{1}{2}f(t_0)^2.$$

Posons $[\alpha, \beta] = [0,1] \cap [t_0 - \eta, t_0 + \eta]$: c'est un segment de longueur $\ell > 0$, car il contient t_0. Comme f^2 est positive sur $[0,1]$:

$$\int_0^1 f(t)^2\,\mathrm{d}t \geq \int_\alpha^\beta f(t)^2\,\mathrm{d}t \geq \ell \cdot \tfrac{1}{2}f(t_0)^2 > 0.$$

Donc $\langle f, f\rangle \neq 0$. Par contraposée, $\langle f, f\rangle = 0$ entraîne $f = 0$. □

Remarque

Le même calcul vaut sur $C^0([a,b],\mathbb{R})$ pour n'importe quel segment avec $a < b$. En revanche, l'hypothèse de continuité est essentielle pour le caractère défini : la fonction nulle partout sauf en $t = 0$ est d'intégrale nulle sans être nulle.

Exemple 2.84 (Produit scalaire intégral sur les polynômes)

Sur $E = \mathbb{R}[X]$, on pose :

$$\langle P, Q\rangle = \int_0^1 P(t)\,Q(t)\,\mathrm{d}t.$$

C'est un produit scalaire sur $\mathbb{R}[X]$, et donc aussi sur chaque $\mathbb{R}_n[X]$ par restriction.

Démonstration

La symétrie, la bilinéarité et la positivité se démontrent exactement comme dans l'exemple précédent : seules interviennent la commutativité du produit, la linéarité et la croissance de

l'intégrale.
Pour le caractère défini, on ne peut pas invoquer directement l'exemple précédent : rien ne dit *a priori* que le polynôme P est nul lorsque la fonction $t \mapsto P(t)$ est nulle sur $[0,1]$. Supposons donc $\langle P, P\rangle = 0$. La fonction polynomiale associée à P est continue sur $[0,1]$, donc, par l'exemple précédent, elle est identiquement nulle sur $[0,1]$. Le polynôme P admet ainsi une infinité de racines ; or un polynôme non nul de $\mathbb{R}[X]$ a au plus $\deg P$ racines. Donc $P = 0$.
□

Remarque

La même démonstration fonctionne avec un *poids* : si w est continue et strictement positive sur $[0,1]$, alors $\langle P, Q\rangle = \int_0^1 w(t)P(t)Q(t)\,\mathrm{d}t$ est encore un produit scalaire.

Définition 2.85 (Espace euclidien)

Un *espace euclidien* est un $\mathbb{R}$-espace vectoriel de dimension finie muni d'un produit scalaire.

Exemple 2.86

- $\mathbb{R}^n$ muni du produit scalaire canonique est un espace euclidien de dimension n.
- $\mathbb{R}_n[X]$ muni de $\langle P, Q\rangle = \int_0^1 P(t)Q(t)\,\mathrm{d}t$ est un espace euclidien de dimension $n+1$.

En revanche, $\mathbb{R}[X]$ et $C^0([0,1],\mathbb{R})$, munis du même produit scalaire intégral, ne sont pas euclidiens : ils sont de dimension infinie.

Définition 2.87 (Norme associée)

La norme associée au produit scalaire est :

$$\|x\| = \sqrt{\langle x, x\rangle}.$$

Théorème 2.88 (Inégalité de Cauchy-Schwarz)

Soit E un $\mathbb{R}$-espace vectoriel muni d'un produit scalaire, de dimension finie ou non. Pour tous $x, y \in E$:

$$|\langle x, y\rangle| \leq \|x\| \cdot \|y\|,$$

avec égalité si et seulement si x et y sont colinéaires.

Démonstration

Si $y = 0$, les deux membres sont nuls et x et y sont colinéaires : l'énoncé est vrai. On suppose donc $y \neq 0$, c'est-à-dire $\|y\| > 0$.
Pour $t \in \mathbb{R}$, développons $\|x + ty\|^2$ par bilinéarité et symétrie :

$$\varphi(t) = \|x + ty\|^2 = \langle x + ty, x + ty\rangle = \|y\|^2\, t^2 + 2\,\langle x, y\rangle\, t + \|x\|^2.$$

Par positivité du produit scalaire, $\varphi(t) \geq 0$ pour tout $t \in \mathbb{R}$. Comme $\|y\|^2 \neq 0$, φ est un trinôme du second degré de signe constant sur $\mathbb{R}$: son discriminant est négatif ou nul, soit

$$4\,\langle x, y\rangle^2 - 4\,\|x\|^2\,\|y\|^2 \leq 0.$$

D'où $\langle x, y\rangle^2 \leq \|x\|^2\,\|y\|^2$, et en prenant la racine carrée $|\langle x, y\rangle| \leq \|x\|\,\|y\|$.

Cas d'égalité. Supposons l'égalité. Le discriminant de φ est alors nul, donc φ admet une racine (double) $t_0 \in \mathbb{R}$. Ainsi $\|x + t_0 y\|^2 = 0$, et le caractère défini du produit scalaire donne $x = -t_0 y$: les vecteurs x et y sont colinéaires.
Réciproquement, si x et y sont colinéaires avec $y \neq 0$, écrivons $x = \lambda y$. Alors $|\langle x, y\rangle| = |\lambda| \|y\|^2 = \|x\| \|y\|$. □

2. Bases orthonormées

Dans toute cette sous-section, E désigne un $\mathbb{R}$-espace vectoriel muni d'un produit scalaire.

Définition 2.89 (Orthogonalité)

Deux vecteurs x et y de E sont *orthogonaux*, ce que l'on note $x \perp y$, lorsque $\langle x, y\rangle = 0$.
Une famille $(x_1, \ldots, x_p)$ de vecteurs de E est :

- *orthogonale* si $\langle x_i, x_j\rangle = 0$ dès que $i \neq j$;
- *orthonormée* (ou *orthonormale*) si elle est orthogonale et si de plus $\|x_i\| = 1$ pour tout i.

Autrement dit, la famille est orthonormée lorsque $\langle x_i, x_j\rangle = \delta_{ij}$ pour tous i, j.

Théorème 2.90 (Théorème de Pythagore)

Soient $x, y \in E$. Alors :

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

Plus généralement, si $(x_1, \ldots, x_p)$ est une famille orthogonale :

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

Démonstration

En développant par bilinéarité et symétrie :

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

L'égalité annoncée équivaut donc à $2\langle x, y\rangle = 0$, c'est-à-dire à $x \perp y$.
Pour le cas général, on développe de même :

$$\left\| \sum_{i=1}^{p} x_i \right\|^2 = \sum_{i=1}^{p} \sum_{j=1}^{p} \langle x_i, x_j\rangle = \sum_{i=1}^{p} \|x_i\|^2 ,$$

puisque tous les termes croisés $\langle x_i, x_j\rangle$ avec $i \neq j$ sont nuls. □

Remarque

Attention : contrairement au cas de deux vecteurs, la réciproque est fausse pour $p \geq 3$. L'égalité $\|x_1 + x_2 + x_3\|^2 = \sum \|x_i\|^2$ signifie seulement que $\sum_{i<j} \langle x_i, x_j\rangle = 0$, ce qui n'oblige pas chaque terme à s'annuler.

Proposition 2.91 (Une famille orthogonale de vecteurs non nuls est libre)

Soit $(x_1, \dots, x_p)$ une famille orthogonale de vecteurs *tous non nuls* de E. Alors cette famille est libre. En particulier, toute famille orthonormée est libre.

Démonstration

Soient $\lambda_1, \dots, \lambda_p \in \mathbb{R}$ tels que $\sum_{i=1}^{p} \lambda_i x_i = 0$. Fixons $j \in \{1, \dots, p\}$ et prenons le produit scalaire de cette relation avec x_j :

$$0 = \left\langle \sum_{i=1}^{p} \lambda_i x_i,\ x_j \right\rangle = \sum_{i=1}^{p} \lambda_i \langle x_i,\ x_j \rangle = \lambda_j \left\| x_j \right\|^2,$$

car tous les autres produits scalaires sont nuls par orthogonalité. Comme $x_j \neq 0$, on a $\left\| x_j \right\|^2 \neq 0$, d'où $\lambda_j = 0$. Ceci valant pour tout j, la famille est libre. □

Théorème 2.92 (Algorithme d'orthonormalisation de Gram-Schmidt)

Soit $(e_1, \dots, e_p)$ une famille libre de E. Il existe une unique famille orthonormée $(\varepsilon_1, \dots, \varepsilon_p)$ telle que :

$$\forall k \in \{1, \dots, p\}, \quad \mathrm{Vect}(\varepsilon_1, \dots, \varepsilon_k) = \mathrm{Vect}(e_1, \dots, e_k) \quad \text{et} \quad \langle e_k,\ \varepsilon_k \rangle > 0.$$

Elle s'obtient par récurrence en posant $\varepsilon_1 = \dfrac{e_1}{\|e_1\|}$ puis, pour $k \geq 2$:

$$f_k = e_k - \sum_{i=1}^{k-1} \langle e_k,\ \varepsilon_i \rangle\ \varepsilon_i, \qquad \varepsilon_k = \frac{f_k}{\|f_k\|}.$$

Remarque

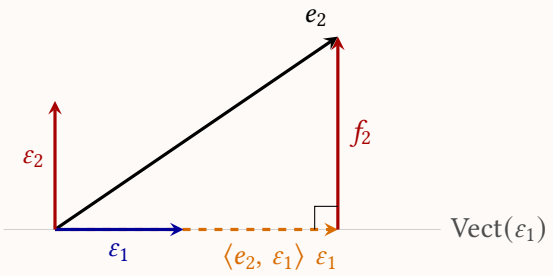
Il n'y a pas lieu d'apprendre ces formules par cœur : on refait le calcul à chaque fois. À l'étape k, on retranche à e_k ses composantes selon les vecteurs déjà construits, ce qui rend le reste orthogonal à $\mathrm{Vect}(e_1, \dots, e_{k-1})$, puis on normalise.

Exemple 2.93 (Orthonormalisation de $(1, X, X^2)$)

Munissons $\mathbb{R}_2[X]$ du produit scalaire $\langle P,\ Q \rangle = \int_0^1 P(t)Q(t)\,\mathrm{d}t$ et orthonormalisons la base canonique $(1, X, X^2)$.

Premier vecteur. Comme $\|1\|^2 = \int_0^1 \mathrm{d}t = 1$, le polynôme constant 1 est déjà unitaire : on prend $\varepsilon_1 = 1$.

Deuxième vecteur. Il faut retirer à $e_2 = X$ sa composante selon ε_1, de sorte que le reste soit orthogonal à ε_1 :



Ici $\langle X, \varepsilon_1\rangle = \int_0^1 t\,\mathrm{d}t = \frac{1}{2}$, qui n'est autre que la valeur moyenne de t sur $[0,1]$. On pose donc $f_2 = X - \frac{1}{2}$: retrancher sa moyenne à un polynôme, c'est exactement le rendre orthogonal aux constantes, puisque $\langle P, 1\rangle = \int_0^1 P$. Il ne reste qu'à normaliser :

$$\|f_2\|^2 = \int_0^1 \left(t - \tfrac{1}{2}\right)^2 \mathrm{d}t = \left[\tfrac{1}{3}\left(t - \tfrac{1}{2}\right)^3\right]_0^1 = \frac{1}{12},$$

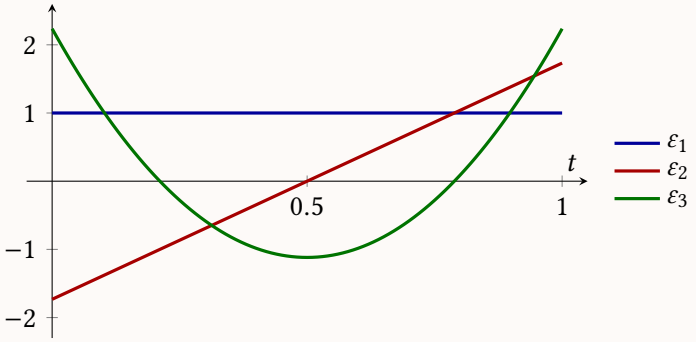
d'où $\varepsilon_2 = 2\sqrt{3}\left(X - \frac{1}{2}\right) = \sqrt{3}\,(2X - 1)$.
Troisième vecteur. Même principe, avec cette fois deux composantes à retirer. On a $\left\langle X^2, \varepsilon_1\right\rangle = \int_0^1 t^2\,\mathrm{d}t = \frac{1}{3}$ et

$$\left\langle X^2, \varepsilon_2\right\rangle = \sqrt{3}\int_0^1 t^2(2t-1)\,\mathrm{d}t = \sqrt{3}\left(\tfrac{1}{2} - \tfrac{1}{3}\right) = \frac{\sqrt{3}}{6},$$

donc

$$f_3 = X^2 - \frac{1}{3} - \frac{\sqrt{3}}{6}\sqrt{3}\,(2X-1) = X^2 - X + \frac{1}{6}.$$

Un calcul analogue donne $\|f_3\|^2 = \frac{1}{180}$, et finalement $\varepsilon_3 = 6\sqrt{5}\left(X^2 - X + \frac{1}{6}\right) = \sqrt{5}\,(6X^2 - 6X + 1)$.
La famille $\left(1,\ \sqrt{3}(2X-1),\ \sqrt{5}(6X^2-6X+1)\right)$ est donc une base orthonormée de $\mathbb{R}_2[X]$:



Les deux derniers polynômes sont d'intégrale nulle sur $[0,1]$: les aires situées au-dessus et au-dessous de l'axe se compensent, ce qui traduit précisément leur orthogonalité à ε_1.

Théorème 2.94 (Existence de bases orthonormées)

Tout espace euclidien E non réduit à $\{0\}$ possède une base orthonormée.
De plus (*théorème de la base orthonormée incomplète*), toute famille orthonormée de E peut être complétée en une base orthonormée de E.

Démonstration

Pour le premier point, il suffit d'appliquer l'algorithme de Gram-Schmidt à une base quelconque $(e_1, \ldots, e_n)$ de E : on obtient une famille orthonormée $(\varepsilon_1, \ldots, \varepsilon_n)$, libre et de cardinal $n = \dim E$, donc une base.
Pour le second, soit $(\varepsilon_1, \ldots, \varepsilon_k)$ une famille orthonormée. Elle est libre, donc le théorème de la base incomplète permet de la compléter en une base $(\varepsilon_1, \ldots, \varepsilon_k, e_{k+1}, \ldots, e_n)$ de E. En appliquant Gram-Schmidt à cette base, les k premiers vecteurs sont inchangés : en effet, si $(\varepsilon_1, \ldots, \varepsilon_j)$ est

déjà orthonormée, alors $f_j = \varepsilon_j$ à l'étape j et $\|f_j\| = 1$. On obtient donc une base orthonormée dont les k premiers vecteurs sont ceux de départ. □

Proposition 2.95 (Calculs dans une base orthonormée)

Soit $\mathcal{B} = (e_1, \ldots, e_n)$ une base orthonormée de E. Alors pour tous $x, y \in E$:

1. $x = \sum_{i=1}^{n} \langle x,\, e_i\rangle\, e_i$: les coordonnées de x dans $\mathcal{B}$ sont les $\langle x,\, e_i\rangle$;
2. $\langle x,\, y\rangle = \sum_{i=1}^{n} \langle x,\, e_i\rangle \langle y,\, e_i\rangle = X^\top Y$, où X et Y sont les colonnes de coordonnées ;
3. $\|x\|^2 = \sum_{i=1}^{n} \langle x,\, e_i\rangle^2$.

Démonstration

Écrivons $x = \sum_{i=1}^{n} x_i e_i$ dans la base $\mathcal{B}$. Pour tout j :

$$\langle x,\, e_j\rangle = \sum_{i=1}^{n} x_i \langle e_i,\, e_j\rangle = x_j,$$

ce qui établit le premier point. Avec $y = \sum_j y_j e_j$, la bilinéarité donne alors :

$$\langle x,\, y\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} x_i y_j \langle e_i,\, e_j\rangle = \sum_{i=1}^{n} x_i y_i,$$

d'où le deuxième point, et le troisième en prenant $y = x$. □

Remarque

Ces formules expliquent l'intérêt des bases orthonormées : dans une telle base, tout produit scalaire se calcule comme le produit scalaire canonique de $\mathbb{R}^n$, et une coordonnée s'obtient par un simple produit scalaire, sans résoudre de système.

3. Projections orthogonales

Définition 2.96 (Orthogonal d'une partie)

Soit A une partie de E. L'*orthogonal* de A est :

$$A^\perp = \{x \in E \mid \forall a \in A,\ \langle x,\, a\rangle = 0\}.$$

Proposition 2.97

Pour toute partie A de E, $A^\perp$ est un sous-espace vectoriel de E, et $A^\perp = (\operatorname{Vect} A)^\perp$.

Démonstration

Le vecteur nul appartient à $A^\perp$. Si $x, x' \in A^\perp$ et $\lambda \in \mathbb{R}$, alors pour tout $a \in A$: $\langle \lambda x + x',\, a\rangle = \lambda \langle x,\, a\rangle + \langle x',\, a\rangle = 0$, donc $\lambda x + x' \in A^\perp$. C'est bien un sous-espace vectoriel.
L'inclusion $(\operatorname{Vect} A)^\perp \subset A^\perp$ est immédiate puisque $A \subset \operatorname{Vect} A$. Réciproquement, si $x \in A^\perp$, tout élément de $\operatorname{Vect} A$ s'écrit $\sum_k \lambda_k a_k$ avec $a_k \in A$, et alors $\langle x,\, \sum_k \lambda_k a_k\rangle = \sum_k \lambda_k \langle x,\, a_k\rangle = 0$.
□

Remarque

En pratique, pour déterminer $F^\perp$ où $F = \mathrm{Vect}(a_1, \ldots, a_p)$, il suffit d'écrire les p équations $\langle x,\, a_k \rangle = 0$: c'est un système linéaire homogène.

Théorème 2.98 (Supplémentaire orthogonal)

Soit F un sous-espace vectoriel de E *de dimension finie*. Alors :

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

Si de plus E est euclidien, $\dim F^\perp = \dim E - \dim F$ et $(F^\perp)^\perp = F$.

Démonstration

Somme directe. Si $x \in F \cap F^\perp$, alors x est orthogonal à lui-même, donc $\langle x,\, x \rangle = 0$, et le caractère défini du produit scalaire donne $x = 0$.
Somme égale à E. Comme F est de dimension finie $k \geq 1$ (le cas $F = \{0\}$ étant immédiat), il possède une base orthonormée $(e_1, \ldots, e_k)$. Soit $x \in E$; posons :

$$y = \sum_{i=1}^{k} \langle x,\, e_i \rangle\, e_i \in F, \qquad z = x - y.$$

Pour tout $j \in \{1, \ldots, k\}$:

$$\langle z,\, e_j \rangle = \langle x,\, e_j \rangle - \sum_{i=1}^{k} \langle x,\, e_i \rangle \langle e_i,\, e_j \rangle = \langle x,\, e_j \rangle - \langle x,\, e_j \rangle = 0.$$

Ainsi z est orthogonal à tous les vecteurs de la base de F, donc à F tout entier : $z \in F^\perp$. On a écrit $x = y + z$ avec $y \in F$ et $z \in F^\perp$.
Dimension. Si E est euclidien, la formule des dimensions d'une somme directe donne $\dim F + \dim F^\perp = \dim E$. Enfin $F \subset (F^\perp)^\perp$ de manière évidente, et ces deux espaces ont la même dimension $\dim E - \dim F^\perp$, donc ils sont égaux. □

Remarque

L'hypothèse « F de dimension finie » est essentielle et ne peut pas être remplacée par « E de dimension finie » : dans un espace préhilbertien de dimension infinie, un sous-espace de dimension infinie n'admet pas nécessairement de supplémentaire orthogonal.

Remarque (Hyperplans)

Soit H un hyperplan d'un espace euclidien E, c'est-à-dire un sous-espace de dimension $n - 1$. Alors $\dim H^\perp = 1$: il existe un vecteur $a \neq 0$, unique à un facteur près, tel que $H^\perp = \mathrm{Vect}(a)$. On dit que a est un *vecteur normal* à H, et :

$$H = \{x \in E \mid \langle x,\, a \rangle = 0\}.$$

Définition 2.99 (Projection orthogonale)

Soit F un sous-espace vectoriel de dimension finie de E. La *projection orthogonale sur* F, notée p_F, est la projection sur F parallèlement à $F^\perp$ dans la décomposition $E = F \oplus F^\perp$.

Autrement dit, $p_F(x)$ est l'unique vecteur de F tel que $x - p_F(x) \in F^\perp$.

Proposition 2.100 (Expression dans une base orthonormée)

Si $(e_1, \dots, e_k)$ est une base orthonormée de F, alors :

$$\forall x \in E, \quad p_F(x) = \sum_{i=1}^{k} \langle x,\, e_i \rangle\, e_i.$$

Démonstration

C'est exactement le vecteur y construit dans la démonstration du théorème précédent : il appartient à F et $x - y \in F^\perp$. Par unicité de la décomposition, $y = p_F(x)$. □

Remarque

La formule exige une base *orthonormée* de F. Avec une base quelconque, elle est fausse : il faut d'abord orthonormaliser par Gram-Schmidt, ou bien résoudre le système donné par les conditions $\langle x - p_F(x),\, a_i \rangle = 0$.

Théorème 2.101 (Distance à un sous-espace)

Soit F un sous-espace de dimension finie de E et $x \in E$. Alors, pour tout $y \in F$:

$$\|x - y\| \geq \|x - p_F(x)\|,$$

avec égalité si et seulement si $y = p_F(x)$. Autrement dit, $p_F(x)$ est l'unique élément de F le plus proche de x, et :

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

Démonstration

Soit $y \in F$. Décomposons :

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

Ces deux vecteurs sont orthogonaux, donc le théorème de Pythagore donne :

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

L'égalité a lieu si et seulement si $\|p_F(x) - y\|^2 = 0$, c'est-à-dire $y = p_F(x)$. □

Exemple 2.102 (Projection sur une droite)

Soit $a \in E$ non nul et $D = \mathrm{Vect}(a)$. La famille $(a/\|a\|)$ est une base orthonormée de D, donc :

$$p_D(x) = \frac{\langle x,\, a \rangle}{\|a\|^2}\, a, \qquad d(x, D)^2 = \|x\|^2 - \frac{\langle x,\, a \rangle^2}{\|a\|^2}.$$

On retrouve au passage l'inégalité de Cauchy-Schwarz : le membre de droite étant positif, $\langle x,\, a \rangle^2 \leq \|x\|^2\, \|a\|^2$.

Exemple 2.103 (Distance de X^2 aux polynômes de degré ≤ 1)

Munissons $\mathbb{R}_2[X]$ du produit scalaire $\langle P, Q\rangle = \int_0^1 PQ$ et posons $F = \mathbb{R}_1[X]$. On a vu que $\left(1, \sqrt{3}(2X-1)\right)$ est une base orthonormée de F. Avec $\left\langle X^2, 1\right\rangle = \frac{1}{3}$ et $\left\langle X^2, \sqrt{3}(2X-1)\right\rangle = \frac{\sqrt{3}}{6}$:

$$p_F(X^2) = \frac{1}{3} + \frac{\sqrt{3}}{6}\sqrt{3}\,(2X-1) = X - \frac{1}{6}.$$

Le polynôme de degré ≤ 1 qui approche le mieux X^2 au sens de cette norme est donc $X - \frac{1}{6}$, et :

$$d(X^2, F) = \left\|X^2 - X + \tfrac{1}{6}\right\| = \sqrt{\int_0^1 \left(t^2 - t + \tfrac{1}{6}\right)^2 \mathrm{d}t} = \frac{1}{6\sqrt{5}}.$$

Exemple 2.104 (Projection sur un plan de $\mathbb{R}^3$)

Dans $\mathbb{R}^3$ muni du produit scalaire canonique, soit $P = \{(x, y, z) \mid x + y + z = 0\}$. C'est un hyperplan de vecteur normal $a = (1, 1, 1)$, donc $P^\perp = \mathrm{Vect}(a)$ et, pour tout $u \in \mathbb{R}^3$:

$$p_P(u) = u - p_{P^\perp}(u) = u - \frac{\langle u, a\rangle}{3}\, a.$$

Par exemple, $p_P(1,0,0) = (1,0,0) - \frac{1}{3}(1,1,1) = \left(\frac{2}{3}, -\frac{1}{3}, -\frac{1}{3}\right)$, et la distance de $(1,0,0)$ au plan P vaut $\frac{|1|}{\|a\|} = \frac{1}{\sqrt{3}}$.